

Chap 1: A single currency for Europe

Key issue: We review the genesis, political motivations and implications for the international monetary system.

- C. Wyplosz, "EMU: Why and How It Might Happen", *The Journal of Economic Perspectives* Vol. 11, No. 4 (Autumn, 1997), pp. 3-21
- B. Eichengreen and J.A. Frieden, "An analytical Introduction", in *The Political Economy of European Monetary Unification*, Oxford : Westview Press, 2001
- F. Van Esch (2012) Why Germany Wanted EMU: The Role of Helmut Kohl's Belief System and the Fall of the Berlin Wall, *German Politics*, 21:1, 34-52

EMU: Why and How It Might Happen

Charles Wyplosz

The adoption of a single currency has long been a holy grail for Europe. Since the late 1950s, various plans had been devised and shelved, as Mundell (1993) describes in a brief and insightful history. But in a few sharp steps between 1988 and 1991, bewildered Europeans saw their governments agree to what is now known as the Maastricht Treaty.

The story begins auspiciously in 1986. The European Community emerges from a decade-long period of little institutional progress, high inflation and rising unemployment following the oil shocks. This is the year when three new countries (Greece, Spain and Portugal) join the European Community and when the Single European Act (frequently dubbed “1992,” the year when it came into effect) is adopted as an extension of the founding Treaty of Rome. The aim of the Single Act is to plug the loopholes which limited the full mobility of people, goods and capital within Europe. In the process, all restrictions to capital movements were eliminated.¹

This last innocuous-seeming step made a move to monetary union unavoidable. The reason is a straightforward implication of the Mundell-Fleming textbook model of an open economy, known in Europe as the “impossible trilogy” principle.² This

¹ Oddly, the implementation date for this part of the act was set on July 1, 1990, a year and a half ahead of the other provisions. Recent European Community members—Greece, Ireland, Portugal and Spain — were given grace periods.

² The implications for Europe of this general principle, also known in Europe as the inconsistent trinity, were first articulated by Padoa-Schioppa (1985). For a textbook presentation of the Mundell-Fleming model, see, for example, Burda and Wyplosz (1997).

■ *Charles Wyplosz is Professor of Economics, Graduate Institute of International Studies, Geneva, Switzerland, and Research Fellow, Center for Economic Policy Research, London, United Kingdom. His e-mail address is <wyplosz@hei.unige.ch>.*

principle asserts that only two of the three following features are mutually compatible: full capital mobility, independence of monetary policy, and a fixed exchange rate. The problem arises because, under full capital mobility, a nation's domestic interest rate is tied to the world interest rate (at least for a country too small to influence worldwide financial conditions). More precisely, any difference between the domestic and world interest rate is equal to the expected rate of depreciation of the exchange rate; that is, if interest rates are 5 percent in the domestic market and 3 percent in global markets, this must reflect that global currency markets expect the currency to depreciate by 2 percent this year. This is known as the interest parity condition: it implies that integrated financial markets equalize expected asset returns, and so assets denominated in a currency expected to depreciate must offer an exactly compensating higher yield.

A country that wants to conduct an independent monetary policy, raising or lowering interest rates for the purpose of its domestic economy, must allow its exchange rate to fluctuate in the market. Conversely, a country confronted with full capital mobility that wants to fix its exchange rate must set its domestic interest rate to be exactly equal to the rate in the country to which it pegs its currency; since monetary policy is now determined abroad, the country has effectively lost monetary policy independence.³ The alternative option of letting exchange rates float was never acceptable to Europeans. The perception is that markets are too integrated to allow for sizable relative price changes. The exchange rate and trade wars from before World War II are still remembered as an example of a jack that must absolutely be kept in the box.

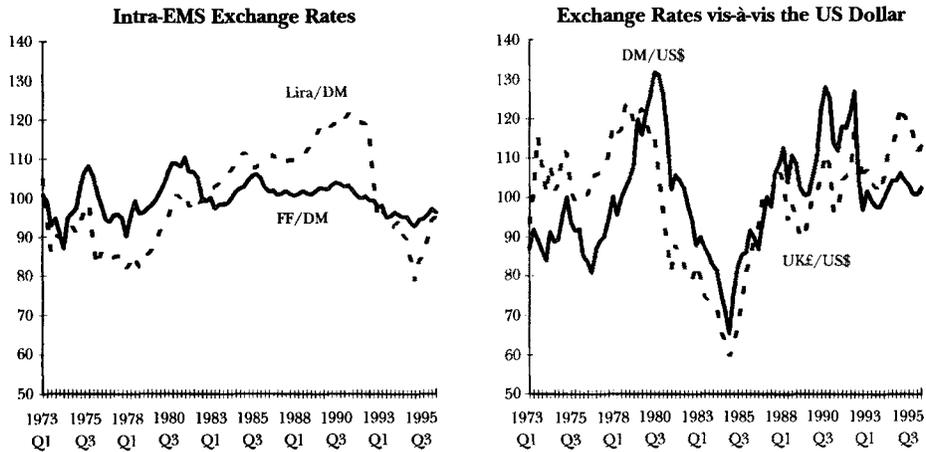
By the time it was decided to free capital flows, the European Monetary System (EMS) had been in place for nearly ten years. Most European Community members had agreed in early 1979 to set up a system of fixed bilateral exchange rates with fluctuation bands of ± 2.25 percent around the declared central parity (± 6 percent for Italy and, briefly, the United Kingdom). Member central banks were committed to intervene jointly to defend the parities, in principle with no limit. When it was felt that existing parities had to be changed, the decision had to be taken by consensus. By the late 1980s, the EMS was commonly hailed as a major success, credited with the relative stability of intra-European real exchange rates during the turbulent post-Bretton Woods period (Begg and Wyplosz, 1993). This is illustrated in

³ In algebraic terms, the interest parity condition, where i is the domestic interest rate, i^* is the global rate, e is the expected rate of depreciation of the exchange rate (in logs), and t is an index of time periods, is:

$$i_t - i_t^* = E_t(e_{t+1}) - e_t$$

A small country which pegs its exchange rate $E_t(e_{t+1}) = e_t = e_0$, where e_0 is the peg, can no longer choose the level of its own interest rate. Only by letting the exchange rate fluctuate can it control the interest rate, and then e_t becomes endogenous in the interest parity equation. This reasoning ignores risk aversion which gives rise to a risk premium term. Among developed countries the risk premium is known to be small and volatile.

Figure 1
Bilateral Exchange Rates



Source: International Financial Statistics

Notes: 100 = average over sample period. EMS started in March 1979. Italian Lira operated under a margin of fluctuation of $\pm 6\%$ until January 1990 and left EMS in September 1992. Sterling was in the system from October 1990 to September 1992, with a $\pm 6\%$ margin of fluctuation.

Figure 1 which presents bilateral exchange rates deflated by consumer price indices (the following conclusions emerge irrespective of the choice of country pairs and price indices). Contrast the left panel which shows intra-European real exchange rates with the right panel which shows real exchange rates vis-à-vis the U.S. dollar. Currencies with normal allowed fluctuations, like the French franc and the deutsche mark, displayed remarkably low volatility in comparison to floating exchange rates. Even where larger margins were allowed, like in Italy, quarter-to-quarter real exchange rates are still less volatile than with floating rates. There is no economic reason for the real exchange rate to be constant in the long-term, of course. However, among the OECD countries, multi-year fluctuations around the long-term trend suggest that most of the observed changes are temporary and do not correspond to structural shifts.

Perhaps blinded by the success of the EMS, leading European policymakers did not perceive that the freeing of capital flows meant the end of monetary policy independence in all but one EMS country. By the late 1980s it had become obvious that the Bundesbank, Germany's central bank, was setting monetary policy for Europe as a whole. One reason for this evolution was relative economic size (further increased by unification following the fall of the Berlin Wall in late 1989). In addition, the Bundesbank had acquired a strong reputation for fighting inflation and keeping its currency strong. For countries where inflation was the number one target, adopting tough monetary conditions under the Bundesbank leadership was in fact welcomed. Small countries, like the Netherlands, had already given up mon-

etary independence. Among the larger ones, the United Kingdom was outside the fixed exchange rate mechanism and therefore could retain monetary policy independence.

However, other larger European nations like France, Italy, and Spain, gradually realized that they had lost control of their domestic monetary policy. They concluded that the only way through which they could regain some influence over their monetary policies was to create a broader European monetary institution which would supersede the Bundesbank, and in which they would have a voice. Naturally, since Germany was being asked to sacrifice one of its most valued institutions for the sake of Europe, it was going to ask a lot in return. In particular, Germany was bound to require that this new European monetary institution offer strong guarantees of price stability. From the very beginning, Europe's future currency would have to be as strong as the deutsche mark. This would mean explicit institutional safeguards and exacting startup conditions. The negotiations leading to the Maastricht Treaty would bear the birthmark of this situation: what Germany asks, Germany gets, provided that it gives up the Bundesbank.

The Maastricht Treaty

The Maastricht Treaty updates and incorporates the 1957 Treaty of Rome, the founding act of the European Community, and incorporates the Single European Act implemented in 1992 (free movement of goods, people, and capital). The treaty has been formally ratified by all member countries. With the Maastricht Treaty, Europe ceases to be called the European Economic Community and becomes instead the European Union or EU, which involves both economic and political union. The economic component of the treaty mainly involves the adoption of a single currency. The political component has been left rather vague, hinting at an evolution towards joint defense and foreign affairs.

The treaty includes a detailed timetable for the adoption of a single currency. It sets in motion a gradual convergence process, espousing the view that the adoption of a common currency is just the cherry on the sundae, the last step in a process through which national currencies become indistinguishable from the deutsche mark. It is formally structured around three stages (Thygesen, 1993). The first stage began in 1992 with the formal ratification of the treaty. During the second stage, started in January 1994, national central banks must be given formal independence and cease to grant direct loans to their nation's treasuries. The shift to the second stage also coincides with the establishment of the European Monetary Institute (EMI), with two main functions. One is to prepare the creation of the European Central Bank, whose statutes and mission are actually laid out in the Maastricht Treaty. The other function of the EMI is to oversee the "convergence criteria" which will be used to decide which countries are ready to enter the monetary union, marking the beginning of Stage III. This may happen as soon as a sufficient number of countries meet the convergence criteria, and must happen by January 1, 1999.

The first formal review which took place in December 1996 concluded that a majority of countries did not satisfy the criteria.

What are these criteria? The underlying notion is that unless countries enter the single currency with similar inflation rates and fiscal positions, the single currency will be unsustainable. Three conditions deal with monetary convergence. First, the inflation rate of any country joining the single currency must be within 1.5 percentage points of the average of the three lowest rates in Europe. Second, the long-term interest rate in a country joining the single currency must not exceed by more than 2 percentage points the interest rates observed in the three countries with the lowest inflation rates, on the grounds that high long-term rates reflect high expected inflation. Third, the exchange rate must have remained within the normal bands of the existing EMS “without severe tensions” for at least two years. Two other criteria concern fiscal policy. They set ceilings on the ratios of debt/GDP (60 percent) and deficit/GDP (3 percent) ratios. At the time of the signing of the Maastricht Treaty in 1991, only Luxembourg—which does not have a currency of its own—could meet the five criteria.

Yet the wording of the treaty leaves some room for flexibility. For example, the 60 percent ceiling can be interpreted as a target if “the ratio is sufficiently diminishing and approaching the reference value at a satisfactory pace” (art. 104c, b). In addition, compliance will be decided by the heads of state upon receiving reports from the European Commission and the EMI, and a recommendation (not a decision) by the Commission, which is notoriously supportive of economic and monetary union.

The Maastricht Treaty had left a number of issues pending. Most of them concern the political side, but some also concern the actual operation of the monetary union. The “excessive deficit procedure” issue has been settled in June 1997. This procedure makes permanent one of the entry convergence criteria, the 3 percent deficit/GDP ceiling. It defines the “exceptional conditions” under which a country may be temporarily allowed to breach the ceiling, and it specifies how noncompliant countries will face first private, and then public reprimands, before being fined. Progress has also been made on symbolic matters: the new currency’s name will be “euro” and the European Central Bank will be established in Frankfurt, Germany.

On all of these issues, the German view has prevailed. The excessive deficit procedure is the one presented by Germany, and initially rejected by a vast majority of countries as excessively restrictive. The name of the currency itself reflects the German rejection of ECU, acronym for European Currency Unit and the name of an ancient French currency, although it is explicitly referred to in the Maastricht Treaty. German influence has not only affected the currency name and location; it is also Germany that insisted on the long transition process and the controversial convergence criteria. Moreover, the statutes and objectives of the European Central Bank remarkably resemble those of the Bundesbank: strong independence from government, responsibility clearly limited to price stability, no explicit involvement in bank supervision, and no lender-of-last-resort function.

Germany will again prevail when it comes to selecting the countries which qualify for membership in the single currency. That decision will be taken by the heads of state in spring 1998, with voting weights determined by country size (a combination of population and GDP). Once chosen as fit, a country must join the Economic and Monetary Union (EMU), even if it does not wish to, with the exception of Denmark and the United Kingdom who made opting out a condition of ratifying the treaty. Thus the Maastricht Treaty envisions a “two-speed” Europe, with a “core” of EMU members and a “periphery” of countries either rejected or opting out. Much of the debate revolves around the initial list of members. Will EMU start as a narrow deutsche mark zone (Germany, France, Belgium, Luxembourg, the Netherlands, Austria, Ireland)? Will the “Club Med” countries (Italy, Spain, Portugal) also join, despite a reputation for tolerating inflation and deficits? Will the Nordic countries (Sweden, Denmark, Finland) want to join? The UK has already let it be known that it will stay out and Greece is not really trying.

Is Europe an Optimal Currency Area?

The decision to adopt a single currency is the outcome of constrained optimization. The constraint is the impossible trilemma: given the freedom of capital flows, the choice is between freely floating exchange rates and monetary union. The assessment is that monetary union dominates a free float. This assessment is based on the experience with floating exchange rates since 1973: wide and long-lasting fluctuations (20 to 50 percent over three to five years) are just not compatible with fully open markets and the complete removal of border posts. While that assessment is open to debate (but seldom challenged so far), the discussion on the intrinsic desirability of the monetary union is moot as long as it ignores the constraint.

Yet, it is probably unavoidable that the question be asked whether EMU is welfare-increasing *per se*.⁴ This question has led to a revival of the literature on optimum currency areas following seminal works by Mundell (1961), McKinnon (1963) and Kenen (1969). Recent efforts have gone into providing formal models which confirmed the early insights (Bayoumi, 1994; Ricci, 1996). However, most of the work has attempted to size Europe up against the Mundell-McKinnon-Kenen criteria. By these particular standards, the case for Europe as an optimal currency area is lukewarm at best.

The (unconstrained) optimum currency area literature establishes the conditions under which two or more countries could share the same currency without seriously adverse consequences. It assumes that the nominal exchange rate has real effects; otherwise, there is no cost in a nation's giving up its own currency. In

⁴ Some studies have attempted to measure directly the costs and benefits from EMU. Bean (1992) concludes that these attempts have failed to come up with tangible answers. The recent report by the Swedish Government Commission on the EMU (1997) provides an excellent and exhaustive review; it concurs that current knowledge prevents any sharp conclusion one way or the other.

particular, the exchange rate is a policy instrument which can affect relative prices such as the real wage paid by producers, the ratio of traded to nontraded goods prices, or the ratio of export to import goods prices. As one example of where this tool could be useful, consider the case where some exogenous shock requires that relative domestic to foreign prices change. Such an adjustment can plausibly be made easier and faster through the exchange rate, rather than by changing nominal prices throughout the economy or through migration of the factors of production from one sector to another.

The three criteria proposed in the literature are those features which make adjustment through exchange rates less effective or less compelling. One criterion is openness to mutual trade; greater openness means that most prices are being determined on markets at the area level, which reduces the ability of the exchange rate to alter significant relative prices. A second criterion is diversification of individual economies; a more diversified economy is less likely to suffer country-specific shocks, which makes its own exchange rate a less useful tool. Finally, the third criterion is mobility of inputs across the area, especially labor. Greater mobility allows an economy to deal with asymmetric shocks through migration, lessening the need for adjustment through exchange rate changes.

On the openness criterion, Europe scores rather well. Measuring openness by looking at exports as a share of GDP, the United States and Japan score 11 percent and 9 percent, respectively. Larger European economies like Germany, Italy, France, and the United Kingdom all have export/GDP ratios above 20 percent, and smaller EU economies like Ireland and Belgium have export/GDP ratios above 70 percent. It makes sense that the smallest European countries are traditionally warm supporters of monetary union. Because of their extreme openness to foreign trade, relative prices in their economy are set on world markets, and the exchange rate is a less useful policy tool.

As to the second criterion, European economies are found usually to be well-diversified. Countries with important endowments in natural resources, like the Netherlands and the United Kingdom with their oil and gas resources, stand apart, but only slightly so. A wide body of research looks at the risk of country-specific (asymmetric) shocks. One set of studies investigates co-movements of key macroeconomic variables like GDP, unemployment, inflation, or the current account balance across European countries (Cohen and Wyplosz, 1989; Weber, 1990). Other studies compare shocks across regions with shocks across countries (de Grauwe and Vanhaverbeke, 1993; von Hagen and Neumann, 1994). The general message is that there is more co-movement in macroeconomic variables among European countries than between individual European countries and the United States or Japan. Further studies attempt to separate out domestic from external shocks, and demand from supply shocks. The underlying argument is that demand shocks are at least partly due to divergence in monetary policy which will be less prevalent in EMU—so attention should focus on supply shocks. Bayoumi and Eichengreen (1993), for example, find more asymmetric supply shocks across Europe than across the United States, although they identify a more coherent group of

core countries around Germany. They suggest that Europe is less-suited to be an optimum currency area than the United States.

Work on the labor mobility criterion clearly suggests that Europe is not an optimum currency area. For example, looking at the United States as a prototype monetary union, Blanchard and Katz (1992) find that when a particular region is hit by an adverse shock, a large proportion of the subsequent drop in employment is matched by labor migration. Applying the same approach to European regions, Decressin and Fatas (1995) find less labor mobility and longer-term effect on regional unemployment, confirming a similar finding by Eichengreen (1991). Two caveats are in order, however. First, the evidence is that the lack of labor mobility is not a national but a regional phenomenon in Europe (Eichengreen, 1993). It affects regions within existing nations of Europe, and there is no reason why monetary union would make things worse. Second, both the occurrence of shocks and labor mobility may change as economic integration proceeds. Krugman (1993) conjectures that economic integration leads to increased regional specialization. In that case, the situation will worsen as the incidence of asymmetric shocks will increase. Frankel and Rose (1996) empirically reject Krugman's conjecture as they find that integration leads to more diversification. In that case, the criteria for an optimum currency area are endogenous. It then comes as no surprise that the United States, which has shared the same currency for a century, appears better suited for a single currency than does Europe.

In the end, we need not be impressed by the result that Europe is not as much an (unconstrained) optimum currency area as the United States. The choice is not between EMU and heaven. It is between EMU and freely-floating exchange rates, with possibly poorly coordinated monetary policies, within an area gradually becoming as tightly integrated as the United States. Would the United States have passed the currency area tests a century ago? And had it failed, all things considered, was it a mistake for the country to adopt a single currency?

Convergence: Will Tough Criteria Backfire?

One striking feature of the Maastricht Treaty is that it anticipates a long eight-year phase from the passage of the treaty in 1991 to the deadline for a single currency by 1999. This long phase-in was the result of a conflict between two competing views.

One view argued that monetary union would be sustainable only if those countries that joined had first achieved a low level of inflation and had resolved fiscal imbalances. This position is commonly referred to as the "economist's view," although it does not seem to have been fully articulated in the professional literature. However, it was popular among the monetary authorities; for example, the Bundesbank championed it under the name of "coronation approach," seeing the shift to monetary union as the last step of successful efforts to eradicate inflationary

behavior. Economic and monetary union was to be born in a land dedicated to a culture of price stability.

The opposing view, generally referred to as the “monetarists’ view,” had the favor of most academic economists.⁵ Their argument was that the creation of a new currency with its own independent central bank would radically alter the wage and price mechanisms, inflation trends, and the incentives of national governments when they decide on fiscal policies. In this view, which is rooted in the Lucas critique, pre-monetary union behavior of both the public and private sectors is a bad predictor of their behavior once the single central bank is in place. Instead, what is needed in the monetarist view are solid institutions, chiefly central bank independence. Other convergence criteria create pain with no assured gain.

Predictably, the “economist” view favored by central bankers won out over the “monetarist” views of academic economists. It is impossible to say what would have happened if EMU had started fairly promptly after ratification of the Maastricht Treaty in 1991. However, what is known is that the period dedicated to convergence has been especially agitated. Even before the Maastricht Treaty could be ratified, a series of exchange rate crises forced Italy and the United Kingdom out of the EMS. After severe currency realignments, the “normal” ± 2.25 percent bands of fluctuation were widened to ± 15 percent, effectively marking the end of the system as initially designed and intended. By mid-1997, about one year before the scheduled selection of the countries which will start EMU, there is considerable pressure to postpone the starting date. The surrounding debate well illustrates the view that there is never likely to be a time when all countries can meet the exacting convergence criteria.

Of the criteria set in Maastricht, those mandating inflation convergence have proven relatively easy to achieve.⁶ However, the budgetary criteria—that the debt/GDP must not be above 60 percent nor the deficit/GDP exceed 3 percent—are more challenging, as Table 1 documents. Why after such a long period of convergence are the budget criteria still some way off? Part of the problem is that the tight monetary policies aimed at meeting the inflation criteria have helped create a slow-growth climate for Europe in the 1990s, with double-digit unemployment rates and no net job creation since the beginning of the decade. While this effort has made it possible to achieve inflation convergence, it has also reduced tax revenues, causing deficits that will not go away and forcing governments to adopt

⁵ For a statement of the “monetarist” view, see Begg et al. (1991). The rationale of the terminology of “economists” versus “monetarists” is unclear. It goes back to earlier debates on economic and monetary union in the 1970s, well summarized in Mundell (1993).

⁶ However, the jury is still out for the criterion concerning the long-run interest rate, which is not to exceed the average of the three lowest-inflation countries by more than 2 percentage points. Since long rates incorporate market expectations of inflation, they are affected by the probability of joining the monetary union. This opens up the possibility of multiple equilibrium: if the markets believe that a country will not join, they may expect monetary policy relaxation and rising inflation, and set high interest rates which indeed rule out EMU membership. Conversely, an expectation that a country will join may bring down long-term rates, thus allowing the country to meet this criteria for convergence.

Table 1
The Maastricht Budget Criteria as of Mid-1997
(percent of GDP)

	<i>Budget Deficit</i> <i>(Limit = 3 percent)</i>	<i>Public Debt</i> <i>(Limit = 60 percent)</i>
Austria	3.0*	71.3
Belgium	2.8*	127.2
Denmark	0.0*	66.8
Finland	2.0*	58.1*
France	3.2	57.8*
Germany	3.2	61.8
Greece	5.2	106.9
Ireland	1.2*	69.0
Italy	3.2	123.3
Luxembourg	N.A.	N.A.
Netherlands	2.3*	74.5
Portugal	2.9*	66.3
Spain	3.0*	69.8
Sweden	2.1*	77.3
United Kingdom	2.8*	54.1*

Source: OECD Economic Outlook, June 1997.

*Note: * denotes compliance.*

further policies of fiscal contraction. This vicious cycle is jeopardizing monetary union both by making the fiscal targets more difficult to achieve and by undermining public support. The situation is now a gamble: either a country reaches EMU and is able to relax after having indeed put its fiscal house in order, or it fails entry (or EMU does not take place at all) because excessively restrictive economic policies have deepened the budget deficit.

Monetary Union and Fiscal Discipline

The inclusion of restrictions on fiscal policy in a treaty which, after all, aims at monetary union, is a source of considerable debate. Before the Maastricht Treaty, most academic analyses emphasized that national fiscal policy would have to become more active to compensate for the loss of the exchange rate instrument.⁷ The opposite approach, that monetary union requires fiscal policy restraint, is grounded in the view that excessive budget deficits may lead to eventual monetization of the debt (Sargent and Wallace, 1981). Monetary authorities were clearly concerned by

⁷ For example, see the papers by Begg, Masson and Melitz, and Wyplosz in *European Economy*, Special Edition No.1, 1991.

high debts in some countries, especially in Italy, whose public debt represents some 18 percent of Europe's GDP. They feared that an explicit or implicit lender-of-last-resort function might force the European Central Bank to step in and indirectly monetize a country's public debt if banks faced a financial crisis in the wake of a default. This concern is reflected in the budgetary criteria for EMU membership and in the "excessive deficit" procedures designed to enforce fiscal rectitude once in the monetary union.

While it is difficult to disagree with the view that fiscal policy ought not to jeopardize monetary and financial stability, how to provide the incentives for appropriate fiscal policy is open to debate. The debate implicitly revolves around one's view of the ability of fiscal policy to play a macroeconomic stabilizing role. It also hinges on the ability to define at the time a deficit is enacted that it is "excessive." In principle, the proper answer must be in terms of "sustainability," since by definition, unsustainable debt buildup will eventually have to be reversed. Fiscal policy sustainability is often associated with stationarity of the debt, usually defined as a stable debt/GDP ratio. In fact, the proper definition of sustainability would hold only that the state will remain solvent, a definition that emphasizes the future behavior of fiscal authorities. By emphasizing future behavior, this view of sustainability also implies that information from the past does not reveal what a country will do after it is inside EMU, and that rules for fiscal rectitude must affect future fiscal policies. A workable definition of sustainability along these lines is a tall order.

The Maastricht approach, relying on arbitrary quantitative limits, is quite unsophisticated.⁸ The 3 percent annual debt/GDP rule corresponds to what is called the "golden rule" in Germany: governments may only borrow to pay for investment spending, and it turns out that governments usually dedicate about 3 percent of GDP to such spending. Even if one ignores doubts about the 3 percent estimate itself, the rule is naïve at best; it ignores socially productive spending like education which is classified as consumption, while it may include ill-designed investment spending. The 60 percent debt/GDP rule was chosen because it was the average of EU countries when the Maastricht Treaty was being negotiated, with not even the pretense of any deeper economic justification.

Yet Europe is not alone in adopting quantitative limits for fiscal policy. How does it work elsewhere, where a unique central bank coexists along with several fiscal authorities? In the United States, for example, states must operate under balanced budgets, borrowing money only by issuing bonds for explicit capital projects. But the comparison must be handled quite carefully. In true federations, the central government is as large as the lower-level governments, and is in charge of macroeconomic stabilization. In Europe, in contrast, the equivalent of a central government is the European Commission, which is not allowed to run deficits and whose spending represents a mere 2 percent of the Europe Union's gross domestic product.

⁸ For a critique of the entry criteria ceilings, see Begg et al. (1991) and Buiter et al. (1993).

The size and role of a powerful central government matters for two main reasons. First, several studies have shown that in federal states, the center smooths out income fluctuations through redistribution from regions in good economic shape to regions undergoing a recession. This function operates automatically through the federal budget, the result of a combination of welfare support and income taxes (Sachs and Sala-i-Martin, 1992; Bayoumi and Masson, 1995; Pisani-Ferry et al., 1995). In this setup, it can make sense to limit the stabilization role of sub-central authorities. Second, quantitative fiscal restraints at some levels of government can actually encourage the buildup of debts at other levels, according to evidence from von Hagen and Eichengreen (1996). The problem occurs when fiscally irresponsible lower-level governments refuse to borrow and can bait the federal authorities into rescuing them. In Europe, a central government with powerful redistribution and stabilization authority is not likely within the foreseeable future. Consequently, Europe needs national-level stabilization policies much more than individual U.S. states do, and there is no risk that national governments will conduct irresponsible fiscal policies in an attempt to extract transfers from a penniless center.

Are there less coarse methods than quantitative limits of providing governments with effective incentives against fiscal irresponsibility? One attractive approach would be to rely on financial markets to impose discipline. In a single currency area, interest rates no longer reflect a country's sovereign risk. Instead, they reflect the risk category of borrowers, be they fiscal authorities (a municipality in the United States, a province in Canada, or a government in Europe) or private borrowers. To the extent that markets price risk correctly, the demand for public debt of various governments could act as both a barometer and a constraint. If a country lets its debt grow and there is an enhanced risk of default, markets should react by downgrading their evaluation and by increasing the interest rate at which new debt is being financed, until fiscal authorities see it to be in their best interest to curtail the deficit.

However, history suggests skepticism about the ability of markets to impose discipline in this way. For one, markets tend to throw good money after bad for a time (Eichengreen and Portes, 1987). When markets do react, it is often too late and too violently. They abruptly cut financing, making it impossible for the government to borrow further and bankrupting large bondholders, among them commercial banks and other financial institutions. This leads to a scenario where central banks may feel compelled to monetize (part of) the debt.⁹

This is presumably why the Maastricht Treaty includes a no-bailout clause which explicitly forbids the rescue of one government either by its fellow members or by community institutions, including the European Central Bank. In this way, fiscal misbehavior becomes a strictly national issue with no union-wide implication and fiscal restraint is unnecessary. Yet Germany has argued that the no-bailout clause cannot be fully credible, that any rule can always be circumvented.

⁹ That this mechanism bears strong resemblance to the events that provoked hyperinflation in Germany in 1922–23 is not irrelevant for an understanding of the Maastricht Treaty.

In the end, the explicit fiscal restraints embodied in the excessive deficit procedure can be seen as insurance against a remote risk that European institutions would be compelled to monetize some nation's out-of-control debts. **This insurance scheme may turn out to be very costly in terms of the ability to run countercyclical policies.**

EMU and the Rest of the World

The potential for the euro to replace the U.S. dollar as the world's premier currency is one of the understated motivations of EMU. In part, the desire is a symbolic one; it is the belief that "great powers have great currencies" (Mundell, 1993, p. 9). In part, it is a hope to reap seigniorage, although U.S. benefits from seigniorage are worth only about 0.2 percent of GDP (Alogoskoufis and Portes, 1992). The usual criteria for becoming the world's lead currency are measures like size (GDP or the share of world trade). By these measures, the prospects for the euro to challenge the dollar are favorable but not overwhelming. For example, Europe's international trade with non-European nations will not exceed by much Germany's current level of foreign trade—once intra-European trade is netted out (Hartmann, 1996). Also, history teaches that it takes time for a reserve currency to change (Eichengreen, 1989; Mundell, 1993). To overcome its handicap relative to the incumbent U.S. dollar, the euro must discover some absolute advantage.

One potential advantage is likely to be greater price stability. As a currency expected to follow a long-run trend of appreciation, the euro will be a currency that stores value better than the alternatives. This prediction derives from the constitution of the European Central Bank, which makes it more independent and more focused on price stability than the U.S. Federal Reserve. If anything, the constitution is even stricter than that of the Bundesbank, so that Europe's economy will be more stable than Germany's (Masson and Turtelboom, 1997). A counterargument is based on politico-economic considerations. The board of the European Central Bank will be composed of representatives of all member countries. With the one-man one-vote principle, Germany's weight will be no larger than that of Belgium or Italy. The constituencies of the European Central Bank will not share the German allergy to even moderate inflation.¹⁰ In theory the outcome may differ from the wishes of the median European voter, and the bias can go in either direction (Alesina and Grilli, 1992). Ultimately, this counterargument is not fully convincing.

¹⁰ This is another reason why the Bundesbank has advocated a long convergence process: to provide for a period of deep conversion to a "stability culture." In a perceptive comment on Buiter et al. (1993, p. 97), Frankel interprets the Maastricht convergence process as a "test of will," referring to Buddhist traditions: "A meditating neophyte is supposed to learn to refrain from responding to a flea by scratching it, just as a political region is supposed to learn to refrain from responding to a local downtick in demand by lowering interest rates."

A second potential advantage for the euro could be the depth and cost-efficiency of financial markets. The market for the euro and euro-denominated assets could be the world's largest, depending on whether the city of London shifts to the euro. Yet the location and prominence of markets relies increasingly less on regional considerations and more on the regulatory environment. Europe will have to fight its own heavy-handed approach and powerful lobbies if it wants the euro to become the world's currency.

Thus, the best bet is that, for a long while at least, the dollar's supremacy will remain. Still, the creation of the euro is bound to affect international monetary relations. Will it lead to more or less instability on exchange markets? Two arguments suggest more instability. First, if the U.S. dollar has been acting as a market leader on exchange rate markets, the shift to a situation of bargaining between more equal partners is likely to create greater volatility. Second, while the fairly open economies of Europe are now keenly interested in stabilizing world currencies, a euro zone would join the United States and Japan as giant economies less inclined to give up domestic policy objectives for the sake of exchange rate coordination. However, the opposite view is that moving from G-7 to G-3 should make it easier to negotiate methods for reducing volatility in exchange rates (Alogoskoufis and Portes, 1992; Goodhart, 1993; Kenen, 1996). In the end, little should change when the European Central Bank steps in the shoes of the Bundesbank as the master of the EMS exchange rate.

Finally, what will be the impact of economic and monetary union on the International Monetary Fund? One view is: nothing much. Each country will retain its existing role. In its annual review exercise, the IMF will have to take account of the fact that monetary policy is no longer a national responsibility, but that is already the case for other monetary unions in Africa and the Caribbean. However, a more entertaining scenario, if unlikely, envisions EMU countries merging as a single IMF member. Not only would Europe cast the largest number of votes and challenge U.S. dominance, but it could invoke the agreements' article that states "the principal office of the Fund shall be in the territory of the member having the largest quota" and request that the IMF move from Washington to Madrid, Frankfurt, Paris or Amsterdam.

The Early Steps: What to Watch For

The Treaty of Maastricht sets a clear timetable: a single currency will come into being no later than 1999. It may seem that all that remains is to watch the countdown before lift-off. Nothing is further from the truth. Power in the boosters is not assured; last minute checks reveal a number of blinking red lights; and politico-economic pressures are building up to dangerous levels. Public support for the euro is lukewarm at best. It is largely incomprehensible. As a symbol of national belonging, it is desirable to some and threatening to others. As the time to launch

draws near, popular anxiety is tending to rise. In virtually every country, politicians are making capital out of their opposition to monetary union.

Must EMU start by January 1999? Several loopholes exist for sidestepping the deadline. First, it is understood that monetary union will not exist without both Germany and France. This gives each of these countries veto power that they can exercise by missing the convergence target. In fact, it appears that both are likely to miss the targets narrowly, which will inevitably lead to negotiations about their situation. Second, certain provisions of the treaty could be twisted to postpone startup beyond 1999, although it would be a farfetched interpretation of the treaty.

By June 1998, Europe's heads of state must agree on the list of the passengers of the first mission. Many countries will not fulfill the formal criteria, so the decision will have a degree of arbitrariness relying on flexibility in the precise wording of the treaty. In anticipation, adversaries of economic and monetary union are calling for a postponement. In fact, any delay would feed doubts that convergence can be achieved and reduce chances of success. In that case, speculators could well unleash new attacks on exchange rates, which might make any transition to a single currency even more difficult to achieve.¹¹

Immediately after the list of members is drawn up, final preparations will start. At least one unresolved issue has been identified. *Legal restrictions* imply that the rate at which currencies will be converted into euro on January 1, 1999, must be those observed at the closing of markets on December 31, 1998. This creates the risk of major exchange market instability in the time leading up to that date, as traders will need to form a view of what the authorities are trying to achieve. Moreover, at a time of high unemployment and with policy settings driven by the need to meet contractionary convergence criteria, some countries may be tempted to secure a temporary competitive advantage by entering monetary union with an undervalued currency. Solutions for tying down the issue ahead of time still remain to be adopted (Begg et al., 1997; Obstfeld, 1997).

According to the Maastricht Treaty, the European Central Bank will come into existence soon after July 1998. It will have to coexist for six months with national central banks due to become its subsidiaries. From January 1999, the European Central Bank will operate only in euros, as will the financial markets. At the retail level, national currencies will continue to circulate and remain sole legal tender until July 2002, but will be legally considered as (horrendous six-digit) fractions of the euro. The euro itself will be finally introduced for retail transactions in January 2002 (probably), opening up a switchover period of six months. Thereafter national currencies will be redeemed in euros for periods to be set by national legislation. The three-year overlap is bound to raise endless practical issues, not the least of which is that it may be difficult for governments and citizens to realize that the

¹¹ In any case, speculative attacks are expected against those countries which are not admitted to the single currency. Such attacks could be minimized if information is gradually leaked to the markets well in advance and if new dates for entry are clearly set along with a clear signal that the next decision will be positive.

European Central Bank is solely in charge after 1999, and all surviving currencies are mere subdivisions of the euro with a fixed and irrevocably set conversion rate.

Conclusion

Currencies and nations normally coincide. Europe is set to attempt an original experiment. Is it going to work? Is it even going to happen? The fact that a year before lift-off, doubts remain about the likelihood that EMU will start, or will start on time, is testimony to the fact that there can be no firm answers to these questions. Yet some simple observations can take us a long way.

The Maastricht Treaty is the fundamental act on which Europe rests. It is an international treaty, formally ratified by all European Union countries, and it supersedes national legislation. Giving up EMU would throw up more than just monetary union. It would create a situation of deep political crisis with unpredictable consequences. For that reason alone, the bet is that EMU will be on, on time.

Is the logic behind monetary union only political? Quite the contrary. The political aim of a single currency has been pursued relentlessly by its advocates since the late 1950s; several explicit attempts failed because economic conditions were not ripe. The Maastricht Treaty only came about because the lifting of capital controls had reduced the alternate options to just two unpalatable extremes: either allow exchange rates to float freely or accept the complete domination of Germany's Bundesbank over Europe's monetary policy.

Freely floating exchange rates are not compatible with a completely borderless economic area. They carry the germs of protectionist pressure and financial instability which threaten economic integration. As for dominance by the Bundesbank, it has been largely beneficial over the last decade, chiefly because inflation has been eliminated. Yet there have been costs: lasting double-digit unemployment, major policy mistakes that led to the currency crises of 1992–93, and continuing disagreements over the objectives of the Bundesbank. The current situation is not sustainable because it entails a fundamental contradiction. On one hand, the Bundesbank derives its leadership from a reputation of undeterred commitment to price stability in Germany. On the other hand, long-lasting leadership requires that all of Europe's economic conditions be taken into account, which is against the Bundesbank's constitutional duty to Germany. Tinkering with the Bundesbank's constitution is not only politically impossible, but doing so would also undermine its credibility and its ability to lead. In this setting, EMU emerges as the best possible economic solution.

Assessing the costs and benefits of monetary union quantitatively is both frustrating and useless. It is frustrating because, frankly, as economists we are unable to compute them with any precision, and we owe it to the profession to admit so in public. Our understanding of monetary and exchange rate policy is regrettably limited, and the lack of a precedent leaves us with more conjectures than certainties. Moreover, quantitative estimates are useless unless they are sized up against the

costs and benefits of the relevant alternatives, which is equally beyond our current ability. The best that can be done in this situation is to gain an understanding of where the costs and benefits are likely to reside.

The direct benefits come in the form of reduced transaction costs and reduced uncertainty, possibly including additional transparency in competition. Such effects are likely to be small, but not trivial. Direct benefits also include lower real interest rates for countries where a sizable currency risk premium exists. Indirect benefits come from the institutional arrangements that accompany EMU. The broadening of central bank independence from political control would not have happened without EMU, and with it comes the realization that international competition is not achieved through lobbying for exchange rate manipulation.

More ambiguous is the role of the fiscal restraints, both the entry conditions and the excess deficit procedure. In most countries, these restraints have promoted long-needed efforts at coming to grip with unsustainable deficits. At the same time, the insistence on price stability along with the adoption of rigid and arbitrary criteria of fiscal rectitude have already played a role in deepening and lengthening Europe's phase of slow growth, with huge costs in terms of unemployment and social suffering. The risk now is of more of the same in the early EMU years. As already noted, these costs are the consequence of EMU's parenthood: Germany could not be expected to give up its famed deutsche mark without extensive guarantees. These demands could not be turned down and have probably become excessive. However, once monetary union exists, many arrangements can be changed. Right now, Europeans are biting the bullet and looking beyond the 1999 horizon.

■ *For useful comments and advice, and without any implication, thanks to the editors, Alan Krueger, Brad De Long, and Timothy Taylor, as well as to Benoît Coeuré, Barry Eichengreen, Hans Genberg, Paul de Grauwe, Paul Masson, Jacques Melitz, Maury Obstfeld and Richard Portes.*

References

Alesina, Alberto and Vittorio Grilli, "The European Central Bank: Reshaping Monetary Politics in Europe." In Canzoneri, Matthew B., Vittorio Grilli, and Paul R. Masson, eds., *Establishing a Central Bank: Issues in Europe and Lessons from the U.S.* Cambridge: Cambridge University Press, 1992, 49-85.

Alogoskoufis, George, and Richard Portes, "European Monetary Union and International

Currencies in a Tripolar World," In Canzoneri, Matthew B., Vittorio Grilli, and Paul R. Masson, eds., *Establishing a Central Bank: Issues in Europe and Lessons from the U.S.* Cambridge: Cambridge University Press, 1992, 273-302.

Bayoumi, Tamim, "A Formal Model of Optimum Currency Areas," *International Monetary Fund Staff Papers*, December 1994, 41, 537-54.

Bayoumi, Tamim, and Barry Eichengreen,

"Shocking Aspects of European Monetary Union," In F. Torres, Francisco and Francesco Giavazzi, eds., *The Transition to Economic and Monetary Union*. Cambridge: Cambridge University Press, 1993, 193–240.

Bayoumi, Tamim, Morris Goldstein, and Geoffrey Woglom, "Do Credit Markets Discipline Sovereign Borrowers? Evidence from the U.S. States," *Journal of Money, Credit, and Banking*, November 1995, 27, 1046–59.

Bayoumi, Tamim and Paul R. Masson, "Fiscal Flows in the United States and Canada: Lessons for Monetary Union in Europe," *European Economic Review*, February 1995, 39, 253–74.

Bean, Charles R., "Economic and Monetary Union in Europe," *Journal of Economic Perspectives*, Fall 1992, 6, 31–52.

Begg, David, and Charles Wyplosz, "The European Monetary System: Recent Intellectual History," in: *The Monetary Future of Europe*. London: CEPR, 1993.

Begg, David, Jean-Pierre Danthine, Francesco Giavazzi, and Charles Wyplosz, "The East, the West, and the Deutschemark," *Monitoring European Integration*, 1990, 1.

Begg, David, Francesco Giavazzi, Luigi Spaventa, and Charles Wyplosz, "European Monetary Union—The Macro Issues," *Monitoring European Integration*, 1991, 2, 3–68.

Begg, David, Francesco Giavazzi, Jürgen von Hagen, and Charles Wyplosz, "EMU, Getting the End-Game Right," *Monitoring European Integration*, 1997, 7, 1–75.

Blanchard, Olivier J. and Lawrence F. Katz, "Regional Evolutions," *Brookings Papers on Economic Activity*, 1992, 1, 1–75.

Buiter, Willem, Giancarlo Corsetti, and Nouriel Roubini, "Excessive Deficits: Sense and Nonsense in the Treaty of Maastricht," *Economic Policy*, April 1993, 16, 57–100.

Burda, Michael and Charles Wyplosz, *Macroeconomics, A European Text*, 2nd ed., Oxford, UK: Oxford University Press, 1993.

Cohen, Daniel and Charles Wyplosz, "The European Monetary System: An Agnostic Evaluation." In Bryant, Ralph, David Currie, Jacob Frenkel, Paul R. Masson, and Richard Portes, eds., *Macroeconomic Policies in an Interdependent World*. Washington: International Monetary Fund, 1989, 311–37.

Decressin, Jorg, and Antonio Fatas, "Regional Labor Market Dynamics in Europe and Implications for EMU," *European Economic Review*, December 1995, 39, 1627–55.

de Grauwe, Paul and Wim Vanhaverbeke, "Is Europe an Optimum Currency Area?: Evidence from Regional Data." In Masson, Paul R. and

Mark P. Taylor, eds., *Policy Issues on the Operation of Currency Unions*. Cambridge: Cambridge University Press, 1993, 111–29.

Eichengreen, Barry, "Hegemonic Stability Theories of the International Monetary System." In Cooper, Richard N., et al. (eds.) *Can nations agree? Issues in International Economic Cooperation*. *Studies in International Economics series*. Washington, DC: Brookings Institution, 1989, 255–98.

Eichengreen, Barry, "One Money for Europe? Lessons from the U.S. Currency Union," *Economic Policy*, April 1990, 10, 119–86.

Eichengreen, Barry, "Is Europe an Optimum Currency Area?," Cambridge, Mass: National Bureau of Economic Research Working Paper 3579, January 1991.

Eichengreen, Barry, "Labor markets and European Monetary Unification." In Masson, Paul R. and Mark P. Taylor, eds., *Policy Issues on the Operation of Currency Unions*. Cambridge: Cambridge University Press, 1993, 130–62.

Eichengreen, Barry, and Richard Portes, "The Anatomy of Financial Crises," In Portes, Richard and Alexander K. Swoboda, eds., *Threats to International Financial Stability*. Cambridge: Cambridge University Press, 1987, 10–58.

Eichengreen, Barry, and Charles Wyplosz, "The Unstable EMS," *Brookings Papers on Economic Activity*, 1993, 1, 51–144.

Folkerts-Landau, David, and Peter M. Garber, "The ECB: A Bank or a Monetary Policy Rule?," In Canzoneri, Matthew B., Vittorio Grilli, and Paul R. Masson, eds., *Establishing a Central Bank: Issues in Europe and Lessons from the U.S.* Cambridge: Cambridge University Press, 1992, 86–123.

Frankel, Jeffrey A., and Andrew Rose, "The Endogeneity of Optimum Currency Areas," London: Centre for Economic Policy Research Discussion Paper No. 1473, September 1996.

Goodhart, Charles A. E. "The European System of Central Banks after Maastricht." In Masson, Paul R. and Mark P. Taylor, eds., *Policy Issues on the Operation of Currency Unions*. Cambridge: Cambridge University Press, 1993, 215–39.

Grilli, Vittorio, Donato Masciandaro, and Guido Tabellini, "Political and Monetary Institutions and Public Financial Policies in the Industrial Countries," *Economic Policy*, October 1991, 13, 341–92.

Hartmann, Philipp, "The Future of the Euro as an International Currency, A Transactions Perspective," unpublished paper, 1996, London: London School of Economics.

Kenen, Peter B., "The Theory of Optimum Currency Areas." In Mundell, Robert A., and Alexander A. Swoboda, eds., *Monetary Problems of the*

International Economy. Chicago: Chicago University Press, 1969, 41–60.

Kenen, Peter B., "Sorting Out Some EMU Issues," *Reprints in International Finance*, No. 29, December 1996, International Finance Section, Princeton University.

Krugman, Paul, "Lessons of Massachusetts for EMU," In Torres, Francisco and Francesco Giavazzi, eds., *The Transition to Economic and Monetary Union*. Cambridge: Cambridge University Press, 1993, 241–69.

McKinnon, Ronald, "Optimum Currency Areas," *American Economic Review*, 1963, 53, 717–725.

Masson, Paul R. and Mark P. Taylor, "Currency Unions: A Survey of the Issues." In Masson, Paul R. and Mark P. Taylor, eds., *Policy Issues on the Operation of Currency Unions*. Cambridge: Cambridge University Press, 1993, 3–54.

Masson, Paul R. and Bart Turtelboom, "Characteristics of the Euro, the Demand for Reserves, and Policy Coordination Under EMU," unpublished paper, 1997, Washington, DC: International Monetary Fund.

Mundell, Robert A., "A Theory of Optimum Currency Area," *American Economic Review*, September 1961, 50, 657–665.

Mundell, Robert A., "EMU and the International Monetary System," In *The Monetary Future of Europe*. London: Centre for Economic Policy Research, 1993.

Obstfeld, Maurice, "A Strategy for Launching the Euro," unpublished paper, March 1997, University of California, Berkeley.

Padoa Schioppa, Tommaso, "Squaring the Circle, or the Conundrum of International Monetary Reform," *Catalyst, a Journal of Policy Debate*, Spring 1985, 1:1.

Pisani-Ferry, Jean, Alexander Italianer, and Roland Lescure, "Stabilization Properties of Budgetary Systems: A Simulation Analysis," *European Economy*, 1995, 5, 511–38.

Ricci, Luca, "A Model of an Optimum Currency Area," Working Paper 97/76, Washington, DC: International Monetary Fund, 1997.

Sachs, Jeffrey D., and Xavier Sala-i-Martin, "Fiscal Federalism and Optimum Currency Areas: Evidence for Europe from the United States." In Canzoneri, Matthew B., Vittorio Grilli, and Paul R. Masson, eds., *Establishing a central bank: Issues in Europe and lessons from the U.S.*. Cambridge: Cambridge University Press, 1992, 195–219.

Sargent, Thomas J. and Neil Wallace, "Some Unpleasant Monetary Arithmetic," *Federal Reserve Bank of Minneapolis Quarterly Review*, 1981, 5, 1–17.

Swedish Government Commission on the EMU, EMU—A Swedish Perspective, Amsterdam: Kluwer, 1997.

Thygesen, Niels, "Economic and Monetary Union: Critical Notes on the Maastricht Treaty Revisions." In Torres, Francisco and Francesco Giavazzi, eds., *The Transition to Economic and Monetary Union*. Cambridge: Cambridge University Press, 1993, 9–45.

von Hagen, Jürgen, "Monetary Union and Fiscal Union: a Perspective from Fiscal Federalism." In Masson, Paul R. and Mark P. Taylor, eds., *Policy Issues on the Operation of Currency Unions*. Cambridge: Cambridge University Press, 1993, 264–96.

von Hagen, Jürgen and Barry Eichengreen, "Federalism, Fiscal Restraints, and European Monetary Union," *American Economic Review, Papers and Proceedings*, May 1996, 86, 134–38.

von Hagen, Jürgen and Ian J. Harden, "National Budget Processes and Fiscal Performance," *European Economy*, 1994, No. 3, 311–418.

von Hagen, Jürgen, and Manfred J. M. Neumann, "Real Exchange Rates within and between Currency Areas: How Far Away Is EMU?," *Review of Economics and Statistics*, May 1994, 76, 236–44.

Weber, Axel A., "EMU and Asymmetries and Adjustment Problems in the EMS: Some Empirical Evidence," London: Centre for Economic Policy Research Discussion Paper: No. 448, August 1990.

The Political Economy of European Monetary Unification: An Analytical Introduction

Barry Eichengreen and Jeffrey Frieden

European monetary unification — the process that led to creation of a single European currency (the euro) and a European Central Bank (the ECB) — is both an economic and a political phenomenon. It is economic in that monetary unification has far-reaching consequences for economic policy and performance Europe wide. Transactions costs have been reduced by the advent of the single currency, stimulating intra-European trade and capital flows. Interest-rate differentials have narrowed now that the separate monetary policies of the founding member states have been replaced by the single policy of the ECB.¹ European finance is being transformed by the explosive growth of euro-denominated bond issues, strategic alliances among national stock exchanges, and a continent-wide wave of bank mergers as the advent of the euro creates for the first time a truly continental financial market.

But EMU is also a political phenomenon. The decision to create the monetary union, the decision of whom to admit, and the decision of whom to appoint to run the ECB are political decisions, taken by political leaders, subject to political constraints, not the social-welfare maximizing decisions of some mythical social planner. They result from a political process of treaty negotiation, parliamentary ratification, and popular referendum. Individuals and interest

¹While "Euroland," as the area comprised of the founding members is fondly known, is made up of 11 countries, they had only ten monetary policies, Belgium and Luxembourg having long since formed a monetary union.

groups support or oppose monetary unification — not just in EU member states that have not yet joined the euro area, but even now in the founding members — on the basis of how they perceive it as affecting their individual welfare, not the welfare of the nation as a whole, much less the welfare of the entire European Union.²

Despite the outpouring of research prompted by EMU, few accounts have systematically analyzed both its political and economic aspects. That is the goal of the present volume. The contributors describe both the political and economic dimensions of the process. They demonstrate how political constraints have shaped the design and operation of Europe's monetary union at the same time that the changes in economic structure brought about by monetary integration continue to transform European politics.

A Short History of European Monetary Unification

Monetary unification has always been at the center of the larger process of European integration. Economically, the creation of a single currency was long seen as necessary for forging a truly integrated European market. Politically, monetary unification has been seen as a practical and symbolic step toward the development of a capacity to formulate social and foreign policies at the European level. Both advocates and opponents of further European political integration have long regarded monetary integration as the thin end of the wedge. For all these

²Throughout this essay, we use the term "European Union" even where the organization was at the time going by another, earlier, name.

reasons, the desirability of European monetary unification has been contested since the idea was first mooted.

Serious discussion of monetary unification goes back to the 1960s.³ In 1969 the Werner Report set forth an ambitious plan for a three-step transition to monetary union to be completed within a decade. In the event, its blueprint was rendered obsolete within weeks by the slow-motion collapse of the Bretton Woods international monetary system. When the Bretton Woods system of par values, which had provided a framework for holding European exchange rates within fluctuation bands of plus or minus 2 per cent, finally collapsed in 1971, the European Community's six founding member states had to focus their energies on limiting the volatility of their exchange rates, lest complaints of arbitrary and capricious changes in competitiveness undermine European solidarity. They resolved to hold their currencies within 2 1/4 per cent bands (only slightly wider than those they had operated under Bretton Woods). They were joined in the "Snake," as this arrangement was prosaically known, by the UK, Ireland and Denmark as those three countries prepared to become members of the European Union.⁴

³Even at this early date, plans for monetary union were linked to broader political problems associated with French ambivalence about the European Community in the aftermath of the franc crisis of 1969. On early monetary plans, see Tsoukalis (1977) and Ypersele (1985). For surveys of the process of monetary integration more generally, see Fratianni and von Hagen (1992), and Gros and Thygesen (1998).

⁴Norway and Sweden also tied their currencies to the mechanism. On the snake, see Tsoukalis (1977), Ludlow (1982), and Coffey (1987).

Soon, however, divergent economic conditions and policies, reflecting the impact of the first oil shock in 1973-4, rendered the Snake unworkable. The least committed members suffered repeated balance-of-payments crises, forcing them to alter or abandon their currency bands. By 1975 only Germany, the Benelux countries, and Denmark remained in the snake. The European countries whose policies diverged most from Germany's — the UK, Ireland, France and Italy — simply left, and the Danes were able to remain only by virtue of serial devaluations.

Discussions of monetary unification resumed once this turbulence had passed. The outcome was the establishment of the European Monetary System (EMS) and its Exchange Rate Mechanism (ERM) in March 1979. All EU member states except the UK participated in the ERM, linking their currencies via a multilateral parity grid which again allowed for fluctuations of plus or minus 2 1/4 per cent.⁵ Provision was made for realignments, although these were expected to be rare. Financing facilities were provided for countries attempting to stabilize their exchange rates in the face of balance-of-payments shocks, and capital controls were relied upon to limit speculative pressures.

Conventional wisdom at the time was that the EMS was unlikely to succeed. The inflation rates of the participating countries differed widely. High-inflation countries had demonstrated an inability to put in place the measures needed to stabilize their currencies against

⁵6 per cent for the Italian lira.

the deutschmark. Faith meant believing that the creation of the EMS itself would strengthen the willingness of high-inflation countries to pursue painful policies of austerity.

Initially, skepticism seemed more than justified. In the first four years of the EMS, balance-of-payments pressures were intense. Exchange rates were realigned seven times, and there were few signs of monetary convergence. Then, however, the outlook began to brighten. Rates of price increase in the high-inflation countries began to decline. From April 1983 to January 1987 there were only four realignments, generally smaller than those which had come before. And from January 1987 to September 1992 there were no major realignments within the ERM. Attracted by its improved performance, Spain, Portugal and the UK all joined the mechanism in this period.

This transformation was stimulated by, and stimulated in turn, progress on the larger project of European integration. The Single European Act of 1986 called for the removal of controls on the movement of goods, capital and persons within the Union. A true common market, in which only countries with well-behaved exchange rates would be permitted to participate, created the prospect of additional rewards for ERM participation. And by mandating the removal of capital controls, the Single European Act pointed up the need for further policy convergence in order for the stability of exchange rates to be maintained.

In this context, and specifically in response to calls by the French and German foreign ministries, the European Council appointed a committee in 1988 headed by European Commission President Jacques Delors to investigate the prospects for further monetary

integration. The Delors Committee recommended that the EU begin moving immediately toward the creation of a single currency. The next step in this process came at Maastricht in the Netherlands in December 1991, when the member states agreed to a sweeping treaty on economic union, giving diplomatic content to the recommendations of the Delors Report.

The Maastricht Treaty, echoing the Delors Report and the Werner Report before it, sketched a transition in three stages. Stage I involved the elimination of Europe's remaining capital controls, the accession of all EU members to the ERM, and hardening of the exchange-rate commitment. In Stage II, with the EMS credible and encompassing, member states would reinforce the independence of their national central banks and strive to satisfy a set of "convergence criteria" designed to facilitate the harmonization of their economic policies and to distinguish member states prepared to live with the consequences of a single monetary policy from those lacking the requisite commitment. A European Monetary Institute would be created to lay the groundwork for the establishment of the ECB. Finally, in Stage III, to commence no later than the beginning of 1999, the European Central Bank would begin operations, to be followed in three years by the issuance of euro banknotes and coins.

With a plan in place and all EU members but Greece participating in the ERM, it appeared that the single currency was only a matter of time. But while Europe's political leaders had had their say, the markets and voters were still to be heard from. The backdrop for their intervention was German reunification, underway since 1990. The costs of reunifying the two Germanys gave rise to huge budget deficits for the Federal Republic, which excited fears of

inflation in the corridors of the Bundesbank. The response of the latter was to raise interest rates. But high interest rates were uncomfortable for Italy, whose debt and deficits were large. They were uncomfortable for the UK, where mortgage interest rates were indexed, whose business cycle was not synchronized with that of the rest of Europe, and where the commitment to European integration was less than firm. More generally, chronically high unemployment rates, which by the early 1990s had become a fact of Europe's economic and political life, made it difficult for governments to stomach high, German-style interest rates. The result was pressure on sterling, the lira and other weak European currencies in the summer of 1992.

In the midst of this gathering storm, public opposition to the Maastricht Treaty materialized. The failure of Danish voters to ratify the treaty in that country's June 1992 referendum and fears that a subsequent referendum in France might also fail raised the possibility that the entire Maastricht process might be derailed. An indefinite postponement of EMU would have rendered it all the more unlikely that high-unemployment countries would be prepared to stay the austerity course. Knowing this, the markets pounced. Despite committing billions of dollars to their battle with currency-market speculators, the Bank of England and Bank of Italy were forced to surrender to the markets. On September 16th, 1992, their respective governments withdrew from the ERM and allowed their currencies to depreciate.

Having toppled two of Europe's larger currencies, speculators turned their attention to the smaller ones. Yet more crises and realignments then drove the currencies of Spain, Portugal and Ireland downward.

Unsettled conditions persisted into 1993. In the summer, with unemployment continuing to rise continent wide, the pressure intensified for interest rate reductions. None of this deterred the Bundesbank, still preoccupied by inflation, from maintaining a tight monetary stance that prevented other EMS members from reducing their interest rates. The resulting dilemma was most serious in France, where a new government took office amidst the recession. The Bank of France attempted to lead by example, reducing interest rates in the hope that the Bundesbank would follow. In this they were disappointed. Inferring that the French authorities were unprepared to hold the line, currency speculators turned their fire on the franc. Massive intervention by the Bank of France failed to repel them. With pressure continuing to mount, European leaders were forced to acknowledge that the old narrow-band ERM had been rendered unworkable by the removal of capital controls, the liquidity of the markets, and the existence of other political and economic imperatives. But rather than abandoning the mechanism, they adopted the stopgap of widening its currency fluctuation bands from 2 1/4 to 15 per cent.⁶ This removed the one-way bets that speculators had found irresistible (since currencies could now appreciate rather than depreciate if speculators turned out to be wrong), and gave governments more room for maneuver. At the same time, however, it raised questions about the capacity of EU member states to keep their exchange rates stable and, more fundamentally, their willingness to subordinate other social goals to a single monetary policy.

⁶The bilateral band for the deutschmark and the Dutch guilder, a currency whose stability was never in question, was kept at 2.25 per cent.

In this, Europe's darkest hour, fears were widespread that the Maastricht process was doomed. With only a weakened exchange rate commitment to bind them, member states might fail to make progress on convergence. And if other EU member states failed to solve their inflation, debt and deficit problems, Germany would be unwilling to embrace them as partners in a monetary union. Most alarmingly of all, if exchange rates grew more volatile as policies diverged, the Single Market might be placed at risk.

This pessimism turned out to be exaggerated. Economic developments help explain why the convergence of policies and institutions between 1994 and 1998 turned out to be more successful than anticipated. By 1994 the shock of German unification had begun to recede, allowing the Bundesbank to reduce rates. The 1993 recession passed, and with the resumption of economic growth, fiscal consolidation and policy harmonization became easier to undertake.

Politically, the transition to EMU was eased by the strong demonstrations of support that plans for a single currency elicited following the 1992-1993 crisis. Even governments that had been forced to devalue reiterated their commitment to the completion of the monetary union, and despite the widening of the currency bands, most were indeed able to keep their currencies close to their central rates. Powerful business groups were vocal in support of EMU, arguing that the 1992-1993 devaluations had disrupted progress toward the Single Market and deeper integration. And electoral support for anti-European parties and candidates waned as economic difficulties receded.

Institutionally, inter-state negotiations allayed some of the fears of those who remained wary of EMU. A Growth and Stability Pact committed EMU members to avoid large budget deficits, thus extending the commitment to fiscal retrenchment beyond the creation of the single currency. This helped to pacify fiscal conservatives concerned about the profligacy of Southern European (and other) governments. At the same time, member states exhibited some flexibility in interpreting the fiscal criteria, which reassured those who worried that EMU would be too rigid a policy straitjacket. All in all, negotiations signaled that the architects of the monetary union were aware of the political constraints and willing to work within them.

Together, then, favorable economic conditions, political momentum, and institutional flexibility combined to make it possible for most EU member states to complete the fiscal retrenchment necessary to qualify for monetary union. While the letter of the convergence criteria was not always strictly met, European policy makers concluded that the aspirants had satisfied their spirit, and each of the 11 member states wishing to participate was deemed worthy when the decision was taken in 1998. Right on schedule, on January 1st, 1999, Europe's monetary union came into being.

The Economics of Monetary Unification

Most analyses of the economics of European monetary unification build on the theory of optimum currency areas, one of the contributions for which Robert Mundell was awarded the Nobel Prize in Economics in 1999. In Mundell's model, the benefits of monetary unification,

which take the form of the reduction in transactions costs consequent on replacing distinct national currencies with a single (common) currency, are balanced against the costs of sacrificing monetary and fiscal autonomy. One might think that the savings in transactions costs are considerable. Tourists changing money at airports pay commissions amounting to anywhere from 2 to 5 per cent of the cash they exchange. But banks and firms doing larger volumes of business in wholesale markets pay much smaller commissions. And such costs, as a share of GNP, depend on the openness of the economy. European Commission estimates suggest that conversion costs absorbed about one per cent of national income for the EU's small, low-income countries but as little as one tenth of one per cent of national income for the large member states for which international transactions are less important. Overall, currency conversion costs averaged less than one half of one per cent of EU national income in the late 1980s. This, it would seem, is a modest return on a process riddled with risks and uncertainties.

It can be argued that the real efficiency advantages come not from the single currency but from the Single Market and that the two initiatives are linked. The Single Market allows European producers to exploit economies of scale and scope. By creating a Europe-wide financial market, it promises to stimulate efficiency-enhancing mergers of banks and securities exchanges. By heightening cross-border competition, it forces European producers to shape up or ship out. By intensifying regulatory competition, it compels European governments to remake their policies in market friendly ways. This is only one vision of the intentions of architects of the Single Market, to be sure, but it is a compelling one.

And the single currency is indispensable, the argument continues, if Europe is to reap the benefits of the Single Market. It enhances transparency. It makes it that much harder for automobile producers to charge different prices in different countries when a single unit of account allows consumers to more readily compare those prices across countries. It makes it harder for banks steeped in traditional ways to survive in sleepy national backwaters, insofar as the elimination of currency risk encourages savers to seek out higher deposit rates and investors to seek out lower loan rates abroad. It makes it harder for unions to insist on restrictive work rules, insofar as a single currency makes it easier for employers to compare labor costs and the elimination of currency risk facilitates the establishment of branch plants in member states where labor is less expensive and workers are more productive.⁷

At the same time, the efficiency advantages of a single currency must be balanced against the disadvantages of a single monetary policy. Those disadvantages take the form of the sacrifice of policy autonomy that comes with moving from 10 or more separate monetary policies to a single level of interest rates across the euro-zone. Recall how in 1992-3, when some European countries feared inflation but others were preoccupied by unemployment, the group as a whole found it hard to agree on a one-size-fits-all monetary policy. Asymmetric disturbances, like the

⁷The benefits are likely to be considerable insofar as barriers to full price convergence appear to have persisted well into the 1990s. Indeed, the existence of different currencies appears to have been one of the principal obstacles to the emergence of a truly integrated market. This is demonstrated by Engel (this volume), whose results imply that a common currency could make a major contribution to the development of more fully integrated, and thus more efficient, goods markets in the European Union.

German reunification shock that was the source of the tension in 1992-3, can still occur under monetary union, but the member states now have no choice but to grin and bear them.⁸ A single monetary policy may then mean uncomfortably high inflation for some but uncomfortably high unemployment for others. The question for those seeking to gauge the costs of monetary unification is how frequently such asymmetric disturbances will occur.

The standard way of gauging whether a given correlation of shocks to different national economies is high or low is to compare it with the correlation of the same variables across the various regions within a functioning monetary union, typically the United States.⁹ By this measure, while the core members of the European Union (Germany, France, the Benelux countries and Denmark) are good candidates for monetary union, the same is less true of the EU periphery. The problem with this approach is that the shocks in question are likely to change with the advent of monetary union, rendering history a poor guide to the future. Demand shocks result from erratic changes in demand-management policy, whose formulation will be transformed by EMU. Asymmetric monetary shocks will disappear with the advent of a single monetary policy (although differences in the monetary transmission mechanism will remain). Asymmetric fiscal shocks will be limited by the Growth and Stability Pact. Supply shocks will

⁸The divergence between the euro "core" and "periphery" (Ireland and Iberia on the one hand, Germany on the other) in the opening quarters of Stage III was one indication that this dilemma was real.

⁹A prototypical study of the incidence of aggregate supply and demand disturbances affecting different candidates for EMU, with comparisons to the United States, is Bayoumi and Eichengreen (1993).

be transformed as Europe reorganizes itself to capitalize on the single currency and the Single Market.¹⁰

Other optimum currency area criteria are less likely to be endogenous with respect to the policy regime. Mundell observed that the costs of subjecting several separate economies experiencing asymmetric shocks to a symmetric monetary policy will be less when labor flows freely from depressed to booming regions. His intellectual descendants pointed out that a single monetary policy is similarly less problematic when wages adjust downward in high unemployment regions, obviating the need to relax monetary policy to fight unemployment. So far, there is little sign that the hardening of the EMS constraint and the transition to EMU have transformed these aspects of labor-market performance. The mobility of labor between EU member states remains low, reflecting deeply entrenched cultural and linguistic barriers. Wages remain rigid, reflecting the inheritance of strong unions and generous social programs. Where economic arrangements are embedded in social institutions, as in the case of labor markets, they are slow to change.

The implication for Europe is not a happy one. Insofar as these rigidities are both serious and slow to change, the costs of monetary unification may be considerable. A shock which raises unemployment in one EMU member state but does not elicit a reduction in interest rates

¹⁰Whether this last development makes supply shocks more or less asymmetric, and therefore raises or reduces the costs of monetary unification, is, unfortunately, uncertain. This is the "Krugman versus Krugman" debate, Krugman (1993) having argued both positions.

by the ECB, because it does not produce comparable unemployment elsewhere in the monetary union, may give rise to a problem of a chronically depressed region. This suggests that the politics of EMU may be more compelling than the economics, or at least that the decision to go ahead needs to be understood on political as well as economic grounds.

The Politics of Monetary Unification

A number of political factors are commonly adduced to help explain the course of European monetary integration. As a point of departure we distinguish inter-state bargaining and domestic distribution.

1. Inter-State Bargaining

Even if monetary union does not enhance the welfare of all countries, it still may be in the interest of some, which then cajole, coerce, or bribe others into participating. This approach, generally associated with what political scientists refer to as "intergovernmentalism," interprets observed outcomes as the result of strategic interaction among national governments.¹¹ Most of those who utilize this approach have in mind a process in which governments trade off objectives -- that is, they have in mind a form of "linkage politics." By linkage is meant the tying together of two (or more) otherwise unconnected issue areas, permitting the parties to an

¹¹For summaries and applications of "intergovernmentalism" in the EU context, see Moravcsik (1991), Moravcsik (1998), and Sandholtz (1993). Moravcsik (1998) uses a blend of the first and second approaches to analyze monetary integration, pages 238-313. Andrews and Willett (1997) discuss the analysis of international monetary relations both in general and with regard to Europe.

agreement to make concessions on one in return for concessions on the other(s). Thus, one country might "give" monetary union (which it does not favor inherently) in return for "getting" political union (which it does) if the perceived benefits of the latter exceed the costs of the former.¹²

This approach, while appealing, is not unproblematic. For one thing, it is easy to fall into a vague invocation of a link among policy areas without paying careful attention to governments' preferences; for years, journalists and others invoked unspecified "geopolitical" motives to explain bargains among EU members. Any analysis that relies on implicit links must explain how it is we know that these links exist. While many commentators have argued that full participation in the Single Market might be hampered by non-participation in the EMS, for example, there is no provision in the Single European Act or any other EU document explicitly establishing this tie.

Moreover, there is scope for the tradeoffs on which linkage arguments rely only when different nations place different values on different issues. If all EU members placed similar weight (positive or negative) on EMU, there would have been little room for trading off

¹²Early variants of this analysis spoke of the "spillover" from one area of European integration to another. This was one of the arguments associated with the "neofunctionalist" approach to regional integration, although the approach tended to focus even more on how integration would create or reinforce bureaucratic or social interests in further integration. Moravcsik (1998), Webb (1983), and Keohane and Hoffman (1991) discuss this and other perspectives; an early statement of the approach is Haas (1958).

concessions in different areas, and no room for linked bargaining that might improve the likelihood of agreement.

Finally, effective interstate bargaining requires that governments be able to make credible threats or promises. Otherwise they will fear that their foreign partners will renege on the commitment and refuse to enter into it in the first place. This is more problematic when issues are linked than when bargaining over each issue occurs in isolation, for not only must commitments on each dimension be credible, but the commitment to link dimensions must be credible as well.

Thus, while inter-state bargaining can be important and may involve linkage politics, its use in analysis requires caution and detail. The parties' goals must be specified and analyzed. And given the importance of credibility, particular attention must be paid to how the parties bind themselves to the linkages they create.

2. Domestic Distributional Issues

Just as countries attracted to EMU may bargain with other member states over participation, interest groups that stand to benefit or lose may play an analogous role domestically. While not benefitting a country as a whole, EMU may still enhance the welfare of particular groups, which prevail on their government to support it. EMU, in this view, is just one example of the special-interest politics common to virtually every economic policy arena.

Serious analysis of the distributional implications of EMU is scarce, although there is some suggestive work [such as Giovannini (1993) and Hefeker (1997)]. A few observations are

probably uncontroversial. Those for whom currency volatility is most costly stand to gain the most from EMU. They include banks and corporations with pan-EU investment or trade interests; for them forgoing national macroeconomic policy is a price worth paying for the elimination of currency risk. For those for whom cross-border transactions are inconsequential, on the other hand, predictable exchange rates are of little value, while national autonomy in the formulation of macroeconomic policies may be extremely important.

Many of the distributional concerns raised by EMU have had to do not so much with the desirability of a single currency per se as with the more immediate problems of adjusting macroeconomic policy to the requisites of a fixed exchange rate. In a high-inflation country, fixing the exchange rate typically leads to real appreciation, which puts pressure on producers of import-competing goods. This can cause a broad constituency to develop reservations about both fixed exchange rates and monetary union. In the context of EMU, because qualifying for participation required meeting the Maastricht fiscal criteria, those who worried about the impact of budget cuts or tax increases on them tended to resist making these sacrifices.

One implication of many distributional arguments is that support for EMU will be shaped by the rise of intra-EU capital mobility and trade. As the EU becomes more financially integrated, the choice between monetary policy autonomy and exchange rate stability becomes increasingly stark. Meanwhile, higher levels of intra-EU trade heighten the importance of exchange rate fluctuations for producers and consumers alike. Meanwhile, increased cross-border investment expands the ranks of those for whom exchange rate fluctuations created

problems. Inasmuch as the increased openness of EU economies has involved more economic agents in cross-border economic activity, and these firms and individuals care about reducing exchange-rate volatility, the drive toward the free movement of goods and capital might be expected to have strengthened support for EMU.

Problems with the distributional approach are not so much theoretical as practical. There is almost no empirical work which successfully measures the distributional effects of different international monetary regimes. Even if such work did exist, it would tell us little about outcomes, because interests are mediated by political institutions. Since institutions can magnify the political influence of some groups while diminishing that of others, similar interests may be expressed differently when, for example, parliaments are chosen by proportional representation than when members come from single districts in a first-past-the-post system.¹³ Thus, any rounded account of EMU must pay close attention to domestic political factors, specifically to the role of interest groups with strong views on EMU and how they operate within national political institutions.

While additional variables can undoubtedly be brought to bear, the two we have described would appear to be central. We need a clear picture of the interests at stake and the institutional setting within which they are situated. We then need to understand how national

¹³For a summary of such arguments and their application to the area of trade policy, see Rogowski (1987).

governments with divergent goals interact at the EU level, including an exploration of the ways in which EMU is linked to other EU policy areas.

The Political Economy of Monetary Unification in Practice

The domestic distributional effects of international monetary policies have been crucial for European monetary developments. Support for EMU has come from international banks and corporations with an interest in reducing currency volatility and deepening the integration of the European market, and from those in high-inflation countries who saw EMU as a way to achieve German-style monetary conditions. Opposition has come from domestically-oriented economic actors and those who expected to bear the brunt of macroeconomic austerity measures. This includes those in high-inflation countries where a fixed exchange rate and subsequent real appreciation undermined domestic producers' international competitiveness.

In France and Italy in the early 1980s, for example, opposition to policies aimed at sustaining the commitment to a fixed exchange rate came from workers in import-competing industries like steel and transportation equipment. In the early years of the Mitterrand government, France's commitment to the EMS was weakened by the resistance of Communists and left Socialists—whose constituencies were in declining manufacturing sectors hard hit by imports—to the austerity measures needed to bring French inflation down to German levels. A similar dynamic was evident in Italy, where the Communist party and its supporters in the labor movement—again concentrated in import-competing industries like steel -- were reluctant to agree to real wage reductions needed to keep the lira in line with its ERM partners.

The 1992-3 EMS crisis provides another example of how domestic political factors affected monetary integration, in this case by impeding the coordination of members states' macroeconomic policies. The British government might have raised interest rates to defend sterling except that the higher rates would have been passed on by mortgage lenders, and many within the ruling Conservative party worried about the objections of property owners. The Italian government might have enacted drastic fiscal measures to solidify its commitment to low inflation, but this was difficult to achieve over **the objections of public employees and others who felt threatened by the prospect of cuts.** The French government might have raised interest rates to defend the franc but was reluctant to pursue a policy that ran the risk of raising the country's already **high unemployment.** The German authorities might have loosened monetary policy in order to reduce pressure on their EMS partners but for the Bundesbank's preoccupation with inflation, which was reinforced by strong domestic anti-inflationary constituencies. Distributional considerations also help to explain the breadth of the euro zone as it was ultimately established. Once the 1992-1993 crisis was history, a new round of domestic debates began over the desirability of monetary union. It was argued that the single currency was indispensable for maintaining domestic political support for the Single Market. This logic went as follows.¹⁴ The more integrated European economies became, the more pronounced were the distributional consequences of intra-EU currency swings. With the completion of the Single

¹⁴This argument is elaborated in Eichengreen and Ghironi (1996).

Market, countries that depreciated their currencies would be able to flood other member states with exports – as happened after the southern European devaluations in 1992-1993. Import-competing producers' complaints about opportunistic exchange-rate grew more intense as European integration proceeded. Countries that fail to hold their exchange rates stable, aggrieved import-competing producers insisted, were not good Europeans. They needed to choose, it was said, between getting into the monetary union or getting out of the Single Market. And the Single Market was too significant an economic achievement to placed at risk. Note that the argument here was that the single currency was a necessary concomitant of the Single Market for political economy, not narrowly economic, reasons.

Yet the timing and character of the milestones in European monetary integration remain difficult to explain on the basis of interest-group politics alone. They were almost certainly heavily influenced by inter-state bargaining. For example, it is observed that the Italian government in 1979 was not enthusiastic about the EMS but was presented with a *fait accompli* by the French and Germans. It is similarly argued that the French government in the 1990s was eager to use EMU to reduce German influence over European monetary policy, while Germany was less ardent in its pursuit of EMU.

The core of the complex bargaining process that has characterized European monetary integration has been the threat or fear that a country refusing to go along with EMU might bear costs on other dimensions of EU policy. In other words, the bargaining relationship involved explicit or implicit links between EMU and other issue areas. Linkages are indeed central to

what Garrett (this volume) has in mind when he argues that the German government went along with EMU—which it found unattractive on the merits—in return for assurances that the EU would move forward on political matters, especially a common foreign policy. Martin (this volume) also emphasizes institutionalized linkages. The explanation of French and Italian commitments to the EMS in the early 1980s presented by Frieden (this volume) explores the domestic political impact of such linkage effects. And many analysts have pointed to the connection between the EU's search for exchange rate stability and its **Common Agricultural Policy (CAP)**, the operation of which was severely complicated by fluctuations in EU currency values.¹⁵

Thus, both domestic politics and inter-state bargaining help us understand why the momentum for EMU was sustained, and why it was not initially sufficient to bring Denmark, Sweden, and the United Kingdom into the fold. Domestic support for monetary unification grew over the course of the 1980s and early 1990s, as continued economic integration created

¹⁵The CAP is far too complex to explore in detail here. Suffice it to point out that in the context of major agricultural subsidies, the EC sets Community-wide food prices. When a currency is devalued, the EC reference price would normally be raised in the devaluing country to counterbalance the devaluation—thus "passing through" the exchange rate change in full to food prices. The inflationary impact of this pass-through would mitigate the devaluation's attempt to restore price competitiveness in non-agricultural sectors. For this reason, the Community devised a series of compensatory arrangements and accounting exchange rates. For our purposes, what is important is only that exchange rate fluctuations complicate Community agricultural policy by changing compensatory farm payments in ways that could disrupt the delicate balance within the EC on farm policy. For more details see McNamara (1998).

heightened and widened the interest in European-wide markets and policymaking institutions. Some of the effects of the 1992-1994 crisis, such as the impact of "competitive devaluations" on Northern European producers, strengthened the resolve to work for as rapid and encompassing an EMU as possible. Meanwhile, the growing importance of EU-wide decisions on trade, regulation, social policy, and even foreign policy made governments loathe to risk being second-class European citizens. By the same token, it is not surprising that the three countries that are least enthusiastic about European federalism – the United Kingdom, Denmark, and Sweden – balked at the single currency. Domestic political considerations and inter-state bargaining thus worked together to bring Europe's monetary union into being in 1999.

The Future of EMU

What picture do these perspectives paint of the future of EMU? From a narrowly economic viewpoint, high capital mobility, which will undoubtedly be fact of economic life in the 21st century, effectively leaves most countries only the choice between adopting a flexible exchange rate and pegging it once and for all. The growth and articulation of international financial markets, which is ongoing, suggests that countries that cannot live comfortably with a flexible exchange rate -- small countries with open economies highly dependent on international trade and financial flows in particular -- will be attracted by a policy which promises currency stability. Some countries, in Eastern Europe for example, may solve this problem by joining the EU and its monetary union. The implication is that Europe's monetary union will gain a constituency of faithful members who see EMU as a zone of monetary and financial stability and

their only viable alternative to the unpalatable option of floating rates.

What of countries like the UK, residents of which are apparently happy to participate in the Single Market, but among whom there remains deep and abiding skepticism about the single currency? A political-economy perspective suggests that this skepticism is rooted in the low level of British commitment to European political integration. For many in the UK, there is a fear that the EU's much-vaunted spillovers and linkages suggest that "one damn thing leads to another" — in this case, that monetary integration will reinforce the movement toward political integration. The decision for the UK may therefore be whether to join Europe's monetary union, and thereby to sign on to the larger project of creating a politically as well as economically integrated Europe, or else to withdraw from the European Union. Given that neither of these options is particularly appealing, one appreciates why the British see themselves as facing a dilemma.

All this assumes that EMU is here to stay. What of the scenario where it falls apart? Feldstein (1997) observes that EU member states with very different preferences are shackled together by a single monetary policy. There is no political union at the outset. National leaders will continue to plump for policies that reflect the differing preferences of their constituencies. Inevitably, some will be disappointed. **And since Europeans are unlikely to agree to the extension of large-scale cross-border transfers prior to the creation of a true political union, there is no way of compensating the losers.** Disagreement over the stance of monetary policy could

then mean serious dispute. Exit, not just voice and loyalty, is one conceivable way for the disaffected to respond.

Technically, exiting the monetary union is straightforward: the government of the participating member state needs only to restart the printing press and reissue the national currency.¹⁶ If a country left the monetary union because it felt that the ECB was following excessively inflationary policies, its "good" domestic currency would drive out the "bad" European currency. If it country instead left because it felt that the ECB's overly restrictive policies were aggravating unemployment, it would in addition have to declare that the euro would no longer be accepted as legal tender within its borders.

But **the decision to continue to support or choose to abandon the monetary union will, like its predecessors, be made on political economy, not economic, grounds.** The principal interstate bargaining consideration is that a country that withdraws from the monetary union could face considerable costs in the form of its standing in the EU on other fronts. Just as issue linkage contributed to the establishment of Europe's monetary union, issue linkage stands in the way of its demise. The principal domestic distributional consideration is that support for EMU is in fact substantial in its member countries, and likely to grow over time. Creating a common currency leads to the formation of groups and lobbies with an interest in its persistence. Investments are planned and made, contracts are signed, firms expand and merge, all with the Single Market and

¹⁶Our discussion of the exit question follows Eichengreen (1998).

the single currency in mind. The economic realities thereby created generate political pressure to maintain the economic policies upon which they were built. As European banks and firms become more European, their stake in the single currency becomes greater, in turn lessening the likelihood that monetary unification will be reversed subsequently.

About This Volume

The contributors to this volume analyze European monetary integration from both the political and economic points of view. The first chapter, by Jeffrey Frieden, considers the experiences of France and Italy in the early years of the EMS. Frieden argues that two factors were central to the politics of exchange-rate policy in these countries. First were the distributional effects expected to result from binding the franc and the lira to the EMS. Import-competing manufacturing sectors opposed fixing exchange rates in ways that would erode their ability to compete with foreign producers, while internationally oriented sectors favored currency stability. Second was the linkage constructed between the EMS and the other aspects of European integration. This forced interest groups in both countries that were ambivalent or hostile to the EMS to weigh these concerns against their general attitude toward the EU, and in many cases to go along with monetary policies that they otherwise might have tried to block.

The second chapter, by Matthew Gabel, investigates the domestic political bases of support for EMU in more detail. Gabel considers how socio-economic status affects sentiment toward monetary union. He confirms that members of socio-economic groups that expected to

gain most from the reduction in currency volatility and the removal of barriers to cross-border trade and finance tended to be most enthusiastic.

Even if special interests are important, it would still be puzzling that domestic political support for EMU was so strong if, as much economic analysis has insisted, monetary union has few economic benefits in and of itself. In his chapter, Charles Engel presents evidence that monetary unification might not entail the economic sacrifices many expect. He shows that local prices have not responded quickly to exchange rate movements, which means that the exchange rate has not been an effective tool of macroeconomic policy. To the extent that local-currency pricing mitigates the ability of currency policy to affect national economies, giving up an independent currency may involve fewer costs than is usually anticipated.

The next pair of chapters, by Geoffrey Garrett and Lisa Martin, examine how inter-state bargaining made EMU possible and affects its future. Garrett analyzes the Maastricht Treaty as a bargain between Germany, a country standing to gain little from monetary union per se, and the rest of Europe. EMU must be understood, he argues, as a trade across issue areas. Germany made significant concessions on matters monetary, according to Garrett, in return for general European support for German unification, as well as for a deepening of trade and investment ties with the rest of the EC.

Martin insists that the EMU debate be viewed as embedded in an institutionalized pattern of inter-state cooperation. The same set of countries is engaged in negotiations over defense, social policy, competition policy and, of course, monetary policy, and the way those negotiations

are structured and institutionalized means that failure to reach agreement on one issue places agreement on the others at risk. The fact that countries stand to lose all the gains they have made from cooperation on a variety of issues has thus tended to lock in agreement on monetary matters.

The role of these institutional arrangements is key, Martin argues, to understanding the outcomes thereby achieved. For example, the disproportionate influence of small states in the governance of the monetary union must be understood as a function of formal decision-making procedures such as unanimity requirements used to solidify cross-issue links.

In the next chapter, Paul DeGrauwe, Hans Dewachter, and Yunus Aksoy provide a detailed analysis of one such arrangement: the decision-making structure of the European Central Bank. They show how the monetary policy of the ECB will depend on the structure of EMU institutions and on whether the members of the ECB board remain primarily national in orientation or develop a more pan-European perspective. Only if a more European view prevails, they argue, will EMU improve conditions in the Euro area.

Their analysis, like most of the literature, takes for granted that monetary union is irreversible. The final chapter, by Benjamin Cohen, challenges this assumption. Contrasting the experience of three surviving monetary unions (the CFA Franc Zone, the East Caribbean Currency Area, and the Common Monetary Area) with three that failed (the East African Community, the Latin Union and the Scandinavian Union), Cohen asks whether economic, political or organizational factors explain success or failure. He concludes that appropriate

economic conditions (high factor mobility, symmetric shocks) and organizational structures (an independent central bank, legal provisions for currency issue) have not always been sufficient to hold a monetary union together; rather, political factors broadly defined are key. The presence or absence of a dominant power or sufficient political cohesion are needed to sustain monetary union over time.

Conclusions

The essays in this volume testify to the complexity of Economic and Monetary Union in Europe and to the need for an integrated political economy approach to understanding it. Macroeconomic analyses help clarify the economic costs and benefits of monetary integration. The benefits include reductions in transactions costs and in the uncertainty associated with cross-border trade and payments. The costs are principally due to the loss of a policy instrument with which to respond to unexpected shocks. Political analyses point to the role of domestic interests and to institutionalized relationships among EU member states. The distributional concerns of interest groups have powerful effects on the political constraints facing national governments and monetary policymakers. But these effects operate within an environment of ongoing negotiation among the European governments engaged in the search for mutually acceptable agreements on the road to broader and deeper integration.

These economic and political factors will continue to be central to the evolution of Europe's economic and monetary union. They will determine its membership – whether any current members leave the Union, which of the four EU member states that are not in EMU join

it, and whether candidate members in Central and Eastern Europe abandon their currencies for the euro. They will have a powerful impact on the policy of the European Central Bank as it establishes itself as the monetary authority for Europe and as one of the world's three principal central banks. For all these reasons, the political economy perspectives presented here help explain the past and understand the future of European monetary integration.

References

Andrews, David and Thomas Willett, 1997, Financial Interdependence and the State: International Monetary Relations at Century's End. *International Organization* 51, No. 3, 479-511.

Bayoumi, Tamim and Barry Eichengreen, 1992, Shocking Aspects of European Monetary Unification, in: Francisco Torres and Francesco Giavazzi, eds., *The Transition to Economic and Monetary Union in Europe* (Cambridge University Press, Cambridge).

Coffey, Peter, 1987, *The European Monetary System—Past, Present and Future* (Kluwer, Amsterdam).

Cohen, Benjamin J., 2000, Beyond EMU: The Problem of Sustainability. This volume.

Eichengreen, Barry, 1992, Should the Maastricht Treaty Be Saved? *Princeton Studies in International Finance* (International Finance Section, Princeton).

Eichengreen, Barry 1993, Thinking about migration: Notes on European Migration Pressures at the Dawn of the Next Millenium, in Horst Siebert, ed., *Migration a Challenge for Europe* (University of Michigan Press, Ann Arbor, Michigan).

Eichengreen, Barry, 1998, European monetary Unification: A Tour D'Horizon, *Oxford Review of Economic Policy* 14, pp.24-40.

Eichengreen, Barry and Fabio Ghironi, 1996, European Monetary Unification: The Challenges Ahead, in Francisco Torres (ed.), *Monetary Reform in Europe*, Lisbon: Catholic University Press, pp.83-120.

Eichengreen, Barry, and Charles Wyplosz, 1993, The Unstable EMS. *Brookings Papers on Economic Activity* 1, pp. 51–143.

Feldstein, Martin, 1997, EMU and International Conflict, *Foreign Affairs* 76, pp.60-73.

Fratianni, Michele, and Jürgen von Hagen, 1992, *The European Monetary System and European Monetary Union* (Westview, Boulder, Colo.).

Frieden, Jeffry, 2000, Making commitments: France and Italy in the European Monetary System 1979–1985. This volume.

Garrett, Geoffrey, 1992, International Cooperation and Institutional Choice: The European Community's Internal Market. *International Organization* 46.

Garrett, Geoffrey, 2000, The Politics of Maastricht. This volume.

Giovannini, Alberto, 1993, EMU: What Happened? Exploring the political dimension of optimum currency areas, in: Guillermo de la Dehesa, Alberto Giovannini, Manuel Guitián, and Richard Portes, eds., *The Monetary Future of Europe* (Centre for Economic Policy Research, London).

Goodman, John, 1992, *Monetary sovereignty: The politics of central banking in Western Europe* (Cornell University Press, Ithaca, N.Y.).

Gros, Daniel, and Niels Thygesen, 1998, *European Monetary Integration* second edition (Longman, London).

Haas, Ernst, 1958, *The uniting of Europe* (Stanford University Press, Stanford, Calif.).

Hefeker, Carsten, 1997, *Interest Groups and Monetary Integration: The Political Economy of Exchange Regime Choice* (Westview Press, Boulder, Colo.).

Keohane, Robert, and Stanley Hoffman, 1991, Institutional change in Europe in the 1980s, in: Stanley Hoffman and Robert Keohane, eds., *The New European Community* (Westview Press, Boulder, Colo.).

Krugman, Paul, 1993, Lessons from Massachusetts for EMU, in Francisco Torres and Francesco Giavazzi (eds), *Adjustment and Growth in the European Monetary Union*, Cambridge: Cambridge University Press, pp.241-261.

Ludlow, Peter, 1982, *The making of the European Monetary System* (Butterworth, London).

Martin, Lisa, 2000, International and domestic institutions in the EMU process. This volume.

McKinnon, Ronald, 1963, Optimum currency areas. *American Economic Review* 53, 717-725.

McNamara, Kathleen, 1998, *The currency of ideas : monetary politics in the European Union* (Cornell University Press, Ithaca).

Moravcsik, Andrew, 1991, *Negotiating the Single European Act: National interests and*

conventional statecraft in the European Community. *International Organization* 45, 19–56.

Moravcsik, Andrew, 1998, *The Choice for Europe* (Cornell University Press, Ithaca).

Mundell, Robert, 1961, A Theory of optimum currency areas. *American Economic Review* 51, 657–665.

Rogowski, Ronald, 1987, Trade and the variety of democratic institutions. *International Organization* 41, 203–224.

Sandholtz, Wayne, 1993, Choosing union: monetary politics and Maastricht. *International Organization* 47, 1–39.

Tsoukalis, Loukas, 1977, *The politics and economics of European monetary integration* (George Allen and Unwin, London).

Webb, Carole, 1983, Theoretical Perspectives and Problems, in: Helen Wallace, William Wallace, and Carole Webb, eds., *Policymaking in the European Community* 2nd edition (John Wiley and Sons, Chichester) 1–42.

Why Germany Wanted EMU: The Role of Helmut Kohl's Belief System and the Fall of the Berlin Wall

FEMKE VAN ESCH

In light of the current Euro-crisis and seemingly erratic German response to it, the question of why Germany wanted EMU has gained new relevance. While EMU entailed economic and political benefits for Germany, this article indicates that taking into account the German shared historical consciousness, and the combined effect of Chancellor Kohl's beliefs and 1989 Eastern European revolution provides a more adequate answer to this question. More specifically, such an explanation provides an answer to remaining questions concerning the timing of the German acquiescence, and the ambiguous and inconsistent behaviour of the Kohl government on the road to Maastricht.

Gelegentlich braught man größere Ereignisse, damit etwas gelingt, klar.

[There are times when major incidents are just what is needed for something to succeed].

(Horst Teltschik)

INTRODUCTION

Today's sovereign debt crisis represents without doubt the most significant crisis the European Union has faced in decades. The effects will be felt throughout Europe for years to come. In addition, it is also proving to be a major test of the robustness of European Economic and Monetary Union (EMU), the decisiveness and power of its institutions and the true level of solidarity among the members of the eurozone. For now, it is uncertain whether EMU will collapse under the pressure – as some commentators have predicted – or whether this crisis is a golden opportunity to improve its workings and to reinforce cross-European solidarity, as more optimistic voices expect.

However, one aspect of the European response to the crisis has already been identified in the public debate as rather perplexing: the attitude and actions of its largest member state, Germany, and the role of its leader therein. The German Chancellor, Angela Merkel, has been accused of acting in an irrational, ambiguous and Euro-sceptic manner, and thereby recklessly jeopardising the survival of the euro.¹ This apparently erratic attitude raises the question of why Germany ever agreed to the establishment of EMU in the first place. All the more so, since Merkel's shifts in attitude are highly reminiscent of the shifting and ambiguous position of the Kohl government in the run-up to the Maastricht Treaty. In view of this, shedding light on the question why

Germany wanted EMU, as well as explaining the ambiguous position of the Kohl government in the years preceding Maastricht has become more relevant than ever.

Over the years, several economic and political explanations for Germany's support of EMU have been put forward.² However, scholars generally agree that none of these explanations can fully explain the German position on EMU:³ questions concerning the ambiguity and shifts in Germany's stance during the European monetary unification process and the timing of the German decision to agree to EMU, after years of reticence, in particular, could use further clarification.

This article will argue that in order to get a better understanding of why Germany wanted EMU – as well as the timing and ambiguity of its decisions on the road to Maastricht – one must take into account the inherently dual nature of Germany's collective history when it comes to European economic and monetary issues.⁴ More importantly, the article argues that the combined effect of Chancellor Kohl's beliefs and the fall of the Berlin Wall played a crucial role in tilting the balance towards a German preference for EMU, making a European agreement possible.⁵

More specifically, the evidence presented in this article shows that the revolutions in Eastern Europe in 1989 set two processes in motion. Firstly, rather than change, the glimmer of hope for German reunification strengthened Kohl's pre-existing pro-European beliefs. As a result, after several years of being caught in the middle of a domestic battle between pro-Europeans and *ordo-liberals*, he finally found the motivation, legitimacy and political clout to throw his full weight behind the EMU project. This led him to meet public resistance to EMU head on and overrule the sceptical German financial elite on several crucial points in the EMU negotiations which made agreement possible.

WHY GERMANY WANTED EMU: STATE OF THE ART

Over the years, many students of EMU have sought to explain why Germany agreed to give up its cherished Deutsche mark and Bundesbank in exchange for the euro and the European Central Bank (ECB). Firstly, a number of political economists have stressed that producers in export-oriented, low-inflation countries, such as Germany, prefer institutional arrangements such as EMU that provide stable exchange rates, and reduce exchange rate risks and transaction costs so as to increase their competitiveness.⁶ Moreover, some argue that EMU provided a next step in Europe's long struggle to cope with the destabilising effect of dollar depreciations on the competitiveness of German exports under flexible exchange rates.⁷ As an export-oriented economy, Germany did indeed have an interest in the fixed exchange rates that EMU would provide, but scholars also generally agree that economic motives alone cannot fully explain Germany's reversion to a positive stance on EMU in the late 1980s. For starters, empirical research shows that during the early 1980s, when Germany's relative inflation rates were low, its response to plans for the establishment of EMU was negative, but became more positive in the late 1980s when its relative inflation rates were actually on the rise and 'to some extent, U.S. policy reverted to a more cooperative stance'.⁸ Secondly, little evidence can be found of major German producer groups who supported EMU prior to the Delors Report, or pressurising the German government.⁹

A more political explanation is grounded in hegemonic stability theory. Some authors have argued that EMU was the next step in the German (or Franco-German)

quest for European monetary hegemony,¹⁰ while others have argued that EMU constituted a German effort to release itself of its unwanted hegemonic position under the European Monetary System (EMS), which antagonised its European partners.¹¹ Apart from the obvious problem of generating contradictory expectations, the first explanation raises the question of whether Germany was large and dominant enough to be considered a European hegemon and be able to pressure other states into joining EMU. Moreover, evidence suggests that in fact it was the Germans who were among the most sceptical and had to be convinced to join. In addition, both explanations struggle to explain the timing of the German reversion to a more positive stance to EMU. From the German perspective, the EMS was at that point functioning satisfactorily as a *de facto* DM-zone with the Bundesbank calling the shots.¹² Dissatisfaction among the other member states about Germany's dominant position had existed since the EMS had been set up – so why did it take until the early 1990s before EMU was established?

The third and final angle from which we can explain Germany's attitude is ideational. It can be argued that ideas on the proper conduct of economic and monetary policy were as crucial in building support for EMU as material factors. From this perspective, from the early 1970s onwards, a consensus on traditional German economic values like price stability, budgetary restraint and Central Bank independence spread throughout Europe. This *ordo-liberal* consensus, which was disseminated by a transnational group of Central Bankers and the experience of monetary co-operation within the EMS,¹³ made further monetary unification amongst member states possible.¹⁴

While it is clear that German ideas about 'sound money' have gained a larger following since the 1970s, this argument seems to overestimate the level of consensus on EMU. For instance, while the issue of an independent ECB was settled early on in the negotiations, the level of budgetary restraint demanded by the German financial elite has always remained a contentious issue. This line of argument also underestimates the reluctance of the German financial elite, and the criticism they directed at EMU throughout the process of monetary union.

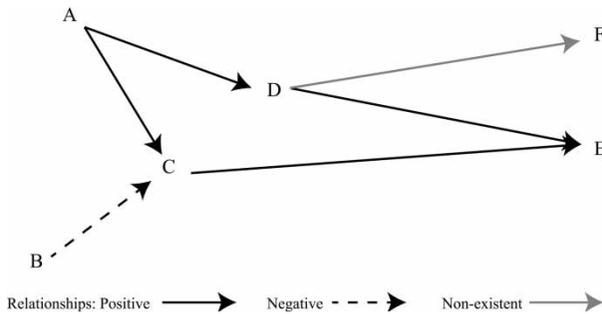
Overall, we can conclude that EMU did offer Germany certain political and economic benefits, and that consensus on *ordo-liberal* economic ideals existed across a section of the Europe's decision-making elite. However, these factors do not provide a conclusive explanation of the German stance on EMU. Two questions in particular deserve closer examination. Firstly, as in the current crisis, the German position in the run-up to Maastricht was puzzlingly ambiguous. Secondly, our understanding of the timing of events remains anecdotal and therefore unsatisfactory.

RESEARCH METHODS

In order to answer these questions, both the evolution of Germany's position on EMU and Kohl's position on further European monetary unification *before* the summer of 1989 and *after* the fall of the Berlin Wall will be studied. To understand the development of the German negotiating position, the bargaining process was traced using primary documents and secondary literature, while the development of the Chancellor's economic and European beliefs was analysed using cognitive mapping.

Cognitive mapping is one of the best developed and systematic methods for studying decision-makers' beliefs used in political psychology.¹⁵ In order to create a

FIGURE 1
EXAMPLE OF A COGNITIVE MAP



cognitive map, a number of (public) speeches, statements or writings are analysed to derive all the causal and normative relationship alluded to in these statements. These relations are subsequently transformed into a graphic map displaying these relationships as arrows between two concepts. The graphic map thus gives a structured overview of which concepts (events, policies, goals) are causally or normatively linked in the mind of the decision-maker (see Figure 1).¹⁶

Moreover, the policy preferences of a decision-maker may be ascertained from the maps by analysing the 'consequent' path from a certain policy to a goal or event: policy or event A in Figure 1, for instance, is valued negatively by the decision-maker, because the 'consequent path' from policy A to goal E shows that it contributes positively to the negative event or goal E.¹⁷ Finally, cognitive mapping allows the researcher to determine the (relative) saliency and centrality of the mentioned concepts in the mind of the decision-maker by analysing the frequency with which it is mentioned, and the number of dyads it is part of (see Tables A1 and A2 in Appendix 2).

To study Kohl's beliefs and trace the developments therein, two cognitive maps were created, each containing up to forty-six concepts and ninety causal or normative relationships. The first map is derived from statements and writings dated before the publication of the Delors Report, and the second map is based on statements from after the summer of 1989.¹⁸ By comparing these maps, changes that took place in Kohl's beliefs are revealed. This article focuses on Kohl's beliefs pertaining to European integration generally and EMU in particular. In addition, his beliefs on three dimensions (dichotomies) central to the decision-making process concerning EMU are analysed: (1) his preferred mode of integration (supranational¹⁹ or intergovernmental);²⁰ (2) his macro-economic convictions (ordo-liberal²¹ or Keynesian);²² and (3) the prioritisation of the economic²³ or political goals²⁴ of integration.

GERMANY'S JANUS-FACE: EUROPATRIOTISM AND SOUND ECONOMICS

To understand the effect and background of Kohl's beliefs, we need to consider the two-faced character of the German position on European monetary integration. This ambiguity can be traced back to two competing and deep-rooted convictions in Germany's collective history: a belief in 'sound' ordo-liberal economic policies on the one hand, and a need to assert itself as a 'good European' on the other hand.²⁵

The first collective memory that influenced the German position on EMU is that of hyperinflation during the inter-war period. Apart from causing great misery amongst the population, the popular German view is that this crisis facilitated the rise to power of the Nazi regime. These experiences led to a preoccupation with inflation and a strong D-Mark in the German collective memory. The post-war strength of the D-mark and the Bundesbank's success in keeping inflation low turned the D-mark and the Bundesbank into symbols of stability and prosperity in the public mind.²⁶ This discourse of *D-mark Patriotism* prioritises the economic goals of European integration and explains the German insistence that the ECB be modelled along the lines of the Bundesbank.

The second important aspect of German collective history was the horror of the Nazi years, and the strained relations with France that this had led to. Under the guidance of the first post-war Chancellor, Konrad Adenauer, the Germans actively began to distance themselves from their own history and to reconstruct their collective identity by turning to Europe as the embodiment of stability, peace and democracy.²⁷ To prevent the country from ever falling back into its old ways, Adenauer developed a policy of reconciliation with France and the *Westbindung* of Germany. According to this line of thinking, the political goals of European integration were more important than Germany's economic goals. By the late 1960s, the discourse of *Europatriotism* had become a broad consensus on the desirability of further European integration, both among the German public and across political parties.

Over time, these two strands of shared German historical memory have influenced policy making in many issue areas, providing a powerful moral compass in German policy making. In the case of EMU, however, the policies inspired by *Europatriotism* clash with those inspired by *D-Mark Patriotism*.²⁸ This is because closer European monetary integration implies sharing the authority to set monetary policy with states that placed a lower value on financial austerity and central bank independence.

The contradiction between these two policy norms resurfaced regularly in the form of rifts and disagreements between their respective adherents.²⁹ The pro-European political elite including the *Bundeskanzler-* and *Auswärtiges Amt*, and the sceptical German financial elite in the Bundesbank and Ministries of Financial and Economic Affairs.³⁰ More importantly, this clash also played a role in the events leading up to the establishment of EMU and the attitude of Helmut Kohl, and partly explains the ambiguity of the German position in the run-up to Maastricht.

GERMANY'S POSITION ON EMU IN THE EARLY 1980S

Given the delicacy of the issue in German domestic politics, Kohl – as a student of history – was well aware that any bold steps towards European monetary unification could result in a stand-off with the country's financial elite and may well lead to electoral problems. So during his early years in office, Kohl abided strictly by an implicit division of labour under which the financial elite were free to determine the German position on European monetary affairs, while he was permitted to engage Germany in further European integration on other issues.³¹ However, this division of labour came under increasing pressure from the French and the European Commission, who started taking initiatives for further European monetary integration, and a reform of the EMS.

Between Preferences and Politics: Der Dicke in a Tight Spot

The first strains in this division of labour became apparent during the Intergovernmental Conference (IGC) on the Single European Act (SEA). Though Kohl was a strong supporter of the SEA, he was publicly critical of the inclusion of a monetary dimension and sided with the highly sceptical Bundesbank. However, during the subsequent European Council, away from the watchful eyes of the Bundesbank and Ministries of Economics and Finance, Kohl followed his naturally pro-European inclinations, giving in to French pressures and agreeing to include the establishment of EMU as a goal in the preamble of the Act.³²

Moreover, when the French subsequently started calling for the reform of the EMS, Kohl agreed to expand the Franco-German Co-operation Treaty to include an economic council that would co-ordinate their respective monetary policies. That agreement was, however, successfully blocked by the Bundesbank. Undeterred, the French renewed their call for the reform of the EMS. Again their plans were heavily criticised by the German financial elite and Chancellor Kohl also 'registered his "coolness" toward the plan almost immediately'.³³

Foreign Minister Genscher, however, spoke out in support of the French proposals and issued his own plan for the establishment of full EMU. The Genscher Plan was careful to observe the line of the German financial elite, emphasising price stability and the independence of the European Central Bank, but nevertheless they fiercely opposed the proposals.³⁴ There was speculation that privately Kohl had agreed to the principles put forward by Genscher, but he refused to support the plans publicly.³⁵ As acting President of the European Council, however, Genscher was able to overrule the sceptical voices and place his proposals high on the agenda of the European Council in Hanover of June 1988.

The Hanover summit was largely devoted to the discussion of the Genscher Plan. Kohl, with the recent experience of the Franco-German Economic Council fresh in mind, was looking for a compromise that would also be acceptable for the 'zurückhaltenden Stimmen in Deutschland' [reticent voices in Germany].³⁶ Ultimately, it was decided to instate a Committee of Central Bank Governors who in a personal capacity would examine the question of how to achieve EMU. The provision that Commission President Delors would preside over the Committee liberated Kohl 'aus der Rolle des Schaniers zwischen Währungsexperten und drängenden EG-Partnern' [from his role as mediator between the monetary experts and pressuring EU-partners].³⁷

Despite the fact that there were furious arguments over the pace of, and preconditions for the process of monetary unification, and the budgetary rules that would apply,³⁸ the Delors Committee managed to agree a report on time. The report largely reflected the preferences of the German financial elite. Even so, the reaction of the German Ministry of Economics still was outright hostile and President of the Bundesbank, Karl-Otto Pöhl immediately went on record that the negotiations on EMU had yet even to begin and that nothing had been decided yet.³⁹ All in all, without significant changes in the political landscape, the chances of EMU ever getting off the ground seemed slim in late 1988.

The Politics of Deference

An analysis of his assertions reveals that Kohl's wavering on the monetary issue constituted a hard choice between political tactics – maintaining party and electoral support and avoiding a clash with the mighty Bundesbank – and his pro-European convictions.

As the first cognitive map demonstrates, Kohl remained, at heart, a proponent of far-reaching European integration as well as the establishment of EMU (see Map 1, author's website).⁴⁰ From a very young age, Kohl had held well-defined and deeply felt convictions on European politics,⁴¹ which are reflected in his first cognitive map. It is clear that European integration played a pivotal role in Chancellor Kohl's belief system. Moreover, his outlook on European integration was very positive.⁴² In his eyes, it would serve the interests of all member states, as well as a whole range of economic goals including the central economic goal of price-stability (P1). In addition, he consistently valued the establishment of EMU (E6) as a positive and relatively important goal.⁴³ As for the Chancellor's preferred mode of integration (dimension 1), it is clear that he had no objections to either intergovernmental or supranational arrangements. However, even though in his eyes European unification required Germany to make sacrifices (M2–G3), Kohl slightly preferred supranational integration over intergovernmental.⁴⁴ See Table 1.

TABLE 1
DOMINANCE OF AND CHANGES IN KOHL'S BELIEF SYSTEM

Dimension	Centrality & saliency	Map 1		Map 2	
		C	S	C	S
1a) Intergovernmental integration	Number of concepts	4	4	3	3
	Aggregate C/S	18	20	10	14
	% of map C/S	9.9	7.5	5.6	5.2
b) Supranational integration	Number of concepts	4	4	4	4
	Aggregate C/S	21	25	34	65
	% of map C/S	11.5	9.4	19.0	24.3
2a) Keynesian economic policies	Number of concepts	3	3	3	3
	Aggregate C/S	7	16	7	8
	% of map C/S	3.8	6.0	3.9	3.0
b) Ordo-liberal economic policies	Number of concepts	4	4	6	6
	Aggregate C/S	22	38	18	32
	% of map C/S	12.1	14.3	10.1	11.9
3a) Priority of the political	Number of concepts	6	6	6	6
	Aggregate C/S	17	25	33	43
	% of map C/S	9.3	9.4	18.4	16.0
b) Priority of the economical	Number of concepts	6	6	2	2
	Aggregate C/S	25	39	4	8
	% of map C/S	13.7	14.6	2.2	3.0
Map total	Number of concepts	46	46	46	46
	Aggregate C/S	182	266	179	268
	% of map C/S	100	100	100	100

Note:

C Centrality

S Saliency

X Dominant belief

On the second, macro-economic, dimension, Kohl's preferences were more pronounced. For someone who reportedly experienced 'an intellectual discomfort with technical economic arguments',⁴⁵ Kohl devoted a great deal of his speeches and writings to economic issues. The map reveals that he internalised the main tenets of the German ordo-liberal outlook on economic and monetary policies.⁴⁶ In agreement with the German financial elite, Kohl was convinced of the need to maintain price-stability (P1), and strict government policies (-L1).

Kohl also had a clear, but less marked preference on the dimension of the priority of the economic or political. Overall, the economic goals of European integration were more central and salient to his belief system, and the ordo-liberal goal of price-stability (P1) ranked as the most important. In political terms, Kohl saw European integration as a vital means of Franco-German reconciliation (F1/F2).⁴⁷ More importantly, however, as early as the early 1980s, Kohl held the belief that a reciprocal relationship exists between German reunification and European integration: German reunification would stimulate European integration (E5-P4; E9-P4) and vice versa (P4-E5; P4-E9). In fact, German reunification was the most central and salient political goal of European integration in his eyes.⁴⁸ In addition, when looking at the cognitive map as a whole, it is clear that, like Adenauer, Kohl ultimately viewed European integration as a political project that served peace amongst nations and the eventual reunification of Germany.⁴⁹

All in all, we can conclude that in line with his Europocentric and Paris-oriented views, 'Kohl was always pro-EMU in principle'.⁵⁰ While Kohl agreed with the ordo-liberal policy position of the German financial elite, this analysis suggests that Kohl's public attitude towards EMU was based on a tactical evaluation of the internal political power configuration within Germany. Until early 1989, Kohl was content to accept the priority of the economic in European monetary affairs but to refrain from meddling in decision making on EMU so as to avoid a political clash with the Euro-sceptical German financial elite and a possible loss of electoral support.⁵¹

THE FALL OF THE WALL AND THE POLITICISATION OF EMU

In May 1989, as EU member states were preparing for the forthcoming discussion on the Delors Report, political changes began taking place in Eastern Europe that would change the face of the whole continent.⁵² The revolutions of 1989 had major repercussions for the European integration project. While in early 1989, Kohl had still been dragging his feet on the European monetary issue, the events in the summer of 1989 had the peculiar effect of strengthening Kohl's pro-EMU beliefs. Moreover, it politicised the issue, dealing the final blow to the already faltering intra-German division of labour, and providing a decisive impetus for the establishment of EMU.

Invoking the Legacy of Adenauer

Even before the Wall had come down, the events unfolding in Eastern Europe affected Chancellor Kohl. With the German Question back on the international agenda, Adenauer's concept of *Westbindung* resurfaced in his mind, enhancing the significance of the European integration project. As a result, Kohl began voicing his conviction that German reunification and European unification were 'two sides of the same

coin' more frequently.⁵³ Moreover, while the rest of the world stood hesitantly by, watching history unfold, the German Chancellor saw the changes underway in the international arena as a window of opportunity. It gave him a discourse and granted him the domestic legitimacy to follow his pro-European convictions in the monetary issue area, to defy the German financial elite and to justify his actions to the voters.

The first issue on which Kohl defied the wishes of the German financial elite was the start of the IGC on EMU. During early summer 1989, Kohl had refused to set a start date for the IGC,⁵⁴ but by August the drastically altered international and domestic political context allowed him to follow his heart. Behind closed doors, he informed his European partners of his wish irrevocably to bind Germany to the Community by means of EMU.⁵⁵ He also confided to the French President, François Mitterrand, that since certain forces 'in the FRG were campaigning against EMU', he wished to prepare the forthcoming summit in close co-operation with France, and had instructed his ministers to refrain from discussing EMU. Finally, Kohl declared that he was prepared to agree to an IGC on EMU, but, to prevent being accused of sacrificing the D-Mark during the election campaign of that autumn, the IGC could not start before the end of 1990.⁵⁶ This agreement was formalised during the Strasbourg Council of 9 December 1989.

After the fall of the Berlin Wall on 9 November 1989, furthering European integration and allaying French fears became even more important to the Chancellor.⁵⁷ However, now the *start date* had been agreed upon, the French demanded the acceleration of the European process by also fixing an *end date* for the IGC. Notwithstanding the fierce opposition of the German Minister of Finance, in April 1990, Kohl and Mitterrand jointly proposed that EMU 'should enter into force on 1 January 1993'.⁵⁸ Since the IGC would have to be concluded before January 1992 in order to allow for the treaty to be ratified in time, the agreement effectively meant that Kohl had conceded an end date for the IGC.⁵⁹

In response to these developments, the German financial elite began to voice their views on EMU more forcefully, focusing on three issues: the independence of the future ECB; the rules on excessive budgetary deficits; and the need for significant economic convergence prior to further monetary unification. Agreement on ECB independence modelled on the Bundesbank was reached early in the process. However, in the case of budgetary rules, the West German financial elite insisted that avoiding excessive deficits should be an obligation laid down in a *quantitative* treaty text, with binding rules and sanctions.⁶⁰ During the subsequent negotiations, however, Kohl was prepared to interpret monetary prudence more flexibly to ease relations with the French.⁶¹

As for the transition to the second and third phase, Kohl felt that the exact timing of the start of the second and third phases was immaterial, but that setting a date was a vital step in ensuring that the move towards EMU was irreversible.⁶² The German financial elite was strongly opposed to setting deadlines. Exactly two weeks after the official reunification of Germany on 3 October 1990, however, Kohl cast the objections of the financial authorities aside, and announced that the second stage would start on 1 January 1994,⁶³ a decision that was formalised at the first Rome summit. Although Kohl made sure that some conditions for the admittance of member states to stage two

were included, the transfer to the second stage would be automatic and irreversible, thus representing 'a triumph of politicians over technical experts'.⁶⁴

Decision making on transition to the third stage followed a similar pattern. On the one hand, the German financial elite demanded strict and binding convergence criteria; on the other, the French continued to press Kohl to agree to a deadline for the transition from second to the third stage.⁶⁵ While Kohl resisted French pressure during the Rome European Council of October 1990,⁶⁶ during the Luxembourg European Council of June 1991, Kohl voiced his support for 1 January 1997 as the end date for the convergence process. Eventually, the European Council decided that if by the end of 1997 the date for the beginning of the third stage had not been set, then it would start on 1 January 1999 for all those member states that fulfilled the quantitative convergence criteria. All in all, Kohl's resolve to bind Germany irrevocably to its European partners had led him to make yet another personal concession to the French.⁶⁷

*Ich wollte Deutschlands und Europas Einheit*⁶⁸

Kohl's new-found determination to establish EMU can be traced back to the peculiar effect that the revolution in Eastern Europe and the prospect – however faint – of German reunification had on the Chancellor's belief system. Rather than experiencing a substantive change in his beliefs, Kohl's pre-existing support for European integration in general, and EMU and supranational modes of integration in particular, were reinforced. Moreover, his macro-economic ideas remained just as *ordo-liberal* as before. Comparing the two maps, however, shows that after the summer of 1989, the priority of the political far outranked the importance of economic goals in the Chancellor's mind.

More specifically, it is clear from the second map that after the summer of 1989, European integration continued to be the pivot of Kohl's belief system.⁶⁹ In fact, his support for EMU (E6) as well as supranational modes of decision making doubled (dimension 1, see Table 1).⁷⁰ The concepts leading into the various modes of European integration in the map indicate that the reinforcement of Kohl's support for European unification was rooted largely in the fortification of his pre-existing belief in the reciprocity of European and German unification. Given this long-term conviction, the increased saliency of his ideas concerning German reunification was a natural corollary to the reinforcement of his ideas concerning European integration.⁷¹

As for macro-economic beliefs, the events of 1989 caused neither a *volte-face* nor a significant reinforcement. Kohl's beliefs remained as *ordo-liberal* as before (see Table 1). However, a significant change did occur in Kohl's ideas about the relative priority of political and economic goals of European integration (dimension 3). After the summer of 1989, Kohl clearly assigned priority to the political: when we compare the maps, there is a major drop in the amount, centrality and saliency of the economic consequences of European integration, while the relative value of political goals almost doubles (see Table 1).⁷² Again, this is largely due to Kohl's increased preoccupation with events in Eastern Europe and the possibility of German reunification.⁷³

As indicated above, Kohl's increasingly political view of European integration does not imply any substantive change in his pro-European convictions, his preference for supranational decision making, or his views on macro-economics. Moreover, the

beliefs that underlay this increase in the priority of the political – the tandem nature of European integration and the likelihood of German reunification – pre-dated the events of 1989. The analysis thus suggests that, rather than cause any fundamental change in Kohl's belief system, the events of 1989 spurred Chancellor Kohl into political action to turn his pre-existing convictions on European economic and monetary unification into a reality.

CONCLUSION

This article has set out to answer the question of why Germany wanted EMU in the first place, a question that has taken on new relevance in the light of the reticent and ambiguous German response to the current sovereign debt crisis. While EMU brought some economic and political benefits for Germany, the existing literature on the subject struggles to account for the precise timing of Germany's acquiescence and the inconsistent behaviour of the Kohl government on the road to Maastricht. This article provides evidence that taking into account Germany's shared historical consciousness, as well as the combined effect of Chancellor Kohl's personal beliefs and the 1989 revolution in Eastern Europe helps to answer these questions.

More particularly, the article shows that the actions and shifting positions of the Kohl government were rooted in the inherent clash between *Europatriotism* and a concern for *Sound Economics* in European economic and monetary policy making. While during the early and mid-1980s, the ordo-liberal German financial elite were allowed to dominate the German position on the issue of European monetary unification, the foreign policy elite pursued more pro-European policies in other areas. However, the evidence presented in this article indicates that this division of labour was upset by the advent of the geo-political revolution in Eastern Europe in the summer of 1989.

Firstly, a systematic analysis of Chancellor Kohl's policy beliefs indicates that – rather than change – the 1989 revolution in Eastern Europe strengthened Kohl's pre-existing pro-European convictions, politicising the issue of European monetary union and inciting him to act on these beliefs. Secondly, the fall of the Berlin Wall instigated a crucial shift in the German domestic balance of power by giving the Chancellor the necessary domestic discourse and legitimacy to overrule the German financial elite on several crucial elements in the EMU negotiations.

Through these two mechanisms, the geopolitical 'critical juncture' of 1989 incited a shift in the German position on EMU: after years of domination by the *Sound Economics* advocated by the German financial elite, this allowed Chancellor Kohl to take charge and guide Germany towards a more *Europatriotic* position. This change in the German position enabled an agreement with her European partners on crucial issues in the Maastricht Treaty, thereby making EMU possible.

In light of Europe's current sovereign debt crisis, these findings are very interesting. It would seem that the reticent and searching attitude of Merkel's government is far from unique but shows great similarities to the attitude of Kohl's government prior to Maastricht. This suggests that – like her predecessors – the Merkel government may be torn between its deep-rooted belief in the importance of *Sound Economics* and its desire to be *Europatriotic*.

Many have puzzled over, and debated the rationale behind the hesitation and prevarication in Germany's position. Some have sought an explanation in the particularities of Merkel's leadership style,⁷⁴ while others have rightly argued that the German government is involved in a very delicate two-level game: on the one hand, Germany's European partners are demanding (financial) solidarity; on the other hand, the German electorate, financial elite and Merkel's coalition partners are demanding budgetary self-control.⁷⁵

However, there may be more to the story. Underlying this political balancing act, Merkel may be wrestling with the same fundamental contradiction in policy goals that Kohl faced. For Germany to continue to defend its belief in *Sound Economics* in a context in which her European partners are questioning the central tenets of ordo-liberalism implies an inherent violation of the goal of *Europatriotism*. 'Good Europeans' come to the aid of fellow member states, but '*Sound Economics*' dictates that the culprits should be responsible for tightening their own belts or learn to abide by the rules the hard way. *Europatriotism* entails working together, give-and-take and allowing each other room for manoeuvre in difficult times, while ordo-liberalism demands automatic sanctions and the revocation of voting rights.

Short of abandoning one of these goals – which the Chancellor does not seem prepared to do at this time – living with, and having to act upon such contradictory goals may give rise to two behavioural patterns. Firstly, the mind may try to come up with ways to deny the inconsistency, legitimise inconsistent behaviour and muddle through. Secondly, depending on the context, different goals will take priority and seek precedence in one's mind at different times.⁷⁶ At first sight, this could explain the erratic stop-go attitude shown by Chancellor Merkel. When confronted with new rescue plans, she responds critically or reluctantly, only to become more yielding later as her European friends invoke the solidarity argument, or when strolling down the beach with President Sarkozy, far away from the critical voices of the ordo-liberals in Frankfurt and Berlin. It may not be that Merkel is being strong-armed into reversing her position; rather, this may be a case of arguments being more persuasive when they are found to be truthful to some extent.

In addition, this pattern is also more consistent with the patterns of the criticism that Merkel has had to cope with than a simple division between national versus European interests. Some domestic voices have argued that Merkel should show more ordo-liberal backbone, while others have accused her of being dangerously nationalistic.⁷⁷ Internationally, she faces the same criticism. Either way, and at either level, she cannot win.⁷⁸ While further research into this question is needed, acknowledging the inherent contradiction between policy goals that the Merkel government is contending with may well improve our understanding of its utterances and actions regarding the crisis.

As for our expectations of the German government in the months to come, we should not forget that it took a full-blown geopolitical crisis for Helmut Kohl really to take the lead on European monetary unification. The historical pattern exposed in this article suggests that possibly the current crisis would need to become even worse to unleash the leadership and vision needed for Germany (temporarily) to free herself from the paradox of *Sound Economics* and *Europatriotism* and lead Europe out of the crisis.

ACKNOWLEDGEMENTS

I thank the participants of the Conference on ‘10 Years of EMU – What Have We Learned?’ of April 2008 at the Hertie School of Governance in Berlin, the anonymous reviewers and Wieger Bakker for their comments on earlier versions of this article. The usual disclaimer applies.

ABOUT THE AUTHOR

Femke van Esch works as an Assistant Professor of European Integration at the Utrecht School of Governance (USG) in The Netherlands. In her dissertation, she studied the role of leaders’ beliefs in the establishment of EMU and published articles on European monetary integration in amongst others the *Journal of Common Market Studies*. She is currently studying the effects of the current sovereign debt crisis on the beliefs of EU political and financial leaders.

NOTES

1. Cf. Charlemagne, ‘The Ultimate EU Slur: Call Your Opponent a Nationalist’, *The Economist*, 3 March 2010, available from www.economist.com/blogs/charlemagne/2010/05/euro_crisis (accessed January 2012); *Der Spiegel*, ‘Merkel’s Greece Deal “Betrays the Concept of Europe”’, 26 March 2010, available from <http://www.spiegel.de/international/europe/0,1518,685840,00.html> (accessed January 2012).
2. Barry Eichengreen and Jeffrey Frieden, ‘The Political Economy of European Monetary Unification: An Analytical Introduction’, *Economics and Politics* 5/2 (2010), pp.85–105; C. Randall Henning, ‘Monetary Integration: The Case of Europe’, *International Organization* 52/3 (1998), pp.537–73; Matthias Kaelberer, ‘Deutsche Mark Nationalism and Europeanized Identity: Exploring Identity Aspects of Germany’s Adoption of the Euro’, *German Politics* 14/3 (2005), pp.283–96; Peter H. Loedel, *Deutsche Mark Politics: Germany in the European Monetary System* (Boulder, CO: Lynne Rienner Publishers, 1999); Andrew Moravcsik, *The Choice for Europe: Social Purpose and State Power from Messina to Maastricht* (Ithaca, NY: Cornell University Press, 1998); Wayne Sandholtz, ‘Choosing Union: Monetary Politics and Maastricht’, *International Organization* 47/1 (1993), pp.1–39; Erich R. Staal, ‘European Monetary Union: The German Political-Economic Trilemma’, *ZEI Discussion Papers* C45 (1999).
3. Eichengreen and Frieden, ‘The Political Economy’; Dorothee Heisenberg, ‘Taking a Second Look at Germany’s Motivation to Establish Economic and Monetary Union: A Critique of “Economic Interests” Claims’, *German Politics* 14/1 (2005), pp.95–109; Martin Marcussen, ‘The Dynamics of EMU Ideas’, *Cooperation and Conflict* 34/4 (1999), pp.383–411; Kathleen R. McNamara, *The Currency of Ideas: Monetary Politics in the European Union* (Ithaca, NY: Cornell University Press, 1998); Thomas Risse, Daniela Engelmann-Martin, Hans-Joachim Knopf and Klaus Roscher, ‘To Euro or Not to Euro? The EMU and Identity Politics in the European Union’, *European Journal of International Relations* 5/2 (1999), pp.147–87; Sandholtz, ‘Choosing Union’, p.25.
4. Karl Kaltenthaler, ‘German Interests in European Monetary Integration’, *Journal of Common Market Studies* 40/1 (2002), pp.69–87.
5. This argument is not uncommon in the literature. Some argue that the fall of the Berlin Wall incited a *quid pro quo* between France and Germany, others that it triggered a further (voluntary) *Westbindung* by Germany. Overall, no agreement exists on whether and how this event played a role. Moreover, there is a marked lack of empirical and systematic research into this matter. Cf. Michael J. Baun, ‘The Maastricht Treaty as High Politics: Germany, France and European Integration’, *Political Science Quarterly* 110/4 (1995), pp.605–25; William M. Chandler, ‘German Influence in Shaping EMU: Still a Tamed Power?’, in Amy Verdun (ed.), *The Euro: European Integration Theory and Economic and Monetary Union* (Lanham, MD: Rowman and Littlefield, 2002), pp.201–14; Staal, ‘European Monetary Union’, p.10.
6. Carsten Hefeker, *Interest Groups and Monetary Integration: The Political Economy of Exchange Regime Choice* (Boulder, CO: Westview Press, 1997); Moravcsik, *The Choice*, pp.38–45; cf. Eichengreen and Frieden, ‘The Political Economy’; Jeffrey A. Frieden, ‘Invested Interests: The Politics of National Economic Policies in a World of Global Finance’, *International Organization* 45/4 (1991), pp.440–51; Sandholtz, ‘Choosing Union’; Staal, ‘European Monetary Union’.

7. Henning, 'Monetary Integration', pp.548, 563–5; cf. Loedel, *Deutsche Mark Politics*, pp.93–108; Staal, 'European Monetary Union'.
8. Henning, 'Monetary Integration', p.566; cf. Eichengreen and Frieden, 'The Political Economy'; Staal, 'European Monetary Union'.
9. Committee for the Study of Economic and Monetary Union (Delors Committee), *Report on Economic and Monetary Union in the European Community* (Luxembourg: Office for Official Publications of the EU, April 1989); Femke A.W.J. Van Esch, 'Mapping the Road to Maastricht: A Comparative Study of German and French Pivotal Decision Makers' Preferences Concerning the Establishment of a European Monetary Union during the Early 1970s and Late 1980s', unpublished Ph.D Thesis, Radboud University Nijmegen, Faculty of Management Sciences, 2007, pp.35–6; Heisenberg, 'Taking a Second Look'; McNamara, *The Currency of Ideas*, pp.33–42; Sandholtz, 'Choosing Union', p.25.
10. Thomas Pedersen, *Germany, France, and the Integration of Europe: A Realist Interpretation* (London and New York: Pinter, 1998); cf. Dorothee Heisenberg, *The Mark of the Bundesbank: Germany's Role in European Monetary Cooperation* (Boulder, CO: Lynne Rienner Publishers, 1999), pp.6–9.
11. Staal, 'European Monetary Union'.
12. Joseph M. Grieco, 'Rule Trajectories: A Neorealist Interpretation of the Maastricht Treaty and European Economic and Monetary Union', in Benjamin Frankel (ed.), *Realism: Restatement and Renewal* (London: Frank Cass, 1996), pp.290–304.
13. Marcussen, 'The Dynamics of EMU Ideas', p.400; McNamara, *The Currency of Ideas*; Amy Verdun, 'The Role of the Delors Committee in the Creation of EMU: An Epistemic Community', *European Journal of Public Policy* 6/2 (1999), pp.308–28.
14. Marcussen, 'The Dynamics of EMU Ideas', p.402.
15. Robert Axelrod, 'The Cognitive Mapping Approach to Decision Making', in Robert Axelrod (ed.), *Structure of Decision: The Cognitive Maps of Political Elites* (Princeton, NJ: Princeton University Press, 1976), pp.3–17; Michael D. Young and Mark Schafer, 'Is There Method in Our Madness? Ways of Assessing Cognition in International Relations', *Mershon International Studies Review* 42/1 (1998), pp.63–96.
16. For an elaborate description of the method used, see Van Esch, 'Mapping the Road to Maastricht', ch.4.
17. A decision-maker prefers policies that are linked to a positively evaluated goal by only positive relationships or an even amount of negative relations, or to a negatively evaluated goal by an uneven amount of negative relations. Decision-makers' preferences are indeterminate when linked to policy by a path containing one or more zero signs.
18. The cognitive maps may be obtained from the author's website: <http://www.uu.nl/rebo/medewerkers/FAWJvanEsch>
19. Including E6/E9/E14/M2/P2/R3.
20. Including E7/E10/E11/M1.
21. Including C1/-L1/P1/S2/C3/H3.
22. Including E2/E3/W1/E13/S3.
23. Including concepts directly following concepts referring to European integration: E1/E2/E4/E8/L2/P1.
24. Including concepts directly following concepts referring to European integration: F1/F2/F4/G4/P4/P5/F7/G8/O1/R2/R3.
25. Kaltenthaler, 'German Interests'.
26. Heisenberg, *The Mark of the Bundesbank*; Kaltenthaler, 'German Interests'; Risse et al., 'To Euro or Not to Euro?', p.168.
27. Risse et al., 'To Euro or Not to Euro?', p.167; Kenneth H.F. Dyson and Kevin Featherstone, *The Road to Maastricht: Negotiating Economic and Monetary Union* (Oxford: Oxford University Press, 1999), pp.260–3.
28. See Van Esch, 'Mapping the Road to Maastricht'; Risse et al., 'To Euro or Not to Euro?'; cf. Kaelberer, 'Deutsche Mark Nationalism and Europeanized Identity'.
29. Van Esch, 'Mapping the Road to Maastricht'; Femke A.W.J. Van Esch, 'The Rising of the Phoenix: Building the European Monetary System on a Meeting of Minds', *L'Europe en Formation* 353/4 (2009), pp.133–48.
30. The distinction between these groups is not clear-cut. Members of the political elite generally supported the Bundesbank's ordoliberal policies, while members of the German financial elite are overall 'Good Europeans' (Risse et al., 'To Euro or Not to Euro?', pp.167–9). However, if forced to choose, their choices will contradict.
31. Eckart Gaddum, *Die deutsche Europapolitik in den 80er Jahren: Interessen, Konflikte und Entscheidungen der Regierung Kohl* (Paderborn: Schöningh, 1994), p.68.
32. Kenneth H.F. Dyson, *Elusive Union: The Process of Economic and Monetary Union in Europe* (New York: Longman, 1994), p.117; Gaddum, *Die deutsche Europapolitik*, pp.322–44; André

- Szász, *The Road to European Monetary Union* (London: MacMillan Press, 1999), pp.89–92. Neither the Bundesbank, nor the Ministries of Economic and Financial Affairs are present during European Council meetings.
33. David R. Cameron, 'Creating Supranational Authority in Monetary and Exchange-Rate Policy: The Sources and Effects of EMU', in Wayne Sandholtz and Alec Stone Sweet (eds), *European Integration and Supranational Governance* (Oxford: Oxford University Press, 1998), p.204, note 28.
 34. Dyson, *Elusive Union*, p.128; Wilhelm Schönfelder and Elke Thiel, *Ein Markt – Ein Währung: Die Verhandlungen zur Europäischen Wirtschafts- und Währungsunion* (Baden-Baden: Nomos Verlagsgesellschaft, 1994).
 35. Agence Europe, *Europe, Daily Bulletin* (Luxembourg: Agence Internationale d'Information pour la Presse), 10 May 1988, p.6; Dyson and Featherstone, *The Road to Maastricht*, p.333; Gaddum, *Die deutsche Europapolitik*, pp.348, 351–2.
 36. Joachim Bitterlich, *Das Europa der Zukunft* (Düsseldorf: Droste Verlag, 2004), p.21; Dyson, *Elusive Union*, p.128.
 37. Gaddum, *Die deutsche Europapolitik*, p.353.
 38. David R. Cameron, 'Transnational Relations and the Development of European Economic and Monetary Union', in Thomas Risse-Kappen (ed.) *Bringing Transnational Relations Back In: Non-State Actors, Domestic Structures and International Institutions* (Cambridge: Cambridge University Press, 1995), p.48; cf. Niels Thygesen, 'The Delors Report and European Economic and Monetary Union', *International affairs* 65/4 (1989), p.640.
 39. DzDPS, Hanns J. Küsters and Daniel Hofmann (eds), *Dokumente zur Deutschlandpolitik, Sonderedition aus den Akten des Bundeskanzleramtes 1989/1990* (München: Oldenburg, 1998), p.308; Schönfelder and Thiel, *Ein Markt*, pp.55–6. Later that year, Kohl cited the December elections as the reason for refusing to commit to an early start date. However, at the time of the June meetings, the date for the elections had not yet been set. DzDPS et al., *Dokumente zur Deutschlandpolitik*, p.454.
 40. For map, see authors website, <http://www.uu.nl/rebo/medewerkers/FAWJvanEsch>
 41. Patricia Clough, *Helmut Kohl: Ein Porträt der Macht* (München: Deutscher Taschenbuch Verlag, 1998), pp.27, 35; Kenneth H.F. Dyson, 'Chancellor Kohl as Strategic Leader: The Case of Economic and Monetary Union', in Clay Clemens and William E. Paterson (eds), *The Kohl Chancellorship: Special Issue of German Politics* (London: Frank Cass, 1998), pp.42–3; Dyson and Featherstone, *The Road to Maastricht*, p.256; Klaus Hofmann, *Helmut Kohl, Kanzler des Vertrauens: Eine politische Biographie* (Bonn: Aktuell, 1984), pp.24, 81.
 42. All concepts referring to European integration are valued positively and have an above-average centrality and saliency of 7.2 and 9.9. Moreover, the two core concepts in the map both refer to European integration (E5/S1).
 43. The Establishment of EMU (E6) has a well above-average centrality of 6 and saliency of 7.
 44. Kohl distinguishes as many concepts referring to intergovernmental as supranational decision making. While valuing all of them positively, the latter are valued higher in terms of both centrality and saliency (see Tables A1 and A2).
 45. Clough, *Helmut Kohl*, pp.62, 138; Dyson and Featherstone, *The Road to Maastricht*, p.257.
 46. Kohl's map includes four concepts that refer to ordo-liberal economic considerations, compared to three 'Keynesian' concepts. Moreover, the saliency and centrality of ordo-liberal considerations are more than twice those of the Keynesian concepts.
 47. Kohl distinguished four concepts referring to friendly Franco-German relations (F1/F2/F4/S4), which he associates positively with various intergovernmental and supranational modes of European integration.
 48. Countering the 'partition of Germany' (P4) is the most important political goal of European integration with the above-average centrality of 6 and saliency of 8.
 49. From the most central and salient economic and political goals served by further European integration, the economic goals (P1, E8) are intermediate goals which lead further into other, political, goals. The main political goals of peace (P5) and German reunification (P4) are the ultimate goals; they are not followed by higher goals.
 50. Dyson, 'Chancellor Kohl as Strategic Leader', p.39; cf. Dyson and Featherstone, *The Road to Maastricht*, p.308.
 51. Dyson and Featherstone, *The Road to Maastricht*, p.334; Gaddum, *Die deutsche Europapolitik*.
 52. Helmut Kohl, (2000) *Ich wollte Deutschlands Einheit* (München: Ullstein, 2000); Philip D. Zelikow and Condoleezza Rice, *Germany Unified and Europe Transformed: A Study in Statecraft* (Cambridge, MA: Harvard University Press, 1997).
 53. Agence Europe, 16 October 1989, p.4; cf. DzDPS et al., *Dokumente zur Deutschlandpolitik*, p.460; Zelikow and Rice, *Germany Unified*, p.95.
 54. DzDPS et al., *Dokumente zur Deutschlandpolitik*, p.308.

55. Ibid., p.298.
56. Ibid., pp.470–3.
57. Ibid., p.683; Helmut Kohl, 'The Community and a United Germany', *Europe Documents* 1607 (1990), pp.1–2.
58. *Agence Europe*, 9 April 1990, p.15; *Agence Europe*, 19 April 1990, p.3.
59. Dyson, *Elusive Union*, p.141.
60. Matthieu L.L. Segers and Femke A.W.J. Van Esch, 'Behind the Veil of Budgetary Discipline: The Political Logic of the Budgetary Rules in EMU and SGP', *Journal of Common Market Studies* 45/5 (2007), pp.1089–109.
61. DzDPS et al., *Dokumente zur Deutschlandpolitik*, p.1545; *Agence Europe*, No. 5149, p.3; Szász, *The Road to European Monetary Union*, pp.159–60.
62. Schönfelder and Thiel, *Ein Markt*, p.104.
63. *Agence Europe*, 18 October 1990, p.6.
64. Colette Mazzucelli, *France and Germany at Maastricht: Politics and Negotiations to Create the European Union* (New York: Garland Publishing, 1997), p.80.
65. *Agence Europe*, 24/25 Sept. 1990, p.1.
66. Dyson and Featherstone, *The Road to Maastricht*, pp.396–7.
67. Mazzucelli, *France and Germany*, p.199. While the Bundesbank may be considered a veto player in European monetary integration, the German Basic Law grants the government the right to determine the scope of its currency area and negotiate international monetary agreements.
68. Cf. Kohl, *Ich wollte Deutschlands Einheit*.
69. Twenty-two rather than eighteen of forty-six concepts refer to European integration, three of which are valued high above-average (S1/E6/P2, see Table A2).
70. The concept of the 'Political Unification of Europe' (P2) became the third 'supranational' concept at the core of Kohl's belief system. The sum of their centrality and saliency values is 99, that of the intergovernmental concepts is 24. In the first map, these were respectively 46 and 38.
71. The centrality and saliency of the concept 'Partition of Germany' (P4) doubled. Moreover, in the second map, in addition to the reciprocal relationship between European integration and German reunification from the first map, several more are identified (E6–/E10–/E11–/G2–P4; G8–G7; G6–/E10–/E11–G8; P4–P2; G7–G8).
72. The number of economic goals drops from six to two, and relative centrality and saliency drop to less than one-fifth of the rates in the first map. At the same time, the relative centrality and saliency of political consequences almost doubles (see Table 1).
73. In addition to the increased centrality and saliency of P4, Kohl now mentions three additional concepts, referring to the reforms in Eastern Europe and (possible) German reunification (F7/G8/R2).
74. Quentin Peel, 'Deliberation by Chancellor Risks Confusion on Several Points', *Financial Times*, 12 May 2011, p.2.
75. Wolfgang Münchau, 'The Political Constraints of the Eurozone', *Financial Times*, 15 Feb. 2010, p.7; Ralph Atkins, 'Greek Dilemma Piles Pressure on Merkel', *Financial Times*, 22 March 2010, p.3; Wolfgang Münchau, 'Germany Pays for Merkel's Miscalculations', *Financial Times*, 11 May 2011, p.9.
76. Jerel A. Rosati, 'A Cognitive Approach to the Study of Foreign Policy', in Laura Neack, Jeanne A.K. Hey and Patrick J. Haney (eds), *Foreign Policy Analysis: Continuity and Change in Its Second Generation* (Englewood Cliffs, NJ: Prentice-Hall, 1995), p.54; Van Esch, 'Mapping the Road to Maastricht'.
77. *Financial Times Magazine*, 'What the Philosopher Saw', 1 May 2010, p.16; *Der Spiegel*, 'Merkel's Greece Deal'.
78. Quentin Peel, 'A Test of Strength', *Financial Times*, 12 April 2011, p.9.

APPENDIX 1: ASSERTIONS USED IN COGNITIVE MAPS

Map 1

- ‘Die Rolle der Bundesrepublik Deutschland in der internationalen Politik’, speech held for the *Chicago Council on Foreign Relations* (23 Oct. 1986).
- ‘Das Erbe Ludwig Erhards – Herausforderung and die Wirtschaftspolitik’, speech held during the *Commemoration of the 90th birthday of Professor Dr Ludwig Erhard* (Bonn, 4 Feb. 1987).
- ‘Für eine gemeinsame europäische Sicherheitspolitik’, speech held at the *Commemoration of the 25th Anniversary of the Franco-German Treaty* (Paris, 22 Jan. 1988).
- ‘Europas Zukunft – Vollendung des Binnenmarktes 1992’, speech held at the *Meeting of German Entrepreneurs* (Bonn, 15 March 1988).
- ‘Zum 40. Jahrestag der Währungs- und Wirtschaftsreform’, speech held at the *Dies of the Ludwig-Erhard Foundation* (Frankfurt, 12 July 1988).
- ‘40 Jahre Bundesrepublik Deutschland – Unsere Verantwortung für Deutschlands Zukunft’, speech held at the opening of the *CDU Congress on ‘40 Years of German Federal Republic’* (Bonn, 18 Jan. 1989).

Map 2

- ‘Aufbruch zu Freiheit und Selbstbestimmung: Die Deutschland- und Europapolitik im Blick auf den Zerfall der Diktaturen des Ostblocks’, declaration on the *State of the Nation in Divided Germany* (8 Nov. 1989).
- ‘Gesicherter Frieden in Freiheit und Wohlstand’, speech held on the occasion of the *Annual Meeting of the Bundesvereinigung der Deutschen Arbeitgeberverbände* (Bonn, 14 Dec. 1989).
- ‘Der Standort Deutschlands in einem künftigen Europa’, speech at the *Conference of the Bureau International de Liaison et de Documentation and the Institut Français des Relations Internationales* (Paris, 17 Jan. 1990).
- ‘The Community and a United Germany’, press conference (Brussels, 23 March 1990).
- ‘A United Germany in a United Europe’, speech delivered to the *American Council on Germany* (New York City, 5 June 1990).
- ‘... im Ziehen von Freiheit, Rechtsstaatlichkeit und Demokratie: Aufgabe deutscher Politik in den neunziger Jahren’, speech at the *Atlantic Council, Georgetown University, Centre for Strategic and International Studies and American Institute for Contemporary German Studies* (Washington, 17 Jan. 1990).

APPENDIX 2

TABLE A1
MAP 1 CENTRALITY AND SALIENCY

Concept	Code	Centrality	Saliency
European integration	E5	22	36
Completion of SEA	S1	16	25
European unification	E9	11	14
Price stability	P1	10	16
General Benefit	U1	9	15
European political cooperation	E11	7	8
Establishment of EMU	E6	6	7
French–German cooperation	F4	6	10
Expansionary governmental policy	L1	6	12
Partition of Germany	P4	6	8
Economic harmonisation	E8	5	6
Further monetary cooperation	M1	5	6
Peace	P5	5	8
(Eur.) Central bank independence	C1	4	6
Liberalisation	L2	4	4
European economic cooperation	E10	3	3
Further development EMS	E12	3	4
Economic development	E4	3	10
EMS	E7	3	3
Majority decisions	M2	3	3
Benefit of Germany	U3	3	2
Increased purchasing power	W1	3	12
Decentralised (European) central bank	C2	2	2
Economic growth	E2	2	2
Employment	E3	2	2
French–German friendship	F1	2	2
French–German reconciliation	F2	2	4
German policy	G1	2	3
(Reunified) Germany in the EC/EU	G2	2	2
German sacrifices	G3	2	2
German strength	G5	2	1
International labour-division	I1	2	5
Making compromises	M3	2	2
Solid government finances	S2	2	3
Social-economic development	S3	2	2
Shared interest Germany and France	S4	2	2
Benefit of France and Germany	U2	2	2
European competitiveness	E1	1	1
European social development	E13	1	1
European economic unification	E14	1	1
Freedom	F3	1	1
German freedom	G4	1	2
History of Europe	H1	1	1
Political freedom of action	P6	1	1
Autonomous role for Europeans	R1	1	3
Trust of the citizens	T1	1	1
Total		182	266
Mean		4	5.8

TABLE A2
MAP 2 CENTRALITY AND SALIENCY

Concept	Code	Centrality	Saliency
Establishment of EMU	E6	14	30
Partition of Germany	P4	12	17
Political unification of Europe	P2	10	17
Reforms in Eastern Europe	R2	10	12
European unification	E9	9	15
European integration	E5	8	11
German strength	G5	8	9
Completion of SEA	S1	8	13
Germany in the EG/EU	G2	6	10
European economic cooperation	E10	5	7
German policy	G1	5	6
German East-German policies	G8	5	5
Hasty steps	H2	5	10
Not fixing early starting-date for IGC	H3	5	11
Benefit of Germany	U3	5	7
European political cooperation	E11	4	6
Expansionary governmental policy	L1	4	5
Credibility of the EMU/EURO	C3	3	4
Germany in Western alliance	G6	3	4
Europe speaking with one voice	O1	3	3
Price-stability	P1	3	7
Social-economic development	S3	3	3
Increased purchasing power	W1	3	4
European strength	E15	2	3
French-German cooperation	F4	2	4
Future architecture (united) Germany	F7	2	3
German European policies	G7	2	2
German monetary structure	G9	2	4
Liberalisation	L2	2	2
Solid government finances	S2	2	2
Sensitivity of markets to mon. issue	S5	2	3
Benefit of Europe	U4	2	2
Benefit of Kohl	U5	2	3
(Eur.) Central bank independence	C1	1	3
Delors report	D1	1	1
European social development	E13	1	1
Economic development	E4	1	1
Economic harmonisation	E8	1	1
Freedom	F3	1	2
Future architecture of Europe	F6	1	2
German sacrifices	G3	1	2
History of Europe	H1	1	5
Further monetary cooperation	M1	1	1
Strengthening of European Parliament	R3	1	3
Self-determination	S6	1	1
General benefit	U1	1	1
Total		179	268
Mean		3.9	5.8

Chap 2: Is the theory of optimal currency areas sufficient?

Key issue: the theory of optimal currency areas provides an analysis of the macroeconomic costs and benefits of a monetary union.

- Krugman, P. (2012). *Revenge of the Optimum Currency Area*. NBER Macroeconomics Annual, 27(1), 439-448
- Eichengreen, B. (1991). *Is Europe an Optimum Currency Area?* NBER N. 3579
- D. Rodrick, "*Economists vs. Economics.*" Project Syndicate. Sep 10, 2015

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: NBER Macroeconomics Annual 2012, Volume 27

Volume Author/Editor: Daron Acemoglu, Jonathan Parker, and Michael Woodford, editors

Volume Publisher: University of Chicago Press

ISSN: 0889-3365

Volume ISBN: cloth: 978-0-226-05277-9; paper: 978-0-226-05280-9, 0-226-05280-X

Volume URL: <http://www.nber.org/books/acem12-2>

Conference Date: April 20-21, 2012

Publication Date: May 2013

Chapter Title: Revenge of the Optimum Currency Area

Chapter Author(s): Paul Krugman

Chapter URL: <http://www.nber.org/chapters/c12759>

Chapter pages in book: (p. 439 - 448)

Revenge of the Optimum Currency Area

Paul Krugman, *Princeton University and NBER*

The creation of the euro was supposed to be another triumphant step in the European project, in which economic integration has been used to foster political integration and peace; a common currency, so the thinking went, would bind the continent even more closely together. What has happened instead, however, is a nightmare: the euro has become an economic trap, and Europe a nest of squabbling nations. Even the continent's democratic achievements seem under threat, as dire economic conditions create a favorable environment for political extremism. Who could have seen such a thing coming?

Well, the answer is that lots of economists could and should have seen it coming, and some did. For we have a long-established way to think about the prospects for currency unions, the theory of optimum currency areas—and right from the beginning, this theory suggested serious concerns about the euro project.

These concerns were largely dismissed at the time, with many assertions that the theory was wrong, irrelevant, or that any concerns it raised could be addressed with reforms. Recent events have, however, very much followed the lines one might have expected given good old-fashioned optimum currency area theory, even as they have suggested both that we need to expand the theory and that some aspects of the theory are more important than we previously realized.

In what follows, I'll start with a very brief and selective review of what I consider the key points of optimum currency area theory, and what that theory seemed, some two decades ago, to say about the idea of a single European currency. Next up is the crisis, and the continuing refusal of many leaders to see it for what it is. Finally, some thoughts on possible futures.

I. Mundell, Kenen, and Currencies

The advantages of a common currency are obvious, if hard to quantify: reduced transaction costs, elimination of currency risk, greater transparency, and possibly greater competition because prices are easier to compare. Before the creation of the euro, some statistical work on the limited number of country pairs sharing a currency suggested that the common European currency might produce an explosion in intra-European trade; that has not happened, but trade does seem to have risen modestly as a result of the single currency, and presumably that corresponds to an increase in mutually beneficial and hence productive exchanges.

The disadvantages of a single currency come from loss of flexibility. It's not just that a currency area is limited to a one-size-fits-all monetary policy; even more important is the loss of a mechanism for adjustment. For it seemed to the creators of optimum currency area (OCA), and continues to seem now, that changes in relative prices and wages are much more easily made via currency depreciation than by renegotiating individual contracts. Iceland achieved a 25 percent fall in wages relative to the European core in one fell swoop, via a fall in the krona. Spain probably needs a comparable adjustment, but that adjustment, if it can happen at all, will require years of grinding wage deflation in the face of high unemployment.

But why should such adjustments ever be necessary? The answer is "asymmetric shocks." A boom or slump everywhere in a currency area poses no special problems. But suppose, to take a not at all hypothetical example, that a vast housing boom leads to full employment and rising wages in part, but only part, of a currency area, then goes bust. The legacy of those boomtime wage increases will be an uncompetitive tradable sector, and hence the need to get at least relative wages down again.

So the advantages of a single currency come at a potentially high cost. Optimum currency area theory is about weighing the balance between those advantages and those potential costs.

Now, what we need to say right away is that this "weighing" takes place only in a qualitative sense: at this point nobody says that the benefits of joining the euro are x percent of GDP, the costs y , and $x > y$, so the euro it is. Instead, it is more along the lines of arguing that Florida is a better candidate for membership in the dollar zone than Spain is a

candidate for membership in the euro zone. This does not necessarily say that Spain made a mistake by joining the euro—nor does it necessarily refute the argument that Florida would be better off with its own currency! But the theory does at least give us some insight into the trade-offs.

We also need to say that in practice very little of optimum currency area theory is concerned with the benefits of a single currency area. Obviously these benefits depend on potential economic interactions; there would be no point in sharing a currency with, say, a colony on Mars that did almost no trade with Earth, and joining the euro makes a lot more sense for, say, Slovakia than it would for Mongolia. But almost all the interesting stuff comes from looking at factors that might mitigate the costs arising from the loss of monetary flexibility that comes with adopting someone else's currency—which brings us to the two big ideas of OCA.

First up, Mundell, whose classic 1961 paper argued that a single currency was more likely to be workable if the regions sharing that currency were characterized by high mutual labor mobility. (He actually said factor mobility, but labor is almost surely the one that matters.) How is this so?

Well, suppose—to take a not at all hypothetical example—that the state of Massachusetts takes a major asymmetric hit to its economy that sharply reduces employment (which is, in fact, what happened at the end of the 1980s). If Massachusetts workers cannot or will not leave the state, the only way to restore full employment is to regain the lost jobs, which will probably require a large fall in relative wages to make the state more competitive; a fall in relative wages that is much more easily accomplished if you have your own currency to devalue. But if there is high labor mobility, full employment can instead be restored through emigration, which shrinks the labor force to the jobs available. And that is what actually happened. Table 1 shows snapshots of the Massachusetts economy at three dates: 1986, the height of the “Massachusetts miracle” centered on minicomputers; 1991, after the shift to PCs and the bursting of a housing bubble had brought a severe local recession; and 1996. Notice that Massachusetts never regained the employment share it lost in the late 1980s bust. Nonetheless, by the mid-1990s it once again had an unemployment rate below the national average, because workers moved elsewhere.

So that is one main theme of optimum currency area theory. But it is

Table 1
Labor Mobility in Action

	MA Share in US Employment	MA Unemployment Rate	US Unemployment Rate
1986	2.70	4.0	7.0
1991	2.48	8.8	6.8
1996	2.43	4.6	5.4

not the only one. There's also Peter Kenen's argument (1969) that fiscal integration—a large “federal” component to spending at the regional or local level—can help a lot in dealing with asymmetric shocks.

Let's once again take a not at all hypothetical example: Florida, after the recent housing bust. America may have a small welfare state by European standards, but it is still pretty big, with large spending in particular on Social Security and Medicare—obviously both are a big deal in Florida. These programs are, however, paid for at a national level. What this means is that if Florida suffers an asymmetric adverse shock, it will receive an automatic compensating transfer from the rest of the country: it pays less into the national budget, but this has no impact on the benefits it receives, and may even increase its benefits if they come from programs like unemployment benefits, food stamps, and Medicaid, which expand in the face of economic distress.

How big is this automatic transfer? Table 2 shows some indicative numbers about Florida's financial relations with Washington in 2007, the year before the crisis, and 2010, in the depths of crisis. Florida's tax payments to DC fell some \$33 billion; meanwhile, special federally funded unemployment insurance programs contributed some \$3 billion, and food stamp payments rose almost \$4 billion. That's about \$40 billion in de facto transfers, some 5 percent of Florida's GDP—and that's surely an understatement, since there were also crisis-related increases in Medicaid and even Social Security, as more people took early retirement or applied for disability payments.

You might argue that since Florida residents are also US taxpayers, we really should not count all of this as a transfer. The crucial point, however, is that the federal government does not currently face a borrowing constraint, and has very low borrowing costs. So all of this is a burden that would be a real problem if Florida were a sovereign state, but it is taken off its shoulders by the fact that it is not.

Wait, there's more: Florida banks benefit from federal deposit insur-

Table 2
Florida and the Feds

	2007	2010
Revenue paid to DC	136.5	111.4
Special unemployment benefits	0	2.9
Food stamps	1.4	5.1

ance; many mortgage losses fell on Fannie and Freddie, the federally-sponsored lending agencies. (More on this financial backing shortly.)

In summary, optimum currency area theory suggested two big things to look at—labor mobility and fiscal integration. And on both counts it was obvious that Europe fell far short of the US example, with limited labor mobility and virtually no fiscal integration. This should have given European leaders pause—but they had their hearts set on the single currency.

Why did they believe it would work? I won't try for a detailed historiography; let me just say that what I recall from discussions at the time was the belief that two factors would make the adjustment problems manageable. First, countries would adopt sound fiscal policies, and thereby reduce the incidence of asymmetric shocks. Second, countries would engage in structural reforms that would make labor markets—and, presumably, wages—flexible enough to cope with such asymmetric shocks as occurred despite the soundness of the fiscal policies.

Even at the time, this sounded to many American economists like wishful thinking. After all, asymmetric shocks do not have to arise from unsound policies—they can come from shifts in relative product demand or, of course, such things as real estate bubbles. And European leaders seemed to believe that they could achieve a degree of wage flexibility that would be more or less unprecedented in the modern world.

Nonetheless, the project went ahead. Exchange rates were locked at the beginning of 1999, with the mark, the franc, and so on officially becoming just denominations of the euro. Then came actual euro notes—and they all lived happily ever after, for values of “ever after” < eleven years.

II. The Euro Crisis

As I just suggested, the architects of the euro, to the extent that they took optimum currency area theory at all seriously, chose to believe that

asymmetric shocks would be a relatively minor problem. What happened instead was the mother of all asymmetric shocks—a shock that was, in a bitter irony, caused by the creation of the euro itself.

In essence, the creation of the euro led to a perception on the part of many investors that the big risks associated with cross-border investment within Europe had been eliminated. In the 1990s, despite the absence of formal capital controls, capital movements and hence current-account imbalances within Europe were limited. After the creation of the euro, however, there was massive capital movement from Europe's core—mainly Germany, but also the Netherlands—to its periphery, leading to an economic boom in the periphery and significantly higher inflation rates in Spain, Greece, and so forth, than in Germany.

This movement was itself a large asymmetric shock, but a relatively gradual one, and one that the European Central Bank (ECB) was willing to accommodate with slightly above-target inflation. Matters were quite different, however, when private capital flows from the core to the periphery came to a sudden stop, leaving the peripheral economies with prices and unit labor costs that were well out of line with those in the core. Suddenly the euro faced a major adjustment problem.

This was the kind of problem optimum currency area theory warned would be very difficult to handle without currency devaluation; euro optimists had believed that reforms would make labor markets sufficiently flexible to deal with such situations. Unfortunately, the pessimists were right. "Internal devaluation"—restoring competitiveness through wage cuts as opposed to devaluation—has proved extremely hard. Table 3 shows hourly labor costs in the business sectors of several peripheral economies that, by common account, entered the crisis with very flexible labor markets; even so, and despite very high unemployment, they have achieved at best small declines.

So optimum currency area theory was right to assert that creating a single currency would bring significant costs, which in turn meant that Europe's lack of mitigating factors in the form of high labor mobility and/or fiscal integration became a very significant issue. In this sense, the story of the euro is one of a crisis foretold.

Yet there have been some surprises—unfortunately, none of them favorable.

First, as far as I know nobody, or almost nobody, foresaw that countries hit by adverse asymmetric shocks would face fiscal burdens so large as to call government solvency into question. As it turned out, the adjustment problems of the euro area quickly turned into a series

Table 3
Hourly Labor Costs in the Business Sector, 2008 = 100

	2006	2007	2008	2009	2010	2011
Estonia	73.1	87.8	100.0	98.2	96.2	100.7
Ireland	91.5	95.7	100.0	103.1	102.4	100.7
Latvia	62.8	81.7	100.0	99.9	97.1	100.3

of fiscal emergencies as well. In this sense, Kenen has turned out to dominate Mundell: lack of labor mobility has not played a major role in euro's difficulties, at least so far, but lack of fiscal integration has had an enormous impact, arguably making the difference between the merely bad condition of America's "sand states," where the housing bubble was concentrated, and the acute crises facing Europe's periphery.

Second, traditional optimum currency area theory paid little attention to banking issues; little thought was given to the importance of national as opposed to regional bank guarantees in the United States. In retrospect, however, we can see just how crucial such guarantees have actually been. Deposits in US banks are guaranteed at the federal level, so that bank bailouts have not been a burden on state governments; in Europe, bank bailouts have helped cause sudden jumps in government debt, most notably in Ireland, where the government's assumption of bank debts abruptly added 40 points to the ratio of public debt to GDP.

The combination of concerns about sovereign debt and the absence of federal bank backing have produced the now-famous phenomenon of "doom loops," in which fears of sovereign default undermine confidence in the private banks that hold much sovereign debt, forcing these banks to contract their balance sheets, driving the price of sovereign debt still lower.

Then there's the lender of last resort issue, which turns out to be broader than even those who knew their Bagehot realized. Credit for focusing on this issue goes to Paul DeGrauwe (2011), who pointed out that national central banks are potentially crucial lenders of last resort to governments as well as private financial institutions. The British government basically cannot face a "rollover" crisis in which bond buyers refuse to purchase its debt, because the Bank of England can always step in as financier of last resort. The government of Spain, however, can face such a crisis—and there is always the risk that fears of such a crisis, leading to default, could become a self-fulfilling prophecy.

As DeGrauwe has pointed out, Britain's fiscal outlook does not look notably better than Spain's. Yet the interest rate on British ten-year

bonds was 1.7 percent at the time of this writing, whereas the rate on Spanish ten-year bonds was 6.6 percent; presumably this liquidity risk was playing an important role in the difference.

An even more striking comparison is between euro area countries and those nations that have pegged to the euro but not actually adopted the currency. Denmark, Austria, and Finland are all, by common agreement, in pretty good fiscal shape. But where Austria and Finland are euro nations, Denmark is merely pegged to the euro. You might have thought that this lack of full commitment on Denmark's part would exact a price in the form of higher interest rates—after all, someday Denmark might choose to devalue. In fact, however, Danish borrowing costs are significantly lower than those in Finland and Austria. To be fair, this could reflect fears that all euro countries will end up being contaminated by the problems of the periphery—say, by suffering large losses on loans between central banks. But a more likely explanation is that Denmark is seen as a safer bet because it could, in a liquidity squeeze, turn to its own central bank for financing, ruling out the self-fulfilling crises that pose risks even to relatively strong euro area governments.

The bottom line here would seem to be that concerns about the euro based on optimum currency area theory were actually understated. Members of a currency area, it turns out, should have high integration of bank guarantees and a system of lender of last resort provisions for governments as well as the traditional Mundell criterion of high labor mobility and the Kenen criterion of fiscal integration. The euro area has none of these.

III. Making the Euro Workable

I will not attempt here to project the likely outcome of the euro crisis, since any such discussion will surely be overtaken by events. Instead, let me ask what it might take to make the euro workable even if it is not optimal.

One answer would be full integration, American-style—a United States of Europe, or at least a “transfer union” with much more in the way of automatic compensation for troubled regions. This does not, however, seem like a reasonable possibility for decades, if not generations, to come.

What about more limited fixes? I would suggest that the euro might be made workable if European leaders agreed on the following:

1. Europe-wide backing of banks. This would involve both some kind of federalized deposit insurance and a willingness to do TARP (Troubled Asset Relief Program)-type rescues at a European level—that is, if, say, a Spanish bank is in trouble in a way that threatens systemic stability, there should be an injection of capital in return for equity stakes by all European governments, rather than a loan to the Spanish government for the purpose of providing the capital injection. The point is that the bank rescues have to be severed from the question of sovereign solvency.
2. The ECB as a lender of last resort to governments, in the same way that national central banks already are. Yes, there will be complaints about moral hazard, which will have to be addressed somehow. But it is now painfully obvious that removing the option of emergency liquidity provision from the central bank just makes the system too vulnerable to self-fulfilling panic.
3. Finally, a higher inflation target. Why? As I showed in table 3, euro experience strongly suggests that downward nominal wage rigidity is a big issue. This means that “internal devaluation” via deflation is extremely difficult, and likely to fail politically if not economically. But it also means that the burden of adjustment might be substantially less if the overall Eurozone inflation rate were higher, so that Spain and other peripheral nations could restore competitiveness simply by lagging inflation in the core countries.

So maybe, the euro could be made workable. This still leaves the question of whether the euro even *should* be saved. After all, given everything I said, it looks increasingly as if the whole project was a mistake. Why not let it break up?

The answer, I think, is mainly political. Not entirely so—a euro breakup would be hugely disruptive, and exact high “transition” costs. Still, the enduring cost of a euro breakup would be that it would amount to a huge defeat for the broader European project I described at the start of this paper—a project that has done the world a vast amount of good, and one that no citizen of the world should want to see fail.

That said, it is going to be an uphill struggle. The creation of the euro involved, in effect, a decision to ignore everything economists had said about optimum currency areas. Unfortunately, it turned out that optimum currency area theory was essentially right, erring only in *understating* the problems with a shared currency. And now that theory is taking its revenge.

Endnote

For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see <http://www.nber.org/chapters/c12759.ack>.

References

- De Grauwe, Paul. 2011. "Managing a Fragile Eurozone." VoxEU, May. Available from <http://www.voxeu.org/article/managing-fragile-eurozone>.
- Kenen, Peter. 1969. "The Theory of Optimum Currency Areas: An Eclectic View." In *Monetary Problems of the International Economy*, edited by R. Mundell and A. Swoboda, 41–60. Chicago: University of Chicago Press.
- Mundell, Robert. 1961. "A Theory of Optimum Currency Areas." *American Economic Review* 51 (4): 657–65.

NBER WORKING PAPERS SERIES

IS EUROPE AN OPTIMUM CURRENCY AREA?

Barry Eichengreen

Working Paper No. 3579

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1991

Prepared for the conference on European Economic Integration and External Relations, University of Basel, August 23-25. I thank Kris Mitchner and Carolyn Werley for research assistance, Statistics Canada for data, and the National Science Foundation, the German Marshall Fund of the United States, and the Institute of International Studies of the University of California at Berkeley for financial support. This paper is part of NBER's research program in International Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #3579
January 1991

IS EUROPE AN OPTIMUM CURRENCY AREA?

ABSTRACT

An optimum currency area is an economic unit composed of regions affected symmetrically by disturbances and between which labor and other factors of production flow freely. The symmetrical nature of disturbances and the high degree of factor mobility make it optimal to forsake nominal exchange rate changes as an instrument of adjustment and to reap the reduction in transactions costs associated with a common currency. This paper assesses labor mobility and the incidence of shocks in Europe by comparing them with comparable measures for Canada and the United States. Real exchange rates, a standard measure of the extent of asymmetrical disturbances, remain considerably more variable in Europe than within the United States. Real securities prices, a measure of the incentive to reallocate productive capital across regions, appear considerably more variable between Paris and Dusseldorf than between Toronto and Montreal. A variety of measures suggests that labor mobility and the speed of labor market adjustment remain lower in Europe than in the United States. Thus, Europe remains further than the currency unions of North America from the ideal of an optimum currency area.

Barry Eichengreen
Department of Economics
University of California
Berkeley, CA 94720

I. Introduction

An optimum currency area (OCA) is an economic unit composed of regions affected symmetrically by disturbances and between which labor and other factors of production flow freely (Mundell, 1961). Insofar as regions within the OCA experience the same shocks, there is no obvious advantage to altering relative prices between them. Insofar as localized concentrations of unemployment nonetheless remain, the free mobility of labor from high- to low-unemployment regions can eliminate the problem. Hence it is optimal to dispense with one of the principal instruments -- changes in the exchange rate -- traditionally used to effect relative price adjustments, and to reap the benefits, in terms of convenience and efficiency, of a common currency.

The question of whether Europe is an optimum currency area is not one, unfortunately, that admits of a simple yes or no answer. Given the rapid progress of the 1992 program and the timeliness of the question, it is all the more unfortunate that the OCA literature does not provide a formal test through whose application the hypothesis can be accepted or rejected. Whatever evidence is considered, some standard of comparison is required.

A number of authors have used other continental economies already possessing a common currency and a free internal market as precisely such a standard. In Eichengreen (1990a) I analyzed balance of payments adjustment and regional labor market dynamics within the United States. Boltho (1989) compared regional income and growth rate disparities within the U.S. and the EC. Poloz (1990) contrasted the variability of relative prices across Canadian regions and the variability of real exchange rates across four European countries.

This paper presents further variations on this theme. I ask whether Europe is (and is likely to remain) further than the United States and Canada from satisfying Mundell's (1961) criteria for an OCA: free mobility of labor within the area and stability of relative prices.

Previous comparisons along these lines have been surprisingly ambiguous. Poloz (1990) found that real exchange rates between Canadian provinces are actually more variable than real exchange rates between France, Italy, the U.K. and Germany. In Eichengreen (1990a), in contrast, I found evidence of faster labor-market adjustment between U.S. regions than between E.C. members, although the difference was not large.

The evidence presented in this paper is less ambiguous. It uniformly points to the conclusion that Europe is less of an optimum currency area than its North American counterparts. Arguing that real exchange rate variability among Canadian provinces, and for that matter among France, Italy, the U.K. and Germany, is a special case, I instead analyze real exchange rate variability among all E.C. members and among the principal regions of the United States. I find that real exchange rates within the E.C. have been more variable than real exchange rates within the U.S., typically by a factor of three to four. In a second approach to analyzing the extent to which disturbances affect regions symmetrically, I examine the comovement of securities prices on the Paris and Dusseldorf stock exchanges with the the prices of shares traded in Toronto and Montreal. Once again, the comparison points to the existence of a much higher correlation of shocks in North America than in Europe.

Finally, direct evidence points to significantly lower labor mobility within Europe than within the United States. Of course, with the removal

of legal restrictions in conjunction with the 1992 program, it is likely that labor mobility within Europe will increase significantly. It is important to bear in mind, however, that the absence of legal restrictions is necessary but not sufficient for high levels of labor mobility. I use a case study of the U.S. North and South, between which a high degree of labor mobility has not always prevailed, to shed light on factors that help break down persistent regional labor market segmentation.

Thus, the bulk of the evidence suggests that the establishment of a currency union in Europe will be associated with non-negligible regional problems. This makes it all the more essential to develop the political and economic institutions necessary for the smooth operation of a currency union. Sachs and Sala-i-Martin (1989) and Eichengreen (1990) have considered the role of fiscal federalism in the U.S. as a regional shock absorber. Whether the absence of comparable institutions in Europe is a serious challenge to the case for an OCA turns out to be a complicated question. I focus on this issue in the penultimate section of the paper, approaching it both abstractly and using a case study approach.

II. Relative Price Disturbances and Region-Specific Shocks

A. Real Exchange Rates

In the OCA literature it is argued that exchange rate changes may be desirable to facilitate adjustment between regions experiencing large changes in relative prices, assuming that wages and other nominally-denominated costs are slow to adapt. A rise in German productivity relative to French productivity or a shift demand from French to German goods will require a fall in French costs and prices relative to

German, or unemployment will result. Devaluation of the franc may circumvent that problem of coordination failure that impedes the adjustment of costs and thereby accelerate the transition to the new steady state.

Thus, the more variable real exchange rates, the stronger the case for exchange rate flexibility. Poloz (1990) recently showed that regional real exchange rates within Canada are more variable than national real exchange rates between France, the U.K., Italy and Germany. The implication is that Europe is every bit as much an OCA as Canada. Quebec nationalists aside, few observers question that Canada is an OCA. Hence, the inference runs, Europe must be one as well.

There are good reasons to argue, however, that the U.S. versus the EC is a more appropriate standard of comparison than Canada versus France, the U.K., Italy and Germany. Canadian provinces are highly specialized in production. Alberta and Saskatchewan specialize in primary commodities, Ontario in manufactured goods. It is not surprising that real exchange rates between them are highly variable. France and Germany are diversified economies. Both possess substantial manufacturing, agricultural and service sectors. It is not surprising that real exchange rates between them are relatively stable.

Moreover, any case on these grounds for a floating exchange rate for Alberta or Saskatchewan is undermined by the small size of provincial populations and the thinness of provincial financial markets. Models in the OCA literature balance the benefits of devaluation by a region suffering a deterioration in its terms of trade against the loss of liquidity services it suffers with an independent currency and a variable exchange rate. The loss of liquidity services is modeled as a decreasing

function of the size of the domestic economy and the depth of its financial markets. Even if Alberta has a more variable real exchange rate vis a vis Ontario than France has vis a vis Germany, such models do not suggest that it would be more desirable for Alberta than for France to maintain a flexible exchange rate.

It may be more illuminating, therefore, to compare the different regions of the United States with all 10 EC members. Population size and the average degree of sectoral diversification are more directly comparable. So are the depth and breadth of regional financial markets in the U.S. with national financial markets in Europe.

The results of such a comparison appear in Table 1. Regional consumer price indices are calculated by the Bureau of Labor Statistics for the North East, North Central, South and West of the United States. The resulting real exchange rates can be compared with relative CPIs within the EC, converted into DM by period average market exchange rates. For the 1970s, the standard deviations of European real exchange rates, on a quarterly basis, range from 5.4 to 14.0 per cent, averaging 8.9 per cent for the period. For the four U.S. regions, standard deviations for the same period range only from 2.0 to 2.7 per cent. For the 1980s, with the decline of oil and commodity price shocks, the variability of U.S. regional real exchange rates fell to still lower levels, to the range of 1.3 to 1.5 per cent. The variability of intra-EC real exchange rates fell as well, to 1.0 to 9.6 per cent, but still averaged 5.7 per cent.

This comparison is likely to be biased by the variability of nominal exchange rates in Europe in the 1970s and 1980s. Edwards has shown for developing countries, as have Mussa (1986) and Eichengreen (1989) for

Table 1

Summary Statistics for Regional Real Exchange RatesOther EC Members Against Germany

1971.1-1979.4, 1971.1=100

	<u>Minimum</u>	<u>Maximum</u>	<u>Standard Deviation</u>
Belgium/Germany	92.71	111.62	5.55
France/Germany	99.84	122.28	5.40
Greece/Germany	100.00	122.28	5.40
Ireland/Germany	100.00	138.39	10.75
Italy/Germany	100.00	145.40	14.02
Netherlands/Germany	88.76	104.31	4.62
Portugal/Germany	86.11	118.06	9.46
Spain/Germany	83.77	117.68	7.42
U.K./Germany	100.00	150.94	13.74

1980.1-1987.4, 1971.1=100

Belgium/Germany	92.99	108.29	4.88
France/Germany	97.75	122.15	3.64
Greece/Germany	105.07	133.54	9.57
Ireland/Germany	87.91	114.83	6.28
Italy/Germany	104.79	126.19	5.67
Netherlands/Germany	89.45	93.48	1.05
Portugal/Germany	89.60	114.87	5.95
Spain/Germany	78.48	95.22	4.62
U.K./Germany	82.40	116.04	9.22

Other U.S. Regions Against the U.S. North East

1973.12-1979.12, 1977.12=100

	<u>Minimum</u>	<u>Maximum</u>	<u>Standard Deviation</u>
North Central/North East	97.253	103.730	2.06
South/North East	96.016	102.653	2.02
West/North East	94.024	103.731	2.74

1980.1-1987.12, 1977.12=100

North Central/North East	98.926	104.835	1.54
South/North East	99.195	104.444	1.32
West/North East	100.805	106.174	1.30

Notes: U.S. data are computed as quarterly averages of monthly consumer prices. Consumer prices are gathered by the U.S. Bureau of Labor Statistics for roughly 100 countries in each of the 4 regions of the U.S.

Source: see text.

industrial countries, that the variability of real exchange rates increases with the variability of nominal rates. The exceptional variability of the sterling-DM real exchange rate in the 1980s is consistent with this presumption. If the U.K., Ireland, Portugal and Greece are excluded on the grounds that they were members of the European Monetary System for at most part of the period, the average variability of intra-EC real rates in the 1980s falls to 4 per cent.

Another way to think about this point is that real exchange rates between European countries in the 1970s and 1980s have been perturbed by both real and monetary disturbances; in the United States, in contrast, monetary disturbances are common to the nation as a whole and should not have an equally dramatic effect on real exchange rates between U.S. regions. The data for the 1970s and 1980s are best interpreted, therefore, as an upper bound on the U.S.-European differential that would obtain if Europe possessed a common currency.

B. Real Security Prices

A second comparison is based on regional stock price differentials. In theory, the prices of equities should reflect the present value of current and expected future profits. If shocks are asymmetrical, profits will rise in one region relative to the other. Hence the more closely real share prices move across regions, the more symmetrical the disturbances and the more rapid the reallocation of factors of production from regions experiencing negative shocks to regions experiencing positive ones.

I compare the differentials between averages of the prices of securities traded on two regional Canadian stock exchanges (Toronto and

Montreal) with differentials between Paris and Dusseldorf. Consistent with arguments presented in Section II.A above, it would have been preferable to conduct this analysis for the U.S. instead of Canada. Though there are stock exchanges in a number of different regions of the United States (the most prominent subsidiary exchanges including Chicago, San Francisco, Philadelphia and Boston), the shares of many of the same companies are bought and sold on each of them, contaminating their share-price indices with common observations (Berlin, 1990). The two Canadian exchanges, in contrast, have nonoverlapping listings, the Montreal index specializing in enterprises located in Quebec, the Toronto index listing firms headquartered elsewhere in Canada. If the dispersion of regional shocks is smaller within Canada, we would expect prices in Toronto and Montreal to move together more closely than prices in Paris and Dusseldorf.

Share price indices were gathered for the last Friday in each quarter from issues of the Financial Times. (The Commerzbank and Herstat Bank Index for Dusseldorf is used in lieu of other German share price indices because it is the only index provided by the Financial Times for the entire period.) Since stock prices are nominally denominated, they must be adjusted for international price and exchange rate differentials. Share prices in Toronto are deflated by the Toronto CPI, share prices in Montreal by the Montreal CPI. (Unpublished CPI data were provided by Statistics Canada.) For Europe I provide two versions of the calculations, one in which franc prices are deflated by the French CPI and German prices are deflated by the German CPI, and a second, which is more appropriate if purchasing power parity does not hold, in which real French securities prices are converted into DM by the nominal exchange rate.

Table 2
Summary Statistics for Real Share Price Indices
Canada and Europe
 (1980.4 = 1.00)

	<u>Coefficient of Variation</u>		
	<u>Toronto/Montreal</u>	<u>Paris/Dusseldorf</u>	<u>Paris/Dusseldorf</u> (Exchange Rate Corrected)
1971.1 - 1987.4	.0451	.2314	.3421
1971.1 - 1979.4	.0305	.2851	.3901
1980.1 - 1987.4	.0350	.1435	.1841

Notes: Real share price indices are constructed as share price indices normalized by consumer price indices for the relevant region. Coefficient of variation is standard deviation divided by the mean. Constituent series are all normalized to 1980.4 = 1.

Source: see text.

This is not a test of the degree of capital mobility between regions. If we thought that perfect capital mobility equalized the return on baskets of securities traded on the exchanges (which might not be an appropriate assumption if the two baskets had different risk characteristics), we would expect holding period returns, or the rate of change of prices plus dividends, to be equal across exchanges. Price levels on different exchanges would move independently, reflecting changes in expected future profitability, so as to permit the preceding condition to obtain.

Table 2 displays the results for the last two decades and for the same subperiods considered Table 1. Share prices in Toronto and Montreal move much more closely together than share prices in Dusseldorf and Paris. Since the respective indices are deflated by domestic prices, inflation differentials do not account for the difference. Adjusting for exchange rate changes between France and Germany does not alter the finding. (The exchange rate adjustment increases the variability of the Paris/Dusseldorf ratio because the exchange rate is more variable than the ratio of real share prices and its covariance with the share price ratio is virtually zero.) There is strong evidence of convergence between Paris and Dusseldorf over time when the 1970s is compared with the 1980s. But even in the 1980s, the ratio of real share prices between Paris and Dusseldorf is five times as variable as the comparable ratio between Toronto and Montreal.

The strong implication of this analysis is that region-specific shocks are greater in Europe than in Canada. There are good reasons, however, to treat this comparison, like the previous one focusing on real exchange rates, with considerable caution. Firms headquartered in Quebec do

business in Ontario, just as firms headquartered in Ontario do business in Quebec. The same is true of firms headquartered in France and Germany, but the degree of interpenetration is likely to be greater at the moment in Canada than in Europe. European commodity prices will move more closely together as border taxes are eliminated, and interest rates and other financial determinants of share prices will move more closely together with the elimination of capital controls. Hence real share prices in different European markets are likely to move more closely together in the future than they do now. It is appropriate to assume that these results provide an upper bound on the North American-European differential.

II. Labor Mobility

A. The Argument and the Evidence

The more mobile factors of production within a region, the more likely that region is, *ceteris paribus*, to constitute an OCA. Consider again the mental experiment of a decline in labor productivity in France relative to Germany, or a shift in demand from the products of French firms to those of their German competitors. Assume that neither a decline in French labor costs nor a change in the nominal exchange rate is feasible. It is still possible for unemployment to be avoided if French labor can migrate freely to Germany, where a notional excess demand for labor exists.

Direct evidence on the extent of interregional labor mobility is hard to obtain. The one systematic comparison of which I am aware (OECD, 1986) concluded that mobility within the U.S. was two to three times as high as mobility within European states. Table 3 shows that in 1980, for example, 6.2 per cent of the U.S. population changed its county of residence, 3.3

per cent its state of residence. In contrast, 1.1 of Englishmen and Welshmen moved between regions, and 1.3 per cent of Germans moved between states. These comparisons must be treated cautiously in light of the very different definitions of regional units used in different countries. But the contrast seems to be too pronounced to be explicable on these grounds. Nor is it plausible that the difference reflects legal barriers to movement, since such barriers do not exist within European countries. Public policy (the council house problem in the UK, or the need to establish residence before qualifying for unemployment benefits, for example) may play a role, but the dominant explanation is that America's shared immigrant past, in contrast to the tradition of ties to one's locality in Europe, continues to influence behavior.

The problem with this evidence is that relatively low levels of labor mobility within Europe may reflect a lesser incentive to move rather than a lower level of intrinsic mobility. At the international level, less labor may move between European countries not only because of border controls but also because adjustment can take place along a number of other margins (by changing nominal exchange rates, for example). At the national level, less labor mobility may occur within European countries not because Europeans are less mobile intrinsically but because a lower incidence of asymmetric regional shocks. To address this possibility, a number of authors have considered the behavior of variables that contain information about the incentive for migration. Boltho (1989), for example, examined evidence on regional income differentials in the U.S., in the EC, and within various European countries. For 1983, the coefficient of variation of per capita incomes was 0.25 for 12 EC members, but only 0.10 for 9 U.S. census

Table 3 (continued)

GEOGRAPHIC MOBILITY -- PROPORTION OF POPULATION WHO CHANGED REGION OF RESIDENCE
(Percentage)

Country	Regional units (number of regions)	Reference population	Multi-year period data 1965-70 1970-75 1975-80	1970	1975	1980	1981	1982	1983
Sweden	Inter-counties(24)	Total popu- lation		2.4(g)	2.4	2.1	1.7	1.7	1.7
	Inter-communes	Total popu- lation		4.8(g)	4.7	4.0	3.5	3.5	3.5
	Inter-cantons(26)	Total popu- lation	7.6	6.3					

a) Employed population at the time of the survey who changed jobs during the previous year and changed usual residence when changing jobs.

- b) 1972
- c) 1976
- d) 1962-68
- e) 1968-75
- f) 1975-82
- g) 1973

Source: OECD (1986)

regions. This would appear at first glance to be strong evidence of the effects of greater factor mobility within the U.S. When the same statistic is calculated for only 9 EC members (excluding Greece, Portugal and Spain), however, it falls to 0.16. Still, a noticeable differential remains.

It is not obvious, however, whether this evidence for 1983 reflects legal barriers to migration between EC countries or cultural impediments. Here evidence on inequality within European countries is useful. The standard deviation of per capita incomes in 1983 was 0.21 for 31 regions of Germany, 0.25 for 20 regions of Italy, 0.21 for 19 regions of Spain, but only 0.16 for 48 U.S. states, as if factor mobility was greater in the U.S. than within any of these European countries. On the other hand, the comparable measures for 21 regions of France and 35 regions of the UK were only 0.15 and 0.12, respectively. (So much for the council house explanation.) The argument that the less footloose nature of Europeans leads to greater income inequality in Europe does not appear to apply uniformly.

The problem with such evidence is that simple tabulations still do not distinguish the disturbances from the response. Interregional income differentials reflect both the extent of asymmetrical shocks affecting incomes in different regions differently, and the elasticity of factor flows with respect to regional income differentials. Tabulations of migration rates reflect changes over time or across locations in the shocks that provide the incentive to migrate as well as the speed of the migratory response. Disentangling the impulse from the response requires a model. In Eichengreen (1990a) I therefore estimated time-series models of regional unemployment differentials for both Europe and the United States.

I examined the speed with which unemployment in various EC countries, when perturbed, converged to its long-run equilibrium relationship to EC-wide unemployment, and compared that with the speed which regional unemployment rates in the U.S. converged to the national average. (No assumption was imposed about the nature of the long-run equilibrium relationship.) The results suggest that regional unemployment rates adjust to one another about 20 per cent more rapidly in the United States than national unemployment rates adjust to one another within the EC. While this conclusion points in the same direction as the evidence cited above, it is still surprisingly weak evidence of slow adjustment in Europe.

Thus, it appears that greater labor mobility leads to faster adjustment to regional shocks in the U.S. than in Europe. But the differential is surprisingly small. A possible interpretation is that the mobility of other factors of production, such as capital, substitutes for labor mobility.

B. Breaking Down Barriers to Labor Mobility: An Historical Interlude

A presumption in this discussion, as in much current policy analysis, is that, with the removal of legal restrictions, labor mobility within Europe is sure to increase. By how much is a matter for debate. The absence of legal restrictions is necessary but not sufficient for labor to move freely between regions. The historical experience of the U.S. South, documented by Wright (1986), from whose analysis my discussion is drawn, illustrates the point and identifies factors that help to overcome a legacy of regional labor-market segmentation.

The origins of a separate Southern labor market are not difficult to

understand. Slavery was only the most visible manifestation of the social, cultural, political and economic institutions that differentiated South from North in the United States. After the Civil War, race relations continued to take on very different forms in the American South and North. Southern labor was provided with significantly lower levels of education than its Northern counterpart.

The result was a strikingly low level of labor mobility between the U.S. North and South in the 75 years from the American Civil War to the second world war. For fully three quarters of a century, farm wage rates without board, a good proxy for the wages of unskilled labor, in states like Mississippi and North Carolina averaged only half their equivalent in states like Ohio and Iowa.

It is important to note what does not explain these differentials. Low Southern wages were not due to the absence of a properly functioning regional labor market. Wage rates for unskilled workers in different Southern states converged steadily over the period. Wage differentials within the South were never significantly larger than wage differentials within the North. Nor do low Southern wages appear to have been due to racial discrimination. Though there is ample evidence of firm-level and occupational segregation, competitive pressures drove the wages of black and white farm laborers to equality. Given the size of the agricultural sector, this dictated the wages that industrial employers could pay for unskilled labor. If they attempted to pay less, workers would simply return to agriculture. Hence the competitiveness of the unskilled agricultural labor market equalized wages for unskilled black and white workers in industry.

Low Southern wages are sufficient to explain why neither Northerners nor Europeans migrated to the South. What then prevented low-paid Southern workers from migrating to the North? In part, the region's history of labor market segmentation perpetuated itself. The information and reception migrants require is provided typically by family or neighbors who made the trip in years past. Southerners lacked transplanted relatives and friends in the North to extend these services. In contrast, European migrants followed their relatives and former neighbors to ports of entry like New York and then to cities in the Middle and Far West. When additional employment opportunities appeared in the North, these were filled not by Southerners but by European immigrants. Wright concludes that the Northern labor market was more integrated with that of Europe than with that of the South.

One would think nonetheless that a few hearty souls would have somehow travelled north, paving the way for others. Additional factors must have contributed, therefore, to the isolation of the Southern labor market. Those additional factors, Wright suggests, were political as well as economic. Large Southern employers and landowners discouraged Northern labor recruiters who might have wished to appropriate their low-wage labor. These same individuals discouraged the provision of education on the grounds that educated workers were more likely to emigrate than others. Since literacy and numeracy enhance mobility, the existence of a substantial wage gap meant that that the South would have been unable to appropriate the benefits of additional educational spending. Agriculture and low wage industries such as textiles and timber benefitted from the elastic supply of low wage labor, and the disproportionate political power

of large landowners and industrialists prevented institutions and markets from responding so as to arbitrage the wage gap between North and South.

If Southern labor failed to move out so as to eliminate interregional wage differentials, why did Northern capital fail to move in to take advantage of cheap Southern labor? To some extent it did, as Wright shows. But Northern capital had to hurdle three barriers. First, capital and labor mobility were complementary, so barriers to one also posed barriers to the other. The difficulties of effectively monitoring investment from afar meant that capital tended to migrate across states only when its owners accompanied it or followed quickly. Hence obstacles to the immigration of persons also impeded the immigration of capital. Second, the predominance of unskilled, relatively uneducated labor in the South dictated the adoption of technologies and production processes very different from those appropriate to skilled labor in the North; Northern investors had little prior opportunity to acquire familiarity with Southern methods. Finally, wealthy Southerners discouraged outside investment, which threatened to drive down the rate of return on their own capital and undermine their political control.

What was responsible ultimately for breaking down the barriers between Southern and Northern labor markets? Wright points to simultaneous supply and demand shocks in the 1940s. On the demand side, World War II created new employment opportunities in the North and West. That the demand for labor rose in the North during wartime meant that, for once, the supply of immigrants from Europe was relatively inelastic. But similar opportunities for Southerners had opened up in the North during War War I without permanently eliminating regional labor market segmentation. Wright

suggests that World War II had more profound effects because its demand-side shock reinforced equally profound supply-side disturbances. The NIRA had reduced labor hours and established minimum wages that were binding for much of Southern industry. The Fair Labor Standards Act of 1938 made wage minima permanent. Federal incentives for agricultural mechanization further reduced opportunities for farm employment for unskilled labor. Unskilled blacks priced out of employment naturally began to seek opportunities elsewhere. The result was massive outmigration by unskilled workers once employment opportunities opened up in the North.

What are the implications of this tale for labor mobility in Europe? A first implication is that the removal of legal restrictions does not automatically produce an integrated labor market. Regional labor market segmentation can be remarkably persistent, especially if distinctive cultural and social factors are embedded in a political system that vests power in individuals with an interest in the maintenance of segmentation. A second implication is that investment in education is important for promoting interregional mobility. A third implication is that breaking down barriers to worker mobility requires policies targeted at both the demand and supply sides of the labor market.

III. Regional Self-Insurance

A. The Argument and the Evidence

A popular explanation for the tolerance in currency areas like the U.S. and Canada of region-specific shocks is that their federal fiscal systems provide regional insurance. If incomes in a U.S. state decline by \$1, federal tax payments by residents of that state decline by 30 cents,

while transfers from Washington, D.C., mostly in the form of federally-funded unemployment insurance benefits, rise by 10 cents (Sachs and Sala-i-Martin, 1989). The impact of regional shocks on inter-regional income differentials is thereby attenuated. Insofar as the locus of regional shocks shifts over time, all regions are rendered better off by risk sharing achieved via the federal fiscal system (Eichengreen, 1990b).

It is important to be clear on the nature of this argument. It is not that fiscal federalism is a necessary prerequisite for monetary unification. Historically, most federal unions established common currencies before adopting extensive systems of fiscal federalism. The United States and Canada are two obvious cases in point. The argument rather is that monetary union accompanied by fiscal federalism is likely to operate more smoothly than monetary union without it, insofar as regional problems that otherwise might arise are mitigated by interregional transfers.

Interregional transfers accomplished through federal taxes and expenditures are justifiable only if insurance cannot be provided by the market. In principle, a lumberjack or an aerospace worker in Washington state should be able to write a contract selling part of his expected labor income to an auto worker in Michigan or to an investment banker in New York City. In practice, problems of moral hazard and adverse selection prevent such diversification of human capital portfolios. Alternatively, individuals should be able to diversify away regional risk by purchasing financial assets, the returns on which are imperfectly correlated with their income streams. Most individuals seem to do so only to a limited extent, a fact for which there are two plausible explanations. The first

one is liquidity constraints: for most workers, financial wealth is a small share of undiversifiable, largely illiquid human capital. The second one is that much of the financial wealth workers possess may be tied up in their homes, the epitome of an indivisible, regional-specific asset. The large literature on state and local public finance is predicated in part on the presumption that there are intrinsic reasons why markets fail to resolve the problem, creating a role for government intervention.

But it does not follow that intervention can only occur at the federal level. Because they possess powers of taxation, state governments can compel their residents to participate in the regional insurance scheme, solving the adverse selection problem. States can borrow on the OCA-wide capital market when regional incomes decline and repay when incomes rise. This would seem to be a perfectly adequate substitute for a system of fiscal federalism.

The capacity to borrow of members of a currency area may be limited, however. The debt they can incur today is limited by the present value of the taxes they can collect tomorrow (taxes which will be used to service the accumulated debt). This is evident in the experience of U.S. states, which are forced to pay sharply rising interest rates as they continue to borrow. Given the high mobility of factors of production within the U.S., individual states cannot credibly promise to raise future taxes significantly above those prevailing elsewhere in the currency and customs union, since footloose factors of production will flee to lower tax jurisdictions. Moreover, problems of moral hazard remain. States that borrow on the OCA-wide capital market have an incentive to default when the time comes to repay the loans. As Bulow and Rogoff (1989) have noted,

reputational considerations may not help. Hence states that run budget deficits are likely to face sharply rising supply curves of external funds. As the costs of fiscal self-insurance rise, state governments may find themselves rationed out of the capital market. These factors are likely to be particularly important for EC members already burdened by high levels of public debt. Belgium, Ireland and Italy all possess public debts that approach or exceed 100 per cent of GNP. These are large debts by Latin American standards. In a recession, when the budget deficit grows and GNP shrinks, this debt-to-income ratio may rise dramatically, exacerbating difficulties of borrowing.

These, then, are the grounds for institutionalizing interregional transfers at the federal level. Table 4 summarizes the extent of fiscal transfers among governments in the United States. Clearly there does not exist the possibility of fiscal federalism on this scale in Europe, where the Community budget is on the order of 1 or 2 per cent of GNP.

Skeptics counter that factor mobility is lower in Europe than in the United States. Hence members of the EC have more latitude to vary future taxes relative to those prevailing elsewhere in the currency union. As noted above, this may be a mixed blessing: while it enhances a country's capacity to borrow, it also increases the need to borrow in a recession.

A second counterargument to the case for fiscal federalism is that fiscal transfers into a depressed region from elsewhere in the federal system discourage factors of production from moving out - that is, from reallocating themselves to other areas where their productivity is higher. This is not an argument against fiscal transfers, however, but a caution against transfers so generous as to seriously distort economic incentives.

Table 4

**U.S. INTER-GOVERNMENTAL REVENUE AS
PERCENTAGE OF RECIPIENTS' EXPENDITURES**

<u>Year</u>	<u>State Receipts from Federal Government as Percentage of State Expenditures</u>	<u>Local Receipts from Federal and State Governments as Percentage of Local Expenditures</u>
1902	1.6	5.8
1922	7.4	7.1
1932	8.0	12.8
1942	16.6	25.4
1958	18.3	24.9
1964	21.2	27.0
1967	23.2	30.3
1972	24.5	33.5
1974	23.9	39.0
1976	23.2	38.4
1978	24.6	39.8
1980	24.0	39.3
1982	21.3	37.2
1984	21.7	35.3
1986	21.8	34.4
1987	20.9	33.7

Source: Break (1967), p. 5, for 1902 - 1964.
Advisory Committee on Intergovernmental Relations,
Significant Features of Fiscal Federalism 1989, vol. 1,
for 1967-1987.

Here the optimal adjustment assistance literature, in which the marginal utility households derive from income transfers is balanced against the marginal costs of discouraging adjustment, provides guidance on how to structure a tax and transfer program.

A final counterargument (the idea for which I owe to Jacques Melitz) is that fiscal federalism, like any form of insurance, creates still other problems of moral hazard which are likely to manifest themselves in labor militancy. Consider the following example. National labor unions seeking to maximize the wage bill set the level of real wages, subject to which firms then choose the level of employment. Assume that there exist transfers from employed workers to their unemployed brethren (unemployment insurance benefits, for example). In general, the union will set wages that are above market-clearing, socially-efficient levels. If the union is region specific (a French union within a single European market, for example), and if the cost of financing unemployment benefits is shifted from French taxpayers to the EC as a whole, the French union has an incentive, *ceteris paribus*, to raise the wage it sets, creating more socially inefficient unemployment. The same holds, *ceteris paribus*, for unions in other countries. Not only does the provision of insurance thereby encourage the outcome, unemployment, whose effects it is designed to mitigate, but the magnitude of the distortion increases with the extent of fiscal federalism.

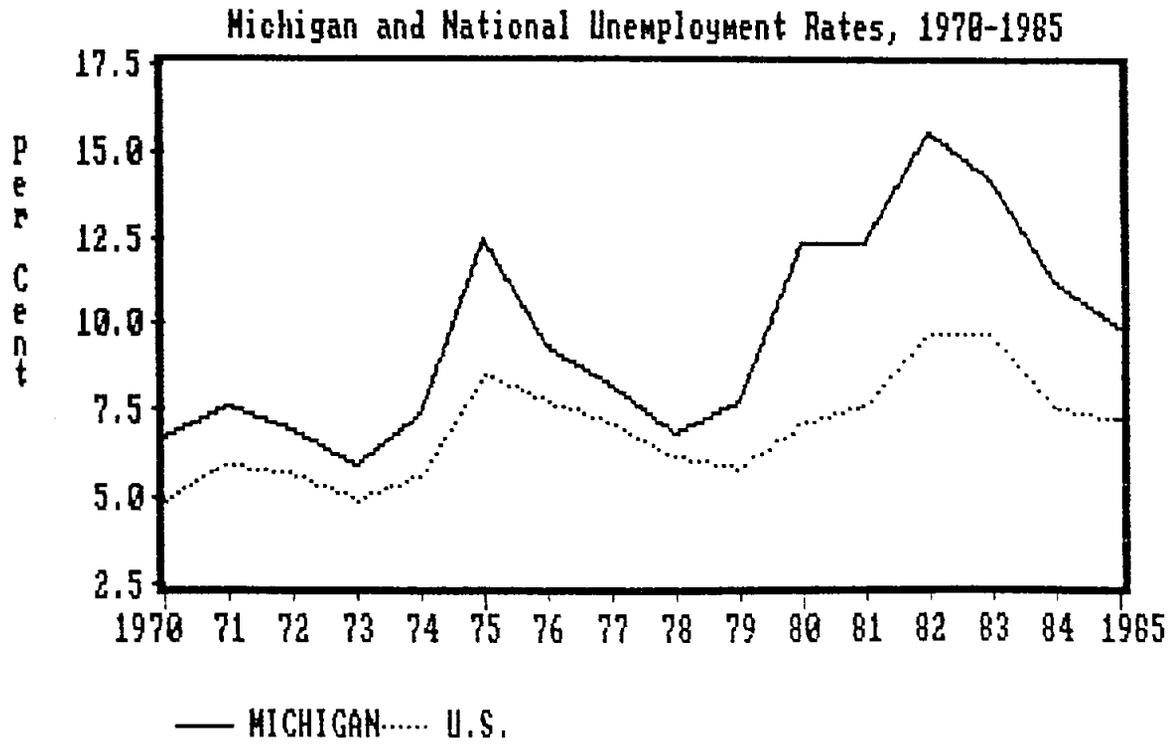
In the United States, a variety of incentive mechanisms built into the administration of unemployment insurance minimize these forms of moral hazard (Rejda, 1984). Each state administers its own unemployment insurance trust fund. In addition, states pay a fraction of the payroll

taxes levied to finance the program into a Federal Unemployment Trust Fund, which is administered by the secretary of the Treasury. States whose own trust funds that move into deficit are able to borrow from this federal fund. Significantly, however, states must pay interest on the monies they borrow from the federal trust fund. Except insofar as those interest rates are set below market levels, states are unable to shift the burden of financing their unemployment programs. Proposals for federal reinsurance of state unemployment insurance programs have been mooted in recent years; under these proposals states would pay unemployment-insurance-related payroll taxes into a federal trust fund in proportion to the value of state payrolls but draw from that fund in proportion to the level of state unemployment. Such a program might well reintroduce the moral hazard problems of which some observers warn.

B. Fiscal Federalism in Practice: An Historical Interlude

To illustrate the importance of the mechanisms described above in adjustment to regional shocks in the United States, I consider the case of Michigan's adjustment to a region-specific shock at the end of the 1970s. Michigan is the most cyclically-sensitive state economy in the U.S. (Bretzfelder, 1973). When America sneezes, the popular saying goes, Michigan catches pneumonia. The case of pneumonia I consider here is the recession that followed the 1979 oil shock. Unemployment rose nationwide following the oil shock and the adoption of disinflationary policies, but as Figure 1 makes clear it rose especially dramatically in Michigan. At its peak in 1982, the differential between unemployment in Michigan and the national average approached six percentage points. The rise in energy

Figure 1



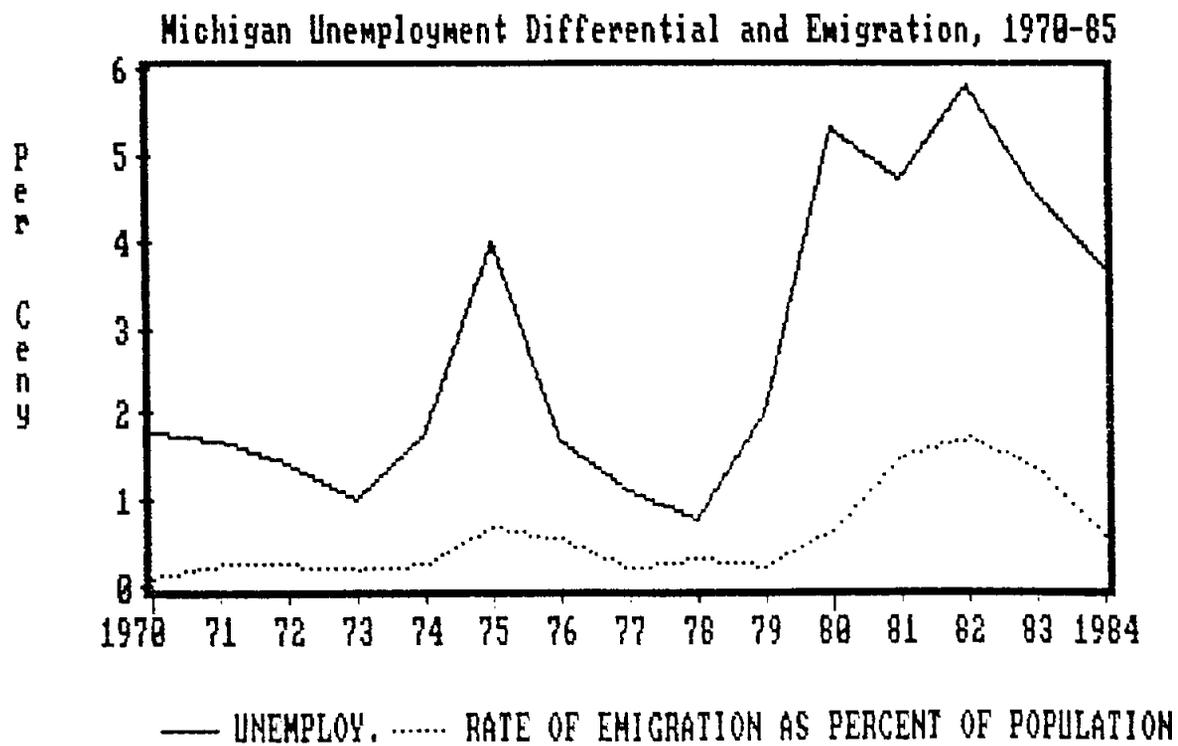
prices had a disproportionate impact on production costs in cold-winter states heavily reliant on space heating. It depressed the demand for motor vehicles as consumers substituted toward more fuel-efficient Japanese imports. Rising interest rates on consumer installment loans and lagging incomes reinforced the slump in the automobile industry.

Figure 2 displays one mechanism by which Michigan adjusted to this shock, namely outward labor mobility. The differential between the Michigan and national unemployment rates is compared with the rate of emigration from Michigan. The two lines in the figure must be compared cautiously, since their numerators differ. (Persons unemployed are expressed as a percentage of the labor force, while emigration is expressed as a percentage of state population.) Nonetheless, the figure shows that interregional labor mobility was one significant form of regional adjustment.

Figure 3 shows the swing in the state budget balance and in net federal transfers to Michigan. Since the state is bound by its constitution to run a balanced budget, the government accumulates a reserve in its Budget Stabilization Fund in good times in order to incur expenses in excess of current revenues in slumps without showing a deficit on its books. It is not the level of the state deficit or surplus that is relevant but the swing between peak and trough.

The series shown is total state revenues including those transferred to local governments minus state government expenditures. (Were transfers to local government netted out and revenues for which the state government is final recipient used instead, the line would shift down but its contours would remain the same.) At its peak, the state deficit measured on this

Figure 2

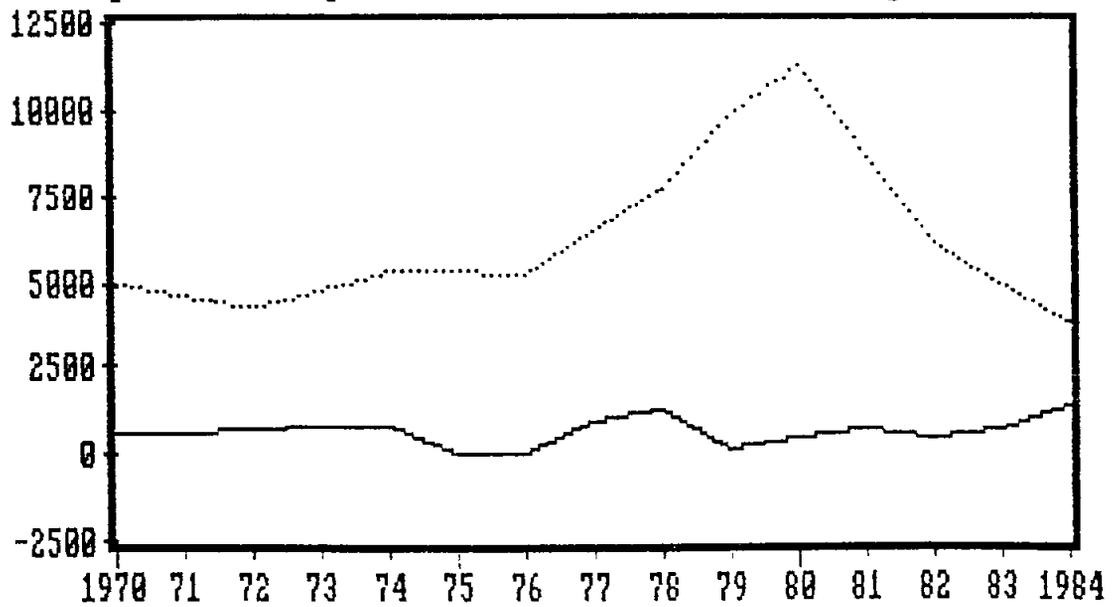


basis would have been \$422 million rather than \$1 million in 1975.

Also displayed is a measure of net transfers to Michigan from the federal government. This series is estimated by the Tax Foundation, a nonprofit research organization, and published in the Michigan Statistical Abstract (Verway, 1987). Constructing it requires assumptions about the incidence of federal taxes. To obtain a continuous time series, I have interpolated linearly where there are missing data. The series shows that the swing in net federal transfers after 1976 was large compared to the shift in the state government's budgetary position. Federal expenditures in Michigan fell short of federal tax payments by Michigan residents according to these calculations. The main reason for the disparity is the low rate of federal defense spending in the state (Erdevig, 1986). A dramatic decline in the differential is evident after 1980. Most of the swing is on the disbursement side: federal expenditure in Michigan rose by 12 per cent between 1979 and 1980 and by an additional 44 per cent between 1980 and 1981, largely reflecting transfers to support the unemployed. (Unemployment insurance and employment training, community and urban development, and Medicaid were the three categories of federal programs to show the largest increases in outlays in Michigan between 1980 and 1981.) Though the federal fiscal shift was large compared to the change in the state's budgetary position, it occurred with a lag. Unemployment started rising in 1979, yet a significant swing in federal transfers began only in 1981, once the position of the state's unemployment insurance trust fund had eroded. Although the largest swing in the state's budgetary position took place in 1978-79, it was another two years before federal transfers respond.

Figure 3

Michigan State Budget Balance and Net Federal Tax Payments (\$ Millions)



— STATE BAL. FEDERAL TAXES MINUS TRANSFER TO MICHIGAN

Though it operates with significant lags, the American system of fiscal federalism plays an important role in regional adjustment within the United States. So does the high level of labor mobility. Neither mechanism can be expected to operate as powerfully in Europe. The implication is that serious thought must be given to the cultivation of other mechanisms to facilitate regional adjustment.

III. Conclusion

This paper has argued that Europe remains further than the United States and Canada from the ideal of an optimum currency area. Real exchange rates are more variable in Europe than in the U.S., suggesting a greater prevalence of region-specific shocks and a case for nominal exchange rate changes to coordinate price-level adjustments between regions. Real securities prices are more variable within Europe, confirming the importance of region-specific shocks. Although regional disparities within Europe are sure to decline with the completion of the internal market, by how far remains a subject for debate. The extent of regional problems within existing currency and customs unions like the United States underscores the need for regional shock absorbers, such as fiscal federalism, to accommodate asymmetrical disturbances.

Rather than simply recapitulating this point, I close with another illustration. The United States currently is grappling with a savings and loan crisis. That crisis is nationwide. Yet its incidence is uneven across regions. It is concentrated in the Southwest, where a depressed regional economy attributable in part to low and falling petroleum prices led to an unusual number of nonperforming real estate loans. The liquidity

needed to prevent widespread failure of financial institutions in the Southwest is transferred from other parts of the United States by the Federal Reserve System. The funds needed to reorganize insolvent institutions and to repay the deposits of residents are transferred into the Southwest through the American system of fiscal federalism. Reportedly, depositors in Texas will receive \$20 billion this year in deposit insurance, while the U.S. Treasury collects only \$1.3 billion in taxes from the savings and loan institutions in that state. In contrast, depositors in failed institutions in Illinois will be paid only \$257 million, while the state's thrifts contribute \$1.4 billion to the government (Mashek, 1990). Complaints by governors and other representatives of states like Illinois, while not unknown, have been few and far between.

Consider, hypothetically, comparable events in Europe. Imagine a wave of involencies among financial institutions in Belgium, for example, due to a shift in demand away from the products of Belgian industry and a rise in Belgian unemployment. To prevent failures from spreading, the Belgian authorities will want to inject liquidity into the banking system. To restore depositor confidence, they will seek to reassure the public that their deposits will be protected, at least to some extent. With a fixed exchange rate and absent capital controls, however, the Belgian National Bank will not be able to increase the money supply unilaterally in order to provide liquidity. Raising taxes to finance the depositor bailout will only exacerbate the problem of insufficient demand. One solution is budgetary transfers from other E.C. members, as in the U.S. system of fiscal federalism, and the injection of liquidity from elsewhere in the currency union, as occurs within the Federal Reserve System. This is

another illustration of the problems that may occur unless existing facilities for financial swaps and fiscal transfers are expanded at an early stage in the transition to European monetary union.

REFERENCES

- Berlin, Howard M. (1990), Handbook of Financial Market Indexes, Averages and Indicators, Homewood, Ill.: Dow Jones-Irwin.
- Boltho, Andrea (1989), "Europe and United States Regional Differentials: A Note," Oxford Review of Economic Policy 5, pp.105-115.
- Break, George (1967), Intergovernmental Fiscal Relations in the United States, Washington, D.D.: Brookings Institution.
- Bretzfelder, Robert B. (1973), "Sensitivity of State and Regional Income to National Business Cycles," Journal of Current Business 53, pp.22-33.
- Bulow, Jeremy and Kenneth Rogoff (1989), "A Constant Recontracting Model of Sovereign Debt," Journal of Political Economy 97, pp.155-178.
- Edwards, Sebastian (1989), Real Exchange Rates in Developing Countries, Cambridge: MIT Press.
- Eichengreen, Barry (1989), "The Comparative Performance of Fixed and Flexible Exchange Rate Regimes: Interwar Evidence," NBER Working Paper no. 3097 (September).
- Eichengreen, Barry (1990a), "One Money for Europe? Lessons from the U.S. Currency Union," Economic Policy 10, pp.117-187.
- Eichengreen, Barry (1990b), "Costs and Benefits of European Monetary Unification," presented to the French Ministry of Finance conference on Stage 2 of the Delors Plan, forthcoming in Pierre Berezgouvoy (ed.), L'interet economique de l'union monetaire.
- Erdevig, Eleanor (1986), "Federal Funds Flow No Bargain for Midwest," Economic Perspectives X (January/February), pp.3-10.
- Mashek, John W. (1990), "Governors Vent Ire at Price Tag," Boston Globe (July 31, 1990), pp.1, 4.
- Mundell, Robert (1961), "A Theory of Optimum Currency Areas," American Economic Review 51, pp.657-665.
- Mussa, Michael (1986), "Nominal Exchange Rate Regimes and the Behavior of Real Exchange Rates: Evidence and Implications," Carnegie Rochester Conference Series on Public Policy 25, pp.117-214.
- OECD (1986), Flexibility in the Labour Market, Paris: OECD.
- Poloz, Stephen S. (1990), "Real Exchange Rate Adjustment Between Regions in a Common Currency Area," unpublished, Bank of Canada.
- Sachs, Jeffrey and Xavier Sala-i-Martin (1989), "Federal Fiscal Policy and Optimum Currency Areas," unpublished, Harvard University.

United States Bureau of Labor Statistics (1988), BLS Handbook of Methods, Washington, D.C.: GPO.

United States Senate (1980), Regional Impact of an Economic Slowdown: The Michigan Picture, Hearings Before the Committee on the Budget, 96th Congress, 1st Session, Washington, D.C.: GPO.

Verway, David, ed. (1987), Michigan Statistical Abstract, Detroit: Wayne State University Press.

Wright, Gavin (1986), Old South, New South, New York: Basic Books.

Chap 3: What lessons can be learned from the US dollar experience?

Key issues: The American monetary unification derived from a turbulent and rough political process.

- R. Sylla, "*The transition to a monetary union in the United States, 1787-1795.*" *Financial History Review*, 13, pp 73-95 (2006).
- Hugh Rockoff, "*How long did it take the United States to become an optimal currency area?*" in Forrest H. Capie and Geoffrey E. Wood (eds.), *Monetary Unions: Theory, History, Public Choice* (London, 2003)
- Henning, C. Randall and Martin Kessler. "*Fiscal federalism: US history for architects of Europe's fiscal union.*" *Bruegel Essays and Lectures Series and Peterson Institute Working Paper 12-1* (2012).

The transition to a monetary union in the United States, 1787–1795

RICHARD SYLLA¹

Stern School, New York University

The US Constitution, written in 1787, adopted in 1788 and implemented by the federal government starting in 1789, created a framework for an American monetary union. It did so by authorising Congress, the national legislature, to coin money and regulate its value, and by prohibiting states of the union from coining money, emitting fiat paper money, or making anything but gold and silver a legal tender in payment of debts. As British colonies before 1776, or as sovereign entities within an American confederation from 1776 to 1788, all US states had exercised a number of monetary powers that the Constitution prohibited to them.

To implement a monetary union, Congress adopted the recommendations of treasury secretary Alexander Hamilton, in his 1791 Mint Report, which drew on earlier discussions in the Confederation Congress of the 1780s. The report defined a new national currency, the US dollar, in terms of fixed weights of both silver and gold in a ratio of 15 to 1. Gold and silver – specie and bullion – thus became by law the monetary base of the USA. During the 1790s, although some individuals and states continued for a time to keep accounts in older state pound/shilling/pence units, the dollar based on the two precious metals became the definitive unit of account for the entire country.

Only a handful of banks existed at the time, but soon there would be many more. These banks issued bank money, notes and deposits, convertible into the specie base. A few weeks after the Mint Report appeared, Congress authorised and President Washington approved legislation embodying another of Hamilton's recommendations, the establishment of a Bank of the United States (BUS). The BUS, a corporation modelled in part on the Bank of England, was up and running by the end of 1791. Like the Bank of England, it would develop central-banking functions.

Defining a unit of account is an essential ingredient of a monetary union. A monetary union can have one medium of exchange or several, as long as the several are defined in terms of the unit of account and have a fixed exchange rate with it and therefore with one another. A single currency, however, is advantageous in

¹ I am indebted to an anonymous referee, Farley Grubb, Ronald Michener, Peter Rousseau, David F. Weiman and Robert E. Wright for comments on earlier drafts of this article. Responsibility for any errors of fact or interpretation is, of course, mine.

representing a stronger commitment to an ongoing monetary union and avoiding the transaction costs of converting currencies. Modern discussions of monetary unions also emphasise the importance, and perhaps the necessity, of common, centralised monetary policies as might be represented by those of a central bank.² The essential reason is to avoid the free-rider problems that can arise when one party to the agreement gains advantage by violating it and others do not retaliate. Since the early USA had multiple note-issuing banks and different state banking systems, a central bank with regulatory powers over note-issuing state banks usually served to make the US monetary union function better in years when it was present (1791–1811 and 1817–35) than in years when it was absent (1811–16 and 1836–63).³

The dollar monetary union, the emerging banking system, and the central bank were three of the components of the US financial revolution that occurred in the years 1789–95. The others were a funded national debt that included the assumption of state debts incurred in the War of Independence, the emergence of organised and active securities markets in leading US cities, and a proliferation of state-chartered business corporations (including banks) that sold tradable shares to investors.

² See the entry on ‘monetary unions’ in *The New Palgrave Dictionary of Money & Finance*, ed. P. Newman, M. Milgate and J. Eatwell (London, 1992), vol. 2, pp. 765–70.

³ See, for example, Bray Hammond, *Banks and Politics in America, from the Revolution to the Civil War* (Princeton, 1957); Jane Knodell, ‘The demise of central banking and the domestic exchanges: evidence from antebellum Ohio’, *Journal of Economic History*, 58 (1998), pp. 714–30; and Knodell, ‘Profit and duty in the exchange operations of the second Bank of the United States’, *Financial History Review*, 10 (2003), pp. 5–30. Within the early US monetary union, convertible notes of individual banks tended to be valued at small discounts from par, with the discount varying in proportion to the distance of the bank from money centres and hence the cost of returning a note for redemption. (Banks sometimes made arrangements for par redemption of their notes in money centres, so a discount might reflect the distance of the note holder from such a centre.) For similar reasons there were domestic exchange rates between cities and regions, reflecting relative supplies of, and demands for, domestic exchange. These phenomena, representing real economic costs, do not mean that a monetary union was not present. The two Banks of the United States tended to make note-issuing banks (rather than the public) bear note-redemption costs, and their nationwide branch networks led them to dominate the domestic exchange business, which are two of the reasons for their political unpopularity. Discounts on notes away from the issuing bank as well as costs of domestic exchange typically were lower when the central banks were in existence. Discounts on banknotes have led some to contend that the early USA was not really a monetary union, but such arguments are tautological since monetary union is defined as the absence of such discounts. In a banking system with multiple note issuers, there are always costs of redeeming notes, just as there are costs of redeeming cheques drawn on multiple banks in banking systems today. It seems unreasonable to say that a monetary union exists when banks and central banks rather than recipients of notes and cheques bear the costs of redemption, and does not exist when the opposite is the case, and thus make the existence of a monetary union turn on the distributional issue of whether money issuers or users pay the costs of operating the payments system. Nonetheless, that is the contention of Arthur J. Rolnick, Bruce D. Smith and Warren E. Weber, ‘Establishing a monetary union in the United States’, in David E. Altig and Bruce D. Smith (eds.), *Evolution and Procedures in Central Banking* (Cambridge, 2003), pp. 227–55. One wonders whether the same sort of reasoning would lead to a conclusion that today there are no monetary unions because merchants accepting credit and debit cards as payments receive a discounted payment from the card-issuing companies.

Although compressed into a far shorter time, the US financial revolution had many similarities to England's, which began a century earlier and took decades to complete.⁴ Both financial revolutions were quantum leaps toward modernity. Both helped to make their countries into leading world economic and political powers. And both were politically controversial at home. Hamilton came to be viewed by his critics as an American Walpole, a sort of 'prime minister' leading the USA down a path toward, in Thomas Jefferson's words, 'a monarchy bottomed on corruption' instead of the virtuous republic intended to result from the American Revolution's successful overthrow of British rule.⁵

The dollar monetary union was the least controversial component of the US financial revolution. Oddly, its origins seem to be more controversial now, two centuries later, than they were when it came into being. No one then argued that banknotes and deposits convertible into a specie-based US dollar were a currency inferior to the heterogeneous state-issued fiat paper currencies that were displaced by the Constitution. Although banks in the 1780s and 1790s were novel institutions to Americans, no one then attempted to persuade others that the monetary clauses of the Constitution were a successful attempt on the part of self-interested parties to usurp the rights of states and transfer the profits of paper money issuing from state treasuries to bank owners and stockholders. Yet Farley Grubb now makes such arguments and allegations.⁶

This article is an attempt to understand why the dollar monetary union was less controversial at its origins than Farley Grubb would make it today. It proceeds in three steps, each of which answers a question or a set of related questions. First, what was the nature of the colonial and early US monetary system? And why has controversy continued for more than two centuries over colonial and state fiat monies issued before 1789 by American governments? Why did the Constitution ban state-issued fiat paper, opening the way to monetary union?

Second, what were the objectives of the US founding fathers in creating the dollar monetary union? Why did Hamilton and Jefferson, political leaders who agreed on few matters of public policy, agree so easily and support one another on the dollar monetary union?

Third, why did the states, which seemingly had the most to lose in the transition from state-issued fiat paper to the convertible US dollar, not register vociferous complaints – indeed, there were barely any complaints at all – in response to the

⁴ P. G. M. Dickson, *The Financial Revolution in England: a Study in the Development of Public Credit, 1688–1756* (London, 1967); John Brewer, *The Sinews of Power: War, Money, and the English State, 1688–1783* (New York, 1989).

⁵ Jefferson as quoted by Ron Chernow, *Alexander Hamilton* (New York, 2004), p. 398.

⁶ Farley Grubb, 'Creating the US dollar currency union, 1748–1811: a quest for monetary stability or a usurpation of state sovereignty for personal gain?' *American Economic Review*, 93.5 (Dec. 2003), pp. 1778–98. For a counterargument, see Ronald W. Michener and Robert E. Wright, 'State "currencies" and the transition to the US dollar: clarifying some confusions', *American Economic Review*, 95.3 (June 2005), pp. 682–703.

constitutional invalidation of state monetary sovereignty? Did the dollar monetary union specifically, and the US financial revolution in general, offer them benefits that outweighed the costs?

Given that we have come full circle, with the US dollar and other currencies of the world now being fiat paper detached from a specie base, a concluding section suggests that the answers to these questions are of more than historical interest. If fiat paper money is generally accepted now, it is certainly fair to ask, as Farley Grubb does, why it was not thought to be acceptable when the founders wrote the US Constitution.

I

In the British colonial era before 1776, Americans became practised monetary innovators.⁷ They adopted and extended the use of wampum, the token money of native Americans made from sea shells which were strung together like beads to make smaller and larger money units. They gave monetary status to a variety of commodities, including tobacco, corn, rice and beaver pelts. And probably for the first time in the history of the western world, they made widespread use of fiat paper currencies.

Fiat paper currency, or bills of credit, took two forms, tax-anticipation script and loan-office notes. The first of these arose when a colonial government, short of funds to meet current demands for expenditure, would issue bills of credit and promise to levy taxes in the future to redeem the notes. Massachusetts made the first issue in 1690. Seven other colonies adopted the innovation by 1712, and the remaining five of the thirteen colonies followed between 1723 and 1755.

Colonial loan-office notes were a form of financial intermediation. They were issued as loans to landholders against mortgages on their lands and other assets. Interest on the loans became an important source of colonial government revenues, reducing the need to resort to taxation. The notes returned to the loan office as borrowers repaid their loans. Often they were reissued. South Carolina began the loan-office innovation in 1712, and all colonies apart from Virginia had adopted it by 1755.

Practically, bills of credit and loan-office notes were indistinguishable within a colony, with both circulating as paper currency. When received by colonial governments in tax or loan payments, they were often reissued for purposes similar to or different from the original purpose for issuing them. In a premodern economic world, the colonial economy, or economies, was unusual in growing at a rate of 3 per cent or more a year – we can infer that from a 3 per cent rate of population

⁷ This and subsequent paragraphs draw on two earlier articles: Richard Sylla, 'Monetary innovation in America', *Journal of Economic History*, 42 (1982), pp. 21–30; and Sylla, 'Monetary innovation and crises in American economic history', in Paul Wachtel (ed.), *Crises in the Economic and Financial Structure* (Lexington, MA, 1982), pp. 23–40.

growth – and it needed money stock expansion to allow that growth to occur and avoid the possibly depressing economic effects of deflation.⁸

Colonial American fiat paper money can best be viewed as a growth-accommodating local monetary supplement to the acknowledged monies of the world, gold and silver in the form of coins of various nations and empires, plate and bars of bullion. These precious-metal monies also circulated in the colonies, and were especially valuable as foreign exchange, although trade among colonies and with other countries typically was financed with bills of exchange drawn on colonial mercantile and foreign financial centres.⁹ Although colonial governments did not formally promise to convert their fiat currencies into specie, in times of peace the issues and reissues of fiat paper were usually managed well enough to keep them roughly at par with specie values. In the New England and southern colonies, however, over-issuing of fiat paper in wartime exigencies led to large depreciations in terms of specie, unless – as happened in the Seven Years' War (called the French and Indian War in America) – the British greatly increased their specie spending in America at the same time. In the middle colonies, fiat paper held its value quite well in relation to specie for long periods. The conclusion of Farley Grubb and many earlier writers that colonial American fiat paper money worked tolerably well in many (but not all) contexts and served useful functions is thus well taken.

Problems with fiat paper money became more severe during the War of Independence and continued in the postwar confederation of the 1780s. The Continental Congress that formed an army and made George Washington its commander in 1775, and then declared independence from Britain in 1776, proceeded to issue and over-issue paper money, so-called Continental currency, until by 1780 the money had become virtually worthless. At the same time, the independent state governments that succeeded British colonial governments did the same. All this is understandable. Congress had no power of taxation, and revolutionaries almost never have good credit with arms-length lenders, leaving paper-money issues and forced loans as the remaining domestic options of revolutionary national public finance. The states did have tax powers, but had learned as colonies that financing wars by issuing fiat paper currency made more political sense than draconian tax increases. Politically motivated loans from other nations, particularly France, supplemented the American revolutionaries' domestic sources of finance.

Here should be raised a rather telling point that seems oddly overlooked in the voluminous literature of colonial and early US monetary history. The paper money of the American Revolution, whether issued by Congress or by states, at its inception was not fiat paper money at all. It explicitly stated that it was intended to be a convertible paper money. A typical early Continental note on its face said, for example: 'This Bill entitles the Bearer to receive EIGHT Spanish milled DOLLARS, or

⁸ Sylla, 'Monetary innovation and crises'.

⁹ John J. McCusker, *Money & Exchange in Europe and America, 1600–1775: a Handbook* (Chapel Hill, 1978).

the value thereof in Gold or Silver, according to the Resolutions of the CONGRESS, held at Philadelphia, the 10th of May, 1775.’ Later emissions said virtually the same thing, even after depreciation had set in: ‘This Bill entitles the Bearer to receive Fifty Spanish milled DOLLARS, or the value thereof in Gold or Silver, according to a Resolution passed by CONGRESS at Philadelphia, September 26th, 1778.’ A typical state note, this from South Carolina, said: ‘This Bill intitles the Bearer to receive FOUR Spanish milled DOLLARS, or the Value thereof in Gold or silver, according to an Act of the General Assembly passed at CHARLES-TOWN, the 23rd Day of December, 1776.’

The South Carolina note also said on its face and on the obverse side that \$4 were equivalent to £6 10s. in the South Carolina unit of account.¹⁰ All the states, like their colonial predecessors, had local units of account expressed in pounds, shillings and pence. Each state’s unit of account differed from others and from sterling, in terms of gold and silver. The Spanish silver dollar, however, was a common hard-money coin circulating in the USA before and after independence, and it became the model for the US dollar – see below, Section II.

That the monies of the American Revolution were not intended to be fiat paper is quite suggestive in the context of current debates on the transition to the dollar and a US monetary union. It suggests that in 1775 and 1776 American revolutionary leaders were in no sense wedded to the virtues of fiat paper money, as could be and has been inferred from colonial experience. It also suggests that they could gain credence about what had been only implicit with colonial paper, by being explicit about the intended convertibility of their paper money to a hard-money base. It suggests that the Spanish dollar, the most familiar coin, or its equivalent in other forms of silver and gold was to be that base, just as these precious metals had been the *de facto* monetary base of the colonial era. And it suggests that a US monetary union with convertible paper was not an invention of Constitution writers a decade later, but rather an idea that was prevalent at the time of the Declaration of Independence and probably even antedated it. During much of the colonial era, the fiat paper currencies issued by American governments, even when not explicitly stated to be convertible to specie, were in practice convertible at customary prevailing rates.¹¹

The problem, of course, was that in practice neither the Congress nor the states were able to fulfill the convertibility promises they made in their Revolution-era paper-money issues. They lacked sufficient specie reserves and ways of gaining such reserves to allow them to honour their promises. Therefore, the currencies that they had meant *de jure* to be convertible currencies became *de facto* fiat paper currencies that lost much of their value in terms of hard money. Recognising the depreciation, Congress and state authorities wrote down the fiat currencies to small fractions

¹⁰ The examples of the currencies quoted here are portrayed in Robert Garson, ‘Coining money: the US dollar and American nationhood’, *Journal of American Studies*, 35 (2001), pp. 25, 27.

¹¹ See Ronald Michener and Robert E. Wright, ‘State “currencies” and the transition to the US dollar: clarifying some confusions’, *American Economic Review*, 95 (2005).

of their original values. At those reduced rates, Continental dollars became a component of the national debt.

In 1781, the wealthy merchant-revolutionary Robert Morris became Congress's Superintendent of Finance. One of his first recommendations was that Congress charter a bank that would issue convertible banknotes and serve as the government's bank. The Bank of North America (BNA) opened at the start of 1782, after the decisive American–French victory over Lord Cornwallis's British army at Yorktown ended the land phase of warfare, but almost two years before Britain recognised American independence in the Treaty of Paris. Most of the BNA's capital came from a specie loan from France, part of which Morris on behalf of Congress invested in BNA shares and then borrowed back from the bank. The BNA was the first US bank in the modern sense of the term. Since the power of Congress to charter a bank was unclear, several states also granted it a charter. After receiving wartime financial services from the BNA, Congress and Morris repaid BNA loans by selling the government's shares to private investors. By 1783, when the war officially ended, the BNA became a local commercial bank instead of a national public bank.

At the very time that the BNA was opening, Morris and his assistant, Gouverneur Morris (no relation), proposed to Congress a plan for a US monetary unit and a mint to produce an American coinage. The unit was based on the Spanish dollar, but was to be only one–1440th of the Spanish coin, a fraction selected because the pound/shilling/pence values of the various state currencies in terms of Spanish dollars were divisible into it. Robert Morris also proposed a decimal coinage, an idea that had been suggested by Jefferson in 1776. The plan went nowhere because of its complexity and Congress's weakness.¹²

Jefferson maintained an interest in monetary units as well as weights and measures. In 1784 he proposed a simpler plan than the Morris plan, making the US dollar much the same as the familiar Spanish dollar, but with a decimal coinage. Congress adopted Jefferson's plan in 1785, and authorised a national mint. According to Ferguson, 'Jefferson's admitted aim was "to get rid of" the state currencies, and he proposed to do so by introducing the Spanish milled dollar as the money unit.'¹³ Again the weakness of the Confederation Congress led to inaction on implementation. The Morris and Jefferson proposals and congressional resolutions indicate, however, that before the Constitution was written, national leaders with a wide range of political views had agreed that the USA should have a national currency based on specie, that it should be a decimal currency, and that the confusion of separate state currencies ought to be ended rather than retained. A specie-based

¹² E. James Ferguson (ed.), *The Papers of Robert Morris*, vol. 4 (Pittsburgh, 1978), pp. 25–40.

¹³ *Ibid.*, p. 28. In correspondence with me, Farley Grubb interprets the 'to get rid of' phrasing of Jefferson as implying that Jefferson wanted the Spanish dollar to be only the unit of account, and would not have any reservations about states issuing paper money denominated in that unit. The passage from Jefferson may be ambiguous. He was a proponent of states' rights and yet he never went on record as objecting to the Constitution's ban on state fiat money issues.

national currency was by no means a new idea, much less a sinister new idea, at the Philadelphia convention of 1787, although the convention appears to be the first occasion on which the prohibition of state currencies was voiced.

What changed between the early 1780s and 1787 was further experience with state fiat money issues. Seven states – Rhode Island, New York, New Jersey, Pennsylvania, North Carolina, South Carolina and Georgia – made further issues of fiat paper after 1783, when peace returned. Much but not all of the experience with these new issues was unfavourable, reminding Americans of the disastrous recent experiences with fiat paper depreciation during the revolution. Edwin Perkins reports that Pennsylvania's paper depreciated by 30 per cent relative to specie by 1788, and that New Jersey's depreciated by 25 per cent. North Carolina's two issues depreciated by 25 to 50 per cent, and Georgia's, by 75 per cent. New York and South Carolina issues, on the other hand, held their value at issue fairly well. Rhode Island's paper had the worst record, losing six-sevenths of its value within 15 months of its issue in 1786.¹⁴

The Rhode Island experience was particularly nettling to the writers of the Constitution and participants in the debates over ratification. A rural faction controlled the state legislature that issued the loan-office paper in 1786. The legislature declared its fiat paper to be a full legal tender, and then attempted to sink the state's debt in 1787 by requiring debt holders to exchange their securities for the fiat paper at par. When debt holders resisted this plan, the legislature cancelled the debts that were not presented for redemption in depreciated paper.¹⁵ Others refusing to accept the paper were threatened with fines and other legal sanctions. These arbitrary measures of a democratically elected state legislature seemed a misguided if not illegal invasion of property rights to the writers of the Constitution, and it is likely that Rhode Island's actions had much to do with the document's ban on further state issues of fiat paper. The Rhode Island legislature's monetary measures were successfully challenged in Rhode Island courts as contrary to the state's Constitution, but not before numerous creditors had been deprived of their property. In the pamphlet

¹⁴ Edwin J. Perkins, *American Public Finance and Financial Services, 1700–1815* (Columbus, OH, 1974), pp. 142–65. Perkins says the 'record of those monies as a store of value was reasonably good', although his account in most instances contradicts that contention. He appears to mean that the issues were ultimately redeemed when accepted by states in tax and loan payments after the Constitution went into effect. That is true, but it is a separate matter from the depreciation that occurred before the paper issues were redeemed.

¹⁵ *Ibid.*, pp. 154–57. A more detailed account of the Rhode Island experience, as well as those of Pennsylvania, New Jersey and New York, is given by John P. Kaminski, *Paper Politics: the Northern State Loan-Offices During the Confederation, 1783–1790* (New York and London, 1989). Like Perkins, Kaminski thinks that the experience with fiat paper money of the four northern states he surveys was reasonably good, although not on the criterion of currency stability, that is, lack of depreciation. Kaminski introduces a new criterion for 'success', namely, 'whether or not it [fiat paper] met the purposes of the politicians who advocated it' (p. 281). On that criterion, which says little about the *legitimacy* of politicians' purposes, perhaps most of history's fiat moneys might be considered successes.

literature about ratification and in ratification debates, Rhode Island was cited numerous times in justifying the ban on state-issued fiat paper.¹⁶ Those who sympathised with states' rights to issue such currency could do little more than remain silent. The ban, the ratification debates indicate, was not a very contentious issue, although some individuals and towns went on record as favouring state paper money. Ironically, Rhode Island, a state opposed to ratification of the Constitution to the bitter end, with its extreme monetary policies probably did more than any other state to help federalists achieve ratification.¹⁷

Grubb excuses what many saw as Rhode Island's excesses on the ground that merchant speculators by 1786 had purchased the state's wartime debt securities at deeply discounted prices. Therefore, little or no accounting loss resulted from the state forcing the speculators to accept deeply depreciated paper money in payment of them.¹⁸ But if these actions were only fair play, one might ask why Rhode Island, shortly after it ratified the Constitution in May 1790, decided to reverse its earlier

¹⁶ Governor Edmund Randolph of Virginia, speaking in the Virginia ratification convention on June 4, 1788, in response to a concern that only nine states needed to approve the Constitution for its ratification, said, 'Ought we have required, the concurrence of all the thirteen? Rhode-Island, in rebellion against integrity; Rhode-Island plundering all the world by her paper money, and notorious for her uniform opposition to every federal duty, would then have it in her power to defeat the Union, and may we not judge with absolute certainty from her past conduct, that she would do so?', *The Debate on the Constitution*, ed. Bernard Bailyn (New York, 1993), part 2, p. 603. In Massachusetts, 'Cassius VI' (*Massachusetts Gazette*, 18 Dec. 1787) said of the proposed constitutional ban on state fiat paper, 'The absolute necessity of power of this nature being vested in a federal head is indisputable. For want of such a power, what vile proceedings have of late disgraced almost every legislative measure of Rhode Island', *The Documentary History of the Ratification of the Constitution*, ed. Merrill Jensen, John P. Kaminski and Gaspare J. Saladino, 18 vols. to date (Madison, 1976–), vol. 4, p. 482. 'A Native of Virginia', 2 April 1788, said, 'The making of any thing but gold and silver a tender in payment of debts, and the impairing of the obligation of contracts is so great a political injustice, that the Constitution here requires the States, that they will forever relinquish the exercise of power so odious. This part of the clause would probably never have been introduced, had not some of the States afforded too frequent instances of unjust laws upon these subjects', *Documentary History*, vol. 9, p. 676. In Connecticut, 'The Landholder XII' (*Connecticut Courant*, 17 March 1788) excoriates Rhode Island's tender acts, saying, 'The rogues of every other state blush at the exhibition, and say you have betrayed by carrying the matter too far. The very naming of your measures is a complete refutation of antifederalism, paper money and tender acts, for no man chooses such a company in argument', *Documentary History*, vol. 17, p. 406. The *Pennsylvania Gazette*, 13 August 1788, citing a private letter noting that North Carolina had voted to reject the Constitution, said, 'North-Carolina and Rhode-Island have probably been permitted to withdraw from the union, on purpose to shew the other states the miseries they have escaped by adopting the government. An attachment to paper money and tender laws, appear in both these corrupted and deluded states, to be the cause of their opposition to the new constitution', *Documentary History*, vol. 18, p. 401. Unfortunately, the *Documentary History* volumes of the ratification debates in North Carolina and Rhode Island, states that experienced severe depreciation of their fiat paper and were the last two states to ratify, have yet to be published.

¹⁷ See Kaminski, *Paper Politics*, p. 258.

¹⁸ Grubb, 'Creating the US dollar currency union', fn. 17, pp. 1789–90.

actions and declare all of its earlier acts to redeem state securities with fiat paper money null and void. It then restored the redeemed securities to their owners after deducting the real (not nominal) value of the fiat paper they had received, and cancelled the forfeiture that had extinguished state debts not presented for redemption in state paper. The reason, of course, was that Congress was about to adopt Hamilton's proposal to assume state debts into the national debt, and Rhode Island, now a member of the union, desired to have as much as possible of its debts so assumed. In the end, most of the Rhode Island debt was so assumed, leading Kaminski to say that Rhode Island state creditors were 'treated justly'.¹⁹

The foregoing account of colonial and early US monetary experience leaves out much, but I think it is one upon which most recent scholars of the subject would broadly agree. Why then has early American fiat money always been a contentious subject, and still is, judging by Farley Grubb's recent work? One reason is that the data available from that early era are simply insufficient to answer many of the questions financial historians would like to answer. There are tolerably good records of the amount of fiat paper money outstanding at most dates, and the same could be said for price and exchange rate data. But data on the amount of specie in the country are not there, so it is not possible, at least without making heroic assumptions, to come up with estimates of the total money stock and its composition. Good output statistics are also lacking, and we have no idea of how much output passed through market exchanges as opposed to being consumed or invested by those who produced it. Cliometricians nonetheless have tried to use what data there are to test, for example, various theories of money.²⁰ But the data are simply insufficient to do that. One product of deficient data bearing on issues of economic and historical importance is endless controversy.

A second reason for continuing controversy is that there were many differences among colonies and states in the provisions authorising fiat money issues, and in the experiences that resulted from it. In some places and at some times, mostly in the middle colonies, fiat paper held its value in relation to specie for long periods; in others, mostly New England and southern colonies, it did not. In some places and at some times, it was a full legal tender; in others it was only a legal tender in public but not in private payments. Sometimes various commodities as well as fiat paper were granted legal-tender status, which truly complicates any attempt to estimate money stocks. Sometimes fiat paper was issued for long periods with little thought given to redeeming it; this was less a problem than it might seem because the colonial economy was expanding at high modern rates, leading to a growing demand for money. Sometimes the paper was redeemed by taxation relatively quickly and then reissued. Sometimes it was redeemed and not reissued. The record is so rich that, by

¹⁹ Kaminski, *Paper Politics*, pp. 255–8.

²⁰ For an attempt to estimate one colony's money stock, see Farley Grubb, 'The circulating medium of exchange in colonial Pennsylvania, 1729–1775: new estimates of monetary composition, performance, and economic growth', *Explorations in Economic History*, 41 (2004), pp. 329–60.

selecting parts of it, one can tell almost any story one wants to tell. And scholars like to generalise. One could, for example, contend that colonial fiat issues were not a full legal tender, and the good experience several colonies had with it in contrast with bad experience in, say, Rhode Island during the 1780s when fiat issues were made a full legal tender, demonstrates that full legal-tender status, not fiat money per se, was the problem. But this argument is suspect because many colonial fiat issues were a full legal tender, with both good and bad results. Some bad results provided a reason for the British to ban further legal tender issues in the Currency Acts of 1751 (which applied to the New England colonies) and 1764 (which applied to the rest of the colonies).

A third reason for controversy is that the dominant political economy changes from one historical era to another. When the dominant political economy supported the specie standard and convertible bank money from the 1790s to the 1930s, scholars selected from bad colonial and confederation experiences, ignoring good experience, to argue against any tampering with the specie standard.²¹ After the Keynesian revolution in economic thought came along, scholars selected from good colonial experience, downplaying much of the bad experience, to argue that managed fiat money was good for prosperity and an antidote to depression.²² When as now all money is fiat money and democracy has become something of an end itself instead of a means to the ends of good government and human happiness, a scholar such as Farley Grubb can question why the Constitution banned fiat money two centuries ago, and find in the ban a self-interested, anti-democratic plot. If anything, study of the debates over the Constitution and related writings from that era reveals that the views of the US founders were more balanced in weighing the pros and cons of fiat and convertible paper moneys than were the views of many subsequent interpreters.

II

As noted above, the Confederation Congress in 1785 adopted Jefferson's proposal for a decimal US dollar based on the Spanish dollar. No action, such as establishing a mint, was taken then to implement it because Congress lacked the resources. It remained for Hamilton, the architect and guiding force of the US financial revolution, to complete the work after the Constitution went into effect. This he did in his Mint Report transmitted to Congress on 28 January 1791. Hamilton provided Jefferson with an advance copy, which Jefferson wrote Hamilton on 24 January that he had 'read over with a great deal of satisfaction'.²³ Jefferson's satisfaction derived

²¹ Charles J. Bullock, *Essays in the Monetary History of the United States* (New York, 1900); Horace White, *Money and Banking* (Boston and New York, 1911).

²² Richard A. Lester, *Monetary Experiments – Early American and Recent Scandinavian* (Princeton, 1939).

²³ Harold C. Syrett (ed.), *The Papers of Alexander Hamilton* (New York, 1962–86), vol. 7, p. 451. The final version of Hamilton's Mint Report is in the same volume, pp. 570–607.

from Hamilton's adopting, extending and refining the essence of Jefferson's earlier plan for the US dollar. With minor revisions Congress adopted Hamilton's plan in 1792, and established a US mint. Hamilton and Congress defined the dollar as 24.75 grains of pure gold and 371.25 grains of pure silver, adopting bimetallism at a silver-gold mint ratio of 15:1, roughly the market ratio at that time. None of this was controversial. Gold and silver had always been the *de facto* monetary base of the country, in the colonial period and the Confederation. In the federal union, they became the monetary base by law.

A monetary union was an explicit goal of Hamilton's plan, which implicitly served the further goal of national unification. He wrote at the beginning of the report:

The unequal values allowed in different parts of the Union to coins of the same intrinsic worth; the defective species of them, which embarrass the circulation of some of the States; and the dissimilarity in their several Monies of account, are inconveniences, which if not to be ascribed to the want of a National Coinage, will at least be most effectually remedied by the establishment of one ...²⁴

Since the Constitution had already taken from the states the right to issue new fiat paper money, once the US dollar was adopted as the national unit of account in 1792, all that remained to achieve monetary union was for the old, pre-1788 state issues to be retired. In some states it happened quickly. In others it took two to three decades, longer than the recent transition to the euro, although amounts outstanding during those decades were not large. Moreover, because the US mint ran into problems in supplying the country with an adequate supply of dollar coins, the legal-tender status of foreign gold and silver coins lasted until the 1850s, far longer than the few years Hamilton had contemplated in his report, and foreign coins continued to be a significant part of the circulating medium. Nonetheless, for most American governments, businesses and individuals, the new dollar in the 1790s became the preferred unit of account.

The US monetary union was a relatively non-controversial act of national sovereignty on which most national leaders of the 1780s agreed. To give a monetary union full sway, the Constitution also granted Congress powers to regulate domestic as well as international commerce, and removed from the states the power to lay duties on imports from or exports to other US states and other countries. Congress could lay taxes on imports to the USA, but not exports. All import duties (and also internal taxes) had to be uniform throughout the country. The USA was to be a common market, an internal free-trade area with free movement of economic resources, as well as a monetary union.

For Hamilton, these measures were parts of a more comprehensive plan of financial, economic and political unification. Other parts of Hamilton's plan, though related to the monetary union, were more controversial. One was the assumption of

²⁴ Ibid., vol. 7, p. 571.

state debts. Another was the Bank of the United States. What role did these measures play in Hamilton's overall plan? And how did they relate to the monetary union?

In the 1780s, the national government as well as all the states had large debts left over from the Revolution. Under the Confederation the states had the power to tax and the power to emit fiat paper. Both powers were applied to debt reduction. Shays's rebellion – a taxpayers' revolt in Massachusetts in 1786 – demonstrated the hazards of a state using its taxing powers in a heavy-handed way to redeem its debts. Rhode Island in 1786, as already mentioned, furnished an example of the heavy-handed use of fiat paper essentially to redeem debts by repudiating them. Rhode Island's actions were likely instrumental in the Constitution's ban on further state fiat issues. Shays's rebellion, apart from adding to the impetus for constitutional reform, did not result in any action at Philadelphia to address the financial problems of indebted states.

Under the Confederation, Congress could do little to service its own debts. Chastened by the collapse of its wartime paper money, Congress decided to issue no more of it. And it had no power to tax. Congress could only levy requisitions – requests – for funds from state governments and hope that the states would comply. Hamilton later, in his unfinished 1795 'Defence of the Funding System', described the requisition system as one of 'imbecility and injustice' because it neither provided the national government with the funds it needed nor (essentially because of free-rider problems) allocated national revenue burdens equitably among states.²⁵

The Constitution solved the national revenue problem by giving the new federal government the power to tax. It did so by giving federal and state governments, apart from import taxes reserved for the federal government, concurrent access to the same tax bases. This was a plan, Hamilton said, that while 'most expedient . . . involved inherent and great difficulties. It may not unaptly be styled the Gordian Knot of our political situation. To me there appeared but one way of untying or severing it, which was in practice to leave the states under as little necessity as possible of exercising the powers of taxation.'²⁶ Though little remarked on by later historians, the primary motivation for Hamilton's controversial plan of assuming state debts into the national debt was not to enhance federal power. Rather, it was designed to prevent a potentially ruinous competition of the federal and state governments for the same tax bases. If states had their fiscal burdens lightened by debt assumption, they would not miss being unable to issue fiat paper money, and they would not have to raise state taxes in ways that might threaten the union just as it came into being.

In fact, the controversial assumption of state debts did more than remove much of the obligation of states to service their debts. In the early 1790s settlement of Revolutionary War accounts that attempted to equalise the costs of independence across states according to population, the so-called creditor states received balances

²⁵ *Ibid.*, vol. 19, p. 22.

²⁶ *Ibid.*, vol. 19, p. 23.

due from so-called debtor states in the form of federal debt securities. These assets provided the states with interest revenues, reducing both their need to tax and any regrets they might have had over losing the authority to issue fiat paper money. Strictly speaking, the amounts due to creditor states should have come from debtor states that had paid less than their fair shares of the costs of defeating Great Britain. But Hamilton expediently did not enforce that obligation, so even the debtor states had their fiscal burdens lightened by federal assumption of state debts. 'It is curious fact which has not made its due impression', Hamilton commented in the 1795 'Defence', 'that in every state the people have found relief from assumption while an incomparably better provision than before existed has been made for the state debts'.²⁷

If the pre-constitutional, pre-assumption situation had remained intact, Hamilton argued that the first threat of war 'would have convinced us of the ineligibility of our situation, of the weakness and embarrassment incident to fifteen or perhaps to 50 different systems of finance'.²⁸ Thus the assumption of state debts by the federal government eased the path to monetary union. And the monetary union made national debt management, with three federal securities all denominated in dollars, much easier than it would have been if the new nation had a different unit of account in every state instead of one national unit.

In Hamilton's overall plan, the relationship of the Bank of the United States to the US monetary union is more easily understood than the relation to it of assumption. Hamilton had studied financial history and was one of the few Americans of that era who understood the role of banks and bank money in a modern economy. While Washington's aide de camp and a colonel in the continental army, Hamilton had called in 1780 for the establishment of a 'Bank of the United States' to stabilise the currency and aid the war effort. That was before any modern bank had appeared in the USA. He made a similar recommendation to Robert Morris, the superintendent of finance, in 1781, even outlining a complex bank charter. Morris and Congress then founded the BNA, a much less ambitious institution than the one proposed by Hamilton. In 1784, Hamilton was a co-founder of the Bank of New York, the country's second or possibly third bank, one in Boston starting at virtually the same time. His first actions after becoming secretary of an empty US treasury in September 1789 were to arrange loans from the BNA and the Bank of New York to begin to pay the new government's unpaid bills.

Little more than a year later, his debt funding and assumption proposals enacted, Hamilton in December 1790 proposed to Congress that it establish a Bank of the United States as a public or national bank. His proposal followed the model of the Bank of England, although with innovations such as partial government ownership and the possibility of opening branches around the country. He used the proposal to tutor Congress and the country on the nature of bank money and how it differed

²⁷ *Ibid.*, vol. 19, p. 35.

²⁸ *Ibid.*, vol. 19, pp. 24–5.

from fiat paper money. In the context of this article, his words are worth quoting at some length:

The establishment of Banks in this country seems to be recommended by reasons of a peculiar nature. Previously to the revolution circulation was in a great measure carried on by paper emitted by the several local governments . . . This auxiliary may be said to be now at an end. And it is generally supposed, that there has been for some time past, a deficiency of the circulating medium . . .

If the supposition of such a deficiency be in any degree founded, and some aid to circulation be desirable, it remains to inquire what ought to be the nature of that aid.

The emitting of paper money by the authority of Government is wisely prohibited to the individual States, by the National Constitution. And the spirit of that prohibition ought not to be disregarded, by the Government of the United States.²⁹ Though paper emissions under a general authority, might have some advantages, not applicable, and be free from some disadvantages, which are applicable, to the like emissions by the States separately; yet they of a nature so liable to abuse, and it may even be affirmed so certain of being abused, that the wisdom of the Government will be shewn in never trusting itself with the use of so seducing and dangerous an expedient. In times of tranquility, it might have no ill consequence, it might even perhaps be managed in a way to be productive of good; but in great and trying emergencies, there is almost a moral certainty of its becoming mischievous. The stamping of paper is an operation so much easier than the laying of taxes, that a government, in the practice of paper emissions, would rarely fail in any such emergency to indulge itself too far, in the employment of that resource, to avoid as much as possible one less auspicious to present popularity . . .

Among other material differences between a paper currency, issued by the mere authority of Government, and one issued by a Bank, payable in coin, is this – That in the first case, there is no standard to which an appeal can be made, as to the quantity which will only satisfy, or which will surcharge the circulation; in the last, that standard results from the demand. If more should be issued, than is necessary, it will return upon the bank. Its emissions . . . must always be in a compound ratio to the fund [specie reserves] and to the demand: Whence it is evident, that there is a limitation in the nature of the thing: While the discretion of the government is the only measure of the extent of the emissions, by its own authority.

This consideration further illustrates the danger of emissions of that sort, and the preference, which is due to Bank paper.³⁰

Hamilton's Bank Report thus makes the case for a paper currency consisting of banknotes and deposits convertible into gold and silver, 'the money of the world' as he described them elsewhere in the report, as a better alternative than an inconvertible paper currency issued by government.³¹

²⁹ This sentence indicates that Hamilton, one of its writers, was not among those who believed that the Constitution removed the right of the federal government to issue fiat money, although some readers of the records of the 1787 debates in Philadelphia, for example Bray Hammond, *Banks and Politics in America from the Revolution to the Civil War* (Princeton, 1957), ch. 4, infer that the delegates intended to do that. Farley Grubb in his contribution here agrees with Hammond's interpretation of the Philadelphia convention debates.

³⁰ Syrett, *Papers of Hamilton*, vol. 7, pp. 320–2.

³¹ *Ibid.*, p. 316.

The Bank Report was controversial, but not because it held that bank money convertible into a specie base was preferable to fiat currency issued by government. No one objected to that aspect of the report. What did seem controversial was whether the federal government had a right to charter the bank or indeed any other corporation, an issue that arose with respect to the BNA a decade earlier and was solved by having several states also charter the BNA. James Madison in Congress in 1791 argued that the Constitution did not give Congress the power to charter a bank. Like Madison, other members of the constitutional convention served in the first Congress of 1789–91; unlike Madison, many of them voted for the BUS, which handily received the approval of both the Senate and the House. President Washington thereupon asked Edmund Randolph, the attorney general, and Jefferson, the secretary of state, for their opinions. Each agreed with Madison's position that the BUS was not explicitly authorised by the Constitution. Washington then asked Hamilton for his opinion, sharing with him the Randolph and Jefferson opinions. Hamilton's landmark response, a sweeping discussion of the Constitution's explicitly stated powers and of related powers implied but not stated by it, persuaded the president to approve the bill.³² Hamilton got his bank, and the doctrine of implied powers that he used to win his case would come to be regarded as a fundamental principle of constitutional law.

The BUS was a crucial element of Hamilton's plan for a financial revolution, as well as an immediate source of loans to meet interest payments on the national debt before revenues became adequate for that purpose. Private investors subscribing to the \$8 million in shares of BUS stock reserved for them could pay with \$6 million of national debt securities – shades of the Bank of England – and \$2 million of specie. In that way, the BUS supported the national debt, and the interest on the debt supported the Bank. The IPO sold out quickly on 4 July 1791, and the BUS opened in Philadelphia in December, with branches established in Boston, New York, Baltimore and Charleston the following year. The example of the BUS encouraged states to charter more banks. The numbers of state banks rose from three in 1790 to 20 by 1795. Thus, by 1795 the USA had a banking system of 25 banks including the

³² A full account of all the reports, debates and opinions regarding the BUS is in M. St Clair Clark and D. A. Hall, *Legislative and Documentary History of the Bank of the United States* (Washington, 1832; reprint New York, 1967), ch. 2. Forrest McDonald, *Alexander Hamilton* (New York, 1979), ch. 9, argues that the Virginians Madison, Randolph and Jefferson did not have strong feelings about the constitutionality issue, but based their opposition to the BUS on a presumed threat by congressmen from Pennsylvania, where the BUS would be based in Philadelphia, to use its presence there as a lever to prevent the national capital from moving from Philadelphia in 1800 to a new federal district carved out of Virginia and Maryland, as had been agreed in 1791 in order to secure votes needed to pass the federal assumption of state debts. Support for McDonald's argument is perhaps furnished by the closing paragraph of Jefferson's opinion against the BUS: 'unless the President's mind, on a view of everything, which is urged for and against this bill, is tolerably clear that it is unauthorized by the constitution; if the *pro* and the *con* hang so even as to balance his judgment, a just respect for the wisdom of the Legislature, would naturally decide the balance in favor of their opinion . . .', Clark and Hall, *Legislative and Documentary History*, p. 94.

BUS and its branches. All of them would keep their accounts in the new national currency, the US dollar.

The federal government subscribed for \$2 million of BUS stock, paying for it with a loan from the BUS. It was the first of many loans to the government. The government paid the BUS 6 per cent interest on the loan and agreed to repay it in equal instalments over ten years. The main intent of the government in taking a stake in the BUS was to launch the national bank and to circulate its banknote currency as quickly as possible, for the benefit of the treasury and the country. But Hamilton also noted that to the extent dividends on BUS shares exceeded the interest on the government's loan from the BUS, which they invariably did because of expanding loan demand and the typical balance-sheet leverage of banks, the government would make a 'positive profit'. Hamilton was teaching by example. He wanted state governments, which previously had been suspicious of banks and reluctant to charter them, to learn that chartering and investing in more banks and other corporations could well be profitable activities. The lesson was not lost on the states – see Section III below.

By 1792, the US monetary union that most of the founding fathers desired was more or less a reality. Few of them may have understood that more controversial measures such as the federal assumption of state debts and the launching of a BUS were integral components of making the convertible-dollar monetary union work. But the gears of the machine that was the financial revolution were beginning to mesh. There was no longing for the old system of state-issued fiat paper money. Moreover, if states wanted to establish loan offices along with banks, nothing stopped them from doing that, as long as the loans were made in specie or convertible bank notes rather than fiat bills.

After a speculative bubble inflated early in 1792, the USA experienced its first financial panic and securities-market crash. These were embarrassments to the new financial system and to Hamilton. They produced much railing against banks, speculators and stockjobbers. But acting as a central banker, Hamilton in cooperation with existing banks deftly managed the crisis with timely liquidity injections, and it quickly passed with minimal damage to the booming US economy.³³ The new financial system had been tested for the first time, and it passed the test.

III

Why did US states and state leaders remain strangely silent about the loss of monetary sovereignty that came with the ratification of the Constitution? They lost the profits that came from stamping paper, declaring it to be money, and purchasing goods and services with the paper. They lost the interest income that came from stamping pieces of paper as loan-office notes and lending them to citizens on

³³ David J. Cowen, *The Origins and Economic Impact of the First Bank of the United States, 1791–1797* (New York, 2000).

mortgage security. They lost the ability, if indeed they thought they had it, to fight state depressions by expanding their monetary issues, thus exercising 'democratic' control over state economic and political destinies. That was a lot to give up, so it is surprising that anti-federalists, who expressed many other reservations about the cession of state sovereignty to a federal union, made few protests about ceding monetary sovereignty.

Monetary sovereignty was not the only area in which states ceded authority to the federal union under the Constitution. States lost the right to tax imports from and exports to other states and other countries. The right to tax imports was exclusively the federal government's. That loss of states was offset, of course, by the federal commitment to take over the costs of defending the states and the country against foreign and other incursions. Before the Constitution, such defence costs periodically had loomed large in state budgets. And longer term, of course, there were economic advantages of a state having unimpeded access to an expanding nationwide market area.

Other advantages the states gained by acceding to a monetary union were mentioned earlier. The controversial assumption of state debts by the federal union, not usually connected with the goal of monetary union although it was so intended in Hamilton's thinking, reduced states' revenue requirements. And for the creditor states, assumption added to revenues by providing interest income from federal securities received in the settlement of state accounts from the War of Independence.

A more important reason why states did not protest the loss of monetary sovereignty is that the loss did not amount to much, if anything at all, in terms of state revenue. That possibility was suggested in two studies that John Legler, John Wallis and I carried out more than a decade ago.³⁴ Both studies indicated that in the first half of the nineteenth century, a number of states relied on banks for significant proportions – at times up to half – of state revenue. Most state revenues derived from banks came from taxes on them or investment income from them. Tax revenues came from annual taxes on bank capital, and from the fees and bonuses state legislatures charged banks in return for charters of incorporation and charter renewals. Investment income came from dividends on bank stock, interest on state deposits in banks, and sometimes from the income of banks wholly owned and operated by states.

The studies also found that states benefited from banks in a version of what today would be called 'off-budget financing', that is, states would demand as a condition of granting a charter or a charter renewal that a bank seeking the charter would have to

³⁴ Richard Sylla, John B. Legler and John Joseph Wallis, 'Banks and state public finance in the new republic: the United States, 1790–1860', *Journal of Economic History*, 47. 2 (June 1987), pp. 391–403, and John Joseph Wallis, Richard Sylla and John B. Legler, 'The interaction of taxation and regulation in nineteenth-century US banking', in Claudia Goldin and Gary D. Libecap (eds.), *The Regulated Economy: a Historical Approach to Political Economy* (Chicago, 1994), ch. 4, pp. 121–44.

provide financial aid to institutions and enterprises that states wanted to support. By inducing banks to provide such support, states avoided the need to support them through state budgets, thus keeping taxes lower than they otherwise would have been. The fact that most early US banks were state-chartered corporations, in contrast with banks in other countries where non-corporate proprietorships and partnerships were the rule, made such practices possible. A charter of incorporation, whether in Europe (where there were fewer of them) or America, was a potentially profitable privilege granted by the state for which it could demand compensation in a number of ways. The fact that the US economy was expanding more rapidly than other economies meant that demand for banking services expanded faster than states could meet it by chartering more banks. That made banking a highly profitable activity. So new banks were willing to pay states for their corporate banking privileges. Less often, established banks were willing to pay states not to charter new competitors. Over the early decades of US history, democratic political pressures increased access to bank charters. By the 1830s, these pressures culminated in free banking, a term in the US context connoting the chartering of banks under general incorporation laws rather than by special legislative acts.

State legislators obviously learned that bank chartering could become a revenue source. But how soon? The evidence is that they were not aware of the possibilities in the 1780s, but became aware of them in the 1790s when Hamilton delivered his lesson about the profit possibilities of government investment in BUS stock and other securities created during the financial revolution.

New York was the first state to act. From 1784, when the Bank of New York was organised, until 1791, the state had rebuffed several requests from the bank for a charter, so it operated for seven years as an unincorporated joint-stock bank. In 1791, shortly after Congress and the president approved the legislation creating the BUS, New York granted a corporate charter to the Bank of New York. This was for defensive reasons. The BUS had the power to open branches nationwide and would certainly open one in New York City. Unless the state wanted to cede the ground of banking to the federal government, it needed to have its own bank – hence the change of heart about the charter.³⁵ Despite such worries, New York (along with several states) invested in the IPO of the BUS in 1791, and in 1811, when Congress narrowly voted down renewal of the BUS charter, New York still held \$60,800 of BUS stock.³⁶

New York invested in its own state banks as well as in the BUS. In 1792, New York exercised a charter option by making an investment of \$50,000 in the stock of the Bank of New York, and it authorised a \$20,000 investment in the newly chartered Bank of Albany, an option it exercised in 1797. The Bank of New York became the depository for state funds. The bank held and collected the interest on

³⁵ Robert E. Wright, *Banking and Politics in New York, 1784–1829* (Ann Arbor, 1996), ch. 4.

³⁶ Joseph Stancliffe Davis, *Essays on the Earlier History of American Corporations* (Cambridge, MA, 1917), vol. 2, p. 107.

nearly \$2 million of the new US debt that the state owned, and it made loans to the state. Thus, the Bank of New York became the state's fiscal agent in the 1790s, just as the BUS was the fiscal agent of the federal government. When New York state chartered the Bank of Manhattan in 1799, it wrote into the charter an option to buy \$50,000 of stock in it, exercising that option after 1800.³⁷

There were similar developments in other states. In Massachusetts, the handsome profits of the Massachusetts Bank chartered in 1784 aroused republican opposition in the early 1790s. The state in 1791 tried to tax the bank, but that was overturned in the courts. A more significant action was chartering a new bank, the Union Bank, in 1792. In an unsuccessful attempt to stave off that potential competition, the Massachusetts Bank, taking its cue from Hamilton's charter for the BUS, offered the state the opportunity to purchase a 20 per cent stock interest in it. The Union Bank charter instead gave the state the option to invest \$200,000–400,000, a quarter to a half of the capital initially authorised, in the new bank's shares. Massachusetts exercised those options, taking the first \$200,000 during 1793–5, and the second \$200,000 in 1795–6. The Union Bank became the state's depository, and the state wrote into the charter that it could borrow \$100,000 from it on demand at 5 per cent interest, repayable in annual installments for up to five years. Also in the charter was a directive to the bank to lend up to one-fifth of its capital outside of Boston, preferably to Massachusetts farmers, in one-year loans upon real estate security.³⁸ In short, a state bank issuing convertible bank money could be made to function in very much the same manner as a colonial or state loan office issuing fiat paper currency. An added advantage of banks was that their lending for profit was likely to result in lending decisions that were more economically efficient than those of politicians operating colonial and state loan offices.

Not to be outdone by New York or Massachusetts, Pennsylvania chartered the Bank of Pennsylvania in 1793 with an authorised capital of \$3 million. By the charter, the state took \$1 million of that, paying for it with \$375,000 in specie, \$375,000 in the new US debt that it owned, and – taking a cue from Hamilton's BUS – the balance with a loan from the bank. In an echo from the colonial and confederation eras, the charter required the bank to lend to the state an additional \$500,000 for the establishment of a loan office enacted in subsequent separate legislation. To the relief of the bank, the state abolished the loan office a year later on the grounds that it had been found 'inexpedient and not to answer the purposes intended by the legislature'.³⁹ Like the Bank of New York and the Union Bank of Boston, the Bank of Pennsylvania became the state's fiscal agent.

On smaller scales, similar actions occurred when banks were chartered in other states. New Hampshire in 1792 invested in its first bank, located in Portsmouth. Maryland chartered the Bank of Columbia in 1793 to aid in the formation of the

³⁷ *Ibid.*, vol. 2, ch. 2.

³⁸ *Ibid.*, vol. 2, ch. 2.

³⁹ *Ibid.*, vol. 2, p. 96, citing and quoting Pennsylvania statutes.

new federal district and national capital being laid out on the banks of the Potomac. The state did not invest in the bank, but the charter authorised the Washington commissioners to subscribe for one-fifth of the stock, and the commissioners actually did subscribe for more than a tenth of it. Connecticut authorised a state subscription to the Hartford Bank in 1792, as did Maryland in the case of the Bank of Baltimore chartered in 1795; both subscriptions were made in 1803. In Rhode Island, the organisers of the Providence Bank chartered in 1791 reserved 10 per cent of the shares for the state, but the state did not subscribe.⁴⁰

States' fiscal interests in banks became greater in the early decades of the nineteenth century, but the precedents were established in 1791. The catalyst was the BUS, which represented to state leaders almost an invasion of their territory by an agent of a new, superior, and somewhat suspect federal government. The BUS, however, also demonstrated to them how they might benefit financially by imitating the federal government's investment in it, and therefore suffer no loss from the cession of monetary sovereignty to the federal union. It is likely that Hamilton intended the BUS to have that effect, for part of his genius was to create incentives that would induce others – state governments chartering banks and other corporations, entrepreneurs forming corporations and other enterprises, and private investors organising securities markets – to complete the financial revolution he had initiated.

Hamilton was not the only one to understand that states could use banks to accomplish the purposes of colonial and state fiat currencies. Writing to Rufus King on 7 August 1791, a month after the IPO of the BUS had been heavily oversubscribed in one hour, Boston federalist and business leader Christopher Gore also saw it clearly:

Some who were shut out from the national bank will be very solicitous of reaping the profits heretofore experience'd from the State Bnk [sic]. The State legislatures are all organiz'd bodies & disciplin'd to counteract the operations of the federal government – it is now a war of words only – we may have to deprecate a different kind of warfare. We well know the State Legislatures are generally dispos'd to discern & oppose the power & influence of the nation; but their meanness & parsimony interfere with their views & besides they are without a military check. Give them a State Bank to recur to and this great evil will be remov'd. The State Bank will continue for a long time to afford great profits to the proprietors of their stock. The State Legislatures will be proud of imitating the conduct of Congress in having a bank under their patronage & having a continuance of this badge of Sovereignty, the emission of money, & the interest of the Commonwealth & the Bank will be made to combine . . .⁴¹

State legislatures, as Gore predicted, were quick to grasp that, with a strong say in how US banking developed, they had not given up the profits of money issuing when the US dollar monetary union came into being.

⁴⁰ Ibid., vol. 2, ch. 2.

⁴¹ Quoted in *ibid.*, vol. 2, pp. 54–5.

IV

The transition to the US dollar monetary union more than two centuries ago was relatively non-controversial. Before and after 1789, the acknowledged monetary base of the colonies and the USA was gold and silver. Colonial and state fiat currency issues were local supplements to that base, prompted for the most part by monetary strains that rapid population growth – 3 per cent or more per year in the colonial era – placed on economic expansion. Fiat money issues were also induced by British colonial policies that frowned on the establishment of colonial mints and banks, and on the accumulation of specie in America. Fiat supplements were rather easily abandoned in the constitutional settlement. That settlement, as it was implemented, reduced the revenue needs of states while at the same time increasing state revenues derived from interest on the national debt and from banks.

The US founding fathers believed that a national currency was an expression of sovereignty. It would as well promote national unity by providing Americans with one currency for all. A common currency reduced transactions costs, made price comparisons among different states and regions easier, and facilitated capital mobility across regions unhindered by the possibility that exchange rate changes would reduce or eliminate returns to investment. With one currency, and freed by the Constitution from political constraints on interstate commerce, the US economy in time became one of the largest free-trade areas in world history. In defining the US dollar in terms of gold and silver, the founders were basing US currency on the acknowledged moneys of the world. That facilitated the new nation's international commerce as well as a persistent inflow of foreign capital. Judging by a recent survey of the effects of monetary unions, which finds that, other things held constant, 'currency union seems typically to be associated with a significant increase of trade, ranging from over 30% up to 90%', the economic intuition of the US founders, tutored as they may have been by Adam Smith, was not misplaced.⁴²

US states on net did not lose much of consequence when the Constitution took away their power to issue fiat currencies. They soon were to realise that the power to incorporate banks and other companies, which they retained, could serve just about every purpose that fiat currency served. And they gained the advantages of a monetary union. That in essence is why the loss of monetary sovereignty by states was non-controversial.

States did lose something, however, at least potentially. They lost the ability to use independent monetary policy and exchange-rate changes to alleviate local distress produced by economic shocks of various kinds. If, for example, the agricultural terms of trade turned against a state or region because of such a shock, being part of a monetary union meant that currency devaluation could not be used to

⁴² Andrew K. Rose, 'A meta-analysis of the effect of common currencies on international trade', NBER Working Paper 10373 (Cambridge, MA, March 2004), p. 13.

cushion losses of bank reserves. Banking panics might result, and we know the USA had such panics in the nineteenth and early twentieth centuries, especially during the periods when there was no central bank. Financial crises themselves could have been independent sources of shocks. States in a monetary union lacked the power to alleviate such shocks with state monetary expansions and lender-of-last-resort actions, as Farley Grubb and others have argued.⁴³ But would US states, if they had had their own currencies from the 1790s to the early twentieth century, also have had the knowledge, ability and will to undertake such modern economic-policy actions with any great effectiveness? That is debatable, and probably to be doubted.

Since fiat currency issued by American colonies and states before 1788 was a supplementary local currency in a monetary system based on gold and silver, not a fiat-based monetary system, it hardly represented a leap toward modernity that was set back by the US constitutional settlement for a century and a half. Modern financial technologies – insurance of bank liabilities and the tools of fiat monetary management by independent central banks, for example – make fiat-based monetary systems in principle, if not always in practice, effective. Those technologies were either not available or just beginning to emerge two centuries ago. Such considerations make it unlikely that the USA would have been better served by having every state issuing its own fiat currency instead of joining the US dollar monetary union. None of the leading founders – Franklin, Washington, Adams, Jefferson, Madison and Hamilton – although they disagreed among themselves on many things, favoured retaining state monetary sovereignty with fiat paper currencies. They were nation builders, and the monetary union was a component of nation building.⁴⁴ Given their successes and subsequent efforts in history to emulate them, it seems odd two centuries later to be second-guessing the motives for and the results of the US-dollar monetary union.

Author's address:

*Department of Economics
Leonard N. Stern School of Business
New York University
44 West 4th Street, Room 8-65
New York, NY 10012-1126
USA
rsylla@stern.nyu.edu*

⁴³ Hugh Rockoff, 'How long did it take the United States to become an optimal currency area?', in Forrest H. Capie and Geoffrey E. Wood (eds.), *Monetary Unions: Theory, History, Public Choice* (London, 2003).

⁴⁴ On nation building, see Max M. Edling, *A Revolution in Favor of Government: Origins of the US Constitution and the Making of the American State* (Oxford, 2003).

NBER WORKING PAPER SERIES
ON
HISTORICAL FACTORS IN LONG RUN GROWTH

HOW LONG DID IT TAKE THE UNITED STATES TO
BECOME AN OPTIMAL CURRENCY AREA?

Hugh Rockoff

Historical Paper **124**

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2000

I learned a great deal from the discussion of a preliminary draft by Ronald I. McKinnon, Anna J. Schwartz, and other participants in the Conference on Monetary Unions at the City University Business School, London, Friday May 14, 1999, and from the participants in the meeting on the Development of the American Economy, the Summer Institute of the National Bureau of Economic Research, Cambridge MA, July 14 1999. I have also learned a great about these issues from Michael Bordo, Robert Greenfield, and Eugene White. Dongbo Pei provided superb research assistance. The remaining errors are mine. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

© 2000 by Hugh Rockoff. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How Long Did It Take the United States to Become an Optimal
Currency Area?

Hugh Rockoff

NBER Historical Paper No. 124

April 2000

Development of the American Economy

ABSTRACT

The United States is often taken to be the exemplar of the benefits of a monetary union. Since 1788 Americans, with the exception of the Civil War years, have been able to buy and sell goods, travel, and invest within a vast area without ever having to be concerned about changes in exchange rates. But there was also a recurring cost. A shock, typically in financial or agricultural markets, would hit one region particularly hard. The banking system in that region would lose reserves producing a monetary contraction that would aggravate the effects of the initial disturbance. Plots of bank deposits by region show these patterns clearly. Often, an interregional debate over monetary institutions would follow. The uncertainty created by the debate would further aggravate the contraction. During these episodes the United States might well have been better off if each region had had its own currency: changes in exchange rates could have secured equilibrium in interregional payments while monetary policy was directed toward internal stability. It is far from clear, to put it differently, that the United States was an optimal currency area. This pattern held until the 1930s when institutional changes, such as increased federal fiscal transfers (which pumped high-powered money into regions that were losing reserves) and bank deposit insurance, addressed the problem of regional banking shocks. Political considerations, of course, ruled out separate regional currencies in the United States. But thinking about U.S. monetary history in this way clarifies the nature of the business cycle before World War II, and may suggest some lessons for other monetary unions.

Hugh Rockoff

Department of Economics

Rutgers University

College Avenue Campus

New Brunswick, NJ 08901

and NBER

rockoff@fas-econ.rutgers.edu

I. The Troubled History of the American Monetary Union

The U.S. monetary union began with the ratification of the Constitution in 1788. It has remained intact, with the exception of the Civil War years, ever since. And the United States has grown and prospered during that time. But it does not follow that the United States has grown and prospered because it has had a monetary union. The benefits of the monetary union – the relative ease with which interregional movements of capital, labor, and final products could take place – are evident. The costs of the monetary union, that I intend to focus upon here, are less evident, but nonetheless important.

Throughout the first 150 years of the U.S. monetary union, at least, the United States was wracked repeatedly by bitter regional disputes over monetary policies and institutions. On more than one occasion, those disputes contributed to uncertainty about the future of policies and institutions that exacerbated economic disturbances, and contributed to mistakes in national monetary policy. Regional disputes over monetary policy arose because of real differences in regional interests: What was good monetary policy from the point of view of one region, was sometimes bad policy from the point of view of another. The most bitter disputes arose when adverse monetary reactions occurred in a region already suffering from a real shock. A decline in the demand for agricultural products, for example, would depress incomes, leading in turn to a round of bank failures and bank runs, and declining regional money supplies, that reinforced the effect of the initial shock. In short, an economic historian

who is looking for illustrations of the cost of relinquishing monetary autonomy can find them in abundance in the monetary history of the United States.

Before turning to the history, however, I want to briefly summarize the theory of optimum currency areas to provide the necessary background for the remainder of the paper.

II. The Theory of Optimal Currency Areas

The optimum currency area hypothesis grew out of the debate over fixed vs. flexible exchange rates. Milton Friedman (1953), Leland Yeager (1959), and others had argued that a country could be better off by allowing its currency to float, and reserving domestic monetary policy for price or employment stability. The advocates of flexible exchange rates had recognized that a country could be too small to profit from flexible rates. But it was Robert Mundell, who first used the term “Optimum Currency Area” in a famous paper published in 1961, who clarified the circumstances under which a region or country would benefit from joining a monetary union.¹ On the one hand, Mundell argued, there were advantages for a region that joined a monetary union derived from minimizing transaction costs. On the other hand there were disadvantages derived from giving up the exchange rate, and changes in the stock of money, as policy tools. Whether a particular region constituted an optimal currency area or whether it would be better off as part of a larger monetary union depended on the net sum of these costs and benefits.

¹ Other important early contributions were McKinnon (1963), Kennen (1969), and Tower and Willet (1976). Kawai (1992) provides a clear summary.

The benefits of a larger monetary union are usually fairly easy to see, although measuring them can be difficult. People can travel to one part of the country to another without having to convert their money; prices of products sold in distant regions can be compared without having to search for information or perform calculations, and interregional investments can be made without the risk of currency fluctuations.

The costs of joining a monetary union are less obvious, and will depend on a number of factors. Consider first, the case in which economic activity is distributed randomly throughout the monetary union. Then the monetary authority need pay little attention to regions. What is optimal for one part of the country will be optimal for another. But suppose that the monetary union is divided into two regions, say East and West, that specialize in producing different goods, say steel and wheat. Now it is possible for there to be significant shifts in demand between the regions. The demand for steel might go up, while the demand for wheat is going down. To use the modern jargon, the country might be subject to asymmetric shocks. The West will run a balance of payments deficit with the East. Reserves will flow from West to East, and the stock of money in the West will fall, aggravating the recession caused by the decline in the demand for wheat. The stock of money in the East will rise, adding to the boom caused by the increase in the demand for steel. It is no longer clear that what is good for one region is good for all. The West might be better off with a national monetary policy that aimed to restore full employment, while the East might want a policy directed toward price stability.

If labor and capital are mobile across regions, then the impact of asymmetric shocks will be limited. Labor, for example, will respond to the shift in demand by moving from West to East, from wheat production and into steel production. The monetary authority will be able to neglect the problem of unemployment and focus on price stability. But now suppose that barriers to labor and capital mobility exist among regions, and assume further that monetary policy can affect real magnitudes in the short-run, say because certain prices or wages are sticky. Then the monetary authority faces a real dilemma. If the monetary authority follows a policy consistent with price stability in the East, it might aggravate the recession in the West; if it tilts full against the recession in the West, it might produce inflation in the East.

The East and West, to put it slightly differently, would be better off with separate currencies, and floating or at least adjustable rates between them. When demand shifted from the West to the East, the western currency would depreciate, mitigating the effects of the decline in demand. Meanwhile the monetary authority in the East could follow a policy aimed at price stability.

This is essentially the theory as originally developed by Mundell. The story depends on imperfections in capital and labor markets, and price and wage stickiness. As we move from the world of Keynesian or classic Monetarist economics, where monetary policy has important short-run effects, to assumptions of perfect labor and capital mobility and ineffective monetary policy, the case for subdividing economic regions into separate currency areas weakens. Nevertheless, as a number of writers have argued, a case for separate currencies may remain (Willett and Wihlborg, 1999). For example, even if all factor markets cleared, a separate currency area could be

justified if non-optimal policies were being followed outside the area. A flexible exchange rate then would prevent the importation of non-optimal price level movements. In any case, I believe, as I will try to show below, that American monetary history offers important examples of regional monetary trends exacerbating the effects of asymmetric real shocks. Before turning to those examples, however, I want to briefly recount the origins of the American monetary union and the regional conflicts that afflicted it during its formative years.

III. Union

Prior to the Revolution the currency of the United States varied from colony to colony. The British pound, and other forms of hard currency such as the Spanish peso were accepted everywhere. But individual colonies also tried to make their own paper currencies legal tenders in order to provide revenues or to aid debtors. Under the Articles of Confederation (during the interregnum between the end of the Revolution and the Constitution) opinion, especially among the politically sophisticated, turned against paper money. The constitution prohibited the states from issuing “bills of credit” (paper money) and gave to Congress the exclusive right to “coin money” and “establish the value thereof,” thus creating a monetary union based on specie.

Part of the opposition to paper money was based on the experience of the very high rates of inflation under the fiat paper money regime of the Revolution. Tom Paine, for example, went from being an advocate of paper money to an opponent (Schweitzer 1989, p. 315) after witnessing the Revolutionary inflation. Opposition to paper money was increased by the development of fractional reserve banking, and the

hope that bank notes would provide the convenience of government issued paper without the risk of overissue.

There was also an important regional dimension to the opposition to state issued paper money. Rapid deflation after the Revolution had left farmers with heavy debt burdens. In the western counties of many states, where agriculture predominated, demands for debt and tax relief became insistent. States attempted to handle the problem in various ways. Some, such as Massachusetts, followed a get-tough policy with farmers who refused to pay. In many of these states farmers took up arms. The most serious outbreak of violence was in Massachusetts where Shays' Rebellion was crushed in 1787. Other states, such as Rhode Island, tried to help farmers by issuing legal-tender paper money and insisting that creditors accept it, even if they were residents of other states. While such policies pacified western farmers, they increased tensions among the states. It is conceivable that interstate tensions, such as those that arose between Rhode Island and its neighbors, could have been resolved by a clause in the Constitution requiring states to keep their currencies at par. But this would still leave room for some states to increase their seignorage by expanding their currencies and allowing them to circulate outside their own borders. Thus, a monetary union was viewed as a necessary prerequisite for a political union (Rolnick, Smith, and Weber, 1993).

The Constitution appeared to settle the debate between those states that would have used monetary policy to help western farmers and those that would not. But the issue reemerged in the debates over the First and Second Banks of the United States. The First Bank of the United States was chartered in 1791. It was part of Alexander

Hamilton's plan for reorganizing the finances of the new government. Modeled to some extent on the Bank of England, it was intended to be a large bank, with a national branching system, that would help manage the new government's finances, and issue a paper money of uniform value (in part because it would be a legal tender) in all parts of the country. The term of the charter was limited to twenty years. When the charter came up for renewal in 1811 there was substantial opposition. Most of the opposition to the Bank, at least measured by the formal arguments against it, centered on the constitutionality of the Bank. The upshot was that the attempt to renew the charter failed, and the Bank was forced to wind up its affairs.

The monetary disturbances associated with the War of 1812 revived interest in a national bank. Such a bank, it was hoped, would pressure state banks into contracting their note issues and resuming specie payments. The Second Bank of the United States was established in 1816, again with a charter limited to twenty years. The Second Bank was similar to the First, but its capital was larger.

Although financial historians have often written favorably about the Second Bank, its career ended disastrously in the famous "Bank War." The Bank War pitted the Second Bank, led by its President, the aristocratic Philadelphian Nicholas Biddle, against an opposition led by the first President from West of the Alleghenies, Andrew Jackson. Not all eastern politicians sided with Biddle and the Bank. In New York, in particular, it was hoped that the Second Bank of the United States would be replaced by a Third Bank of the United States with headquarters in New York City rather than Philadelphia. And not all western politicians supported Jackson. Davy Crockett, a Congressman who represented poor farmers from western Tennessee (and King of the

Wild Frontier!), supported the Bank. (Shackford, 1986, *passim*).² In general, however, the Bank's opponents were from the West and its supporters from the East.

The exact reasons for Jackson's opposition to the Bank are still a matter of dispute. Jackson was first elected in 1828. His outspoken criticism of the Bank began soon after, and led to an attempt to renew the Bank's charter before Jackson came up for his second election in 1832. The bill to renew the charter passed both houses of Congress, but was vetoed by Jackson. The veto message has been the subject of intense scrutiny and debate by historians. Jackson cited a number of reasons for vetoing the bill to recharter the Bank: foreigners held a considerable amount of stock (although they could not vote, a fact he failed to notice), a competition for the charter would produce more revenues for the Treasury, and so on.

But the interesting point from our perspective is that Jackson stressed that the Bank was controlled by Eastern moneyed interests, and had followed policies harmful to western farming interests. Jackson's reasoning on this issue has been faulted. (Temin 1969). Nevertheless, it is clear that what Jackson perceived to be a conflict between eastern and western monetary interests inspired his opposition to the Bank, and galvanized the country. Jackson won an overwhelming reelection, sealing the fate of the Bank. Although there would be further battles, the war was lost. The charter of the Bank expired in 1836 and it wound up its affairs. In the end the antagonism toward the Bank, rooted in Western antagonism toward moneyed interests, had produced a momentous change in the monetary institutions of the United States. The

² It should be noted that although Crockett's support for the Bank may have been based on general considerations, he owed money to the bank, a debt that was partially cancelled through the personal intervention of Nicholas Biddle.

United States would not have any institution resembling a central bank until the Federal Reserve was established in 1913.

In 1837 the United States experienced a severe banking panic; numerous banks failed, and the banks were forced to suspend specie payments. In 1838 specie payments were resumed, and things began to look up. But a second panic in 1839 inaugurated a long recession marked by falling prices and a contraction in real output, although the degree of contraction in real output has been debated. To what extent did the Bank War contribute to the Crises of 1837 and 1839 and the subsequent recession? Peter Temin (1969) has argued that international forces, largely independent of the Bank War, explain the Jacksonian inflation, and the Crises of 1837 and 1839. Marie Sushka (1976), however, has argued that the Bank War did have an impact by increasing uncertainty about the soundness of the monetary system. Undoubtedly, independent international shocks are part of the story, perhaps the major part. But it seems probable that the uncertainty about the future of monetary arrangements created by the Bank War made holders of bank liabilities more fearful about the soundness of the banking system than they otherwise would have been, at least in some measure, and contributed through this channel to the banking crises and the recession that followed.

IV. Disunion

The Civil War was the result, of course, of the great national division over slavery. Nevertheless, financial factors did play a small role, influencing, perhaps, the timing of the War. Before the War the South had a relatively well developed banking

system, and there is a good deal of evidence for capital market integration (Bodenhorn and Rockoff, 1992). Indeed, Southerners were proud of their banking system, and their economic system based on “King Cotton.” The Crisis of 1857 was an eye-opener in both the South and the North. As southerners saw it, the crisis was largely of Yankee making. It started in New York, with the failure of a branch of the Ohio Life Insurance and Trust Company, and spread through the rest of the country. In the end the South suffered relatively less than did the North. Many radical Southern secessionists seized on the evidence provided by the crisis of 1857 to push their case that the South would be better off as an independent country with its own economic and monetary policies. In the North the newly formed Republican Party tried, with some success, to pin the blame for the Crisis on the Democrats. Thus the Crisis had the ironic result of strengthening the two factions least willing to compromise on the issue of slavery. (Huston, 1987).

The war divided the nation into three monetary regions. In the East and Middle West a fiat money standard prevailed based on the greenback. In the south, another fiat standard held sway based on the Confederate dollar. The Pacific coast, however, remained on gold. The southern monetary system, of course, gradually collapsed with the Confederacy, and came to an end in 1865. Returning to gold, and thus reuniting the two currencies of the United States took until 1879.

V. Reunion

After 1865 the South and Northeast were on the same (greenback) monetary system. There was no central bank. The supply of high-powered money was largely

determined by the policies of congress and the secretaries of the treasury. Republicans dominated. The goal of monetary policy was returning to the prewar price level and gold convertibility. These long-cherished goals were achieved on January 1, 1879 when the United States returned to the gold standard at the prewar parity. Throughout this period the policy of resumption faced determined regional opposition. Republicans in the Northeast favored resumption; Democrats and their allies in the Greenback Party, based in the Middle West and the South opposed resumption and favored monetary expansion.

Southern and Western Opposition to Resumption.

In 1866 Congress passed the Contraction Act, which called for the reduction in the amount of greenbacks in circulation, with a view to early resumption of specie payments. When a recession ensued, considerable opposition to this policy developed. In the 1868 election the currency was a major issue. Western democrats, following a now familiar refrain, demanded that Civil War bonds be paid in greenbacks unless the law specifically required payment in specie. Ulysses Grant, a financial conservative, however, was elected president. In March 1869 Congress voted to pay the Civil War debt in coin. Nevertheless, a more gradual approach toward resumption, referred to as “growing up to the currency,” was adopted. Roughly speaking, the policy was to freeze the stock of high-powered money so that economic growth would produce a gradual decline in the price level.

Even this policy faced a severe political test. After the Panic of 1873, Congress voted for an increase in the stock of greenbacks. Grant vetoed the measure,

triggering formation of the Greenback party – “a combination of middle western farmers, small businessmen, and labor intellectuals.” (Merk 1978, p. 445) In 1875 a lame duck Republican Congress adopted the Resumption Act which called for specie payments to be resumed on January 1, 1879.

Opposition to resumption coming from Greenbackers in the South and West, and Democrats with Greenback sympathies was fierce. In 1876 the House voted 106-86 to repeal the Resumption Act, but repeal died in the Senate. In 1877 the House again voted for repeal and for measures that would have expanded the stock of money. But in the Senate in 1878 a compromise was worked out. The advocates of soft money were persuaded to support a limited expansion of the stock of silver money provided for in the Bland-Allison Act, and the policy of resumption was kept on track. Resumption, as I noted above, was achieved in 1879. But it had been a near thing.

The opponents of resumption could hardly have been expected to give up their cause simply because it produced uncertainty about the final outcome. Nevertheless, by repeatedly placing resumption in doubt, the soft money faction created uncertainty about future exchange rates, that affected nominal interest rates, and that probably created an additional hindrance to international capital flows. (Calomiris 1994).

During this period, it should be noted, the West remained on the gold standard. Indeed, in 1873 the National Banking Act was amended to permit banks in California to issue currency redeemable in gold (yellowbacks). Thus from 1865 to 1879, when the greenback currency became convertible into gold we have a monetary rarity: a strong political union, untouched by war, with two currencies, greenbacks

and yellowbacks, circulating at a floating exchange rate. Interest rates on the Pacific Coast were high, perhaps reflecting some of the exchange rate uncertainty. Rates, however, had been high before the war and would remain high after resumption. In any case, the “need” to reunite the currency, as contemporary observers saw it, strengthened the case for resumption.

VII Misunion?

Middle Western opposition to the gold standard continued to smolder after Resumption. In the late 1880s opposition burst into flame once more. The main problem was farmer unrest aggravated by low farm prices. Table 1 shows prices for the crops that were key in the regions where discontent was at a maximum. The real price of wheat had fallen from 100 at the time of resumption to 86 in 1890. Farmer Alliances were formed, and showed surprising strength in the 1890 congressional elections. Eight Middle Western and one Southern Populists were elected.

Congress took note of the growing pressure for inflation and passed the Sherman Silver Purchase Act in 1890 which required the Treasury to purchase 4.5 million ounces of silver per month, virtually the entire U.S. output, paying with new legal tender currency redeemable in gold or silver at the discretion of the Treasury. The Silver Purchase Act was a compromise, designed to appease southern and western inflationists, and western silver interests, without going all the way to the free coinage of silver-backed currency. Fear of silver, however, produced a reduction in the Treasury’s stock of gold, further increasing fears that the United States might abandon the gold standard.

Although Farm prices rose briefly in 1891, they tumbled again in 1892, and demands for monetary expansion were renewed. In 1892 the Northern Alliance entered the presidential race as the Populist Party. Their platform, the Omaha platform, was a wish list of radical reforms, monetary reforms prominent among them.

The stage was now set for the Great Depression of the 1890s, and the accompanying “Battle of the Standards” (gold vs. silver), the famous debate over monetary policy fought along regional lines. Before describing those events, however, I want to digress briefly and consider the extent to which those regions matched the criteria for optimal currency areas.

A Digression on the Optimal Currency Area Criteria

Were the regions that opposed resumption in the 1870s, and that supported greenbacks and bimetallism in subsequent decades, separately optimal currency areas?

Optimal-currency-area theorists have described several factors that identify an area as a candidate for its own currency: (1) it must be a large area, (2) it must be specialized in the production of certain goods and subject to asymmetric shocks, (3) labor mobility between the candidate region and other regions is limited, (4) capital mobility between the candidate region and other regions is limited, and (5) fiscal transfers between the candidate region and other regions are limited. If the regional economies of the United States were relatively small, then the case for viewing them as candidates for separate currencies would be off to a bad start. It would be hard to

make a case, for example, for a single state as an optimal currency area. The costs of currency conversion mount for a small open economy. But, in fact, the major census divisions of the United States were, by world standards, large economies. By 1900, for example, U.S. national income was about twice that of Britain. (Friedman and Schwartz, 1982, pp. 122, 130). Estimates of the distribution of personal incomes in the United States place the share of the Middle Atlantic region (economically the largest) at about 31 percent. So a back-of-the-envelope calculation might put the income of this region at 62 percent of Britain's.

These regions were subject, moreover, to asymmetric shocks. Regional agricultural specialization had begun in the colonial era. The famous "North Thesis" (North 1961) maintains that the specialization of the South in plantation agriculture, especially cotton, was the driving force behind American economic growth before the Civil War. After the Civil War the South remained a land specialized in the production of cash crops: sugar, tobacco, rice, and cotton. Indeed, because of the changes in the structure of Southern agriculture, the South produced more cotton after the Civil War, and devoted a larger share of its resources to the production of cotton, than it had before the war. The prices of cash crops rose and fell with the business cycle, but seemed to be especially hard hit in certain periods.

In terms of labor mobility the South was clearly a world apart until World War II. Gavin Wright (1996) in *Old South, New South*, has shown that while considerable integration was achieved across labor markets within the South (wages for unskilled white and black workers were almost the same) the southern labor market remained a separate low-wage market. Racism limited mobility. Black southerners, of course,

suffered the most; but even white southerners had to overcome stereotypes and ill will, especially after the Civil War. Once migration patterns were established, moreover, it was hard to change them in response to changing economic conditions, because earlier migrants provided information and support for later migrants. Mobility among other regions was much higher. But moving in response to regional shocks was even into the 1930s a process accomplished with considerable difficulties. The Joads were doing the right thing from an economic point of view – moving from a depressed region to an expanding region during the 1930s – but the *Grapes of Wrath* is an appropriate title.

The extent of capital mobility is more debatable. The standard view is that integration of regional capital markets was not achieved until the turn of the century because of a simple reluctance of capital to migrate, or because of institutional factors such as differences in banking and usury laws. Indeed, Bodenhorn's (1995) data shows that substantial interregional interest rate differentials persisted through the 1930s, although Bodenhorn attributes these differences to risk. One can say that after 1900, if not before, capital market integration served to ameliorate the affects of regional shocks.

In the United States today fiscal transfers tend to offset asymmetric shocks. Unemployment benefits, for example, will rise in regions suffering from high unemployment. In the nineteenth century, however, the federal government was simply too small a share of GDP to offset regional shocks through fiscal transfers. The largest federal transfer program by far was the Civil War pension program. And this program was somewhat responsive to economic conditions. The depression of the

1890s contributed in some measure to the expansion of benefits under the program that occurred at that time. But southerners, immigrants, men who were too old or too young to have served, or who hired substitutes were not eligible. Women had to marry to become eligible. It was not until the adoption of programs such as unemployment compensation and agricultural price supports in the 1930s that one can point to fiscal transfers as a legitimate mechanism for overcoming asymmetric shocks.

The *a priori* case for believing that prices and wages in the separate regions were relatively sticky (thus strengthening the case for an independent monetary policy) appears to be relatively weak. The labor unions or oligopolistic industries that economists often point to as sources of stickiness were unimportant. There was considerable political agitation about the danger of the Trusts, but how much they contributed to price rigidity is debatable. In the South, even a significant portion of agricultural rental contracts was indexed: the famous sharecropping contracts.

Nevertheless, there were elements of institutional rigidity: taxes, mortgages, cash rentals, and so on. And beyond purely institutional sources of rigidity, there were the usual coordination problems. Adjusting to lower world prices for agricultural products meant a coordinated fall in wages, prices, and rents. Along most dimensions (labor and capital mobility, and fiscal transfers) the monetary union was strengthened during the 1930s. But along this dimension, the union was weakened.

Certain regions of the United States clearly exhibited many of the signs of good candidates for separate currencies, at least until the 1930s. But can one identify episodes in which substantial costs were imposed on these regions because they were part of a monetary union? Classic cases of optimal-currency-area dilemmas, a boom

in one region combined with a recession in other regions, were probably rare. More common were differences among regions in the magnitude of cyclical fluctuations, and in the timing of contractions and recoveries.

The Great Depression of the 1890s

The Great Depression of the 1890s, like the Depression of the 1930s, involved two severe recessions in close order. The economy declined for 17 months from January of 1893 to June 1894, and then after a comparatively weak recovery, declined again for 18 months from December 1895 to June 1897. Unemployment figures are necessarily somewhat problematic, but the figures we do have show the rate of unemployment at double-digit levels from 1893 through 1898, with a peak of 18.4 percent in 1894. (U.S. Bureau of the Census 1975, p. 135)

As we noted above, the early 1890s were characterized by concern over the maintenance of the gold standard stemming from the Sherman Silver Purchase Act, the decline in the stock of Treasury gold, and the rise of the Populists. In May 1893 a banking panic was touched off by commercial failures in New York. In June 1893, the administration revealed that it would press for the repeal of the Sherman Silver Purchase Act, and this seemed to ease pressures in financial markets. In July, however, further commercial and bank failures, led to a renewal of the panic. Bank runs and failures occurred in all regions. Starting in New York, banks throughout the country restricted the convertibility of notes and deposits into gold. The restriction on convertibility somewhat eased the situation. High interest rates drew gold into the United States and specie payments were resumed in September.

But the next three years were characterized by continued difficulties. Populists in the West and South continued to agitate for free and unlimited coinage of silver at a bimetallic ratio of 16:1. It was widely believed that adoption of bimetallism at that rate would have created substantial inflation and driven the United States off the gold standard. Uncertainty about the standard was reflected in higher interest rates. The Republicans favored continued commitment to the gold standard, although some Republicans, typically from the western states, called for an international conference aimed at restoring bimetallism, but at a bimetallic ratio that would permit continued circulation of both metals, and that would not produce inflation. The Democrats were badly split. The eastern wing of the party, led by former President Grover Cleveland, favored maintaining the commitment to gold; the western and southern wings favored bimetallism at 16:1. At the Democratic National Convention held in Chicago in 1896, the Westerners overthrew the Easterners. William Jennings Bryan, a Democratic Congressman from Nebraska with strong Populist sympathies, was nominated after a stirring speech, one of the most famous in American history, in which he declared that the Republicans would not be allowed to “Crucify Mankind upon a Cross of Gold.”

Despite his oratorical skills, Bryan lost the election to William McKinley, who favored the Republican brand of international bimetallism. Bryan carried states only in the West and South. Ironically, new flows of gold soon began to reverse the deflation that had persisted since the end of the Civil War. Demands for inflation through bimetallism or fiat paper became superfluous. The United States formalized its commitment to the gold standard with the Gold Standard Act of 1900.

Why was the United States so badly split along regional lines over monetary policy? Historians, traditionally, have seen the issue as one of creditors (eastern bankers) against debtor (western farmers). One problem with this view, as was recognized by Bryan and others at the time, is that any help from inflation would likely be partial and temporary, because interest rates would rise to reflect expected inflation. Frieden's (1997) recent argument that the support for 16:1 came primarily from exporters who looked to devaluation to improve earnings seems more persuasive.

As we saw above, however, the regions at odds with each other were, in many ways, separable currency areas. It makes sense to look at regional stocks of money, or what in fact are available, regional deposits. Figure 1 shows deposits by region from 1875 through 1896. Deposits in each region were set to 100 in 1875 to make it easier to compare regional trends.³ Granted, regional deposits reflected as well as caused changes in regional economic activity. Nevertheless, to the extent that these deposit movements reflected interregional transfers of reserves (balance of payments problems) or bank failures that might have been prevented by lender of last resort operations, they represent an independent influence on economic activity.

In any case, the picture is rather dramatic. The Great Depression of the 1890s left only a small imprint on deposits in New England or the Middle Atlantic region (dominated by New York and Pennsylvania). In New England deposits fell 2.4 percent from 1892 to 1893, but then more than regained their loss, rising 7.2 percent from 1893 to 1894. In the Middle Atlantic region, deposits fell 4.7 percent, and then

rose 7.7 percent. On the other hand, the impact on deposits in the West and South, the centers of the Populist revolt, were dramatic. Deposits in the South fell 18.6 percent from 1892 to 1893 and regained only 8.2 percent between 1893 and 1894; in the West deposits fell 16.4 percent, and only recovered .4 percent. Deposits in both regions were lower in 1896 than they had been in 1892. Is it any wonder that politicians in the West and South were calling for measures to increase the stock of money (remonetization of silver, or more radically, agricultural price supports financed by issues of fiat money) and that politicians in New England and the Middle Atlantic States called for a stand pat policy?

Data for all banks (both national and non-national) is available during this period only for four regions. The Middle Western and Pacific regions were aggregated in the source for non-national banks. This is worrisome because the Pacific Coast (which is dominated by California) was growing rapidly. Deposits in this region might have followed a somewhat different path than in other western states. In addition, the figures on non-national banks may be subject to reporting errors that vary in magnitude across regions and over time, despite the painstaking work undertaken by David Fand (1954) in putting these figures together. I have, however, computed National Bank deposits, which are likely to be more accurate, for five regions, separating the Middle West and the Far West.

These estimates are plotted in Figure 2. Again, the deposits in each region have been set to 100 in 1875. As expected, national bank deposits grew extremely rapidly on the Pacific Coast, rising by a factor of 18 between 1875 and 1900. The

³ The appendix describes how these numbers were computed.

most important point, however, is that the crisis of the 1893 is most evident in three regions: the Pacific Coast and Territories, the West, and the South. Again, the Middle Atlantic region and New England record only small impacts from the Crisis of 1893.

Who was right, the East or the West? I find it hard not to believe that the falling stocks of deposits and bank credit in the South and West did not contribute in some measure to the economic distress those regions were suffering, and would not have been relieved by monetary expansion. Taking a longer-term view, Milton Friedman (Fall 1990, and December 1990) has argued that adoption of bimetallism earlier in the postbellum period would have produced a more satisfactory behavior of the price level. He concludes, however, that by 1896 the time for adopting bimetallism had passed.

In any case, one thing seems clear. An unequivocal commitment to either gold or bimetallism would have avoided the uncertainty which itself was part of the problem. Milton Friedman and Anna J. Schwartz put it this way, in *A Monetary History* (1963, p. 134)

In retrospect, it seems clear that either acceptance of a silver standard at an early stage, or an early commitment to gold would have been preferable to the uneasy compromise that was maintained, with the uncertainty about the final outcome and the consequent wide fluctuations to which the currency was subjected.

Politicians, as Figures 1 and 2 show, were advocating the real interests of their own regions. Deposit growth in the East, although affected by the panic of 1893, was sufficiently close to trend to justify sticking with the current monetary policy; deposit growth in the South and West was depressed and justified a change in policy.

The optimal solution, were it politically feasible, might have been separate currencies, say for the East, the West, the South, and the Pacific Coast. The West and the South would have adopted a silver standard in the 1890s, while the East and the Pacific Coast, given its historical attachments, would have stayed on gold. Money stocks would not have fallen in the West and the South as much as they did. Their currencies, moreover, would have depreciated against gold, making it easier to dispose of wheat, cotton, and other agricultural products on domestic and world markets. The debate over monetary policy, and the resulting uncertainty, which affected banks in all regions would not have happened.

The Panic of 1903 (the Rich Man's Panic) and the Panic of 1907

Figure 3 plots deposits by region for the period 1900 to 1914.⁴ Deposits in each region have been set equal to 100 in 1900. The major events during this period were the panic of 1903 (the Rich Man's Panic) and the Panic of 1907. Here the regional pattern is different from the 1890s.

New York financial markets came under severe stress in late 1902. Short-term interest rates shot up, the stock market crashed, and a number of financial houses went bankrupt. The associated cyclical contraction was relatively long (23 months), from September 1902 to August 1904. In the Rich Man's panic the biggest impact was on the Middle Atlantic (New York) region, although even in this region the panic shows up as a period of relatively slow growth in deposits, rather than as an actual

⁴ All types of deposits in both national and non-national banks are summed.

decline. Deposits in the other regions, by way of contrast, were not affected much at all.

Although a severe jolt, the Rich Man's panic did not produce a banking panic or a severe economic contraction. The Panic of 1907, however, produced a much broader and deeper reaction in the banking system. Pressure began to build in the New York money market in the summer and fall of 1907. A major shock occurred in October 1907 when a run on the Knickerbocker Trust Company forced it to suspend payments.⁵ Other Trust companies soon were in difficulty as well. A banking panic soon gripped the nation, and the banks were forced to restrict the convertibility of bank notes and deposits into gold.

Nearly all regions of the country were affected. Only the plot of deposits for New England fails to show a dramatic imprint from the crisis. But as in 1903, the interesting feature of the data is the impact on the Middle Atlantic States, and the contrast between those states and the Middle West. Deposits in the Middle Atlantic States fell 2.38 percent between 1907 and 1908. Deposits in the South and on the Pacific coast fell by even larger amounts. But deposits in the Middle West fell by only 1.99 percent. The regional impact of the panic of 1907 was clearly very different from the impact of the panic of 1893.

As might be expected the political response to the crises in 1903 and 1907 was also very different than it was in the 1890s. During the 1890s, the eastern establishment was convinced that monetary reform was a foolish idea pushed by

dangerous Mid-western radicals such as Bryan; now the East was convinced that monetary reform was a wise idea advocated by the best scientific minds. In the wake of the Panic of 1907 the Aldrich-Vreeland Act (May 1908) was passed which created an emergency currency that could be issued during panics; and created the National Monetary Commission to investigate the monetary system, and recommend reforms. Senator Nelson W. Aldrich of Rhode Island, a long time Republican leader, and determined opponent of the Populists, headed the Commission. The main recommendation was the creation of a type of central bank. The United States would be divided into districts, and the banks in each district would keep their reserves in a district bank that was owned and controlled by the member banks. There would be a central board, controlled by the district banks, with the power to issue a gold-backed currency. Aldrich introduced a bill in Congress embodying this plan.

By the time that the bill came up for debate, however, the Democrats controlled the Congress and the Presidency. Goaded by the populists among them, the Democrats insisted on changes in the bill. Indeed, William Jennings Bryan, who carried Populist hopes for free silver in 1896, was then Secretary of State, and is said to have played an important role in the negotiations. The final result was legislation that differed from the Republican model in two ways, one that would remain important, and one that would not. First, the new institution was to be run by people appointed by the Federal Government, and not by the banks. Second, the right to issue currency would be the responsibility of the district banks, rather than the central

⁵ The Trust companies were banks that had grown up as ways of getting around the strict asset regulations imposed by the Comptroller of the Currency and the New York

board. The ability of the District banks to issue their own currencies would not prove to be a major feature of the system, it reflected a Populist hankering for a monetary system that would respond to the differing needs of differing regions.

The Great Depression of the 1930s

The Great Depression of the 1930s was the most severe in American economic history. Ever since the publication of *A Monetary History of the United States*, Friedman and Schwartz (1963), changes in the stock of money, and mistakes in monetary policy, have been granted an important role in the economic historian's account of the Depression. The extraordinary impact of the Depression can be seen in Figure 4, which plots all bank deposits by Federal Reserve District from 1922 to 1941. I have switched from broader regions to Federal Reserve districts, primarily because data for the Federal Reserve districts is readily available. Nevertheless, the Federal Reserve districts correspond, roughly, to economic regions, so little is lost in switching from broader regions to Federal Reserve districts. To make the figure easier to read, deposits in each region have been set to 100 in 1929.⁶

No region was immune to the crisis, but there were significant regional differences. Within the whole period I have highlighted two subperiods, 1929-1931 and 1934 to 1936, when there were marked regional differences in the rate of change of deposits, and when important mistakes were made in monetary policy.

The most famous, and probably the most important, error in American monetary history was the failure of the Federal Reserve to act as lender of last resort

banking authority.

for the banking system during the contraction from 1929 to 1933. A great deal has been written about the reasons for the failure of the Federal Reserve to take appropriate actions during this period. The personal and institutional rivalries stressed by Milton Friedman and Anna J. Schwartz (1963); the adherence to misleading doctrines about how policy actions worked, Calomiris and Wheelock (1998), Meltzer (1998); and the weight placed on adherence to the gold standard, Eichengreen (1992) undoubtedly were important in producing the lack of response to the crisis exhibited by the Federal Reserve. Regional loyalties had declined, in part because World War I had boosted nationalism over sectionalism.

Nevertheless, Figure 5, which focuses on 1929 to 1931, suggests that regional differences need to be woven into the traditional story.⁷ The deviations among regions are striking. In June 1931, almost two years into the Great Depression, the stocks of deposits in the San Francisco, Boston, and Philadelphia districts, and in the weighty New York district, were still above the June 1929 levels.

Economists and policymakers from those regions, who tended to look toward events in their own region, whether consciously or not, would have been less likely to stress the need for drastic countermeasures. As evidence that the need for action was recognized by some observers, Friedman and Schwartz (1963, p. 409) cite representative Sabbath of Illinois writing to Federal Reserve Board chairman Eugene Meyer in January 1931: “Does the board maintain there is no emergency at this time? To my mind if ever there was an emergency it is now, and this I feel, no one can successfully deny.” Can it be entirely irrelevant that at the time Sabbath was writing

⁶ All figures are for June dates, usually close to June 30.

deposits had already fallen drastically in the Chicago district and the St. Louis district (which covers southern Illinois)?

Residence in a district hard hit by deposit losses, it must be admitted, was no guaranty of sensitivity to the crisis. James McDougal, President of the Federal Reserve Bank of Chicago, consistently opposed open market security purchases in part because he thought it would be useless for the Federal Reserve to try to offset a natural market process: liquidating bad loans. Marriner Eccles, who became chairman of the Federal Reserve Board in 1935, might have been expected to be a consistent advocate of monetary expansion. He was a banker from the West (Utah), a region with a long tradition of monetary radicalism, and one that had been hard hit by the Depression as shown in Figure 4. He had been appointed, moreover, because of his sympathy for the New Deal. But his own study of the Depression, and his reading of the heretical under-consumptionist William T. Foster, had pushed him toward the view that monetary policy was a relatively impotent tool for controlling the economy. The real action was on the side of fiscal policy

Roy A. Young, President of the Boston Bank, however, at one point based his opposition to open market purchases on a regional argument: open market purchases would lead to a piling up of reserves in the money centers, with little effect on the regions of the country that really needed reserves. Thus, although policy positions in the 1930s do not divide as neatly along regional lines as they did in the 1890s, there is some evidence that differences in regional perspectives contributed, at least in a small measure, to the paralysis that gripped monetary policy making in the early 1930s.

⁷ The figures are for June 1929, June 1930, and June 1931; the cyclical peak was in August 1929.

The second major policy error during the Depression was the decision by the Federal Reserve to raise bank reserve ratios in three steps -- in August 1936, March 1937, and May 1937. Friedman and Schwartz (1963), and more recently Meltzer (1998), have explored many of the intellectual and personal currents that produced the decision. Nevertheless, the diversity in regional experiences shown in Figure 6 is suggestive. By June of 1935 the stocks of deposits in the New York, Richmond, and San Francisco districts, had all recovered their June 1929 levels. Rapid expansion in the ensuing year carried the stocks of deposits in these districts to levels between 20 and 30 percent above the levels of June 1929. Moreover, employment in these regions had also recovered well.⁸ A Federal Reserve President in one of these districts, who based his conclusions solely on conditions in his own district, might well conclude that it was time to adopt a more restrictive monetary policy before things got out of hand and inflation threatened.

In the “heartland” districts, however, conditions were very different. In the Cleveland district deposits in June 1935 were little more than 80 percent of what they had been in June 1929, and although growth was rapid, deposits were still below the June 1929 level in 1936. Unemployment was still high. The president of a district bank in the heartland might well conclude that further monetary expansion was required. To put it differently, the Federal Reserve, at least to judge by deposit growth, faced an optimal-currency-area dilemma in 1936. Some regions needed stimulation; others needed restraint.

⁸ It may seem surprising that the Richmond region followed the path of New York and San Francisco. In fact, however, as is now well understood, the South did relatively well during the depression.

It is interesting to ask what would have happened had the United States been divided into separate currency areas – separate currencies for, say, the East, the South, the Middle West, and the Pacific – during the 1930s. Separate currencies, of course, were ruled out by political considerations. A currency is a symbol of sovereignty, like the flag, and it is as hard to imagine any country deliberately choosing to divide its currency. But thinking about separate currencies can throw light on the economics of the Depression. Even with separate currencies, monetary policies might well have been similar in the East, the South and on the Pacific coast to what they actually were. The central bank of the East, for example, would not have acted a lender of last resort in the early 1930s, but there would have been no need for it to do so. It would have slowed monetary expansion in the mid-1930s, but it would have been logical for it to do so.

The central bank in the Middle West, however, might have acted differently. With hundreds of banks failing in the region, with politicians calling for action, a central bank that had full responsibility for the region might well have acted as lender of last resort in the early 1930s. Moreover, a Middle Western central bank, under intense pressure from local interests, might well have followed a more inflationary policy. Silver interests were potent in the Middle West and might well have demanded additional purchases of silver financed through an increase in the monetary base. The Middle Western currency might well have depreciated relative to other regional currencies. But this would have helped employment in the manufacturing sectors of the Middle West, which were among the hardest hit in the nation, because

they suffered from increasing labor costs due to unionization, as well as from the decrease in demand for consumer durables.

To be sure, if someone like James McDougal, who advocated deflationary measures as President of the Chicago district bank during the 1930s, had been president of the Midwestern Central bank that I am imagining, he might have succeeded in imposing deflationary policies despite political pressures to do otherwise. But not necessarily; the argument that open market purchases were actions taken by bankers in faraway places that influenced banks in faraway places would no longer apply. Long experience with the conduct of monetary policy might have made central bankers in the Midwest more adroit than central bank board members. In short, the monetary union, which in more stable times was a source of strength for the United States, appears to have been a liability during the 1930s.

VII. Communion?

It is generally assumed that the United States became a smoothly functioning monetary union, at least for the purpose of comparison with the European Monetary Union, in the postwar era. (For example, Feldstein 1997; Wyplosz 1997). The comparison may tend to exaggerate how well the U.S. monetary union functions. There have been asymmetric real shocks, such as the oil price fluctuations that hit Texas particularly hard, or the changes in manufacturing that hit the “rust belt.” Perhaps there were regional banking problems that exacerbated these disturbances that we have not paid sufficient attention to because we are accustomed to think in aggregate terms.

But it is true that several institutional changes took place during the Depression and World War II that weakened older divisions. One was the development of federally funded transfer programs, such as unemployment insurance, social security, and agricultural price supports, which cushioned regional shocks, and redistributed reserves lost through interregional payments deficits. Penelope Hartland (1949), using data from the Federal Reserve's Interdistrict Settlement Fund showed that the regions that had been hit by terms of trade shocks during the 1930s lost reserves to other regions through trade deficits, but that government transfers materially offset these losses. Between 1929 and 1933, for example, the Minneapolis Federal Reserve district lost \$247 million in reserves on private transactions. This was offset, however, by a gain of \$229 million on federal government transactions. On the other hand, the Boston Federal Reserve district gained \$644 million in reserves on private transactions, while losing \$575 million on federal government transactions. (Hartland, 1949, p. 397). Seymour Harris (1957, pp.174-192) noted the regional payments problems during the 1930s, and argued that separate currency areas would have ameliorated these problems. Harris also noted similar regional payment problems in the early post-Second-World-War era.

A second institutional change was the breakdown of long-term isolation of the southern labor market. During the war a strong northern labor market and the absence of immigrants pulled workers, white as well as black, from the South, and established networks that provided information and support for later migrants. In addition, federal labor legislation in the form of minimum wages and regulation of hours and

conditions of work, and federal incentives to mechanize agriculture, established during the 1930s added to the postwar flow of migrants from the South.⁹

A third factor that improved the functioning of the U.S. monetary union after the war was the absence of major banking and financial crises emanating from regional shocks. Deposit insurance, and monetary policies that reacted quickly to economic downturns, tended to minimize the regional banking problems that characterized recessions in the prewar era.

VIII. Lessons from the Troubled History of the U.S. Monetary Union

Weighing the costs and benefits of monetary unification is a difficult task. On the one hand, monetary unification means reduced transaction costs, easier comparison of prices in different regions, long-term investment without fears of devaluation, and so on. On the other hand, unification means relinquishing the capacity to use exchange rate changes and monetary policy to prevent monetary problems from magnifying distress originating in other sectors. Frequently, the experience of the United States is cited as evidence that in fact the benefits of a monetary union greatly outweigh the costs. After all, the monetary union of the United States has survived (with a temporary break during the Civil War era) since the adoption of the Constitution in 1788. But the survival of the U.S. monetary union is at best weak evidence that the net effects have been positive. There are many government policies, tariffs for example, which have survived for decades for

⁹ Fiscal federalism and the improvements in the functioning of the labor market are discussed in Eichengreen (1998), chapters 2 and 3. Wright (1996) discusses the breakdown of barriers to labor migration. Libecap (1998) discusses the origins of agricultural price supports.

political reasons, often the support of special interests, even though the claim that these policies contributed positively to the general welfare is dubious.

In truth, the U.S. experience shows that fears about the loss of monetary autonomy are far from baseless. American monetary history provides numerous examples of regional shocks that were magnified by monetary reactions. Typically, a region-specific shock to financial or agricultural markets produced a loss of regional bank reserves through an internal drain, caused by fears about the solvency of the regional banking system, and an external drain, caused by a regional balance of payments deficit. The result would be a regional contraction of bank money and credit that would cause headaches even for businesses not effected by the initial shock. A political battle would often follow. The regions that had experienced the contraction would demand a reform of the whole monetary system. The resulting uncertainty about the future of existing monetary institutions would further aggravate the initial contraction in economic activity.

During these episodes the United States might well have been better off, from a purely economic point of view, had it been divided into separate currency areas. Regions hit by severe asymmetric shocks would have been able to devalue their currencies, which would have reduced interregional losses of reserves. Within the region, expansionary monetary policies would have shored up the banking system, preventing runs or severe contractions of credit. Other regions would have been free to follow more conservative monetary policies, eliminating political battles over monetary institutions. Separate currencies for separate regions were not ruled out by any logical inconsistency. During and after the Civil War (1861-1879) the Pacific

coast had its own currency, the yellowback. But separate currencies for separate regions were ruled out eventually by political considerations. In the course of the nineteenth century currencies came to be seen as symbols of sovereignty, and separate regional currencies became as unthinkable as separate armies.

Nevertheless, speculating about this counterfactual helps us to understand the U.S. business cycle and may suggest some lessons for countries contemplating joining or remaining within a monetary union. For a country that is debating whether to join a monetary union the lesson is that the facile argument that the United States has had a monetary union, and therefore monetary unions must be good things, doesn't stand close scrutiny. Second thoughts are in order. For countries already firmly committed to a monetary union, the lesson is that it is extremely important to adopt the institutions adopted by the United States in the 1930s -- a system of inter-regional fiscal transfers and some form of deposit insurance, or regionally sensitive lender-of-last-resort facilities -- so that asymmetric real shocks are not aggravated by banking crises.¹⁰

Although the Eastern financial centers, and industrial Middle West had been integrated by the turn of the century, it was not until the 1930s that all regions, including the South, could be said to be parts of a single optimal currency area. How long did it take the United States to become an Optimal Currency Area? A reasonable minimum might be one hundred and fifty years! Hopefully, it will not take the European Monetary Union quite so long.

¹⁰ Capie (1998) has drawn a similar lesson from a variety of historical examples.

References

Bodenhorn, Howard and Hugh Rockoff, "Regional Interest Rates in Antebellum America," in *Strategic Factors in American Economic History: A Volume to Honor Robert W. Fogel*, edited with Claudia Goldin (University of Chicago Press, 1992).

_____. "A More Perfect Union: Regional Interest Rates in the United States, 1880-1960." In *Anglo-American Financial Systems: Institutions and Markets in the Twentieth Century*, eds. Michael D. Bordo and Richard Sylla. New York: Irwin Professional Publishing for New York University Salomon Center, 1995, pp. 415-453.

Calomiris, Charles W. "Greenback Resumption and Silver Risk: the Economics and Politics of Monetary Regime Change in the United States." In *Monetary Regimes in Transition*, eds. Michael D. Bordo and Forrest Capie. Cambridge: Cambridge University Press, 1994, pp. 86-132.

Calomiris, Charles W. and David C. Wheelock. "Was the Great Depression a Watershed for American Monetary Policy?" In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, eds. Michael D. Bordo, Claudia Goldin, and Eugene N. White. Chicago: University of Chicago Press, for the NBER, 1998, 23-64.

Capie, Forrest. "Monetary Unions in Historical Perspective: What Future for the Euro in the International Financial System." *Open Economies Review*. Vol. 9: 447-65. Supplement 1 1998.

Eichengreen, Barry J. *Golden Fetters: the Gold Standard and the Great Depression, 1919-1939*. New York : Oxford University Press, 1992.

_____. "European Monetary Unification." *Journal of Economic Literature*, 31 (September 1993): 1321-1357.

_____. *European Monetary Unification: Theory, Practice, and Analysis*. Cambridge, MA: The MIT Press, 1998.

Fand, David I. *Estimates of Deposits and Vault Cash in the Non-national Banks in the Post Civil War Period in the United States: 1876-1896*. Ph.D. dissertation, University of Chicago, 1954.

Feldstein, Martin. "The Political Economy of the European Monetary Union: Political Sources of an Economic Liability." *The Journal of Economic Perspectives* 11 (Fall 1997): 23-42.

Frieden, Jeffrey A. "Monetary Populism in Nineteenth-Century America: An Open Economy Interpretation." *The Journal of Economic History*. 57 (June 1997): 367-95.

Friedman, Milton. "Bimetallism Revisited." *Journal of Economic Perspectives*. 4 (Fall 1990): 85-104.

_____. "The Case for Flexible Exchange Rates." In *Essays in Positive Economics*. Chicago: University of Chicago Press, 1953.

_____. "The Crime of 1873." *Journal of Political Economy*. 98 (December 1990): 1159-94.

Friedman, Milton and Anna Jacobson Schwartz. *A Monetary History of the United States*. Princeton: Princeton University Press for the NBER, 1963.

_____. *Monetary Statistics of the United States: Estimates, Sources, Methods*. New York: Columbia University Press for the NBER, 1970.

_____. *Monetary Trends in the United States and the United Kingdom*. Chicago: University of Chicago Press for the NBER, 1982.

Greenfield, Robert and Hugh Rockoff. "Yellowbacks Out West and Greebacks Back East: Social-Choice dimensions of Monetary Reform." *Southern Economic Journal* 62 (April 1996): 902-915.

Harris, Seymour Edwin. *International and Interregional Economics*. New York, McGraw-Hill, 1957.

Hartland, Penelope. "Interregional Payments Compared with International Payments." *The Quarterly Journal of Economics* 63 (Aug., 1949): 392-407.

Heim, Carol E. "Uneven Impacts of the Great Depression: Industries, Regions, and Nations." In *The Economics of the Great Depression*, ed. Mark Wheeler. W.E. Upjohn Institute for Employment Research: Kalamazoo, Michigan. 1998. Pp. 29-61.

Huston, James L. *The Panic of 1857 and the Coming of the Civil War*. Baton Rouge: Louisiana State University Press, 1987.

Kawai, Masahiro. "Optimum Currency Areas." In *The New Palgrave Dictionary of Money and Finance*, eds. Peter Newman, Murray Milgate, and John Eatwell. London, The Macmillan Press Limited, 1992, pp. 79-81.

Kenen, Peter B. "The Theory of Optimum Currency Areas: an Eclectic View." In *Monetary Problems of the International Economy*, eds., Robert A. Mundell and Alexander K. Swoboda. Chicago: University of Chicago Press, 1969.

Libecap, Gary D. "The Great Depression and the Regulating State: Federal Government Regulation of Agriculture, 1884-1970." In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, eds. Michael D. Bordo, Claudia Goldin, and Eugene N. White, pp. 181-224, 1998.

McKinnon, Ronald I. "Optimum Currency Areas." *American Economic Review* 53 (September 1963): 717-25.

Meltzer, Allan. "Chapter 6, In the Back Seat." Pittsburgh: Carnegie Mellon University, manuscript, March 1998.

- Merk, Frederick. *History of the Westward Movement*. New York: Alfred A. Knopf, 1978.
- Mundell, Robert A. "A Theory of Optimum Currency Areas." *American Economic Review* (September 1961): 657-65.
- Mundell, Robert A. "Currency Areas, Common Currencies, and EMU." *American Economic Review*. Vol. 87 (2). p 214-16. May 1997.
- North, Douglass Cecil. *The economic growth of the United States, 1790-1860*. Englewood Cliffs, N.J., Prentice-Hall, 1961.
- Rolnick, Arthur J; Smith, Bruce D; and Warren E. Weber. "In Order to Form a More Perfect Monetary Union." *Federal Reserve Bank of Minneapolis Quarterly Review*. Vol. 17 (4), Fall 1993, pp. 2-13.
- Shackford, James Atkins. *David Crockett: The Man and the Legend*. Lincoln, Nebraska: University of Nebraska Press, 1986.
- Schweitzer, Mary M. "State-Issued Currency and the Ratification of the U.S. Constitution." *Journal of Economic History*, Vol. 49, No. 2, The Tasks of Economic History. (June, 1989), pp. 311-322.
- Sushka, Marie Elizabeth. "The Antebellum Money Market and the Economic Impact of the Bank War." *Journal of Economic History* 36 (December 1976): 809-35.
- Temin, Peter. *The Jacksonian economy*. New York: Norton, 1969.
- Tower, Edward and Thomas D. Willett. *The Theory of Optimum Currency Areas and Exchange-Rate Flexibility*. Special Studies in International Economics No. 11. Princeton: Princeton International Finance Section, May 1976.
- U.S. Board of Governors of the Federal Reserve System. *Banking and Monetary Statistics*. Washington D.C: The National Capital Press, 1943.
- _____. *All Bank Statistics, 1896-55*. Washington D.C: Government Printing Office, 1959.
- U.S. Bureau of the Census. *Historical Statistics of the United States, Colonial Times to 1970*, Bicentennial Edition, Part I. Washington D.C: Government Printing Office, 1975.
- U.S. Comptroller of the Currency. *Annual Report of the Comptroller of the Currency*. Washington D.C: Government Printing Office, 1920.
- Wallis, John Joseph. 1989. "Employment in the Great Depression: New Data and Hypothesis." *Explorations in Economic History* 26: 45-72.
- Willett, Thomas D. and, Clas G. Wihlborg. "The Relevance of the Optimum Currency Area Approach for Exchange Rate Policies in Emerging Market Economies." In *Exchange-rate Policies for Emerging Market Economies*, eds. Richard J. Sweeney, Clas G. Wihlborg, and Thomas D. Willett. Boulder and Oxford: Westview Press, 1999.

Wright, Gavin. *Old South, New South: Revolutions in the Southern Economy Since the Civil War*. Baton Rouge: Louisiana State University Press, 1996.

Wyplosz, Charles. "EMU: Why and How it Might Happen." *The Journal of Economic Perspectives* 11 (Fall 1997): 3-22.

Yeager, Leland Bennett. "Exchange Rates Within a Common Market." *Social Research* 25 (January 1959): 415-38.

Table 1. Agricultural Prices and the NNP Deflator, 1879-1900 (1879 = 100)			
	Net National Product Deflator	Real Price of Wheat	Real Price of Cotton
1879	100	100	100
1880	110	93	105
1881	108	103	100
1882	112	104	105
1883	110	91	92
1884	105	85	97
1885	98	86	103
1886	96	80	94
1887	97	77	102
1888	99	87	100
1889	100	87	103
1890	98	86	113
1891	97	93	89
1892	93	80	83
1893	95	67	87
1894	89	59	78
1895	88	64	83
1896	85	70	92
1897	86	87	84
1898	88	94	68
1899	91	74	73
1900	95	69	100

Sources: NNP deflator: (Friedman and Schwartz, 1982), pp. 122-123. Table 4.8, col. 4. Price of wheat and price of cotton: (U.S. Bureau of the Census, 1975): pp. 208-209, series E123 and E126.

A Chronology of the U.S. Monetary Union	
1788	The Constitution is ratified. States are prohibited from issuing paper money. The U.S. monetary union is launched.
1791	The First Bank of the United States is chartered.
1811	The First Bank of the United States comes to an end.
1816	The Second Bank of the United States is chartered.
1832	President Andrew Jackson vetoes the bill to recharter the Second Bank, stressing the oppression of the West in his veto message.
1836	The Second Bank of the United States comes to an end.
1837	The nation is hit by a severe banking panic, inaugurating a period of hard times.
1857	The nation is hit by a severe banking panic. Southern firebrands and Northern Republicans both make political capital from the crisis.
1861	The Civil War begins. The United States is divided into three currency areas: Greenbacks in the Northeast, Confederate dollars in the South, and Gold in California.
1865	Lee surrenders. The Confederate dollar ceases to function.
1866	Congress passes the Contraction Act looking to a rapid return to the gold standard.
1873	The silver dollar is omitted from the list of official coins (The Crime of 1873). National Banks in California are permitted to issue notes backed by gold (yellowbacks.)
1879	Resumption of specie payments. The yellowback and greenback are reunited.
1896	William Jennings Bryan, an advocate of free silver, is nominated by the Populists and Democrats; William McKinley, an advocate of international bimetallism, is nominated by the Republicans. Bryan carries only a few states in the West and South
1900	The Gold Standard Act firmly commits the United States to the Gold Standard and symbolizes the end of the "Battle of the Standards."
1907	A Banking Panic leads to the establishment of the National Monetary Commission.
1913	The Federal Reserve System is established. Republican proposals for a currency issued by a privately controlled central bank are defeated. Instead a federation of regional banks, each issuing their own currency, is created. William Jennings Bryan, now secretary of State, plays an active role in fashioning the legislation.
1929-1931	Beginning of The Great Contraction. The Stock market crashes and a severe monetary contraction begins in the nation's heartland.
1936-1937	The Federal Reserve raises required reserve ratios contributing to a sharp recession that prolongs the depression.

Figure 1
All Bank Deposits by Region 1875-1896
1875 = 100

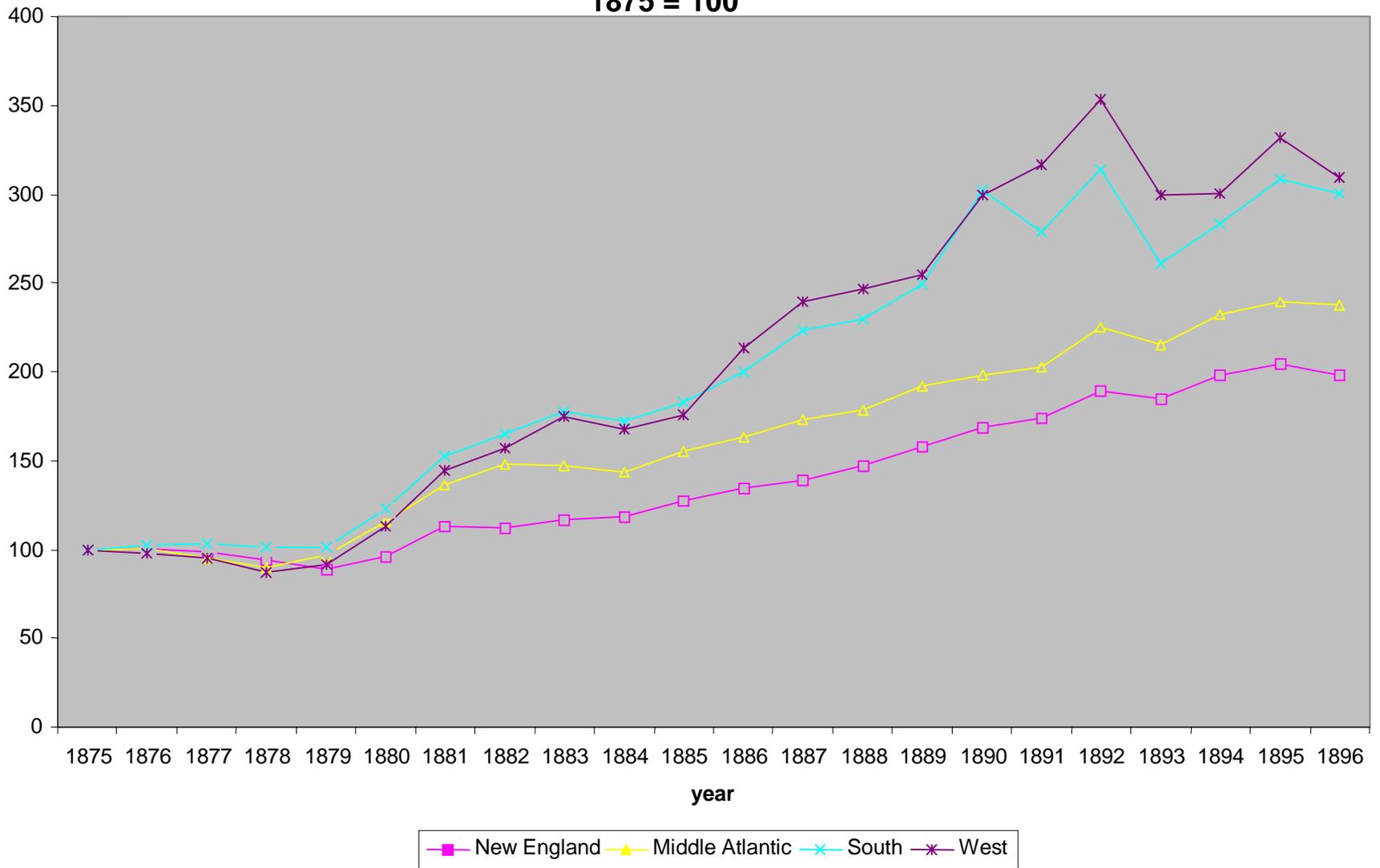


Figure 2
National Banks Deposits by Region, 1875-1900
1875 = 100

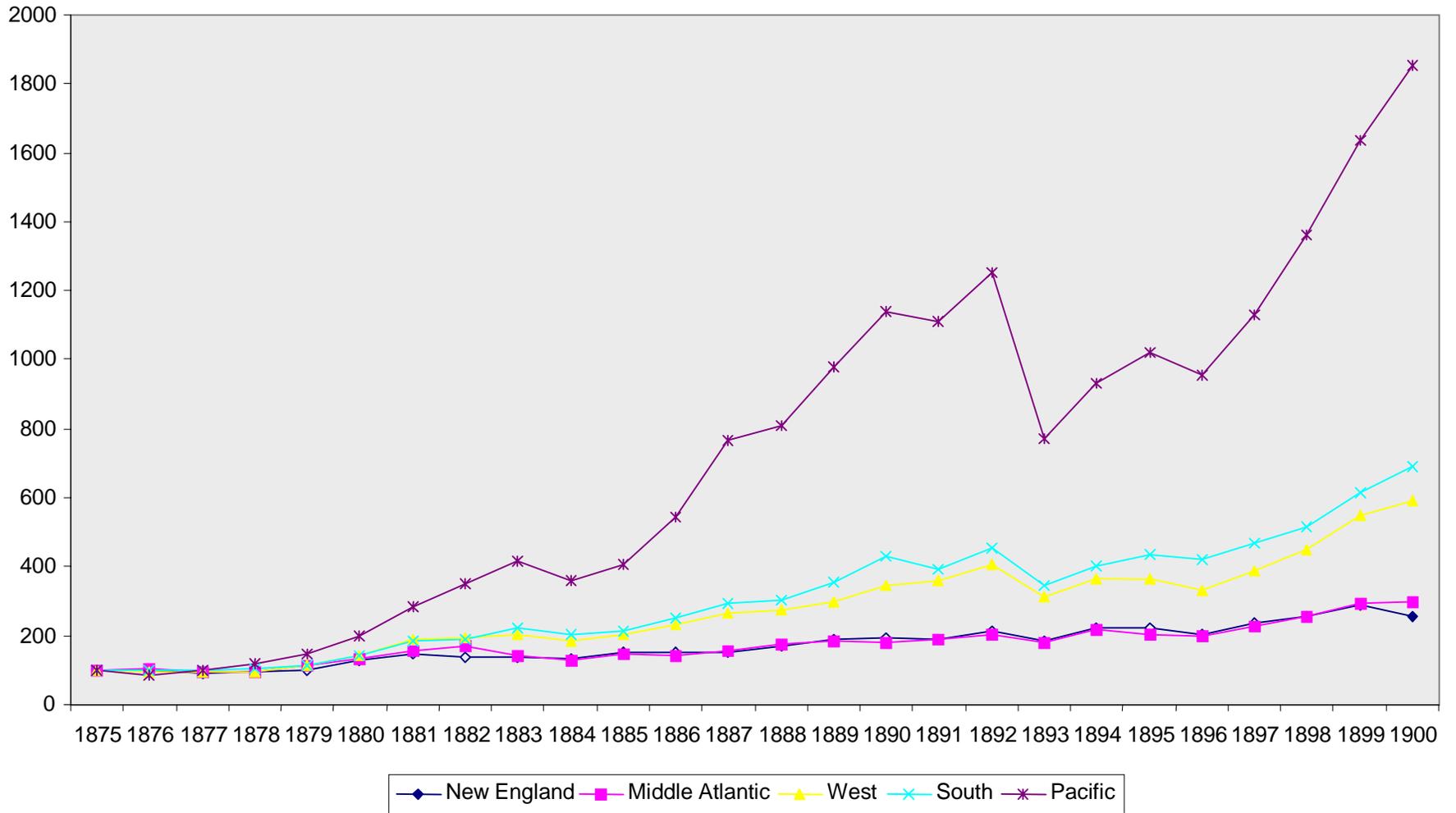


Figure 3
All Bank Deposits By Region, 1900 - 1914
(1900 = 100)

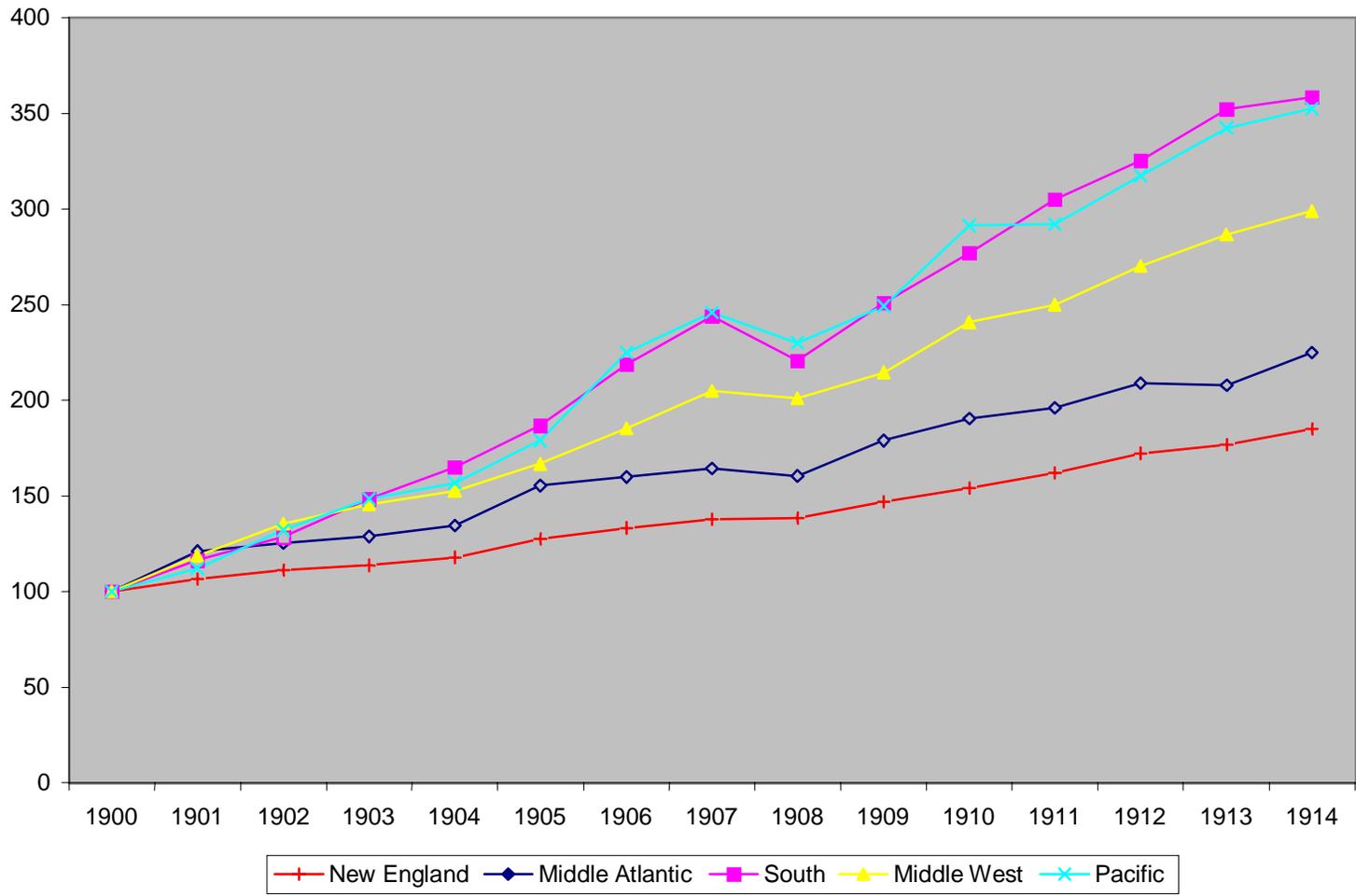


Figure 4
Deposits by Federal Reserve District 1922-1941
1929 = 100

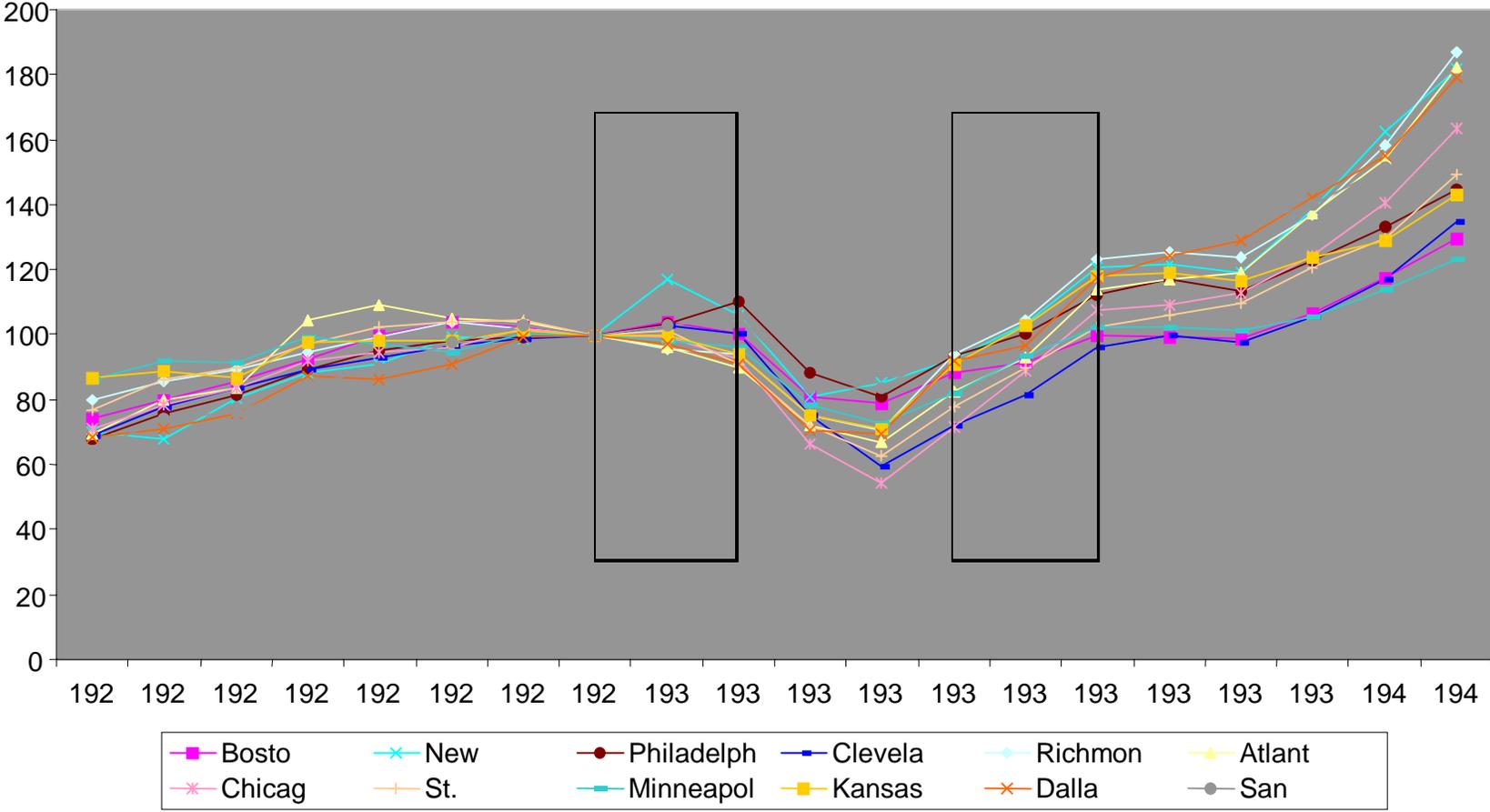


Figure 5
Deposits by Federal Reserve District
The First Two Years of the Depression: 1929-31
1929 = 100

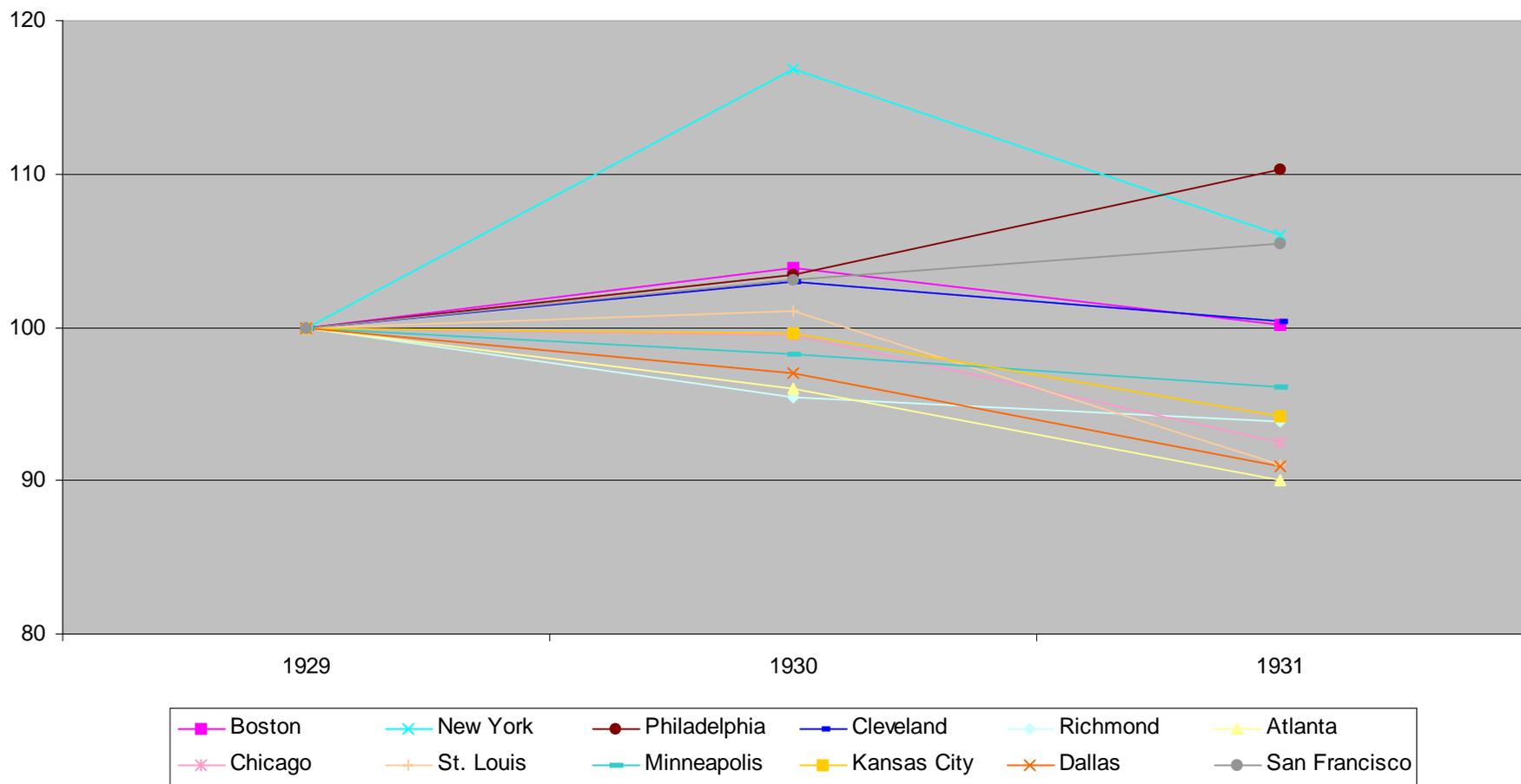
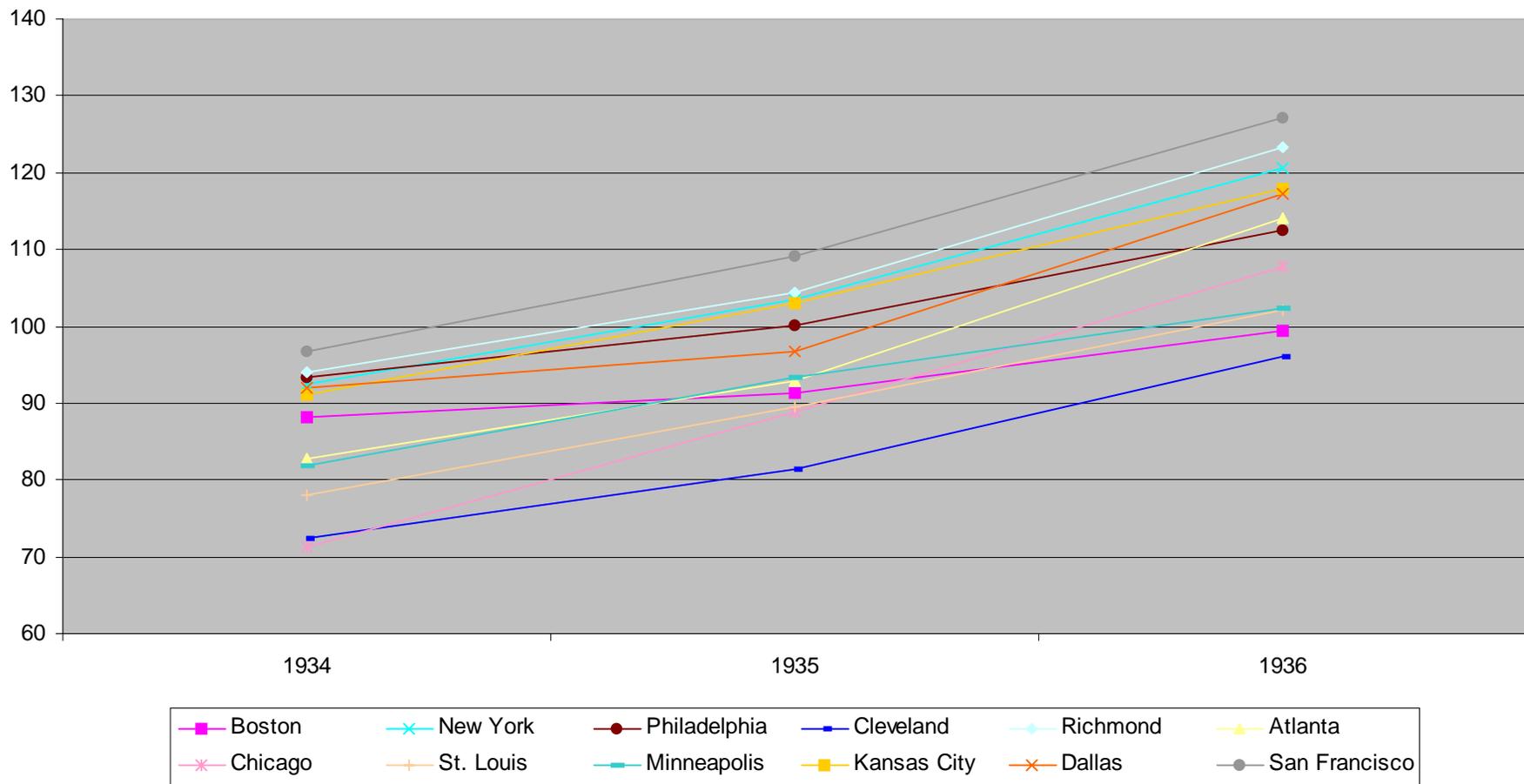


Figure 6
Deposits by Federal Reserve District
Before the Doubling of Reserve Ratios: 1934-36
1929 = 100



Appendix: Sources of Data

Deposits: 1875-1914.

Exact figures on the stock of money by region are not available. The amount of coins within the United States, for example, can be estimated from figures on minting and imports and exports, but the amount within any one region cannot be estimated accurately. But figures on deposits, and in some periods bank notes, by place of issue (although not by where they are held) are available. Fand (1954, pp. 72-76) estimated deposits in Non-national banks in four regions: New England, Middle States, Southern States, and Western states for the years 1875-1896. These regions do not correspond exactly to economic regions. Perhaps the main problem from this perspective is the combination of the Pacific Coast with the other western states. It would be, however, extremely time consuming to build up separate estimates for the non-national banks for this region, so I have relied on the national bank data to provide a picture of the Pacific Coast. To Fand's estimates of deposits in non-national banks, I added deposits of national banks. The source is a table by state that appeared regularly in the *Annual Report of the Comptroller of the Currency*. U.S. Comptroller of the Currency (1920, pp. 307-343).

- (1) **New England:** Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont.
- (2) **Middle Atlantic:** Delaware, District of Columbia, Maryland, New Jersey, New York, and Pennsylvania.
- (3) **South:** Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, and West Virginia.
- (4) **Middle West:** Illinois, Indiana, Indian Territory, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, Oklahoma, and Wisconsin.
- (5) **Pacific Coast, Western states, and Territories:** Alaska, Arizona, California, Colorado, Dakota Territory, Hawaii, Idaho, Montana, Nevada, New Mexico, North Dakota, Oregon, Puerto Rico, South Dakota, Utah, Washington, and Wyoming.

Deposits: 1896-1914. For this period *All Bank Statistics* (U.S. Board of Governors, 1959) gives data for all types of deposits in all classes of banks by states. The data shown in the figure are a sum of demand and time deposits. The regions are defined above.

Deposits: 1914-1941. For this period I have switched to Deposits by Federal Reserve District because the data is readily available; U.S. Board of Governors of the Federal Reserve System (1943, pp. 688-927).

FISCAL FEDERALISM:
US HISTORY FOR
ARCHITECTS OF
EUROPE'S FISCAL UNION

by

C. Randall Henning and Martin Kessler

BRUEGEL ESSAY AND LECTURE SERIES

BRUEGEL ESSAY AND LECTURE SERIES

Fiscal federalism: US history for architects of Europe's fiscal union

C. Randall Henning and Martin Kessler

© Bruegel 2012. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted in the original language without explicit permission provided that the source is acknowledged. The **Bruegel Essay and Lecture Series** is published under the editorial responsibility of Jean Pisani-Ferry, Director of Bruegel. Opinions expressed in this publication are those of the author alone.

Bruegel
33, rue de la Charité, Box 4
1210 Brussels, Belgium
www.bruegel.org

ISBN: 978-9-078910-24-4

FOREWORD

Ever since the blueprints for monetary union in Europe were first drawn up, the United States, considered as a collection of individual states or regions, has served as a benchmark for assessing its feasibility and evaluating alternative policy options. Starting with Robert Mundell's seminal 1961 article on optimal currency areas, countless papers have explored the inner workings of US labour, product and capital markets, and of its public finances, in the hope of learning lessons for Europe.

It could be argued that this US inspiration is mistaken. After all, it is not the only economic and monetary federation in the world. Other federations work on different principles – especially when it comes to public finances – and there is no guarantee that US arrangements are optimal – especially, again, regarding public finances. But we know the US better and we think we understand it better, so success or failure relative to the US test carries much more weight than with the Australian, Canadian, Indian or Swiss tests. For better or worse, the US remains our ultimate policy laboratory.

This essay on US fiscal federalism by Randall Henning and Martin Kessler builds on the established tradition. But unlike many papers that take current US features as a given, they tell us what present arrangements governing responsibility over public debt gradually emerged from, and why. By bringing in the historical dimension and the trial-and-error process that took place over more than two centuries, they help us understand the logic behind alternative arrangements and why the current one has in the end prevailed.

Their careful historical account yields several important lessons. It first recalls that the US system as we know it, with its combination of a large federal budget responsible for the bulk of public debt and limited thrifty state budgets subject to balanced budget rules, emerged gradually from a sequence of events; in fact the initial set-up, as designed and enforced by Alexander Hamilton, was almost exactly the opposite. Second, it makes clear that beyond economic principles, attitudes towards what was in the aftermath of independence called the 'assumption' of state debt were shaped by broader political considerations – not least the aim of building a genuine federal government. Third, it explains how after the US was firmly established as a federation, changing political conditions led to a reversal of the federal government's stance and to the enforcement of a 'no bail-out' principle. An intriguing feature of US history is therefore that the competences and features of federal government grew out of its assumption of state debt, and that the centre imposed a *de-facto* no bail-out regime only after having assumed essential powers.

Another interesting observation by Henning and Kessler is that balanced budget rules were adopted spontaneously by states in response to financial stress and defaults, rather than as a disciplinary device mandated by the centre. Thus, there is still significant variability between states regarding the *modus operandi* and strictness of budget rules. The question remains if what matters is the strictness of the rule, or deeper political preferences at state level, of which the rule is only an expression.

Finally, as Henning and Kessler emphasise, a no less important lesson for Europe is that policy principles and institutions should be looked at as a system rather than in isolation. As the authors point out, it may seem obvious to recall that states in the US can abide by strict budget balance rules to the extent the federal government is responsible for stabilisation and the bail-out of insolvent banks, but this simple lesson is sometimes overlooked in European discussions.

*Jean Pisani-Ferry, Director, Bruegel
Brussels, January 2012*

FISCAL FEDERALISM: US HISTORY FOR ARCHITECTS OF EUROPE'S FISCAL UNION

INTRODUCTION¹

However the euro area navigates its immediate financial crisis, fundamental reform of fiscal governance – institutions, rules, and procedures – will remain a central part of its long-term agenda. The relationship between the monetary union and its member states is at the heart of the debate. The European Council in December 2011 decided on a ‘fiscal compact’ for the euro area that includes the adoption of a new rule restricting deficits by member states in their constitutions or framework laws – ‘debt brakes’. This measure complements an array of other provisions to prevent the emergence of large fiscal deficits and strengthen the sanctions for rule violations. Debate continues over whether the euro area must create a deeper fiscal union, including the introduction of joint and several responsibility for bonds.

These debates about Europe’s future often reference the historical experience of the United States. Some salient features of US fiscal federalism that are frequently addressed in the literature on European monetary integration include the (i) transfers that take place between the federal government and the states, (ii) absence of federal bailouts of the states, and (iii) limitations on the deficits of the individual states and their correspondingly low debt.

This essay briefly reviews American fiscal history in light of questions confronting the future of fiscal union in the euro area. We draw lessons

for the dilemmas that Europe faces and point readers to particularly illuminating episodes and references in the literature on the United States. The fact that states encountered major debt crises and defaulted, yet the union managed to overcome them intact, points to relevant lessons for European policymakers in the current turmoil. There is a deep and varied literature and we have not exhausted it; but we believe that readers will be well rewarded by referring to the contributions listed here. Focusing on the problem of how to grapple with dysfunctional members of a union, and placing the financial challenge in political and historical context, this essay is a hybrid of analytical interpretation and literature review. The euro area will not want to replicate US institutions, but will want to bear in mind the lessons from US successes and mistakes when redesigning its own institutions. We consider first the key phases in US fiscal history, and then examine issues and lessons relating to balanced budget rules that have been adopted by the states.

US FISCAL HISTORY IN A NUTSHELL

The fiscal history of the US federal government and its relationship to state and local governments can be assessed in five phases or episodes: (1) Alexander Hamilton's reforms immediately after the establishment of the new federal government under the US Constitution in 1789; (2) state defaults during the 1840s; (3) a series of defaults at state and local level after the Civil War; (4) the Great Depression of the 1930s; and (5) scattered municipal defaults between 1970-2010. We consider each in turn, devoting somewhat disproportionate attention to the early, formative period.

Hamilton's plan

The first Secretary of the Treasury, Alexander Hamilton, has been credited with creating a modern financial system for the new United States². The magnitude of his achievements becomes clear when the prior condition of the US economy is considered. Before 1790, the United States

was effectively bankrupt, in default on most of its debt incurred during the Revolutionary War, and had no banking system, regularly functioning securities markets or national currency³. The federal government was reliant on the thirteen states to collect and share tax revenue, and was unable to pay war veterans or to service, let alone redeem, debts. Under the Articles of Confederation, the federal government had no executive branch, judicial branch or tax authority. As an 'institutional equilibrium', the United States were decidedly unstable and the financial predicament largely drove the constitutional reform of 1787 in Philadelphia.

After George Washington was inaugurated as President in April 1789, and Hamilton was confirmed, the new treasury secretary began to propose a series of institutional innovations on which he had been cogitating for at least a decade. Based on the secretary's study of British and Dutch financial institutions, these proposals included the establishment of the Bank of the United States, the mint, securities markets, and, most interestingly for our purposes, the assumption of state debts by the federal government. These were to be complementary, mutually reinforcing elements of the new American financial architecture. He delivered the first of his blueprints in the form of the *Report on Public Credit* to Congress in January 1790 (Chernow, 2004, pp297-306; Hamilton, vol. 6, *Report on Public Credit*).

The debt assumption plan involved the transfer of state debt to the federal government in the amount of \$25 million. Added to existing federal debt owed to foreign governments (France) and domestic investors, amounting to \$11.7 million and \$42.1 million respectively, federal debt would then amount to \$79.1 million (Sylla, 2011) – a very large sum compared to nominal GDP in 1790, which is estimated at \$187 million. In at least three prolonged debates over the course of 1790 and 1791, opponents in Congress levelled several arguments against the various elements of the plan. They objected that it (1) would reward speculators who had purchased debt from (distressed) original investors; (2) was unfair, as some states (such as Massachusetts and South Carolina) would be relieved of greater bur-

dens than others (such as Virginia and North Carolina); and (3) would empower the federal executive at the expense of Congress and the states.

Thomas Jefferson, who reluctantly served as Secretary of State in Washington's first administration, levelled particular criticism at the Bank of the United States, which Hamilton proposed should help manage the assumed federal debt, hold federal tax receipts and provide intermediation. Only three banks existed in 1789 and state governments served as financial intermediaries (Savage, 1988, p99). The Bank of the United States was to be capitalised at \$10 million – several times larger than the combined capital of existing banks – with \$8 million held by the private sector and \$2 million held by the US government. The federal government's share was to be paid in Treasury securities, as would three-quarters of the private share, thus harnessing the assumption plan to the development of the new financial system. Jefferson argued strenuously that the bank, not specifically provided for at the Philadelphia convention, was unconstitutional.

Hamilton addressed these objections systematically, arguing against 'discrimination' between the original and present holders of debt on the grounds not only of fairness but also that it was essential for an efficient, unified securities market. He justified 'assumption' on the grounds that state debt was issued for a common purpose: to prosecute the War and secure independence. He defended the constitutionality of his proposals by arguing that, even though the Constitution did not mention the establishment of a new bank, the authority to do so was implied by granting the executive the responsibility for the financial affairs of the federal government – the 'implied powers doctrine', which became a permanent feature of US constitutional law (Chernow, 2004, pp344-61).

More fundamentally, Hamilton saw his plan as not simply a way to secure credit for the federal government, or even to establish a national financial system, but as a grand political project. He envisaged the plan as (1) aligning the issuance of debt with the tax base, securing

from the individual states the federal government's claim to the tax base reserved to it under the Constitution (tariffs), (2) securing for the federal government the allegiance of the holders of federal debt and the financial system revolving around it, and (3) generally binding the states to the union (Hamilton, *Report on Public Credit*).

As part of the plan – and this point is sometimes lost in reviews of this period – Hamilton *restructured* the debt. The restructuring applied to both prior federal debt and the newly assumed state debt. The secretary adopted a menu approach to debt exchange of (a) a bond paying 6 percent, equal to the previous rate, (b) a bond paying 3 percent, and (c) one paying 6 percent but on which interest was deferred for ten years. He sweetened the offer by providing call protection (the Treasury could not withdraw bonds when the interest rate fell) and a sinking fund to assure repayment (Sylla and Wilson, 1999; Ratchford, 1941, pp52-72; McGrane, 1935). Investors were offered a swap of old debt for a combined package of the three new bonds, on a voluntary basis. The new bonds were perpetual, without a redemption date, on the model of British consols. By the time Hamilton left office in early 1795, 98 percent of domestic (federal and state) debt had been exchanged on these terms. Foreign debt, mainly to France, was repaid in full, with accrued interest, with the proceeds of a new loan from Dutch bankers (Sylla, 2011, pp11-13).

The implementation of the assumption plan quickly became embroiled in distributional conflict among the states. The federal government assumed \$18.3 million in state debt in 1790 and charged the states accordingly through accounts that were to be settled in order to equalise the per capita costs of financing the Revolutionary War. The creditor states ended up being owed \$3.5 million by the debtor states and were issued this amount in new federal bonds plus another \$0.5 million to cover interest arrears, raising total state debt assumed to \$22.5 million. The debtor states were forgiven the corresponding balances that they owed⁴. Albert Gallatin, who served as treasury secretary in the Jefferson and Madison administrations, argued that Hamilton assumed \$10.9 million more of state debt than would have

been necessary had the settlement of accounts been concluded before assumption. Hamilton defended the sequencing as important for avoiding a conflict between the state and federal levels over the tax base, which, with the exception of import tariffs, was subject to the concurrent jurisdiction of the two levels⁵.

The contemporary debate over monetary union in Europe appears polite compared to the ferocity with which Hamilton's plan was debated in Congress. The assumption plan was rejected once by Congress and then stitched into a famous compromise relocating the seat of government from New York to the District of Columbia. But the debate effectively reopened the fragile constitutional compromise of 1787 over the balance to be struck between the states and the union, and between Congress and the executive⁶. The politics around these issues were so vituperative in spring 1790 that, in Chernow's (2004, p326) assessment, it would not have been far-fetched to think that the union could break up⁷. As it was, the episode crystallised opposing forces in American politics and forced the irreconcilable split between the 'Federalists', led by Hamilton, and the 'Democratic Republicans', led by Jefferson and James Madison, that was to define American politics for decades (Chernow, 2004, pp320-331). The Bank of the United States became the primary battlefield for this ongoing factional conflict⁸.

Establishing the 'no-bailout' norm in the 1840s

Although a critical part of the US financial system, the debt assumption of 1790 set a precedent that endured for several decades. The federal government assumed the debt of states again after the War of 1812 and then for the District of Columbia in 1836. During this period, the possibility of a federal bailout of states was a reasonable expectation; moral hazard was substantially present. This pattern was broken in the 1840s, when eight states plus Florida, then a territory, defaulted.

Hamilton had wanted to make the federal government the sole creditor of the states with the assumption plan (Rodden, 2006, p57), but was unsuccessful. With the exception of the War of 1812, states issued

relatively little debt during the early nineteenth century and levied correspondingly few taxes. Land sales, bank charters and various investments were the main sources of state revenue, and the role of state government was quite limited. Westward expansion during the 1820s and 1830s, however, gave rise to demands for infrastructure, such as canals and turnpikes – which could only be financed through borrowing. This borrowing was done on the theory that debt would be serviced and repaid by tolls and other project revenue, without raising taxes, dubbed ‘taxless finance’ (Wallis, 2005). The Erie Canal was one of the first and most financially successful of these projects. With the financial panic of 1837 and recession of 1839-43, however, much of the debt incurred became unserviceable.

The indebted states petitioned Congress to assume their debts, citing the multiple precedents. British and Dutch creditors, who held 70 percent of the debt on which states later defaulted, pressed the federal government to cover the obligations of the states. They argued that the federal government’s guarantee, while not explicit, had been implied. Even the prices of bonds of financially sound states fell, and the federal government was cut off from European financiers in 1842. In that year, Lord Ashburton, the main British negotiator for Barings, wrote that the United States was an “ungovernable and unmanageable anarchy” (Roberts, 2010). John Quincy Adams evidently believed that another war with Britain was likely if state debts were not assumed by the federal government⁹.

However, on this occasion Congress rejected the assumption petition and was able to do so for several reasons. First, debt had been issued primarily to finance locally beneficial projects, rather than national public goods. Second, domestically held bonds were not a large part of the US banking portfolio, and default had limited contagion effects, at least through this particular channel. Third, the financially sound states were more numerous than the deeply indebted ones. And, finally, the US economy had matured to the point where it was less dependent on foreign capital. Foreign loans were crucial to Hamilton’s plan in 1790, but they were a minority contribution when investments eventually

resumed in the 1850s (McGrane, 1935, pp21-40; Savage, 1988, pp105-118; Wibbels, 2003; Wallis, 2005).

Eventually, most states repaid all or most of their debt as a condition for returning to the markets. The State of Maryland provides a good example. It had financed the construction of the Chesapeake and Ohio Canal, the Baltimore and Ohio Railroad, which competed with the canal project, and a number of other railroads. When state officials were unable to service the debt with revenues from these projects, they had no system of direct taxation on which to fall back, and the value of real property declined quickly. The state suspended payments on bonds between 1841 and 1847, but resumed payments thereafter, including accrued interest, and accessed the British market through Barings again in 1849. The state paid off its pre-crisis debt by 1851 (McGrane, 1935, pp82-101). However, Maryland and the states that had defaulted returned to markets at a premium, whereas the others were able to borrow at normal rates relatively soon after the crisis (English, 1996).

The rejection of debt assumption established a 'no bailout' norm on the part of the federal government. The norm is neither a 'clause' in the US Constitution nor a provision of federal law. Nevertheless, whereas no bailout request had been denied by the federal government prior to 1840 (Ratchford, 1941), no such request has been granted since, with one special exception discussed below. The fiscal sovereignty of states, the other side of the no-bailout coin, was thereby established.

During the 1840s and 1850s, states adopted balanced budget amendments to their constitutions or other provisions in state law requiring balanced budgets. This was true even of financially sound states that had not defaulted. The adoption of balanced budget rules continued over the course of subsequent decades, so that eventually three-fourths of the states had adopted such restrictions. Because this is a direct analogue to the adoption of constitutional 'debt brakes' in the euro area, it is important to understand the political economy of the adoption of these provisions of the states in the nineteenth century. We devote a section to this topic below. Suffice to say at this point,

however, that several states did not adopt such amendments and in 1860, on the eve of the Civil War, collective state indebtedness stood at \$247.4 million, \$67.5 million greater than in 1841. This sum was four times the size of the federal debt [Savage, 1988, p118].

Reconstruction defaults

After the Civil War, the process of reintegrating the Southern states into the Union and reviving their economies – ‘Reconstruction’ – witnessed an extraordinary degree of corruption and political dysfunctionality¹⁰. “As soon as the military authorities were removed,” writes Ratchford (1941, p170), “the Reconstruction governments rushed to plunder public treasuries. Since those treasuries were usually empty and since the possibilities of taxation were severely limited, the only alternative was to despoil the public credit”. By the end of Reconstruction in 1874, the total debt of the eleven Southern states had risen to \$247.6 million from \$111.4 million in 1865. Most of it took the form of direct state bonds or guarantees of railroad company bonds. Ratchford (1941, p180) estimates that about \$12 million of these bonds was sold abroad. Much of this new debt was regarded as odious by Southern electorates, imposed by “corrupt and hostile governments supported by outside military force”¹¹. Eight states thus repudiated part of their debt or reduced it by other not-so-voluntary means by \$116.3 million over the next sixteen years.

Although the no-bailout position of the federal government *vis à vis* the states had been established prior to the Civil War, the position of the states themselves *vis à vis* their counties and cities remained to be tested. During Reconstruction, many local governments also accumulated debt that they were unable to service during the 1870s. Almost all of this debt was held abroad or by out-of-state residents – a pattern facilitated of course by integrated capital markets – and the taxes necessary to repay would have come from upper-income households. States successfully fended off petitions for bailouts – none were provided – and all states re-wrote their constitutions to prohibit them [Inman, 2003, p58 and p65], and several strengthened limits on

issuance of their own debt¹². Remarkably, Inman (2003) records only one instance of state bailout of a municipality (Camden, New Jersey) in the history of the United States¹³.

Great Depression and fiscal shift

The 1930s saw another wave of defaults by local governments and the last default to be recorded by a state. Between 1920 and 1930 capital investment by local governments doubled, financed by general obligation bonds backed by property tax. Owing to the collapse in the tax base with the Depression, by December 1935 more than 3,200 local governments had defaulted on \$2.4 billion of these debts. This debt was owed to domestic rather than foreign investors, yet neither state nor federal bailouts were provided. Instead, massive fiscal shifting took place as states and the federal government adopted new programmes, taking over some functions from local government, and provided direct assistance as local government cut spending (Inman, 2003, p59). There was a complete reversal in the relative shares of total government spending of the three levels over the course of the Depression. Whereas in 1932 local governments spent 50 percent, states 20 percent and the federal government 30 percent of the total, by 1940 local governments spent 30 percent, states 24 percent, and the federal government 46 percent (Wallis, 1984). The period thus marks the ascendance of the federal government relative to the states and, notwithstanding President Franklin D. Roosevelt's instinctive fiscal conservatism, the introduction of countercyclical demand management at the federal level. Most of the defaulted debt and interest was repaid in full by 1940 (Inman, 2003, p66). The last state default occurred in 1933 when Arkansas suspended payments on its highway bonds. By 1943, the majority of defaulting issues were refinanced and the state returned to good standing in debt markets (Ang and Longstaff, 2011).

1970-2010: scattered municipal bankruptcies

Over the last several decades, until the 2008-09 crisis, the United States did not see a wave of bankruptcies but rather a series of

relatively isolated municipal problems. Between 1970-2009, 54 municipal bond issuers defaulted, while between 1988-2009 about 170 jurisdictions declared bankruptcy (Kasperek, 2011, p16). The case of New York City in 1975 is remembered by the headline in the *New York Daily News*, 'Ford to New York: Drop Dead'. The city's rescue was primarily organised by the Governor of New York through the Municipal Assistance Corporation (MAC). When Philadelphia encountered problems in 1990, the response was organised by the state of Pennsylvania. The cities of Bridgeport and Miami and Orange County defaulted. However, the states did not provide funds to repay bondholders in these cases. The single exception to this rule is the state of New Jersey, which bailed out the city of Camden (Inman, 2003, p60). The bankruptcy filing of Jefferson County, Alabama, in November 2011, is the most recent; it follows the bankruptcy filing by Harrisburg, Pennsylvania, and a couple of other cities¹⁴.

The single exception to the federal government's no-bailout position is the case of the District of Columbia in the 1990s, an exception that proves the rule. In this case, Congress did indeed take control of the District's finances, injected funds, and managed the budget for four years through the District of Columbia Financial Control Board, created in 1995, which left the city in surplus after four years. This was possible because of a special clause in the Constitution giving Congress authority over the administration of the District – authority that does not extend to the 'sovereign' states¹⁵.

In the present crisis, the finances of the states of California and Illinois have captured attention. Both states have large populations and economies – 13 percent and 4.5 percent of national GDP respectively – and have large budget deficits and dysfunctional politics. Both states also have balanced budget amendments in their state constitutions, illustrating the leaky character of these provisions¹⁶. The size of these states and the impact of a default by one of them on US financial markets have generated speculation about the possibility of a federal bailout *in extremis*. While this might be a possibility in the abstract, the absence of a modern precedent places a high institutional bar on such

action. The expectation of such a bailout is hardly detectable in the domestic political wrangling within both states; their political parties seem resigned to resolving these issues independently.

BALANCED BUDGET RULES

In light of the provisions being adopted in Europe, three aspects of the balanced budget rules of the states deserve elaboration: (1) the politics of their propagation; (2) the exact nature of the requirements and their variation among the states; and (3) their effectiveness in limiting deficits.

Emergence

The adoption of balanced budget rules among most of the states during the nineteenth century raises a number of interesting questions. What drove the adoption across disparate states? Did capital markets insist on them by discriminating between states with provisions and those without? What role did voters and elections play? Did the federal government promote these rules? The published work casts some light on these questions; but many questions remain to be addressed by future research.

The first wave of adoptions among nineteen states between 1842-57 is closely linked to the financial panic of 1837 and subsequent economic depression. According to Wallis (2005) and Wallis and Weingast (2008), the emergence of balanced budget rules should be understood as the demand of voters for more transparent and realistic financing rules. All of the states that defaulted in the 1840s except Florida, Mississippi and Arkansas, wrote some kind of deficit restriction into their constitution immediately afterward. The point was not so much to forbid deficits altogether, but to avoid 'taxless finance' and other forms of infrastructure financing which were either not sustainable or easily led to corruption¹⁷. Quite often, state governments were forced to obtain public approval by referendum to issue debt for a project and

simultaneously increase taxes in order to service it. New states admitted to the Union after the Civil War generally included debt limits in their constitutions (Ratchford, 1941, p122, whose explanation is consistent with Wallis and Weingast).

The federal government was passive during the adoption of these provisions by the states. The federal government certainly did not mandate the adoption of these provisions and it does not appear that it was promoting them either. Nor does it appear that states pressed for conformity on the part of their neighbours, or even, by these accounts, that states were competing against one another for access to lower-cost financing by adopting them. Existing treatments suggest that states were acting autonomously, though the financial challenges were common, and the political pressures for adoption were internal. By contrast, the current adoption of 'debt brakes' in the euro area is driven more by the most dominant member states and the euro-area institutions. But internal support is almost surely necessary for the meaningful implementation and perpetuation of these rules. The episode of the 1840s also underscores the importance of crisis as a driver of institutional change, which is reflected in the contemporary European experience.

Rule characteristics¹⁸

Because each state adopted its own balanced budget rule (Vermont is the only state without such a rule in some form), there is a large variation in the way these rules function. According to the National Association of State Budget Officers (NASBO), which surveyed State Budget Controllers, 44 states have a constitutional or statutory rule that requires the governor to submit a balanced budget, 41 require the legislature to pass a balanced budget, while 37 demand that the governor signs a balanced budget. Finally, 43 states simply forbid carrying over a deficit to the next budget plan (annually or biannually). Note however that different readings of state laws or jurisprudence have led some academics or institutions to rank the restrictiveness of these provisions differently.

Balanced budget rules usually apply only to the state's *general fund*, which receives most tax collections and from which most expenditures are made. Grants and reimbursements from the federal government make up most of states' non-general fund: those funds are balanced *de facto* because federal grants are earmarked for specific projects, and the money spent matches the money received. Moreover, most states have separate operating and *capital budgets*; bond finance for capital projects does not fall within most balanced budget rules (NCSL, 2010). Therefore most balanced budget rules are 'golden'¹⁹. Note, however, that unfunded liabilities are often excluded from state debt calculations, and rules sometimes allow the diversion of revenues from pension funds to the general fund, aggravating the underfunding of obligations in the long run.

Finally, it must be noted that, while we have focused on balanced budget rules, rules that apply specifically to expenditure or taxes also have a substantial impact on the budget behaviour of states. Some states, such as California, require a supermajority (or qualified majority) in the legislature to raise taxes. Poterba and Rueben (2001) show that such limitations on tax increases tend to raise state bond yields. Conversely, constraints on the spending side seem to facilitate lower yields. 'Rainy day funds', another common fiscal instrument, allow states to save and smooth the fiscal path, by saving at the crest of the business cycle.

Effectiveness

Balanced budget rules have not prevented states from getting into fiscal trouble – witness the recent experience of California and Illinois. On the other hand, the overall debt of states has been reasonably well contained. In 2009, California's was less than 8 percent of state GDP, New York 11.2 percent, and New Jersey 12.1 percent (see Table 1). Political scientists and economists have tried to untangle the effects of these provisions from factors such as party control of state government, political culture and capital markets (see Poterba 1996; Bayoumi, Goldstein and Woglom, 1995, on market discipline; Rose,

Table 1: Debt and deficits of 'problem states', 2009-11

	Deficit as a % of general fund			State debt as % of GDP	State and local outstanding debt as % of GDP
	2009	2010	2011	2009	2009
Arizona	36.8%	65.0%	39.0%	4.9%	19.4%
California	36.7%	52.8%	20.7%	7.3%	20.2%
Nevada	19.9%	46.8%	54.5%	3.6%	20.7%
Illinois	15.1%	43.7%	40.2%	9.0%	20.3%
New Jersey	18.8%	40.0%	38.2%	12.1%	19.6%
New York	13.2%	38.8%	15.9%	11.2%	26.8%

Source: McNichol, Oliff and Johnson (2011); data for 2011 is estimated from state sources; Bureau of Economic Analysis for debt data; Census Bureau for state GDP up to 2009. Note: The states selected had the largest deficits in 2010 as a percentage of their general fund.

BOX 1: TWO EXAMPLES OF BALANCED BUDGET RULES

1 California:

In 2004, by referendum, California passed Proposition 58, also called the 'Balanced Budget Act', which reinforced the previous balanced budget rule by constitutional amendment. The specific provision, now included in the constitution as article IV, section 12(g) states:

"For the 2004–05 fiscal year, or any subsequent fiscal year, the Legislature may not send to the Governor for consideration, nor may the Governor sign into law, a budget bill that would appropriate from the General Fund, for that fiscal year, a total amount that, when combined with all appropriations from the General Fund for that fiscal year made as of the date of the budget bill's passage, and the amount of any General Fund moneys transferred to the Budget Stabilization Account for that fiscal year pursuant to Section 20 of Article XVI, exceeds General Fund revenues for that fiscal year estimated as of the date of the budget bill's passage. That estimate of General Fund revenues shall be set forth in the budget bill passed by

the Legislature”.

The Act also allows the Governor to proclaim a fiscal emergency in specified circumstances, and to submit proposed legislation to address the fiscal emergency; requires the Legislature to stop other action and act on legislation proposed to address the emergency; establishes a budget reserve; provides that the California Economic Recovery Bond Act is for a single object or work; and prohibits any future deficit bonds.

2 Illinois:

Article VIII, Section 2 of the Constitution of Illinois reads:

- a) The Governor shall prepare and submit to the General Assembly, at a time prescribed by law, a State budget for the ensuing fiscal year. The budget shall set forth the estimated balance of funds available for appropriation at the beginning of the fiscal year, the estimated receipts, and a plan for expenditures and obligations during the fiscal year of every department, authority, public corporation and quasi-public corporation of the State, every State college and university, and every other public agency created by the State, but not of units of local government or school districts. The budget shall also set forth the indebtedness and contingent liabilities of the State and such other information as may be required by law. Proposed expenditures shall not exceed funds estimated to be available for the fiscal year as shown in the budget.
- b) The General Assembly by law shall make appropriations for all expenditures of public funds by the State. Appropriations for a fiscal year shall not exceed funds estimated by the General Assembly to be available during that year.

2010, for a general survey). Our bottom line from a review of this literature is that balanced budget rules probably do have salutary independent effects on debt accumulation but that interaction with the political environment and markets is critically important.

First, there is a great deal of variation in the strictness of rules among states and more stringent conditions lead to less borrowing. Bohn and Inman (1996) estimate that a requirement that the budget be balanced at the end of the year reduces the probability of a deficit from 26 percent to 11 percent. The rule has stronger effects when it is inscribed in the constitution and exceptions require a qualified majority in the legislature. Requiring preparation of a balanced budget is easily circumvented by optimistic economic projections, whereas prohibitions on carrying over a deficit from one year to the next are effective (see, Hou and Smith, 2009, who distinguish between 'political' and 'technical' provisions, and Mahdavi and Westerlund, 2011). Enforcement also varies among states, with Virginia, for example, lacking a binding mechanism (NCSL, 2010). In the 26 states in which Supreme Court judges are elected, rather than nominated by the governor and confirmed by the legislature, the rules are enforced more strictly and deficits thereby limited (Bohn and Inman, 1996).

Second, states with stricter rules are better perceived by the market. Poterba and Rueben (1999) show that weak provisions cost 10 to 15 basis points, when compared to similar states with stricter anti-deficit rules. In a subsequent paper (Poterba and Rueben, 2001), they show that unexpected deficits lead to higher yields for states with weak rules than states with strong rules.

Third, several studies have found the effectiveness of balanced budget rules to be contingent on the politics within the state, such as unified party control of both houses of the legislature (Alt and Lowry, 1994) and unified control of the governorship and legislature (Poterba, 1994, as cited in Briffault, 1996). Inman (1998) admits that scholarship cannot rule out the possibility that balanced budget rules are adopted by states that are fundamentally fiscally conservative and thus have

little independent impact (the endogeneity problem). After reviewing the literature of the mid-1990s, Briffault (1996, p60) concludes: "It seems likely that the real importance of a constitutional balanced budget requirement is that it signals the high value that a state's political culture sets on a balanced budget. A state with such a constitution may be more likely to balance its budget, but that is less attributable to the independent force of the legal requirement than to the political values and tradition that put it in the constitution in the first place". The likely effectiveness of balanced budget rules that are adopted by euro-area member states in the midst of the present crisis should be assessed in this light.

MACROECONOMIC STABILISATION

Because state and local budgets are about 40 percent of total government spending in the United States, fiscal policy is effectively shared by the levels within the federal system. This is often not given due recognition in the discourse about macroeconomic stabilisation, in which the role of the federal government is sometimes an exclusive focus. Balanced budget provisions of the states do not provide for cyclical adjustment of the calculated deficit. These provisions, to the extent that they are effective, require raising taxes and/or cutting spending when revenue falls during recessions. The size of this effect can offset a substantial portion of the countercyclical movement of the federal budget position. Krugman (2008), for example, refers to the states as the 'fifty little Herbert Hoovers', pursuing fiscal contraction when Keynesian measures were in order as the United States was sliding into the 'Great Recession'.

The conventional wisdom has been that the budget positions of state and local governments move procyclically in the United States. Poterba (1994) found strong evidence for this, and that states with stricter balanced budget rules cut spending more than those with looser rules during recessions. The strength of this finding appears to vary over time and over the business cycle, with some studies concluding that

state and local budgets overall are neutral or even weakly countercyclical (Sorensen, Wu and Yosha, 2001; Hines, 2010). Sorensen and Yosha (2001) report that state budget positions are countercyclical at the top of the business cycle but procyclical in recessions. During the Great Recession and slow recovery, most studies find that state and local budgets have acted procyclically (Aizenman and Pasricha, 2010; Follette and Lutz, 2010; and Kasperek, 2011), while some find neutrality (Hines, 2010). These studies agree, however, that in the aggregate²⁰, state and local budgets do not help to stabilise the macroeconomy during recessions; that role is played by the federal government in the United States²¹.

Fiscal transfers from the federal government directly into state budgets, to help them fulfil federal mandates and otherwise alleviate budget pressure, ameliorate the procyclical influence of the states during downturns²². The American Recovery and Reinvestment Act (ARRA) of 2009, for example, provided large amounts of support to the states. According to the Bureau of Economic Analysis (BEA), the level of total grants-in-aid to state and local governments in 2009 was \$482 billion, \$70 billion of which came from the stimulus package. Federal support then rose to \$532 billion in 2010, of which \$100 billion was accounted for by ARRA. A large part of the support was directed through Medicaid to cover the shortfall of revenues at the state level. The rest was either spent in the education sector or earmarked for various investment projects (see Table 2). In 2010, according to the Congressional Budget Office (CBO), 75 percent of the grants to states contained in the stimulus package were used to finance state deficits rather than fund new projects. The high-profile protests in mid-Western states during summer 2011 responded to state expenditure reductions that in large measure were done in anticipation of the phasing out of federal stimulus.

In sum, the federal government (1) is the only level that provides significant stabilisation during recessions, while the states are likely to be procyclical, and (2) injects federal money into state programmes directly. Both roles render the balanced budget rules at the state level

Table 2: Federal support to the states, 2009-11 (US\$ billions)

	2009		2010		2011 (Q1-Q3), annualised	
	Transfers from federal government	Of which, from the stimulus package	Transfers from federal government	Of which, from the stimulus package	Transfers from federal government	Of which, from the stimulus package
Grants-in-aid to state and local governments	482.4	70.2	531.5	100.8	504.3	43.5
Medicaid	264.4	41.3	281.5	45.9	265.3	2.7
Education	59.8	21.3	71.8	30.9	67.9	24.3
Other	158.3	7.6	178.3	23.7	171.1	16.5

Source: Bureau of Economic Analysis.

more sustainable than they would be in the absence of the federal government and its fiscal system.

FINANCIAL MARKETS AND CONTAGION

US banking and capital markets are another element of the context in which budget rules operate and in which the states relate to the federal government on fiscal matters. Consider first the capital markets and then the organisation of banking, its regulation and restructuring.

State and local government debt in the United States totalled \$2.45 trillion at the end of 2010, 16.7 percent of GDP. Of this amount, roughly 40 percent had been issued by states and 60 percent by local entities. This compared to federal debt held by the public of \$9.36 trillion, or 64.4 percent of GDP. Most of the state and local debt is longer than one year, with an average maturity of over 14 years, issued to finance capital expenditures; relatively little has been issued to finance current spending²³. Thus, only 16 percent of outstanding debt is general obligation bonds backed by the full faith and credit of the issuer; the vast majority is secured by the revenue stream of infrastructure and other investment projects. Balanced budget rules thus appear to have been effective in configuring the composition of state debt and the recourse of investors. Half of outstanding state debt is accounted for by seven states – California, New York, Massachusetts, Illinois, New Jersey, Pennsylvania, and Florida – with California and New York representing 12.9 and 11.7 percent respectively. Most of these bonds are sold within the US, the interest on them being exempt from federal income tax. These outstanding obligations comprise the municipal bond market in the US²⁴.

Municipal bonds normally trade at a premium relative to US Treasury securities, owing to the federal tax exemption. From 2005-07, for example, the yield on ten-year municipal bonds was roughly one percent below ten-year Treasury bonds. That relationship was reversed briefly during the acute phase of the crisis at the end of 2008 and restored in

spring 2009, with the spread evaporating at the end of 2010 and remaining quite narrow through 2011. Shocks in state bond markets generally do not appear to have substantial consequences for the markets in Treasury securities and vice versa²⁵.

Do the markets effectively discriminate among the bonds of different states? As of late 2011, Standard and Poor's rates thirteen states AAA, fourteen states AA+, seventeen states AA, four states AA-, one state (Illinois) A+ and one state (California) A-²⁶. Between 2004 and 2007, when European sovereign bond spreads were nearly eliminated, the average spread between the Aaa and Bbb state bonds (using Moody's ranking) were in the range of 58 to 46 basis points. That spread rose to 207 basis points during the crisis in 2009 and declined somewhat in 2010. Credit default swap (CDS) markets emerged for the bonds of eleven states during the Great Recession. Yields on California bonds spiked to 4.55 percent in December 2008, when its CDS rate peaked over 500, and fluctuated in the neighbourhood of 3 percent thereafter. Ten-year CDS rates ranged from 288 for California and 284 for Illinois to about 65 for Texas, Virginia and Maryland in September 2011.

The operation of the US municipal market suggests comparisons to those of the European sovereign market. Markets seem to have been more discriminating among the US states than euro-area members during the quiescent mid-2000s, particularly given the relatively small differences in the debt load of the states compared to the differences in those of the euro-area members. Whether the markets' extreme differentiation of European sovereign bonds between 2010-11 is fully justified is beyond the scope of this paper, in comparison, the markets' moderate differentiation among US state bonds seems to be broadly proportionate to risk²⁷.

How effectively do US institutional arrangements and market structures insulate one state from the fiscal mistakes of another, given that an individual state default might trigger systemic risk? One might expect that macroeconomic contagion should be much stronger in a tightly integrated federation such as the US than in a more loosely

integrated one (Auerbach, 2011). However, spillovers in the US municipal market tend not to be of the 'contagious' type but of the 'flight-to-quality' type. When the large borrowers such as California, Georgia, Maryland and the City of New York experience problems, other issuers see their yields fall (Arezki, Candelon and Sy, 2011). Although contagious linkage is sometimes found, such as between California and New York bonds, studies using CDS data also generally find that insulation is strong (Ang and Longstaff, 2011).

State bond yields could move more independently than the level of macroeconomic integration would lead us to expect for several reasons. First, the existence of a deep and liquid market in US Treasury securities as a 'safe haven' might play a role that has not been sufficiently investigated. Second, US banks do not seem to transmit shocks to states as European banks do to European sovereigns, probably owing to differences in bond ownership and regulatory frameworks (Ang and Longstaff, 2011). Third, setting fiscal rules independently in each state and enforcing them internally might isolate deviant behaviour and protect others, whereas a rule set and enforced centrally (such as the excessive deficit provisions of the Maastricht Treaty and the Stability and Growth Pact) might quickly lose credibility when one sovereign breaches it.

Fundamental characteristics of the US financial system are germane to the operation of the municipal bond market and to the role of balanced budget rules. Banks are less important conduits for finance in the United States than they are in Europe. Only 8.6 percent of state bonds outstanding were owned by US commercial banks in 2010. Although state regulation of banks was the norm in earlier periods, US banking regulation is now far less geographically fragmented than in Europe. During the savings and loan banking crisis of the 1980s, which had a strong regional dimension, the federal government took primary responsibility for the restructuring. During the far larger rescue and restructuring of the banking system during 2008-10, the states played very little role. The risk pool is nationwide in the United States and the bank rescue did not impair the fiscal position or creditworthiness of the

individual states. Although Europe has taken substantial steps toward a common regulatory framework, the contrast with the United States remains stark.

CONCLUSION: RAMIFICATIONS FOR EUROPE

We have reviewed US fiscal federalism, from Alexander Hamilton to the present, to provide insights for the architects of fiscal federalism in Europe. The US federal government has not bailed out state and local governments since the early nineteenth century and the no-bailout norm has been formed politically; there is no clause in the Constitution. States adopted balanced budget rules of varying strength during the nineteenth century and these rules seem to be consequential, sometimes as a constraint, sometimes as a signal to capital markets, for state fiscal policy.

Balanced budget rules among the states seem to parallel the effort – adopted at the March 2011 European Council meeting and affirmed at the December 2011 summit²⁸ – to introduce constitutional rules or framework laws, ‘debt brakes’, in the member states of the euro area²⁹. The fiscal compact agreed at the December 2011 summit specified that under these restrictions members’ annual structural deficits should not exceed 0.5 percent of nominal GDP. Before drawing too heavily on the US experience in concluding that constitutional debt brakes are a key solution to Europe’s debt problems, however, Europeans should consider three essential aspects of the context in which the balanced budget rules of the states operate. The US experience suggests that the particular path through which rules are adopted and enforced is likely to be critical to their implementation, and that introducing such rules for euro-area member states should be accompanied by a federal system of fiscal powers and a common fund for rescuing and recapitalising banks. We consider these three caveats in turn.

Within the US federal system, the states are ‘sovereign’ with respect to debt³⁰. This sovereignty has two facets. On the one hand, the federal

government neither mandates nor enforces balanced budget rules for the states. Although states were responding to similar financial problems in similar ways, these rules were adopted autonomously and are implemented independently from the federal government. In modern parlance, states' 'ownership' of these rules is complete. On the other hand, states have no recourse to the federal government when they have difficulty servicing and repaying debt. The federal government assuages funding problems in current expenditures, such as through the ARRA of 2009, but these transfers are largely discretionary and do not relieve state and local governments of debt obligations.

The American constitutional design is thus very different from what European leaders envisage for the euro area: debt brakes that are mandated by the Union and enforced by the European Commission and the EU Court of Justice. The difference is likely to be consequential in two respects. We suspect that local ownership and enforcement make debt brakes more effective than under central mandates, particularly in the context of credible no-bailout norms, and that rules that are centrally mandated are likely to prove to be more brittle than those adopted in a decentralised fashion. When one state violates the rule, as the experience with the Stability and Growth Pact demonstrated, its applicability to other states is less credible. That is less likely to be the case with rules that have been adopted autonomously.

We acknowledge that some of the impetus for debt brakes comes from within euro-area countries. The present crisis could be sufficiently traumatic and thus politically transformative to produce an autonomous reduction in debt tolerance within some of the most afflicted member states, just as the US states adopted balanced budget rules autonomously from the federal government in the nineteenth century. Such an autonomous change in preferences would augur well for the effectiveness of debt brakes. But the strength of the internal shift in debt tolerance is uncertain and is likely to vary significantly among member states.

The second fundamental caveat is that the federal government's

relationship with the states must be seen within the context of a broader fiscal union³¹. Since Alexander Hamilton's plan was enacted, federal debt has been supported by the full system of federal powers, including a sweeping power to tax³². The federal government's role in public expenditure and taxation is substantial relative to the states. The theory of optimum currency areas has trained attention on the fiscal transfers among different regions of the country that are effected through the federal system of revenue and expenditures as well as through direct budget support to states and local governments. The magnitude of these transfers has been significant, though their exact importance is the subject of some debate.

However, the macroeconomic stabilisation role of the federal government is more important than intra-regional transfers in considering the budget restrictions of the states. Critically, the rigidity brought on by balanced budget provisions at the state level is facilitated by fiscal flexibility at the federal level. Despite the leakiness of these provisions, state and local budgets have behaved procyclically during recessions in the United States. Since the 1930s, the federal budget has helped to stabilise the national economy in countercyclical fashion. Without this, state-level restrictions would have been difficult or impossible to sustain. Although automatic stabilisers might play a greater role in some of the national economies in Europe than in the US states³³, we believe that creating stringent state-level debt brakes in Europe without a capacity for countercyclical stabilisation would be a serious mistake.

Europe faces a choice between designing debt brakes to provide for countercyclical action at the national level and creating a common countercyclical fiscal instrument of considerably larger size than the present EU budget. Each has advantages and disadvantages. Providing for countercyclical action in national provisions raises problems of enforcement and coordination with the fiscal stance of other members. Which institutions would calculate structural budget positions, ensure that these calculations are unbiased, and provide for consistency across member states? Creating a common capacity for countercyclical action requires strong political cohesion and robust institutions for

the monetary union. The need for a countercyclical fiscal capacity at one level or the other is not a new observation, but we believe that it is an inescapable one, the implications of which have not yet been sufficiently incorporated in European deliberations about the fiscal architecture.

US banking and capital markets are the third element of the context in which budget rules operate and the states relate to the federal government on fiscal matters. Compared to Europe, banks are less important conduits for finance relative to capital markets and bank regulation is less fragmented, being more of a federal responsibility. Stabilising the banking system, along with stabilising the macroeconomy, has been the responsibility of the federal government. In the US, the states have not themselves undertaken large-scale bailouts or recapitalisation of banks over the last century. As a consequence, the need to stabilise the banking system did not enter into conflict with balanced budget rules at the state level. In the euro area, by contrast, harmonisation of bank regulation is still young, and the fiscal costs of bank rescues and recapitalisation remain primarily a national responsibility. The introduction of debt brakes threatens to collide with the need to mount large-scale rescues of banking systems at the level of member states. As such provisions are put in place, therefore, it is all the more important that the euro area unifies banking regulation and creates a common pool of fiscal resources for rescuing, restructuring and recapitalising banks [Posen and Véron, 2009; Véron, 2011].

REFERENCES

- Alt, James E. and Robert C. Lowry [1994] 'Divided government, fiscal institutions and budget deficits: evidence from the States', *American Political Science Review* 88(December): 811-828
- Aizenman, Joshua and Gurnain K. Pasricha [2010] 'Net fiscal stimulus during the Great Recession', *Working Papers* 16779, National Bureau of Economic Research
- Amudrsky, Robert S. and Clayton P. Gillette [1992] *Municipal debt finance law: theory and practice*, Aspen Publishers
- Ang, A. and F. A. Longstaff [2011] 'Systemic sovereign credit risk: lessons from the US and Europe', *Working Paper* 16982, National Bureau of Economic Research
- Arezki, Rabah, Bertrand Candelon and Amadou N. R. Sy [2011] 'Are there spillover effects from munis?' *Working Paper* 290 (December), International Monetary Fund
- Auerbach, Alan J. [2011] 'Fiscal institutions for a currency union', paper prepared for a conference on *Fiscal and Monetary Policy Challenges in the Short and Long Run*, sponsored by the Deutsche Bundesbank and the Banque de France, Hamburg, May 19-20
- Baunsgaard, Thomas and Steven Symansky [2009] 'Automatic fiscal stabilizers: how can they be enhanced without increasing the size of government?' *Staff Position Note* SPN/09/23, International Monetary Fund
- Bayoumi, Tamim, Morris Goldstein and Geoffrey Woglom [1995] 'Do credit markets discipline sovereign borrowers? Evidence from the US states', *Journal of Money, Credit and Banking* vol. 27(4): 1046-59
- Bohn, H. and R. P. Inman [1996] 'Balanced-budget rules and public deficits: evidence from the US states', *Carnegie-Rochester Conference Series on Public Policy* 45, pp. 13-76
- Bordo, Michael D., Agnieszka Markiewicz, and Lars Jonung, 2011. A Fiscal Union for the Euro: Some Lessons from History. NBER Working Papers 17380. Cambridge, MA: National Bureau of Economic Research

Briffault, Richard (1996) *Balancing acts: the reality behind state balanced budget requirements*, New York: Twentieth Century Foundation

Chernow, R. (2004) *Alexander Hamilton*, New York: Penguin Press

Dolls, Matthias, Clemens Fuest and Andreas Peichl (2010) 'Automatic stabilizers and economic crisis: US vs. Europe', *Working Paper* 16275, National Bureau of Economic Research

English, W. B. (1996) 'Understanding the costs of sovereign default: American state debts in the 1840s', *American Economic Review* 86, no. 1: 259-275

Fatas, A. and I. Mihov (2006) 'The macroeconomic effects of fiscal rules in the US states', *Journal of Public Economics* 90 (1-2), 101-117

Follette, Glenn and Byron Lutz (2010) 'Fiscal policy in the United States: automatic stabilizers, discretionary fiscal policy actions, and the economy', *Working Paper* 43, Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board

Galligan, Brian (2008) 'Comparative federalism', in R. A. W. Rhodes, Sarah A. Binder and Bert A. Rockman (eds) *The Oxford Handbook of Political Institutions*, pp. 261-280, Oxford University Press

Gelpern, Anna (2012) 'Bankruptcy, backwards: the problem of quasi-sovereign debt', *Yale Law Journal* (forthcoming)

Hallerberg, Mark (2006) 'Fickle fiscal federalism? Fiscal relationships in the US and EU', in Anand Menon and Martin Schain (eds) *Comparative federalism*, Oxford University Press

Hallerberg, Mark (2011) 'Fiscal federalism reforms in the European Union and the Greek crisis', *European Union Politics* 12, no. 1: 127-142

Hallerberg, Mark, Rolf Strauch and Jürgen von Hagen (2009) *Fiscal governance: evidence from Europe*, Cambridge University Press

Hallerberg, Mark and Guntram B. Wolff (2006) 'Fiscal institutions, fiscal policy and sovereign risk premia in EMU', *Public Choice* vol 136, 3-4

Hamilton, A. (1789) *First report on public credit*, available at http://press-pubs.uchicago.edu/founders/documents/a1_8_2s5.html

Hamilton, A. (1987) *Papers of Alexander Hamilton*, edited by J. E. Cooke and H. C. Syrett, New York: Columbia University Press

Hines, James R. (2010) 'State fiscal policies and transitory income fluctuations', *Brookings Papers on Economic Activity*, Fall 2010, pp. 313-337

Hou, Y. and D. L. Smith (2009) 'Do state balanced budget requirements matter? Testing two explanatory frameworks', *Public Choice* 145, 57-79

Iara, Anna and Guntram B. Wolff (2011) 'Rules and risk in the euro area', *Working Paper* 2011/10, Bruegel

Inman, Robert, P. (1998) 'Do balanced budget rules work? US experience and possible lessons for the EMU', *Working Paper* 5838, National Bureau of Economic Research

Inman, Robert P. (2003) 'Transfers and bailouts: enforcing local fiscal discipline with lessons from US Federalism', in J. Rodden, G. Eskelund and J. Litvack (eds) *Fiscal decentralization and the challenge of hard budget constraints*, MIT Press

Johnson Simon, and James Kwak (2010) *Thirteen bankers: The Wall Street takeover and the next financial meltdown*, Pantheon Publishing

Kasperek, Annika (2011) 'Fiscal policy in US states', *Current Issues* December 2011, Deutsche Bank Research

Kirkegaard, Jacob Funk (2011) 'What the European Union did and did not accomplish', *RealTime Economic Issues Watch*, Peterson Institute for International Economics, retrieved on 6 January 2012 from www.piiie.com/realtime/?p=2575

Krugman, P. (2008) 'Fifty Herbert Hoovers', *New York Times*, 29 December

Mahdavi, S. and J. Westerlund (2011) 'Fiscal stringency and fiscal sustainability: Panel evidence from the American state and local gov-

- ernments', *Journal of Policy Modeling* 33 (6), 953-969
- Marzinotto, Benedicta, Andre Sapir and Guntam B. Wolff (2011) 'What kind of fiscal union?' *Policy Brief* 2011/06, Bruegel
- McGrane, C. R. (1935) *Foreign bondholders and American state debts*, Washington DC: Beard Books
- McNamara, Kathleen R. (2002) 'State building, the territorialization of money, and the creation of the American single currency', in David M. Andrews, C. Randall Henning and Louis W. Pauly (eds) *Governing the world's money*, Cornell University Press
- McNichol, E., P. Oliff and N. Johnson (2011) *States continue to feel recession's impact*, Center on Budget and Policy Priorities, Washington DC
- Monkonen, Eric K. (1995) *The local state: public money and American cities*, Stanford University Press
- National Conference of State Legislatures (2010) 'State balanced budget provisions', *NSCL Fiscal Brief*, October, available at www.ncsl.org
- Oates, Wallace E. (1999) 'An essay on fiscal federalism', *Journal of Economic Literature* vol. 37(3), pp. 1120-1149
- Orth, John V. (1987) *The judicial power of the United States: the eleventh amendment in American history*, Oxford University Press
- O'Rourke, Kevin (2011) 'A summit to the death', *Project Syndicate*, 9 December
- Perkins, Edwin J. (1994) *American public finance and financial services, 1700-1815*, Ohio State University Press
- Posen, Adam S. and Nicolas Véron (2009) 'A solution for Europe's banking problem', *Policy Brief* 2009-03, Bruegel
- Poterba, James M. (1994) 'State responses to fiscal crises: the effects of budgetary institutions and politics', *Journal of Political Economy* vol. 102(4), pages 799-821

- Poterba, James M. (1996) 'Do budget rules work?' *Working Paper* 5550, National Bureau of Economic Research
- Poterba, J. M. and K. S. Rueben (2001) 'Fiscal news, state budget rules, and tax-exempt bond yields', *Journal of Urban Economics* 50, 537-562
- Poterba, J. M. and K. S. Rueben (1999) 'State fiscal institutions and the US municipal bond market', in J. M. Poterba (ed) *Fiscal institutions and fiscal performance*, University of Chicago Press
- Randolph, B. C. (1931) 'Foreign bondholders and the repudiated debts of the Southern states', *American Journal of International Law* XXV: 74
- Ratchford, B. U. (1941) *American states debts*, Duke University Press
- Roberts, A. (2010) "An ungovernable anarchy": the United States' response to depression and default, 1837-1848', *Intereconomics* 45 (4), pp. 196-202
- Rodden, J. (2006) *Hamilton's paradox: the promise and peril of fiscal federalism*, Cambridge University Press
- Rose, S. (2010) 'Institutions and fiscal sustainability', *National Tax Journal* 63 (4): 807-838
- Savage, J. (1988) *Balanced budgets and American politics*, Cornell University Press
- Sbragia, A. (2008) 'American federalism and intergovernmental relations', in R. A. W. Rhodes, Sarah A. Binder and Bert A. Rockman (eds) *The Oxford Handbook of Political Institutions*, Oxford University Press.
- Sorensen, Bent E., Lisa Wu and Oved Yosha (2001) 'Output fluctuations and fiscal policy: US state and local governments 1978-1994', *European Economic Review* 45(7): 1271-1310
- Sorensen, Bent E. and Oved Yosha (2001) 'Is state fiscal policy asymmetric over the business cycle?' *Economic Review*, Federal Reserve Bank of Kansas City, issue Q III: 43-64
- Sylla, R. (2011) 'Financial foundations: public credit, the national

- bank, and securities markets', in D. A. Irwin and R. Sylla (eds) *Founding choices: American economic policy in the 1790s*, University of Chicago Press
- Sylla, R. and J. W. Wilson (1999) 'Sinking funds as credible commitments: Two centuries of US national-debt experience', *Japan and the World Economy* 11, no. 2 (April): 199-222
- Véron, N. (2011) 'Testimony on the European debt and financial crisis', *Policy Contribution* 2011/11, Bruegel
- Von Hagen, Jürgen and Charles Wyplosz (2008) 'EMU's decentralized system of fiscal policy', *European Economy Economic Papers* 306, European Commission, Directorate-General for Economic and Financial Affairs
- Wallis, J. J. (2005) 'Constitutions, corporations, and corruption: American states and constitutional change, 1842-1852', *Journal of Economic History* 65, no. 1: 211-256
- Wallis, J. J., R. Sylla and A. Grinath (2004) 'Sovereign debt and repudiation: the emerging-market debt crisis in the US states, 1839-1843', *Working Paper* 10753, National Bureau of Economic Research.
- Wallis, John Joseph and Barry R. Weingast (2008) 'Dysfunctional or optimal institutions: state debt limitations, the structure of state and local governments, and the finance of American infrastructure', in Elizabeth Garrett, Elizabeth Graddy, and Howell Jackson (eds) *Fiscal challenges: an interdisciplinary approach to budget policy*, Cambridge University Press
- Wibbels, Eric (2003) 'Bailouts, budget constraints, and leviathans: comparative federalism and lessons from the early United States', *Comparative Political Studies* 36: 475-508
- Woodward, Comer V. (1971) *Origins of the New South, 1877-1913*, Louisiana State University Press

NOTES

- 1 This essay is adapted from a presentation by Randall Henning at the European Central Bank workshop on *Euro-area governance*, Frankfurt, 29-30 September 2011. The authors wish to thank Jean Pisani-Ferry for his encouragement of this project and Guntram B. Wolff and Nicolas Véron for comments on a previous draft. The final text is the responsibility of the authors alone. This essay is published simultaneously by the Peterson Institute.
- 2 Ron Chernow (2004) has authored the definitive biography of Hamilton, containing his intellectual development in finance, the implementation of his ideas and his place in the formative period in American history.
- 3 Sylla (2011) places the development of Hamilton's ideas in the context of the financial challenges at the time. The most commonly cited versions of Hamilton's papers were collected and published by Harold C. Syrett and Jacob E. Cooke in 27 volumes, cited here as 'Hamilton', by volume.
- 4 Perkins (1994), chapter 9, as cited by Sylla (2011, p19).
- 5 Sylla (2011) judges Hamilton to be the winner in the overall debate with Gallatin.
- 6 Recall that Hamilton authored roughly three-fifths of the *Federalist Papers* and that James Madison, who later split with him over the financial plan, and John Jay were his partners in that ratification campaign.
- 7 Observers who might be dismayed by the present state of US politics, on fiscal matters in particular, can take some measure of consolation from the fact that the American union survived considerably worse divisions over finance in its early years.
- 8 For a lively contemporary treatment, see Johnson and Kwak (2010, pp14-22).
- 9 Rodden (2006, pp55-64) contains a nice treatment of this period. See, as well, Wallis, Sylla and Grinath (2004), English (1996), and McGrane (1935), a seminal contribution that is rich in historical detail.
- 10 See, especially, Woodward (1971), pp51-106.
- 11 Ratchford (1941, p196), quoting B. C. Randolph (1931) 'Foreign bondholders and the repudiated debts of the Southern states', *American Journal of International Law* XXV, 74. Ratchford's book is one of the classic histories on the debt of the states, covering the colonial period up to the Great Depression.
- 12 Monkkonen (1995) examines the treatment of local debt and provides a detailed account of the 1870 revision of the Illinois state constitution.
- 13 A word about nomenclature is in order. In casual usage in the United States, the word 'municipality' usually refers to a city or town. In formal usage in finance, however, 'municipal debt' is a broader category that includes the

- debt of states as well as local entities. 'Local', as distinct from 'state', refers to counties, cities and school and special utility districts.
- 14 'Bankruptcy rarely offers easy answer for counties' (2011) *New York Times*, 10 November.
 - 15 Article I, Section 8, gives Congress exclusive legislative jurisdiction over the District of Columbia. The formal name of the control board was the 'District of Columbia Financial Responsibility and Management Assistance Authority'.
 - 16 The Illinois state comptroller is blunt. When speaking about the budget, she says, "It isn't balanced. It's never balanced. There's always ways to have things off budget . . ." *Financial Times*, 4 November 2011.
 - 17 Wallis (2005) notes that these reforms coincide with changes in the law of incorporation and tax rules.
 - 18 The legal literature on state default and fiscal federalism includes Amdursky and Gillette (1992) and Orth (1987). For an up-to-date review, see Gelpern (2012).
 - 19 The meaning of 'golden' as a modifier for 'rule' has shifted in this discourse over time. As we use the term, a golden rule requires that current expenditures and tax receipts be balanced but allows borrowing for long-term public investment. Despite the prevalence of the golden rule at the state level, proposals to differentiate between operating and capital budgets at the federal level have not gained traction in the United States.
 - 20 There is substantial variation among states (Hines, 2010) and, as discussed, the impact of balanced budget rules is contingent on other factors.
 - 21 Oates (1999) argues that macroeconomic stabilisation is best allocated to the central government.
 - 22 Hines (2010) finds that the main source of discretionary spending is the grants-in-aid from the federal government.
 - 23 Standard & Poor's Municipal Bond Indices at <http://www.standardandpoors.com/indices/sp-municipal-bond-maturity-based-indices/en/us/?indexId=spfi--ish-usd---t----->, Securities Industry and Financial Markets Association (SIFMA) and Thomson Reuters. Over the last four years, only California, Connecticut, Illinois and Arizona have financed current expenditure in significant amounts.
 - 24 Kasperek (2011) provides an excellent up-to-date review.
 - 25 Arezki, Candelon and Sy (2011). Exceptions pertain to the largest state issuers.
 - 26 These ratings have not, at least at time of writing, been affected by S&P's downgrade of US securities to AA+ in August 2011.
 - 27 Kasperek (2011) argues that risks within the US municipal market have been exaggerated.
 - 28 European Council, 'Statement by the euro-area heads of state or govern-

- ment', Brussels, 9 December 2011, available at http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/126658.pdf. For brief evaluation of this agreement, see, O'Rourke (2011) and Kirkegaard (2011), among others. For a blueprint for fiscal union, see Marzinotto, Sapir and Wolff (2011).
- 29 A large number of studies address the effectiveness of fiscal rules in Europe. We will not review that literature here, except to note, by way of example, that Iara and Wolff (2011) show that such rules have a significant impact and might have reduced Greece's interest rates by 100 basis points at the outset of that country's crisis. Hallerberg and Wolff (2008) show that the balance of authority between finance ministers and spending ministers within governments has a significant impact on bond yields. See also, von Hagen and Wyplosz (2008) and Hallerberg (2011).
- 30 Orth (1987) is a classic study of the constitutional law in these respects.
- 31 For good reviews of the political science literature on fiscal federalism, see Sbragia (2008) and Galligan (2008). For comparisons between the US and the EU, see Hallerberg (2006) and Hallerberg, Strauch and von Hagen (2009). On the relevance of the nineteenth century US for the monetary union, see McNamara (2002).
- 32 Federal debt is not a joint and several liability of the states *per se*, as presently discussed with respect to the proposal to issue 'euro bonds', but a common liability of the federal system. The power to tax is critical to the federal government's ability to issue Treasury securities.
- 33 Dolls, Fuest and Peichl (2010) and Baunsgaard and Symansky (2009).

ABOUT BRUEGEL

Bruegel is a European think tank devoted to international economics. It started operations in Brussels in 2005 as a Belgian non-profit international organisation supported by European governments and leading corporations. Bruegel seeks to contribute to the quality of economic policymaking in Europe through open, facts-based and policy-relevant research, analysis and discussion.

Bruegel issues a range of publications. Bruegel Policy Briefs provide concise, topical analysis targeted at an audience of executives and policy decision-makers, with an emphasis on concrete policy orientation. Bruegel Policy Contributions are responses to requests by policymakers or public bodies, including testimonies at hearings or responses to public consultation. Bruegel and its researchers also publish working papers, op-eds, collaborative papers with other organisations, and essays. The Bruegel Essay and Lecture Series provides a compact record of the views of scholars and practitioners as they write for, or lecture at, Bruegel events. *Fiscal federalism: US history for architects of Europe's fiscal union* is the sixth publication in the Bruegel Essay and Lecture Series.

Bruegel's research is independent and does not represent the views of its board or members. For a full picture of Bruegel activities and publications, visit the website at www.bruegel.org.

PREVIOUS VOLUMES IN THE BRUEGEL ESSAY AND LECTURE SERIES

Why Germany fell out of love with Europe, by Wolfgang Proissl

ISBN: 978-9-078910-16-9, July 2010

Politics and trade: lessons from past globalisations, by Kevin O'Rourke

ISBN: 978-9-078910-11-4, January 2009

Testing times for global financial governance, by Ignazio Angeloni

ISBN: 978-9-078910-08-4, October 2008

10 lessons about budget consolidation, by Jens Henriksson

ISBN: 978-9-078910-03-9, July 2007

Will global capitalism fall again? by Jeffrey Frieden

June 2006

For details of these and other Bruegel publications, visit www.bruegel.org

Chap 4: The models of currency crises

Key issue: the economic analysis distinguishes three different types of currency crises, presented in three generations of models

- P. Krugman, "Introduction" in *Currency crises*, University of Chicago Press (2000)
- Goldstein, Itay, and Assaf Razin. *Review of theories of financial crises*. National Bureau of Economic Research, 2013. Introduction and Section 4.
- S. Saxena. "The Changing Nature of Currency Crises." *Journal of Economic Surveys*, Volume 18, Issue 3, Pages: 321-350 (2004).
- Koo, Richard. "Balance sheet recession is the reason for secular stagnation," *Secular Stagnation: Facts, Causes and Cures*, CEPR Press.
- "The sovereign-debt crisis has echoes of the ERM debacle," *The Economist*. July 14, 2011.
- Irving Fisher. "Out of Keynes's shadow," *The Economist*. Feb 12, 2009.
- "Bleeding out," *The Economist*. June 9th 2010.

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Currency Crises

Volume Author/Editor: Paul Krugman, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-45462-2

Volume URL: <http://www.nber.org/books/krug00-1>

Publication Date: January 2000

Chapter Title: Introduction to "Currency Crises"

Chapter Author: Paul R. Krugman

Chapter URL: <http://www.nber.org/chapters/c8687>

Chapter pages in book: (p. 1 - 6)

Introduction

Paul Krugman

There is no generally accepted formal definition of a currency crisis, but we know them when we see them. The key element is a sort of circular logic, in which investors flee a currency because they fear that it might be devalued, and in which much (though not necessarily all) of the pressure for such a devaluation comes precisely from that capital flight. Such crises have been a recurrent feature of the international economy ever since gold and silver coins were replaced by paper; currency crises played a large role in the economic turmoil of the interwar era, in the breakup of Bretton Woods, and in the early stages of the Latin American debt crisis of the 1980s. And since the late 1970s currency crises have also been a major subject of academic study.

But nothing in the past history of the subject prepared economists for the 1990s. Future historians may, in fact, dub this the Age of Currency Crises: never before, not even in the interwar period, have currency crises played such a central role in world affairs. The massive attacks that roiled the European Monetary System in 1992–93 were a high-water mark for currency speculation; but that mark was soon surpassed by the “tequila” crisis of 1995; and *that* mark surpassed by the still-unfolding Asian currency crisis of 1997–98, which at the time of writing seems to be spreading back to Latin America. Currency crises—both crises that actually do happen and the sometimes desperate efforts of national governments and international agencies to head them off before they start—have become a defining force for economic policy in much of the world.

Paul Krugman is professor of economics at Princeton University and a research associate of the National Bureau of Economic Research.

It might seem that the decision to hold a special conference on currency crises at the National Bureau of Economic Research was a natural response to this vastly heightened importance of the subject. In fact, however, this is only partly true. Martin Feldstein approached me with the idea for such a conference in 1996, when the European and Latin crises had subsided and the idea that something even worse could happen to emerging Asia never even entered people's minds. By the time of the conference in February 1998, of course, what had been conceived as a low-key academic research meeting had acquired enormous salience.

Aside from topicality, however, why was another conference on a well-established research field needed? The answer is that despite two decades of research on the causes and consequences of currency crises, important issues are either unresolved or require a fresh look in the face of new experience. Here is a review of the three main issues that the conference revisited.

What Drives Crises?

Despite the fact that everyone agrees in a broad sense about the nature of currency crises—about the circular logic in which expectations of devaluation make devaluation more likely—the details have long been disputed. In general, this dispute divides researchers along two fault lines.

First, through what channel does speculation against a currency make the devaluation of that currency more likely? In the early currency crisis models, the channel was assumed to be essentially a mechanical linkage: speculation led to a depletion of foreign exchange reserves, which would then force the central bank to give up its defense of the original parity. One can still defend this assumption as a useful first pass at the problem—and in some cases even a reasonably good description of events. However, from the early years of currency crisis research many economists have argued that in the real world governments have much more freedom of action than the stylized models recognized, and that as a result the logic of such crises was more subtle and less mechanical than the existing theory.

Until recently the most influential alternative to the classic foreign-reserve-driven crisis model was what Eichengreen, Rose, and Wyplosz (1996) have dubbed “second generation” crisis models, which emphasize not the mechanical exhaustion of foreign exchange but the problems of macroeconomic policy. Loosely, a second-generation model imagines a government that is physically able to defend a fixed exchange rate indefinitely, say, by raising interest rates, but that may decide the cost of defense is greater than the cost in terms of credibility or political fallout from abandoning the defense and letting the currency float. In this case a currency crisis can develop because doubts about the government's willingness to defend the parity force it to raise interest rates, and the need to

keep interest rates high in turn raises the cost of defending the parity to a level the government finds unacceptable. The second-generation model came into its own in the European crises of 1992–93 and remains an important piece of the research agenda. In chapter 1 Barry Eichengreen and Olivier Jeanne argue that a version of second-generation theory does well at explaining not only recent European experience but also one of the important currency crises of the interwar period: Britain's departure from the gold standard in 1931. Allan Drazen's analysis of contagion, in chapter 2, is also predicated on second-generation logic, in which abandonment of the fixed parity brings macroeconomic gains but political costs, costs that are less if other countries have also faltered.

However, the emerging market crises of 1995 and 1997 do not seem to fit either first- or second-generation logic. After all, according to second-generation models devaluing or floating a currency gives the government freedom to follow more expansionary policies; yet in both Latin America and Asia currency crises were followed by severe recessions. Most researchers have therefore concluded that a different type of model—third generation?—is needed, with most work focusing on the issues involving financial intermediaries and liquidity effects. Guillermo Calvo's chapter 3 offers a rich menu of possible third-generation crisis models; Steven Radelet and Jeffrey Sachs offer an insightful comparison of the “real economy” implications of alternative crisis stories in chapter 4 and argue for what may be becoming the canonical story about emerging market crises, the idea that such crises are essentially bank runs that manifest themselves through the foreign exchange market.

While the mechanism of crisis has been one source of dispute, the determinants of whether and when a crisis occurs has been another. Early crisis models were “fundamentalist”: that is, crises happened to countries whose fixed exchange rates were unsustainable in the long run, and happened when underlying economic conditions—such as the level of foreign exchange reserves—deteriorated past some critical point. However, an alternative tradition—particularly associated with Maurice Obstfeld (1994), who unfortunately could not attend the conference—argues that crises involve a strong element of self-fulfilling prophecy, that exchange regimes that could have survived indefinitely can nonetheless collapse if subjected to an essentially random speculative attack. Radelet and Sachs argue this position strongly for emerging Asia: drawing on both economic evidence and the apparent lack of early warning signs from the financial markets, their view is that this was a gratuitous crisis, that nothing in the Asian situation warranted such a sudden reversal of fortune.

It may be worth noting that in the months following the conference, two broad strands of “third generation” modeling did emerge. One strand—building on an idea largely due to Ronald McKinnon (see in particular McKinnon and Pill 1996) but applied to the Asian crisis by a number of

authors—focused on the role of implicit loan guarantees in generating excessive risky investment. Currency crises are then interpreted as a crisis of this moral hazard regime; it is the collapse of the investment that precipitates the macroeconomic setback.

The other strand elaborates on the “bank run” story suggested by Radelet and Sachs, emphasizing self-fulfilling collapse via either literal bank runs—a view promulgated by Chang and Velasco (1998a, 1998b)—or some kind of balance-sheet-driven financial contraction.

While there are dissenters, it seems fair to say that academic opinion has swung fairly strongly toward the self-fulfilling crisis view, largely because of the phenomenon of “contagion”: the way that a crisis in one country seems able to trigger a crisis in another, even when the economic links appear to be minor. Most economists have concluded that this can happen only if believing makes it so—that is, if the nervousness created by a crisis in one country can set in motion a self-fulfilling run on the currency of another. However, the Drazen chapter points out that since in many crisis models the cost of abandoning a peg is essentially political—presumably because of the loss of credibility when a government reneges on a pledge to maintain the exchange rate—the lack of strong economic linkages may be irrelevant. If governments find safety in numbers, if devaluing when other countries have already devalued is less costly to one’s reputation than acting alone, contagion can be consistent with a fundamentalist story about the timing of crises.

Clearly these controversies have not been resolved; but clearly also there has been a deepening of our understanding of the issues, and considerable movement in the views of the main protagonists.

How Should We Model Governments?

One of the main issues in currency crisis modeling—closely tied to the distinction between first-, second-, and third-generation models—is the question of how to think about government behavior. The early models assumed a very passive government, which stolidly doled out foreign exchange until the last dollar was gone; later models have tended to assume a much more sophisticated, activist policy. It also turns out that analyzing crises requires that one specify not only what the government will do *during* the crisis but what policies it will follow if its defense of the exchange rate fails.

Most of the papers in the volume give at least some consideration to these issues. Two focus specifically on government policies. In chapter 5 Robert Flood and Peter Garber take on a currency regime that was at the time of writing still prospective—the “Stage III” regime in Europe, in which the euro exists as a unit of account but not yet as an actual circulating currency. It is widely assumed that the financial arrangements within

Stage III will ensure the invulnerability of that system to speculative attack. Flood and Garber point out, however, that at least in principle those arrangements imply a willingness of national central banks to extend each other completely open-ended credit lines; if one has doubts about whether they will actually do so, one also has doubts about whether the system is really crisis-proof. One need not agree that a Stage III crisis is at all likely to agree that this kind of careful attention to the implications of monetary arrangements can be crucial.

Sebastian Edwards and Miguel Savastano consider the policies followed by the Bank of Mexico after the 1994 crisis—a period during which the peso suffered a megadevaluation, far greater than most analysts had expected, and in which interest rates rose to unexpected heights. In chapter 6 Edwards and Savastano show that the textbook assumption that a central bank pegs until its reserves are gone, then lets the currency float freely, is far from the reality: in actuality the peso was the subject of considerable attempts at short-term management even in the postcrisis float.

What Are the Effects of Crisis?

If there has been one area in which views of currency crises have shifted most since the origins of the academic literature two decades ago, it is probably the consequences of such crises for the real economy. In the early models crises were thought of as monetary events with few real consequences. Second-generation models, which emphasized the macroeconomic payoff to devaluation, suggested that real economic performance should if anything improve following a successful speculative attack—and the folk wisdom has been that this was in fact the case in Europe following the 1992 crisis. More recently, the experience of emerging economies has suggested to many that currency crises, by forcing these countries to move suddenly from current account deficit to surplus, cause severe economic downturns.

On closer examination, however, these generalizations are far less clear. Robert Gordon revisits the postcrisis European landscape in chapter 7 and finds that the conventional view that it was better to fail than to succeed at currency defense is heavily colored by just one comparison—Britain versus France—and that the overall picture is much less clear. (And the subsequent rise of the pound suggests that structural factors may have played a bigger role, and monetary ones a smaller one, in British success than widely believed.) Gian Maria Milesi-Ferretti and Assaf Razin focus on emerging economies in chapter 8, examining a cross section of many crisis episodes. Their surprise conclusion is that currency crises and sudden current account reversals are not the same thing—and that while currency crashes are normally associated with sharp declines in output, current account reversals are not. At the very least this suggests that we need

to rethink the channels through which the adverse effect of currency crises takes place.

A Field in Transition

The study of currency crises is a field in flux, largely because the world itself keeps on throwing up new crises for us to examine at a rate that would be gratifying to scholars if it were not so terrifying to policymakers. The papers collected in this volume are very far from the last word. But they represent the latest thinking, captured at a moment in which some of the best minds in economics were focused on the theory and practice of speculative attack, and will surely serve as the basis for much more work to come.

References

- Chang, R., and A. Velasco. 1998a. Financial fragility and the exchange rate regime. New York: New York University, C. V. Starr School of Business. Working paper.
- . 1998b. Financial crises in emerging markets: A canonical model. New York: New York University, C. V. Starr School of Business. Working paper.
- Eichengreen, B., A. Rose, and C. Wyplosz. 1996. Speculative attacks on pegged exchange rates: An empirical exploration with special reference to the European Monetary System. In *The new transatlantic economy*, ed. M. Canzoneri, W. Ethier, and V. Grilli. New York: Cambridge University Press.
- McKinnon, R., and H. Pill. 1996. Credible liberalizations and international capital flows: The overborrowing syndrome. In *Financial deregulation and integration in East Asia*, ed. T. Ito and A. O. Krueger. Chicago: University of Chicago Press.
- Obstfeld, M. 1994. The logic of currency crises. *Cahiers Economiques et Monétaires* (Bank of France) 43:189–213.

NBER WORKING PAPER SERIES

REVIEW OF THEORIES OF FINANCIAL CRISES

Itay Goldstein
Assaf Razin

Working Paper 18670
<http://www.nber.org/papers/w18670>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2013

The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Itay Goldstein and Assaf Razin. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Review of Theories of Financial Crises
Itay Goldstein and Assaf Razin
NBER Working Paper No. 18670
January 2013
JEL No. E61,F3,F33,G01,G1

ABSTRACT

In this paper, we review three branches of theoretical literature on financial crises. The first one deals with banking crises originating from coordination failures among bank creditors. The second one deals with frictions in credit and interbank markets due to problems of moral hazard and adverse selection. The third one deals with currency crises. We discuss the evolutions of these branches of the literature and how they have been integrated recently to explain the turmoil in the world economy. We discuss the relation of the models to the empirical evidence and their ability to guide policies to avoid or mitigate future crises.

Itay Goldstein
Wharton School
University of Pennsylvania
Philadelphia, PA 19104
itayg@wharton.upenn.edu

Assaf Razin
Department of Economics
Cornell University
Uris 422
Ithaca, NY 14853
and Cornell University
and also NBER
ar256@cornell.edu

1. Introduction

Financial and monetary systems are designed to improve the efficiency of real activity and resource allocation. A large empirical literature in financial economics provides evidence connecting financial development to economic growth and efficiency; see, for example, Levine (1997) and Rajan and Zingales (1998). In theory, financial institutions and markets enable the efficient transmission of resources from savers to the best investment opportunities. In addition, they also provide risk sharing possibilities, so that investors can take more risk and advance the economy. Finally, they enable aggregation of information that provides guidance for more efficient investment decisions. Relatedly, monetary arrangements, such as the European Monetary Union (EMU) and many others in the past, are created to facilitate free trade and financial transactions among countries, thereby improving real efficiency.

A financial crisis – marked by the failure of banks, and/or the sharp decrease in credit and trade, and/or the collapse of an exchange rate regime, etc. – generates extreme disruption of these normal functions of financial and monetary systems, thereby hurting the efficiency of the economy. Unfortunately, financial crises have happened frequently throughout history and, despite constant attempts to eliminate them, it seems unlikely that they will not repeat in the future. Clearly, the last few years have been characterized by great turmoil in the world's financial systems, which even today, more than five years after its onset, does not seem to have a clear solution. Between the meltdown of leading financial institutions in the US and Europe, the sharp decrease in lending and trading activities, and the ongoing challenge to the European Monetary Union, these events

exhibit ingredients from several types of financial crises in recent history: banking crises, credit and market freezes, and currency crises.³

Over the years, many theories have been developed to explain financial crises and guide policymakers in trying to prevent and mitigate them. Three literatures have been developed, more or less in parallel, highlighting the analytical underpinnings of three types of crises: banking crises and panics, credit frictions and market freezes, and currency crises. At a later stage, mainly following the East Asian crisis in the late 1990s, these literatures have become more integrated, as the events in the real world proved that the different types of crises can occur together and amplify each other in different ways.

In this survey, we provide a review of the basic theories of financial crises of the three types described above and the way they interact with each other. Importantly, this is not meant to be a comprehensive survey of the financial-crises literature. The literature is too big to be meaningfully covered in full in one survey. Instead, we attempt to present the basic frameworks and describe some of the directions in which they influenced the literature and the way they relate to recent events. We also address some of the policy challenges and shed light on them using the analytical tools at hand. We hope that this survey will be helpful in highlighting the basic underlying forces that have been studied in the literature for over three decades in a simple and transparent way, and will be an easy and accessible source to the many economists who are now interested in exploring the topic of financial crises following the events of the last few years.

In Section 2, we review the literature on banking crises and panics. Banks are known to finance long-term assets with short-term deposits. The advantage of this

³ Many authors provide detailed descriptions of the events of the last few years. For example, see Brunnermeier (2009) and Gorton (2010).

arrangement is that it enables banks to provide risk sharing to investors who might face early liquidity needs. However, this also exposes the bank to the risk of a bank run, whereby many creditors decide to withdraw their money early. The key problem is that of a coordination failure, which stands at the root of the fragility of banking systems: When more depositors withdraw their money from a bank, the bank is more likely to fail, and so other depositors have a stronger incentive to withdraw. In this section, we describe the theoretical underpinnings behind bank runs and the lessons to policy analysis.

Banking systems have been plagued with bank runs throughout history; see, e.g., Calomiris and Gorton (1991). Policy lessons adopted in the early 20th century led governments to insure banks, which substantially reduced the likelihood of such events. However, runs are still a prominent phenomenon behind financial crises. Many runs happened in East Asian and Latin American countries even in the last two decades. In the recent turmoil, a classic ‘text-book’ type of bank run was seen in the UK for Northern Rock Bank (see Shin (2009)), where investors were lining up in the street to withdraw money from their accounts. Beyond that, there were many other examples of runs in the financial system as a whole. The repo market, where investment banks get short-term financing, was subject to a run (Gorton and Metrick (2012)) when financing has all of a sudden dried up. This led to the failure of leading financial institutions, such as Bear Stearns and Lehman Brothers. This credit squeeze was to a large extent a coordination failure among providers of capital in this market, who refused to roll over credit, expecting a failure of the borrower due to the refusal of other lenders to roll over credit. This is similar to the models of bank runs due to coordination problems that we review. Runs also occurred on money-market funds and in the asset-backed-commercial-paper

market (see for example, Schroth, Suarez, and Taylor (2012)), which were very prominent events in the recent crisis.

While Section 2 emphasizes fragility faced by financial institutions due to coordination failures by their creditors, in Section 3 we review models that analyze frictions in loans extended by financial institutions and other lenders. Broadly speaking, these are models of credit frictions and market freezes. This literature highlights two key problems that create frictions in the flow of credit from lenders to borrowers. When these frictions strengthen, a financial crisis ensues that can even lead to a complete freeze. One problem is that of moral hazard. If a borrower has the ability to divert resources at the expense of the creditor, then creditors will be reluctant to lend to borrowers. Hence, for credit to flow efficiently from the creditor to the borrower, it is crucial that the borrower maintains ‘skin in the game’, i.e., that he has enough at stake in the success of the project, and so does not have a strong incentive to divert resources. This creates a limit on credit, and it can be amplified when economic conditions worsen, leading to a crisis. Another problem is that of adverse selection. In the presence of asymmetric information between lenders and borrowers or between buyers and sellers, credit and trade flows might freeze. Again, this may lead to a crisis if asymmetric information is very extreme.

Such forces were clearly on display in recent years. The credit freeze following the financial meltdown of 2008, and the credit flow freeze in the interbank markets are both manifestations of the amplification of economic shocks due to the frictions in credit provision, brought by the principal-agent models that we review here. As economic conditions deteriorated, borrowers found themselves with less ‘skin in the game’, and so lenders refused to provide credit to them, since doing so would lead borrowers to divert

cash, and take excessive risk. This, in turn, worsened the economic conditions of borrowers, amplifying the initial shock. Similarly, the sharp increase in asymmetric information, following the collapse of Lehman Brothers in 2008, contributed to a total market freeze, where investors were reluctant to trade in assets with each other, due to the heightened uncertainty about the value of assets they trade.

Overall, the models of Sections 2 and 3 highlight fragility on the different sides of the balance sheet of a financial institution. Clearly, both types of fragility have been at work in recent crises, as we mention above. Importantly, such fragilities can reinforce each other, as we point out in Section 3. That is, creditors of a financial institution are more likely to panic and run when problems of moral hazard and asymmetric information reduce the value of its assets or make it more uncertain.

An important aspect of financial crises is the involvement of the government and the potential collapse of arrangements it creates, such as an exchange rate regime. In Section 4, we review models of this kind, focusing on currency crises. Many currency crises, e.g., the early 1970s breakdown of the Bretton Woods global system, originate from the desire of governments to maintain a fixed exchange rate regime which is inconsistent with other policy goals such as free capital flows and flexible monetary policy. This might lead to the sudden collapse of the regime. Like in the bank-run literature, coordination failures play an important role here too. When the central bank tries to maintain a fixed exchange rate regime, it might decide to abandon it under pressure from speculators. Then, speculators again find themselves in a coordination problem, where they attack the regime if and only if they believe others will do so. In such coordination failures, the event of a currency crisis becomes a self-fulfilling belief. This is also similar to debt

crises, where the government may decide to default under pressure from creditors. Then, creditors are facing a coordination problem, where they liquidate their bond holdings if and only if they expect that others will liquidate their claims. Consequently a debt crisis becomes a self-fulfilling expectation.

Such models are highly relevant to the current situation in the European Monetary Union. In the basis of the theory of currency crises is the famous international-finance trilemma, according to which a country can choose only two of three policy goals: free international capital flows (benefitting international risk sharing), monetary autonomy (the ability to employ monetary policy tools to stabilize inflation and output fluctuations), and the stability of the exchange rate (bringing about a reduction in transaction costs associated with trade and investment). Countries in the Euro zone now realize that in their attempt to achieve the first and third goal, they have given up on the second goal, and so have limited ability to absorb the shocks in economic activity and maintain their national debts, triggered by the global financial crisis. Coordination problems among investors and currency speculators aggravate this situation, and may have an important effect on whether individual countries in Europe are forced to default and/or leave the monetary union.

While the traditional literature on currency crises focused on the government alone, in Section 4.3 we review the ‘third-generation’ models of currency crises, which essentially connect models of banking crises and credit frictions (reviewed in Sections 2 and 3, respectively) with traditional models of currency crises (reviewed in Subsections 4.1 and 4.2). Such models were motivated by the East Asian Crises of the late 1990s, where financial institutions and exchange rate regimes collapsed together, demonstrating

the linkages between governments and financial institutions that can expose the system to further fragility. This is again relevant for the current situation in Europe, as banks and governments are intertwined, and the fragility of the system depends to a large extent on the connections between them.

2. *Banking Crises and Panics*

Depository institutions are inherently unstable because they have a mismatch in the maturity structure between their assets and liabilities. In particular, they finance long-term investments with short-term deposits. This exposes banks to a risk of bank runs: when many depositors demand their money in the short term, banks will have to liquidate long-term investments at a loss, leading to their failure. This can lead to a self-fulfilling belief, whereby the mere belief that a bank run will occur causes a bank run, as depositors are better off withdrawing their money if they expect others to do so.

Diamond and Dybvig (1983)⁴ provide a framework for coherent analysis of this phenomenon. In their model, agents may suffer idiosyncratic short-term liquidity needs. If they operate in autarky, consuming the returns on their own endowments, they would not be able to enjoy the fruits of long-term investments. By offering demand-deposit contracts, banks enable short-term consumers to enjoy those fruits. Banks rely on the fact that only a forecastable fraction of agents will need to consume early, and thus offer a contract that transfers consumption from the long-term consumers to the short-term consumers. This contract improves welfare as long as agents demand early withdrawal only if they genuinely need to consume in the short term. Banks thereby enable risk sharing among agents who *ex ante* do not know whether they will have early liquidity

⁴ Another important paper on the topic from that period is Bryant (1980).

uninformed traders there, since then they are exposed to a lesser adverse selection problem. As a result, there is a coordination problem that can lead to sharp changes in market depth, resembling what we see in a financial crisis. Recently, Morris and Shin (2012) show that the amplification becomes even more severe when traders have different information about the extent of the adverse selection problem, i.e., about how many informed traders are present. This leads to a contagious process, by which very small changes can lead to a market freeze.

4. Currency Crises

Historically, financial crises have often been marked with large disturbances in currency markets, which have spilled over to the financial sectors and the real economies of affected countries in various ways. In the events of the recent years, the deepening of the crisis in Europe is strongly associated with the attempt to maintain the common currency area, which also has implications for the financial sectors and real economies of countries in the Euro zone. In this section, we review the development of the theoretical literature on currency crises, and its connection to the literatures on banking panics and crises and credit frictions, which were reviewed in the previous sections.

Currency crises originate from the attempt of governments to maintain certain financial and monetary arrangements, most notably a fixed-exchange rate regime. Their goal is to stabilize the economy. At times, these arrangements become unstable, which leads to a speculative attack on a fixed exchange rate regime and from there to a financial crisis.

The best way to understand the origins of currency crises is to think about the basic trilemma in international finance. A trilemma, as Mankiw (2010) recently wrote in the context of the 2010 Euro crisis, is a situation in which someone faces a choice among three options, each of which comes with some inevitable problems. In international finance, it stems from the fact that, in most nations, economic policy makers would like to achieve the following goals. First, make the country's economy open to international capital flows, because by doing so they let investors diversify their portfolios overseas and achieve risk sharing. They also benefit from the expertise brought to the country by foreign investors. Second, use monetary policy as a tool to help stabilize inflation, output, and the financial sector in the economy. This is achieved as the central bank can increase the money supply and reduce interest rates when the economy is depressed, and reduce money growth and raise interest rates when it is overheated. Moreover, it can serve as a lender of last resort in case of financial panic. Third, maintain stability in the exchange rate. This is because a volatile exchange rate, at times driven by speculation, can be a source of broader financial volatility, and makes it harder for households and businesses to trade in the world economy and for investors to plan for the future.

The problem, however, is that a country can only achieve two of these three policy goals. In order to maintain a fixed exchange rate and capital mobility, the central bank loses its ability to control its policy instruments: the interest rate, or equivalently the monetary base. Because, under free capital mobility, the interest rate becomes anchored to the world interest rate, by the interest rate parity, and the monetary base is automatically adjusted to the pre-determined money demand. This is the case of individual members of the European Monetary Union. In order to keep control over the

interest rate or equivalently the money supply, the central bank has to let the exchange rate float freely, as in the case of the US. If the central bank wishes to maintain both exchange rate stability and control over the monetary policy, the only way to do it is by imposing capital controls, as in the case of China.

Currency crises occur when the country is trying to maintain a fixed exchange rate regime with capital mobility, but faces conflicting policy needs, such as fiscal imbalances or fragile financial sector, that need to be resolved by independent monetary policy. This leads to a shift in the regime from the first solution of the trilemma described above to the second one. The sudden depreciation in the exchange rate is often referred to as a currency crisis. It often has implications for the financial system as a whole and for the real economy, where agents were used to rely on a fixed exchange rate regime, and often have to adjust to the change abruptly and unexpectedly. We elaborate more on this in Subsection 4.3.

The theoretical currency-crises literature is broadly classified into three generations of models, which we now turn to describe in more detail.

4.1 First-Generation Model of Currency Crises

This branch of models, the so-called ‘first generation models of currency attacks’ was motivated by a series of events where fixed exchange rate regimes collapsed following speculative attacks, for example, the early 1970s breakdown of the Bretton Woods global system.

The first paper here is the one by Krugman (1979).¹⁷ He describes a government that tries to maintain a fixed exchange rate regime, but is subject to a constant loss of reserves, due to the need to monetize persistent government budget deficits. These two features of the policy are inconsistent with each other, and lead to an eventual attack on the reserves of the central bank, that culminate in a collapse of the fixed exchange rate regime. Flood and Garber (1984) extended and clarified the basic mechanism, suggested by Krugman (1979), generating the formulation that was widely used since then.

Let us provide a simple description of this model. The model is based on the central bank's balance sheet. The asset-side of the central bank's balance sheet at time t is composed of domestic assets $B_{H,t}$, and the domestic-currency value of foreign assets $S_t B_{F,t}$, where S_t denotes the exchange rate, i.e., the value of foreign currency in terms of domestic currency. The total assets have to equal the total liabilities of the central bank, which are, by definition, the monetary base, denoted as M_t .

In the model, due to fiscal imbalances, the domestic assets grow in a fixed and exogenous rate:

$$\frac{B_{H,t} - B_{H,t-1}}{B_{H,t-1}} = \mu.$$

Because of perfect capital mobility, the domestic interest rate is determined through the interest rate parity, as follows:

$$1 + i_t = (1 + i_t^*) \frac{S_{t+1}}{S_t},$$

where i_t denotes the domestic interest rate at time t and i_t^* denotes the foreign interest rate at time t . Finally, the supply of money, i.e., the monetary base, has to be equal to the

¹⁷ The model by Krugman (1979) builds on an earlier paper by Salant and Henderson (1977) about a speculative attack on gold reserves.

demand for money, which is denoted as $L(i_t)$, a decreasing function of the domestic interest rate.

The inconsistency between a fixed exchange rate regime, $S_t = S_{t+1} = \bar{S}$, with capital mobility and the fiscal imbalances comes due to the fact that the domestic assets of the central bank keep growing, but the total assets cannot change since the monetary base is pinned down by the demand for money, $L(i_t^*)$, which is determined by the foreign interest rate. Hence, the obligation of the central bank to keep financing the fiscal needs puts downward pressure on the domestic interest rate, which, in turn, puts upward pressure on the exchange rate. In order to prevent depreciation, the central bank has to intervene by reducing the inventory of foreign reserves. Overall, $\bar{S}B_{F,t}$ decreases by the same amount as $B_{H,t}$ increases, so the monetary base remains the same.

The problem is that this process cannot continue forever, since the reserves of foreign currency have a lower bound. Eventually, the central bank will have to abandon the current solution of the trilemma – fixed exchange rate regime and perfect capital mobility – to another solution – flexible exchange rate with flexible monetary policy (i.e., flexible monetary base or equivalently flexible domestic interest rate) and perfect capital mobility.

The question is what the critical level of domestic assets $B_{H,T} = B_{H,T}^*$ is and what the corresponding period of time T is, at which the fixed-exchange rate regime collapses. As pointed out by Flood and Garber (1984), this happens when the shadow exchange rate – defined as the flexible exchange rate under the assumption that the central bank's foreign reserves reached their lower bound while the central bank keeps increasing the domestic assets to accommodate the fiscal needs – is equal to the pegged exchange rate.

At this point, there will be a speculative attack on the currency, depleting the central bank's foreign reserves and forcing an immediate devaluation of the domestic currency.

This is depicted in Figure 4. The upper panel depicts the shadow exchange rate schedule and the pegged rate schedule as functions of domestic assets. Their intersection determines the level of domestic asset where the regimes switch occurs. The switch will not occur at a lower level of domestic assets because at that point there is no incentive to launch the speculative attack (it will yield a trading loss to the speculators). The switch will not occur at a higher level either because there is gain to be made at the intersection point, leading all speculators to attack at that point. The lower panel describes the path of foreign assets, which are quickly depleted at the point of the attack.

[Insert Figure 4 Here]

4.2 *Second-Generation Models of Currency Crises*

Following the collapse of the European Exchange Rate Mechanism (ERM) in the early 1990s, the so-called first-generation model of currency attacks did not seem suitable anymore to explain the ongoing crisis phenomenon. The events in Europe at that time featured governments actively making decisions between fighting the declining economic activity level and remaining in the exchange rate management system. Hence, there was a need for a model where the government's choice is endogenized, rather than the first-generation models where the exchange rate regime is essentially on 'automatic pilot'. This led to the development of the so-called 'second generation model of currency attacks,' pioneered by Obstfeld (1994, 1996).

Hence, in this line of models, the government/central bank is setting the policy endogenously, trying to maximize a well-specified objective function, without being able to fully commit to a given policy. An outcome of these models is that there are usually self-fulfilling multiple equilibria, where the expectation of a collapse of the fixed exchange rate regime leads the government to abandon the regime. This feature seemed attractive to many commentators as they thought it captured well the fact that crises where unexpected in many cases. This is related to the Diamond and Dybvig (1983) model of bank runs described in Section 2, creating a link between these two literatures.

Obstfeld (1996) discusses various mechanisms that can create the multiplicity of equilibria in a currency-crisis model. Let us describe one mechanism, which is inspired by Barro and Gordon (1983). Suppose that the government minimizes a loss function of the following type:

$$(y - y^*)^2 + \beta\varepsilon^2 + cI_{\varepsilon \neq 0}.$$

Here, y is the level of output, y^* is the target level of output, and ε is the rate of depreciation, which in the model is equal to the inflation rate. Hence, the government wants to minimize some combination of the rate of inflation and the distance from the target level of output. In addition, the third term is an index function, which says that there is a fixed cost in case the government deviates from the existing exchange rate. The interpretation is that the government is in a regime of zero depreciation (a fixed exchange rate regime), and deviating from it is costly.

Overall, when deciding on the rate of depreciation, the government has to weigh the costs against the benefit of depreciation. The costs are coming from the second and third terms above: There is a cost in operating the economy under inflation and there is a cost

in deviating from the promise of a fixed exchange rate regime. The benefit in depreciation is that it enables reduction in the deviations from the target level of output. More precisely, creating inflation (which is equivalent to depreciation here) above the expected level serves to boost output.

This effect of inflation on output is coming from the Philips Curve. It is demonstrated in the following expression, specifying how output is determined:

$$y = \bar{y} + \alpha(\varepsilon - \varepsilon^e) - u.$$

Here, \bar{y} is the natural output ($\bar{y} < y^*$, i.e., the government sets an ambitious output target level to overcome distortions in the economy), u is a random shock, and ε^e is the expected level of depreciation/inflation that is set endogenously in the model by wage setters based on rational expectations. The idea is that an unexpected inflationary shock ($\varepsilon > \varepsilon^e$) boosts output by reducing real wages and increasing production.

Importantly, the government cannot commit to a fixed exchange rate. Otherwise, it would achieve minimum loss by committing to $\varepsilon = 0$. However, due to lack of commitment, a sizable shock u will lead the government to depreciate and achieve the increase in output bearing the loss of credibility. Going back to the trilemma discussed above, a fixed exchange rate regime prevents the government from using monetary policy to boost output, and a large enough shock will cause the government to deviate from the fixed exchange rate regime.

It can be shown that the above model generates multiplicity of equilibria. If wage setters coordinate on a high level of expected depreciation/inflation, then the government will validate this expectation with its policy by depreciating more often. If they coordinate on a low level of expected depreciation, then the government will have a

weaker incentive to deviate from the fixed exchange rate regime. Hence, the expectation of depreciation becomes self-fulfilling.

Similarly, closer to the spirit of the Krugman (1979) model, one can describe mechanisms where speculators can force the government to abandon an existing fixed-exchange rate regime by attacking its reserves and making the maintenance of the regime too costly. If many speculators attack, the government will lose significant amount of reserves and will be more likely to abandon the regime. A self-fulfilling speculative attack is profitable only if many speculators join it. Consequently, there is one equilibrium with a speculative attack and a collapse of the regime, and there is another equilibrium where these things do not happen.¹⁸

This issue is also strongly related to sovereign debt crises and in particular those currently experienced in Europe. Speculators can attack government bonds demanding higher rates due to expected sovereign-debt default, creating an incentive for the central bank to abandon a currency regime and reduce the value of the debt or alternatively to default. This justifies the initial high rates. In the sovereign-debt literature, authors have studied self-fulfilling debt crises of this kind. For example, Cole and Kehoe (2000) analyze the debt maturity structure under financial crises brought on by a loss of confidence in a government, which can arise within a dynamic, stochastic general equilibrium model.

As we discussed in Section 2, having a model of multiple equilibria creates an obstacle for policy analysis. Morris and Shin (1998) were the first to tackle the problem of multiplicity in the second-generation models of speculative attacks. They first express

¹⁸ Note that self-fulfilling speculative attacks can arise naturally from a first-generation model as demonstrated by Obstfeld (1986). Hence, this is not the distinguishing feature of the second-generation models. Rather, the optimizing government is the distinguishing feature of the second-generation models.

this model in an explicit game theoretic market framework, where speculators are players having to make a decision whether to attack the currency or not. Then, using the global-games methodology, pioneered by Carlsson and van Damme (1993), they are able to derive a unique equilibrium, where the fundamentals of the economy uniquely determine whether a crisis occurs or not. This is important since it enables one to ask questions as to the effect of policy tools on the probability of a currency attack. The global-games methodology, relying on heterogeneous information across speculators, also brought to the forefront the issue of information in currency-attack episodes, leading to analysis of the effect that transparency, signaling, and learning can have on such episodes (e.g., Angeletos, Hellwig, and Pavan (2006) and Goldstein, Ozdenoren, and Yuan (2011)).¹⁹

4.3 Third-Generation Models of Currency Crises

In the late 1990s, a wave of crises hit the emerging economies in Asia, including Thailand, South Korea, Indonesia, Philippines, and Malaysia. A clear feature of these crises was the combination of the collapse of fixed exchange rate regimes, capital flows, financial institutions, and credit.²⁰ As a result, many researchers felt that the first two generations of models of currency crises, which were described in the previous two subsections, were not sufficient for analyzing the events in Asia. There was a strong need to incorporate banking panics and credit frictions into these models. This led to extensive research on the interplay between currency crises and banking crises, sometimes referred

¹⁹ For a broad review of the global-games methodology and its various applications, see Morris and Shin (2003). There is also a large literature that followed the original developments, analyzing conditions under which the unique-equilibrium result fails to hold. See, e.g., Angeletos and Werning (2006) and Hellwig, Mukherji, and Tsyvinski (2006).

²⁰ For a broad description of the events around the Asian Crisis and the importance of capital flows in conjunction with the collapse of the exchange rate see Radelet and Sachs (1998) and Calvo (1998).

to as the *twin crises*, and between currency crises and credit frictions.²¹ Such models are often referred to as the ‘third-generation models of currency crises’. In the context of this survey, it is important to note that such models bring together elements from the early currency crises literature described in Subsections 4.1 and 4.2 with elements from the vast literatures on banking panics and credit frictions described in Sections 2 and 3, respectively.

One of the first models to capture this joint problem was presented in Krugman (1999). In his model, firms suffer from a currency mismatch between their assets and liabilities: their assets are denominated in domestic goods and their liabilities are denominated in foreign goods. Then, a real exchange rate depreciation increases the value of liabilities relative to assets, leading to deterioration in firms’ balance sheets. Because of credit frictions as in Holmstrom and Tirole (1997), described in Section 3, this deterioration in firms’ balance sheets implies that they can borrow less and invest less. The novelty of the Krugman’s paper is that the decrease in investment validates the depreciation in a general equilibrium setup. This is because the decreased investment by foreigners in the domestic market implies that there will be a decrease in the aggregate demand for the local goods, relative to foreign goods (the Keynes-Ohlin “transfer problem” in international trade), leading to real depreciation. Hence, the system has multiple equilibria with high economic activity, appreciated exchange rate, and strong balance sheets in one equilibrium, and low economic activity, depreciated exchange rate, and weak balance sheets in the other equilibrium. Other models that extended and continued this line of research include: Aghion, Bacchetta, and Banerjee (2001),

²¹ For empirical evidence on the twin crises, see Kaminsky and Reinhart (1999).

Caballero and Krishnamurthy (2001), and Schneider and Tornell (2004). The latter fully endogenize the currency mismatch between firms' assets and liabilities.

A different line of research links currency problems with the bank runs described in Section 2. Chang and Velasco (2001) and Goldstein (2005) model the vicious circle between bank runs and speculative attacks on the currency. On the one hand, the expected collapse of the currency worsens banks' prospects, as they have foreign liabilities and domestic assets, and thus generates bank runs. On the other hand, the collapse of the banks leads to capital outflows that deplete the reserves of the government, encouraging speculative attacks against the currency.

Accounting for the circular relationship between currency crises and banking crises is important for policy analysis, as it makes some well-received conclusions much less appealing. For example, traditional banking models may advocate a lender-of-last-resort policy or other expansionary policies during a banking crisis to mitigate the bank-run problem. However, accounting for the circularity between bank runs and currency attacks, it is shown that such policies might backfire as they deplete the reserves available to the government, making a currency crisis more likely, which in turn might further hurt the banking sector that is subject to a currency mismatch problem.

As we mentioned in Section 4.2, there is a strong link between currency-crises models and sovereign-debt models, exemplified by the Cole and Kehoe (2000) framework. Hence, the models reviewed in this subsection, tying banking and credit problems with currency crises, can be very helpful in analyzing the connection between banking crises and sovereign-debt crises. This seems to be a very relevant and timely avenue for research given the current situation in Europe, where the faith of governments

is intertwined with that of banks due to the various connections between banks' balance sheets and governments' balance sheets.

4.4 *Contagion of Currency Crises*

In Section 2.6, we reviewed theories of contagion focused on the contagion of crises across different banks. The forceful transmission of crises across countries generated a large literature of international financial contagion, which is very strongly related to the literature reviewed in Section 2.6. Kaminsky, Reinhart, and Vegh (2003) provide a review of the theories behind such contagion.²² They define contagion as an immediate reaction in one country to a crisis in another country.

As we wrote in Section 2.6, there are several theories that link such contagion to fundamental explanations. The clearest one would be that there is common information about the different countries, and so the collapse in one country leads investors to withdraw out of other countries, see e.g. Calvo and Mendoza (2000). Models of the connections of portfolios across different countries, e.g., Allen and Gale (2000b), Kodres and Pritsker (2002), Dasgupta (2004), and Goldstein and Pauzner (2004) also shed light on such international contagion. An explanation that is more directly related to currency depreciation is proposed by Gerlach and Smets (1995). If two countries compete in export markets, the devaluation of one's currency hurts the competitiveness of the other, leading it to devalue the currency as well.

Empirical evidence has followed the above theories of contagion. The common information explanation has vast support in the data. Several of the clearest examples of contagion involve countries that appear very similar. Examples include the contagion that

²² For a broader review, see the collection of articles in Claessens and Forbes (2001).

spread across East Asia in the late 1990s and the one in Latin America in the early 1980s. A vast empirical literature provides evidence that trade links can account for contagion to some extent. These include Eichengreen, Rose, and Wyplosz (1996) and Glick and Rose (1999). Others have shown that financial linkages are also empirically important in explaining contagion. For example, Kaminsky, Lyons, and Schmukler (2004) have shown that US-based mutual funds contribute to contagion by selling shares in one country when prices of shares decrease in another country. Caramazza, Ricci, and Salgado (2004), Kaminsky and Reinhart (2000) and Van Rijckeghem and Weder (2001) show similar results for common commercial banks.

5. *Concluding Remarks*

The global financial crisis that started in 2007 and has not been resolved yet took much of the economic profession by surprise. Explaining the forces behind the crisis and coming up with suggestions for policymakers on how to solve it and fix the system going forward have become top priorities for many economists, some of whom are new to the topic of financial crises.

As we argue in this paper, many of the forces in play in the current and recent turmoil have been featured in the literature on financial crises for more than three decades now. Hence, it is important to go back to the main streams of this literature and summarize them to better understand the main forces behind crises, how they interact, how they apply to current and recent events, and what they imply for future policy. In this paper, we attempted to achieve these goals.

The Changing Nature of Currency Crises since the 1970s

Sweta Chaman Saxena¹

Abstract: Currency crises that have been observed in recent years are not a new phenomenon, but the main features of the crises in Latin America in the 1970s and early 1980s are quite different from the crisis in Europe in 1992. Theoretical literature has evolved over time to account for the changing nature of these crises. While many theoretical and empirical papers have been written about various episodes of these crises, the change in their mechanism over time has not been demonstrated well. This paper fills the gap in the literature by graphically depicting the main features of these crises. Such a visual analysis should allow the reader to better understand and follow the changes in the mechanisms over time.

JEL Classification: F31, F32, F34, F41

Key Words: currency crises, banking crises, first generation models, second generation models, twin crises, herding, contagion, moral hazard

¹ Assistant Professor, Graduate School of Public and International Affairs, 3E15 Posvar Hall, University of Pittsburgh, Pittsburgh, PA 15260; email: ssaxena@pitt.edu. I would like to thank Alexis Haakensen and Sherri Barrier for excellent research assistance and Deniz Atasoy, Brandon Tracy and Sandra Williamson for editorial suggestions. This paper draws on my survey with Kar-yiu Wong.

1. Introduction

Currency crises that have been observed in recent years are not a new phenomenon—they date back to the gold standard era. However, main features of the crises in Latin America in the 1970s and early 1980s are quite different from the crisis in Europe in 1992. Theoretical literature has evolved over time to account for the changing nature of these crises. While many theoretical and empirical papers have been written about various episodes of these crises, the change in their mechanism over time has not been demonstrated well. This paper fills the gap in the literature by graphically depicting the main features of these crises. Such a visual analysis should allow the reader to better understand and follow the changes in the mechanisms over time.

In the analysis of these currency crises, we note two common features, namely, a fixed exchange rate regime and capital flight and speculative attack. While this paper summarizes much of the literature on currency crises, data is restricted to the first and second generation models and the models of twin crises.^{2,3} The rest of the paper is organized as follows. Section 2 discusses the first generation models' application to Latin America. Section 3 unfolds the European crisis of 1992. Section 4 presents the stylized facts from the twin crises literature. The next three sections summarize the literature on herding, contagion and moral hazard, respectively. Section 8 concludes.

2. First Generation Models and Latin America

The literature on currency crises began with the works of Krugman (1979) and Flood and Garber (1984). These papers have now come to be known as the *First*

² The dates of all crises are taken from Kaminsky and Reinhart (1999).

³ In order to conserve space, we only show data for 5-6 episodes of crises for each generation model. The longer version of the paper is available upon request from the author.

Generation Models of currency crises. They typically depict the scenario of the Latin American countries in the 1970s and early 1980s, where the collapse of the exchange rate regime occurs under the monetization of budget deficits that depletes the foreign exchange reserves of the central bank. The sudden collapse comes about due to perfectly mobile capital that moves to maintain the uncovered interest parity. In a perfect foresight version of the first generation models, instantaneous capital flows ensure that there are no jumps in the exchange rate that would represent a profit opportunity for speculators. When the shadow value of the exchange rate crosses the fixed rate, there is a sudden loss of reserves and an increase in the interest rates and the currency begins to depreciate.

Figures 1 through 5 depict the main characteristics of the first generation crises in Latin America in the 1970s and early 1980s. Budget deficits (mostly exceeding 8-10% of GDP) were financed by soaring growth rates of money and domestic credit. High current account deficits along with massive capital outflows accompanied the loss of foreign exchange reserves and/or an increase in interest rate. The exchange rate eventually depreciated. This is illustrated for the Bolivian crisis from 1982-85, the Brazilian crises in 1983, 1986 and 1989-90, the Chilean crisis from 1971-74, Peruvian crises in 1976 and 1987 and the Uruguayan crisis in 1982.

Although theoretical models explain the empirical occurrence of these crises in a straightforward manner, they suffer from some obvious weaknesses. First, the rule of deficit financing assumed is very rigid, even though it is not sustainable in the long run. Second, while the investors are actively maximizing the returns on their assets, the government is too passive—it knows that the central bank is losing reserves, and hence it can and should abandon parity before exhausting all reserves, but it chooses not to do so. Lastly, theoretical models imply that the exchange rate will adjust smoothly when the

exchange rate regime shifts from a fixed to a flexible one. In reality, countries experience quick, unexpected and huge devaluations.⁴ The perfect information postulate in these models prohibits discrete jumps in the exchange rate, but uncertainty, as introduced by Krugman (1979) and Flood and Garber (1984), may not be sufficient to explain the big devaluations that follow the crises. In Krugman (1979), the incomplete knowledge on the part of the investors about how much of its reserves the government is willing to use to defend the exchange rate produces the possibility of alternating balance of payments crises and recoveries of confidence. In Flood and Garber (1984), uncertainty in domestic credit creation produces the same cycles of crises and recoveries.⁵

3. Second Generation Models and Europe

The first-generation models had linear behavioral functions. Second-generation models focus on non-linearities in government behavior—studying what happens when government policy reacts to changes in private behavior or when the government faces an explicit trade-off between the fixed exchange rate policy and other objectives. Even when policies are consistent with the fixed exchange rate, attack-conditional policy changes can *pull* the economy into an attack. In contrast, the first-generation models generate an attack by having inconsistent policies before the attack *push* the economy into a crisis. Other models show that a shift in market expectations can alter the government's trade-offs and bring about self-fulfilling crises. The newer models admit the possibility that the economy can be at a *no-attack* equilibrium where speculators see but do not pursue available profit opportunities. In such a situation, anything that serves to coordinate the

⁴ The exchange rates in figures 1 through 5 do not show discrete devaluations because the data is period-averaged and is annual.

⁵ These models differ from Obstfeld (1996a), which introduces the cost to holding foreign currency.

expectations and actions of speculators can suddenly cause an attack (Flood and Marion (1998)).

The experience of the European countries in 1992-93 lacked the features of the first-generation models—namely, the inconsistency between continuous creation of domestic credit and a fixed exchange rate. Due to severe speculative attacks on their currencies, the member countries of the European Monetary System allowed more flexibility in their exchange rates, permitting their currencies to move within a band of ± 15 percent instead of ± 2.25 percent for most Exchange Rate Mechanism (ERM) rates in August 1993. However, two years later the exchange rates returned to their pre-crisis level, which shows that these European countries did not have inherent macroeconomic problems and that currency crises can arise even when economies have sound macroeconomic fundamentals.

Looking at the breakdown of the ERM, Obstfeld (1994) suggests the following features of the crises experienced by these European countries:

1. There are reasons why the government wants to abandon the peg (to inflate away the debt burden denominated in domestic currency, and to follow expansionary monetary policies in case of unemployment, etc).
2. There are reasons why the government wants to defend the peg, hence a conflict between the two (to facilitate international trade and investment, to gain credibility if has a history of high inflation, and as a source of national pride or commitment to an international cooperation).
3. The cost of defending the peg rises when people expect the peg will be abandoned, primarily because people *in the past* expected that the exchange rate would be depreciated *now*. Hence, anticipation of devaluation makes the debt-holders and

worker unions in the past demand higher interest rate and wages, making the debt-burden too high and industries uncompetitive at the current exchange rate level.

Obstfeld (1986, 1996a)⁶ demonstrates that the important trigger for a crisis is people's expectation. Even though the economy can be fundamentally strong, if people expect a devaluation in the near future, they could put enormous pressure on the central bank by converting their domestic currency to foreign currency before the devaluation. If a sufficient number of people proceed as such, the central bank could run out of foreign reserves, necessitating a devaluation. In this case, the crisis is self-fulfilling. Often the models that emphasize the above characteristics are called the *Second Generation Models of currency crises*.

In addition, these models show that the timing of the crisis is indeterminate because that too depends on people's expectations. In the indeterminate range, the economy could certainly face *an attack* if a large trader can take a massive position against the fixed exchange rate (like George Soros did against the British pound in 1992). But if the foreign exchange market consists of many small, credit constrained traders and no large trader, then without anything to coordinate their expectations and actions, the investors cannot mount an attack of sufficient size to move the economy from a *no-attack* to an *attack equilibrium*. Hence, the economy can maintain a fixed exchange rate indefinitely unless something coordinates expectations and actions to cause an attack (Flood and Marion, 1998).

Figures 6 through 10 demonstrate the crisis that hit Europe in 1992-3. In the wake of German Unification in 1989, Germany's fiscal policy became expansionary, which

⁶ Obstfeld (1996a) introduces the cost of holding foreign currency which was missing from Krugman (1979) and Flood and Garber (1984).

was counteracted by the Bundesbank's contractionary monetary policy as seen in the rise in interest rate (to prevent inflation). With perfectly mobile capital across the European borders, unfavorable interest differentials with Germany led to outflow of capital from the rest of Europe. In order to maintain the balance of payments equilibrium, the European countries had to raise their interest rates. This would have normally been appropriate, except that all the countries were in a cyclical downturn at that time. Speculators believed that it would be too costly for governments to maintain the peg by raising interest rates, in the face of recession and high unemployment levels. Hence, the speculative attack.

As can be seen in the figures, budget deficits before the crisis were low (less than 5% of GDP). None of these countries had burgeoning current account deficits, like the Latin American countries in the first generation models. Fixed exchange rate under the ERM, along with positive inflation differentials with Germany, led to real appreciation of exchange rates,⁷ which made these economies uncompetitive. Recession or cyclical downturn, along with rising interest rates, led the speculators to believe that the governments would not be able to maintain their respective pegs. This led to speculative attacks on the British pound, Italian lira and French franc, which later spread to other countries.⁸

While these models emphasize multiple equilibria, Krugman (1996) disputes such indeterminacy of the speculative attacks and suggests that timing of the crisis can be determined by tracing the deteriorating fundamentals. In responding to Krugman's criticism, Kehoe (1996) and Obstfeld (1996b) point out that the new crisis models do not

⁷ Note: RNULC stands for real effective exchange rates based on normalized unit labor costs.

⁸ Refer to Buiter, et al (1998) for a day-to-day description of the European crisis.

assert that every fixed exchange rate regime must be subject to a self-fulfilling crisis, and that while it is argued that a self-fulfilling crisis can occur even though a fixed exchange rate can be sustained in the absence of speculation, there is no denial that deteriorating fundamentals can lead to a crisis and can eliminate multiple equilibria and indeterminacy of timing of a crisis. Furthermore, a speculative attack may have nothing to do with the long-run sustainability of a fixed exchange rate regime if it occurs in the face of some temporary problems faced by an economy, as was clearly witnessed by the European economies.

4. Models of Twin Crises

The newer generation models emphasize the importance of the financial sector and capital flows in currency crises, hence, the term *twin crises*. The frequent occurrence of twin crises (Nordic countries in 1990s, Turkey in 1994, Venezuela, Argentina and Mexico in 1994, Bulgaria in 1996, and Asian countries in 1997) has led academicians and policymakers to ascertain the causes and causation of these crises, which is debatable.

1. Stoker (1995) and Mishkin (1996) argue that a *balance of payments crisis leads to a banking crisis*. According to Stoker, an external shock in the face of a fixed exchange rate leads to loss of reserves. If this loss of reserves is not sterilized, then a speculative attack is followed by a period of abnormally high interest rates leading to a credit crunch, increased bankruptcies and financial crisis. Mishkin argues that devaluation could weaken the position of the banks if they have foreign currency denominated liabilities.
2. However, Diaz-Alejandro (1985), Velasco (1987), Calvo (1995) and Miller (1995) argue that *banking crises lead to balance of payments crises*. The argument is that

central banks print money to bailout financial institutions and hence lose their ability to maintain the prevailing exchange rate commitment.

3. But Reinhart and Vegh (1996) and Kaminsky and Reinhart (1999) suggest that the *two crises have some common causes*—an example of “perverse” dynamics of an exchange rate-based inflation stabilization plan. Since prices are slow to converge to international levels, the exchange rate appreciates markedly. Initially, the boom in imports and economic activity is financed by borrowing abroad. This leads to a widening of current account deficits and financial markets infer that the stabilization program is unsustainable, hence the currency is attacked. The increase in bank credit during the boom is financed by foreign borrowings, such that when capital flows out and the asset market crashes, it leads to the collapse of the banking system as well. McKinnon and Pill (1996), Goldfajn and Valdes (1997) and Chang and Velasco (1998) show that the intermediating role of the banking sector (which creates liquidity problems later on) can make the boom-bust cycles more pronounced, leading to the collapse of the currency and the banking sector.⁹

The stylized facts that these models tend to explain are that banking crises are highly correlated to currency crises, capital inflows increase steadily before the crisis and fall sharply during the crisis and banking activity (intermediation) increases some time before the collapse. In particular, Kaminsky and Reinhart (1999) show that:

1. During the 1970s, when markets were highly regulated, there was no apparent link between the balance of payments (BOP) and banking crises. The two crises became more entwined in the 1980s following the liberalization of financial markets across

⁹ These are open-economy versions of the classic Diamond-Dybvig (1983) model of bank runs, where uninsured demand deposit contracts are able to provide liquidity but leave the banks vulnerable to runs.

many parts of the world. Their results show that the collapse of a currency deepens the banking crisis, activating a vicious spiral.

2. While banking crises often precede BOP crises, their results point to common causes. Both crises are preceded by recessions or below-normal economic growth, in part attributed to a worsening of terms of trade, an overvalued exchange rate (affecting exports negatively) and the rising cost of credit. A shock to financial institutions (possibly financial liberalization and/or increased access to international capital markets) fuels the boom phase of the cycle by providing access to financing, but leaves the economy vulnerable as the unhedged foreign liabilities of the banking system rise.
3. They show that crises (external or domestic) are typically preceded by a multitude of weak and deteriorating economic fundamentals and the incidence of crises where the economic fundamentals were sound are rare (as Krugman argues). Compared with single crises, the economic fundamentals tended to be worse, the economies were considerably more frail, and the crises (both banking and currency) were far more severe for the “twin” crises.

Figures 11 through 16 demonstrate the occurrence of the twin crises in Latin American, European and Asian economies in the 1980s and 1990s. Financial indicators (M2 multiplier, ratio of M2 to reserves, domestic credit as a % of GDP, bank deposits, lending-deposit ratio and interest rates) rise rapidly before the crisis. The excessive liquidity fuels growth in the economy. Since domestic inflation is higher than the world inflation, the real exchange rate appreciates, which, along with a deterioration in the

terms of trade, leads to current account deficits¹⁰ and squeezes profit margins, leading to bankruptcies, increase in non-performing loans, deepening of the economic contraction and banking sector problems. This reveals the weaknesses in the economy and leads to capital outflows. Under the fixed exchange rate regime, the peg is initially defended by the loss of foreign exchange reserves and/or high interest rates. But eventually the currency collapses, bringing down the booming banking sector. These stylized facts are depicted for the Chilean crisis of 1981-83, the Danish crisis of 1987-90, the Indonesian and the Korean crisis of 1997, the Norwegian crisis of 1988-91, and the Venezuelan crises of 1989 and 1993-94.

5. Herding

The bank run literature takes us into a similar explanation of currency crises—the *herd behavior*. Herding, which is an example of information cascade, is said to exist when individuals tend to choose actions similar to previous actions chosen by other individuals. In other words, with herding effects, individuals tend to move in conformity, and a small shock to society could lead to a mass shift in the actions of people. In some special cases, people choose to give up their private information or signals they possess and follow the actions of others, even though the private information or signals they have would suggest they should act otherwise.

A famous example is Keynes' beauty contest example.¹¹ Earlier work includes the papers by Leibenstein (1950) on the bandwagon effects. Recently, more rigorous models

¹⁰ This trend gets reversed most of the time following the depreciation of the exchange rate.

¹¹ In a beauty pageant, a judge picks the girl that he thinks others will pick, rather than the one he considers to be the most beautiful.

have been suggested to explain herd behavior.¹² Several models that have been introduced to explain investment behavior are mentioned here.

Froot, et al (1992) show that speculators with short horizons may herd on the same information, trying to learn what the other informed traders know. These could lead to multiple equilibria, and herding speculators may even choose to study information that is completely unrelated to fundamentals. So, the large perceived penalty for missing a bull market leads managers to follow the pack even if fundamentals do not warrant it; conversely, the penalty of losses during the bear market are lower as all other managers are losing money as well.¹³

Krugman (1998a) suggests similar reasons why herding might occur. First, there is a *bandwagon effect*, which is driven by the awareness that investors have private information—where investors ignore their own information and thrive on the information of other investors. Suppose that investor 1 has special information about the Thai real estate market, investor 2 has special information about the financial conditions of the banks and investor 3 has information about the internal discussions of the government. If investor 1 gets some negative information, he may sell, since that is the extent of his information; if investor 2 learns that 1 has sold, he may sell also even if his own private information is neutral or slightly positive. And investor 3 may end up selling even if his own information is positive because the fact that 1 and 2 have sold leads him to conclude that both may well have received bad news, even when in fact they have not. Krugman argues that bandwagon effects in markets with private information create a sort of “hot money” that at least sometimes causes foreign exchange markets to overreact to news

¹² See, for example, Banerjee (1992), Bikhchandani, et al (1992), and Froot, et al (1992).

¹³ As Krugman (1998a) puts it: “I feel worse if I lose money in a Thai devaluation when others don’t, than I will if I lose the same in the general rout.”

about national economic prospects. Second, much of the money invested in crisis-prone countries is managed by agents rather than directly by principals—where the *principal-agent problems* arise. To the extent that money managers are compensated based on comparison with other money managers, they may have strong incentives to act alike even if they have information suggesting that the market's judgement is in fact wrong.

Chari and Kehoe (1997) link debt-default actions of the governments to herding behavior. In their model, investors have private information about the state of the economy and have prior notion about the competence of the government. The credibility of the government is built on its ability to pay its debts—the government is competent if it could repay foreign debt in a crisis state. If the prior that the government is competent is either very high or very low, then the investors ignore their private information and either lend or not lend, respectively. If the prior is in the intermediate range, the possibility of herding arises. In this range, capital flows are very sensitive to small pieces of information and hence volatile.

Although herding is a type of distortion, where certain massive actions by some individuals can hurt the economy, these actions could be entirely rational from individuals' point of view.¹⁴ Such rational behaviors occur when there are payoff externalities (payoffs to an agent adopting an action increases as the number of agents adopting that action increases) or principal-agent problems (managers have an incentive to hide in the herd so that their actions cannot be evaluated).

The above models usually assume sequential actions by individuals, so that those who take actions later will observe what actions others have taken previously. Calvo and Mendoza (1998) introduce a model in which herding can exist even when individuals

have simultaneous decision-making. They find that with informational frictions, herding behavior may become more prevalent as the world capital market grows. With globalization, the cost of collecting country-specific information to discredit rumors increases and managers, facing reputational costs, choose to mimic the market portfolio. Hence, small rumors can induce herding behavior and move the economy from a *no attack* to an *attack equilibrium*.

Of course, as Flood and Marion (1998) argue, in many cases, herding could explain some part of the currency crises, but not the whole. First, individuals are less likely to ignore their own or new information in a world where they can adjust their strategies continuously to new information. Second, in case strategic interactions are important, then the cascade story is unsatisfactory, because the potential capital gains arising from the action of one agent do not depend on actions chosen by others.

6. Contagion

Another way crises propagate is through *contagion*. The currency crises of the 1990s have consisted of 3 regional waves: the ERM crisis (1992-3), the Latin American crisis (1994-5) and the Asian crisis (1997-8). The issue of contagion explains how seemingly different countries can experience the same crisis.¹⁵

Contagion could be explained in terms of real linkages between the countries: a crisis in country A worsens the fundamentals of country B. For example, suppose Thailand and Malaysia export the same products in the world markets. Hence, a Thai devaluation would depress Malaysian exports, and could push Malaysia past the critical

¹⁴ Refer to Devenow and Welch (1996) for a summary on rational herding literature.

¹⁵ Refer to Kaminsky, et al (2003) for a discussion of “fast and furious” crises, where some crises induce contagion effects, while others do not.

point that triggers a crisis. In Europe, there was an element of competitive devaluation: depreciation of the pound adversely affected the trade and employment of France, and increased pressure on the French government to abandon its peg to the deutschemark.

In addition to trade, there could be direct financial linkages as well. Financial institutions in the home country may have a credit exposure or equity stakes in corporations, financial institutions, or real estate in the neighboring countries. A crisis in the neighboring country could then spillover by causing weakness in the home country's financial sector. This seems to explain the propagation of the crisis from Russia to Brazil. The investors liquidate holdings in one part of their portfolio due to losses in another part.

However, Krugman (1998a) argues that trade links in Asia and Europe were weak¹⁶ and in Latin America nil. Mexico is neither an important market nor an important competitor for Argentina; why, then, he asks, should one crisis have triggered another? He offers two rational explanations for contagion between seemingly unlinked economies, as noticed by Drazen (1997). First, countries are perceived as a group with some common, but imperfectly observed characteristics. For example, Latin American countries share a common culture and therefore, a "Latin temperament"; but the implications of this temperament for economic policy may be unclear. Once investors have seen one country with that cultural background abandon the peg under pressure, they may revise downward their estimates of the willingness of other countries to defend their parities—i.e., "wake up call". That is what happened in the case of the Asian Crisis. There was no news—just re-examination and re-evaluation of the already existing

¹⁶ A mere devaluation in the case of Europe would not have meant a unilateral increase in exports, as the European trading partners were all in a recession or at least in a downturn.

information because countries with similar background gave up the parity. Second, the political commitment to a fixed exchange rate itself is subject to herding effects. This was clearly the case in Europe—once Britain and Italy left the ERM, it was politically less costly for Sweden to abandon the peg to the deutsche mark than it would have been had Sweden devalued on its own.

However, not all contagion is negative. During the wave of optimism that followed the Mexican and Argentine reforms in the early 90s, countries that had done little actual reform, like Brazil, were also lifted by the rising tide. The apparent myopia of markets about Asian risks seems to have been fed by a general sense of optimism about Asian economies in general (as emphasized by Radelet and Sachs (1998)).

7. Moral Hazard

The final explanation for currency crises is the problem of *moral hazard*, as proposed by Krugman (1998b) and Corsetti, et al (1998) for the Asian crisis. Moral hazard can occur under asymmetric information because borrowers can alter their behavior after the transaction has taken place in ways that the lender regards as undesirable. In financial markets, however, moral hazard could occur in the absence of asymmetric information; i.e., moral hazard arises from the possibility that investor behavior will be altered by the extension of government guarantees that relieve investors of some of the consequences of risk taking.

Krugman (1998b) considers the case of over-guaranteed and under-regulated financial intermediaries. Since these institutions are not risking any capital ex-ante, and have the liberty to walk away at no personal cost in case of bankruptcy, the economy engages in excessive investment. This economy is made worse by globalization, due to

its increasing this country's access to the world capital market. If it did not have access to the world capital market, then excessive investment demand would show up as high rates of interest, and not as excessive investment. But access to world market allows the moral hazard in the financial sector to translate into real excess capital accumulation.

Corsetti, et al (1998) also recognize moral hazard as a source of over-investment, excessive external borrowing and current account deficits. Unprofitable projects and cash shortfalls are re-financed through external borrowing as long as foreign creditors lend to domestic agents against future bail-out revenue from the government. The government deficits need not be high before the crisis, but refusal of foreign creditors to re-finance the debt (knowing that the government's stock of foreign reserves is below a critical minimum) forces the government to mediate and guarantee the outstanding stock of external liabilities. The government could either raise sufficiently large revenues from explicit taxation or it might recourse to seigniorage revenues (printing money). In the former case, the financial case doesn't coincide with an exchange rate collapse. In the latter case, expectation of inflationary financing leads to expectation about exchange rate depreciation, which creates a wedge between domestic and international interest rates, causing the currency to collapse.

In fact, the argument of moral hazard is not only applicable to the intermediaries, but extends to governments as well. Proponents of moral hazard argue that the IMF creates a bailout for governments or investors in the event of a crisis. However, Radelet and Sachs (1998) do not see the Asian crisis as a result of carelessness on the part of the investors who were confident of a bailout, knowing that only the state-owned enterprises can be bailed out in the event of a crisis. According to Radelet and Sachs, if the creditors feared the risk of a crisis in Asia, then the spread on Asian bonds should have increased,

but it did not. If the creditors felt an increasing risk of a government-led bailout, then ratings of long term government bonds should have gone down, but they did not either. A large part of the investment went into the risky equity market, and bank loans went to the non-financial corporate sector, where a direct government bailout was least possible. Creditors had been aware of weak bankruptcy laws and ineffective judicial systems in Asia. Hence, the foreign investors lent because they anticipated these economies to perform well, and not because they believed that they would be bailed out.

8. Conclusion

This paper attempts to fill the gap in the literature by clearly demonstrating how the nature of currency crises has changed over time. The paper graphically presents the stylized facts for the first and second generation models of currency crises and also on twin crises. The crises in the 1990s have been more frequent and more severe when compared to the crises of the earlier decades. A frequently advanced reasoning is “globalization” and immediate transfer of news due to developments in information technology. Clearly, crises have been a result of poor and problematic policies (as in the first generation models) as well as of mere expectations of speculators (second generation models). In this globalized world, where information transfers so fast, it is in the interest of the policymakers to follow prudent macroeconomic policies in order to avoid speculative activity. Countries need to timely assess the stance of their policies, for example, Germany should have realized early enough the mayhem its tight monetary policy would cause for other nations and should have abandoned that policy. Or policymakers should realize that fixed exchange rates in the face of huge budget deficits and open capital accounts are an invitation to speculative attacks. In fact, Levy-Yeyati

and Sturzenegger (2003) find that less flexible exchange rate regimes are associated with slower growth and greater output volatility for developing countries. Hence, emerging market economies should move towards more flexible exchange rates (not necessarily free floats) and/or should even consider selective capital controls on inflows to change the composition of the flows towards longer maturity (as in Chile). In the absence of good macroeconomic policies, time and again, countries have and will continue to be punished by the markets.

Appendix 1: Data Source

Series	Source
Nominal Exchange Rate	IFS line ae; rf
Real Effective Exchange Rate (based on CPI)	IFS line rec
Real Effective Exchange Rate (based on RNULC)	IFS line reu
Total Foreign Reserves Minus Gold	IFS line 11.d
Reserve Money	IFS line 14
Consumer Price Index	IFS line 64
Bank Deposits	(IFS line 24 + IFS line 25) / IFS line 64
Domestic Credit	IFS line 32
M2 measure of money supply	IFS line 34+ IFS line 35
Discount Rate	IFS line 60
Treasury Bill Rate	IFS line 60C
Deposit Rate	IFS line 60L
Lending Rate	IFS line 60P
Exports	IFS line 70
Imports	IFS line 71
Terms of Trade	(IFS line 74 / IFS line 75)*100
Current Account Balance	IFS line 78ald
Capital Inflows	IFS line 78bjd + IFS line 78cad
Budget Balance	IFS line 80
GDP	IFS line 99b
GDP Volume	IFS line 99bvr
Unemployment Rate	IFS line 67R
Share Price	IFS line 62

References

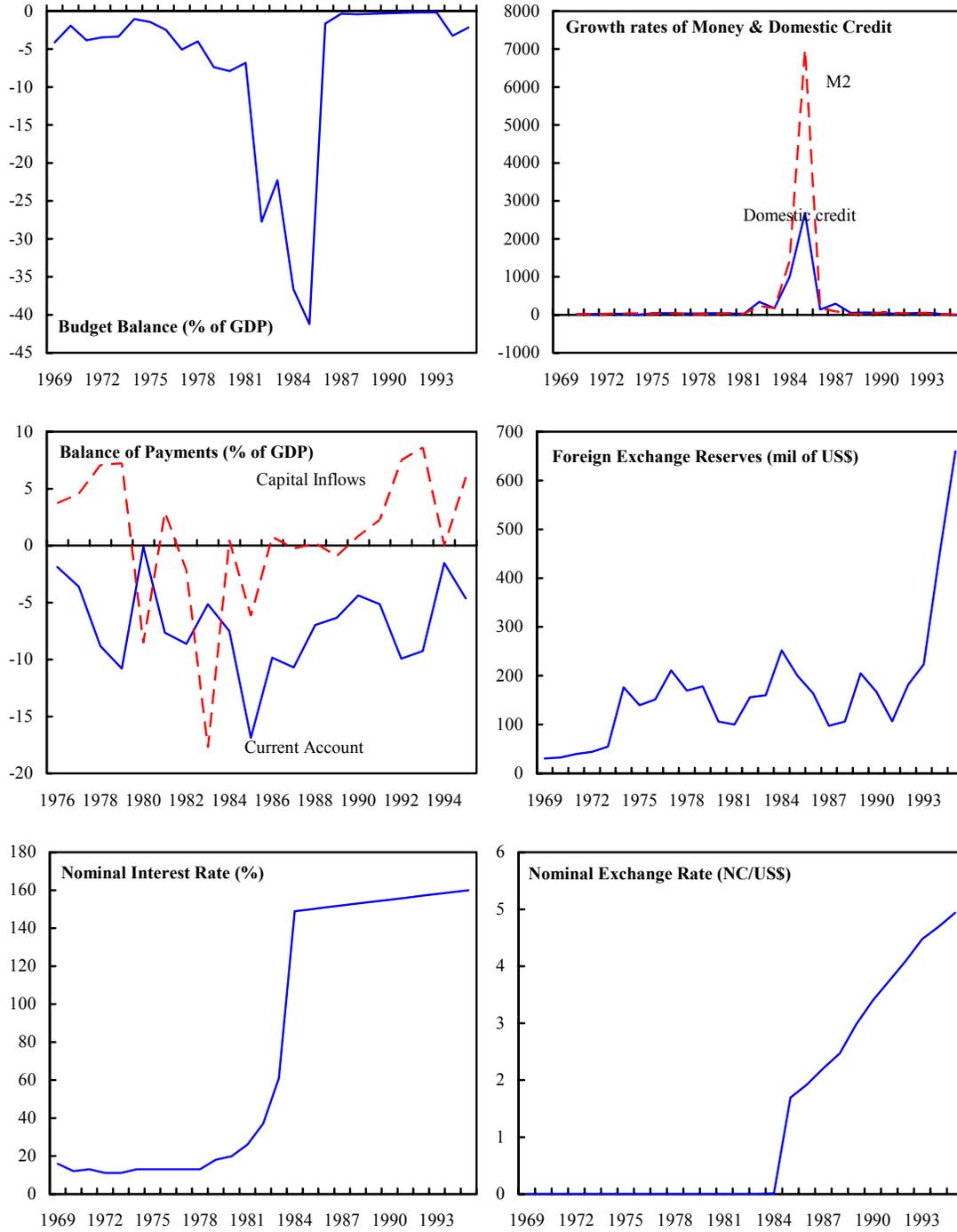
- Banerjee, Abhijit, 1992. "A Simple Model of Herd Behavior", *The Quarterly Journal of Economics*, Vol. CVII, No. 3, August 1992, pages 797-817.
- Bikhchandani, Sushil, David Hershleifer, and Ivo Welch, 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy*, pages 992-1026.
- Buiter, Willem, Giancarlo Corsetti and Paolo Pesenti, 1998. *Financial markets and European monetary cooperation*.
- Calvo, Guillermo, 1995. "Varieties of Capital Market Crises", *mimeo*, University of Maryland.
- Calvo, Guillermo and Enrique Mendoza, 1998. "Rational Herd Behavior and the Globalization of Securities Markets", *mimeo*, University of Maryland.
- Cerra, Valerie and Sweta C. Saxena, 2000. "Contagion, Monsoons and Domestic Turmoil in Indonesia: A Case Study in the Asian Crisis", *International Monetary Fund Working Paper*, WP/00/60, Washington, DC.
- Chang, Roberto and Andres Velasco, 1998. "Financial Crises in Emerging Markets: A Canonical Model", *Federal Reserve Bank of Atlanta, Working Paper*, 98-10.
- Chari, V.V. and Patrick Kehoe, 1997. "Hot Money", *National Bureau of Economic Research Working Paper*, 6007, Cambridge, Massachusetts.
- Corsetti, Giancarlo, Paolo Pesenti and Nouriel Roubini, 1998. "Paper Tigers? A model of the Asian crisis", *mimeo*, New York University.
- Devenow, Andrea and Ivo Welch, 1996. "Rational herding in financial economics", *European Economic Review*, Vol. 40, pages 603-15.

- Diaz-Alejandro, Carlos, 1985. "Good-bye Financial Repression, Hello Financial Crash", *Journal of Development Economics*, Vol. 19.
- Eichengreen, Barry, Andrew Rose and Charles Wyplosz, 1996a. "Contagious Currency Crises: First Tests", *Scandinavian Journal of Economics*, 98(4), 463-84.
- Flood, Robert and Peter Garber, 1984. "Collapsing Exchange Rate Regimes: Some Linear Examples", *Journal of International Economics*, Vol. 17, pages 1-17.
- Flood, Robert and Nancy Marion, 1998. "Perspectives on the Recent Currency Crisis Literature", *IMF Working Paper*, WP/98/130.
- Froot, Kenneth, David Scharfstein, and Jeremy Stein, 1992. "Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation", *Journal of Finance*, 47(4), September 1992, pages 1461-84.
- Goldfajn, Ilan and Rodrigo Valdes, 1997. "Capital Flows and the Twin Crises: The Role of Liquidity", *IMF Working Paper*, WP/97/87.
- Kaminsky, Graciela, Saul Lizondo and Carmen Reinhart, 1997. "Leading Indicators of Currency Crises", *IMF Working Paper*, WP/97/79.
- Kaminsky, Graciela and Carmen Reinhart, 1999. "The Twin Crises: The Causes of Banking and Balance of Payments Problems", *American Economic Review*, 89(3), 473-500.
- Kaminsky, Graciela, Carmen Reinhart and Carlos Vegh, 2003. "The Unholy Trinity of Financial Contagion," *Journal of Economic Perspectives*, 17(4), 51-74.
- Kehoe, Timothy J., 1996. "Comments on Krugman (1996)," in *Macroeconomics Annual*, NBER, Massachusetts, Cambridge: 378-392.
- Krugman, Paul, 1979. "A Model of Balance-of-Payments Crises", *Journal of Money, Credit and Banking*, Vol. 11, No. 3, August 1979, pages 311-325.

-, 1996. "Are currency crises self-fulfilling?" in *Macroeconomics Annual*, NBER, Massachusetts, Cambridge: 345-378.
-, 1998a. "Currency Crises", *mimeo*, MIT.
-, 1998b. "What happened to Asia?" *mimeo*, MIT
- Leibenstein, H., 1950. "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand," *Quarterly Journal of Economics*, 64 (2): 183-207.
- Levy-Yeyati, Eduardo and Federico Sturzenegger, 2003. "To float or fix: Evidence on the impact of exchange rate regimes on growth," *American Economic Review*, 93(4), 1173-93.
- McKinnon, Ronald and Huw Pill, 1996. "Credible Liberalization and International Capital Flows: the Overborrowing Syndrome", in *Financial deregulation and integration in East Asia*, edited by Takatoshi Ito and Anne Krueger, Chicago University Press: 7-42.
- Miller, Victoria, 1995. "Central Bank Reaction to Banking Crises in Fixed Exchange Rate Regimes", *mimeo*, Universite de Quebec a Montreal.
- Mishkin, Frederic, 1996. "Understanding Financial Crises: A Developing Country Perspective", in *Annual World Bank Conference on Development Economics*, World Bank, Washington D.C., pages 29-62.
- Obstfeld, Maurice, 1986. "Rational and Self-Fulfilling Balance of Payments Crises", *American Economic Review*, 76 (1): 72-81.
-, 1994. "The Logic of Currency Crises", *Banque De France, Cahiers Economiques et Monetaires*, No. 43, pages 189-213.
-, 1996a. "Models of currency crises with self-fulfilling features", *European Economic Review*, 40: 1037-47.

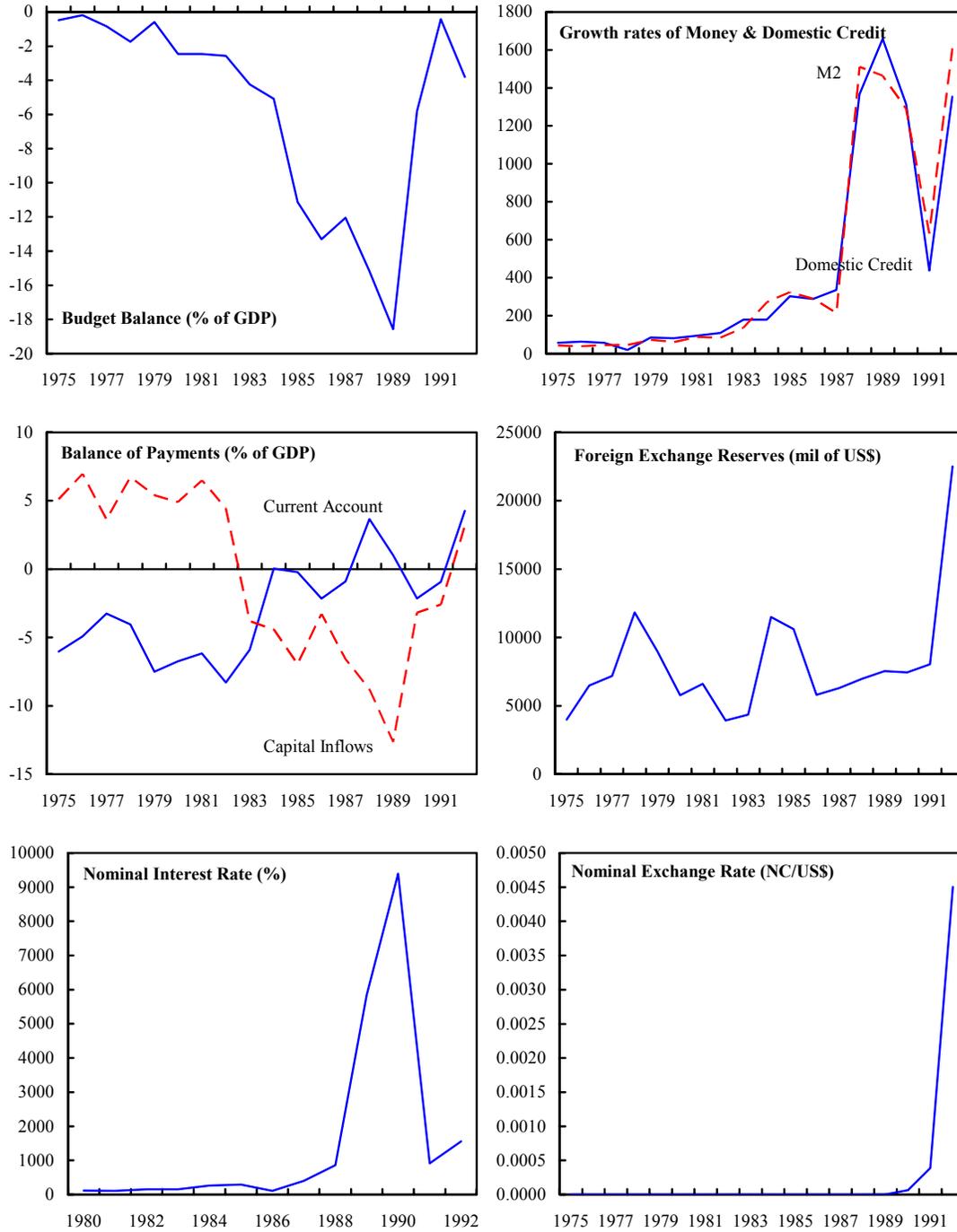
-, 1996b. "Comment on Krugman (1996)," in *Macroeconomics Annual*, NBER, Massachusetts, Cambridge: 393-403.
- Radelet, Steven and Jeffery Sachs, 1998. "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects", *mimeo*, Harvard Institute for International Development.
- Reinhart, Carmen M. and Carlos A. Vegh, 1996. "Do Exchange Rate-Based Inflation Stabilizations Sow the Seeds of their Own Destruction?" *mimeo*, University of Maryland and UCLA.
- Sachs, Jeffery, Aaron Tornell and Andres Velasco, 1996. "Financial Crises in Emerging Markets: The Lessons of 1995", *Brooking Papers on Economic Activity*, No. 1, 147-217.
- Salant, Stephen and Dale Henderson, 1978. "Market Anticipation of Government Policy and the Price of Gold", *Journal of Political Economy*, 86: 627-48.
- Saxena, Sweta C. and Kar-yiu Wong, 1999. "Currency Crises and Capital Controls: A Selective Survey," *mimeo*, University of Pittsburgh and University of Washington.
- Stoker, James, 1995. "Intermediation and the Business Cycle under a Specie Standard: The Role of the Gold Standard in English Financial Crises, 1790-1850", *mimeo*, University of Chicago.
- Velasco, Andres, 1987. "Financial and Balance of Payments Crises," *Journal of Development Economics*, 27: 263-93.

Figure 1. Bolivia: First Generation Model



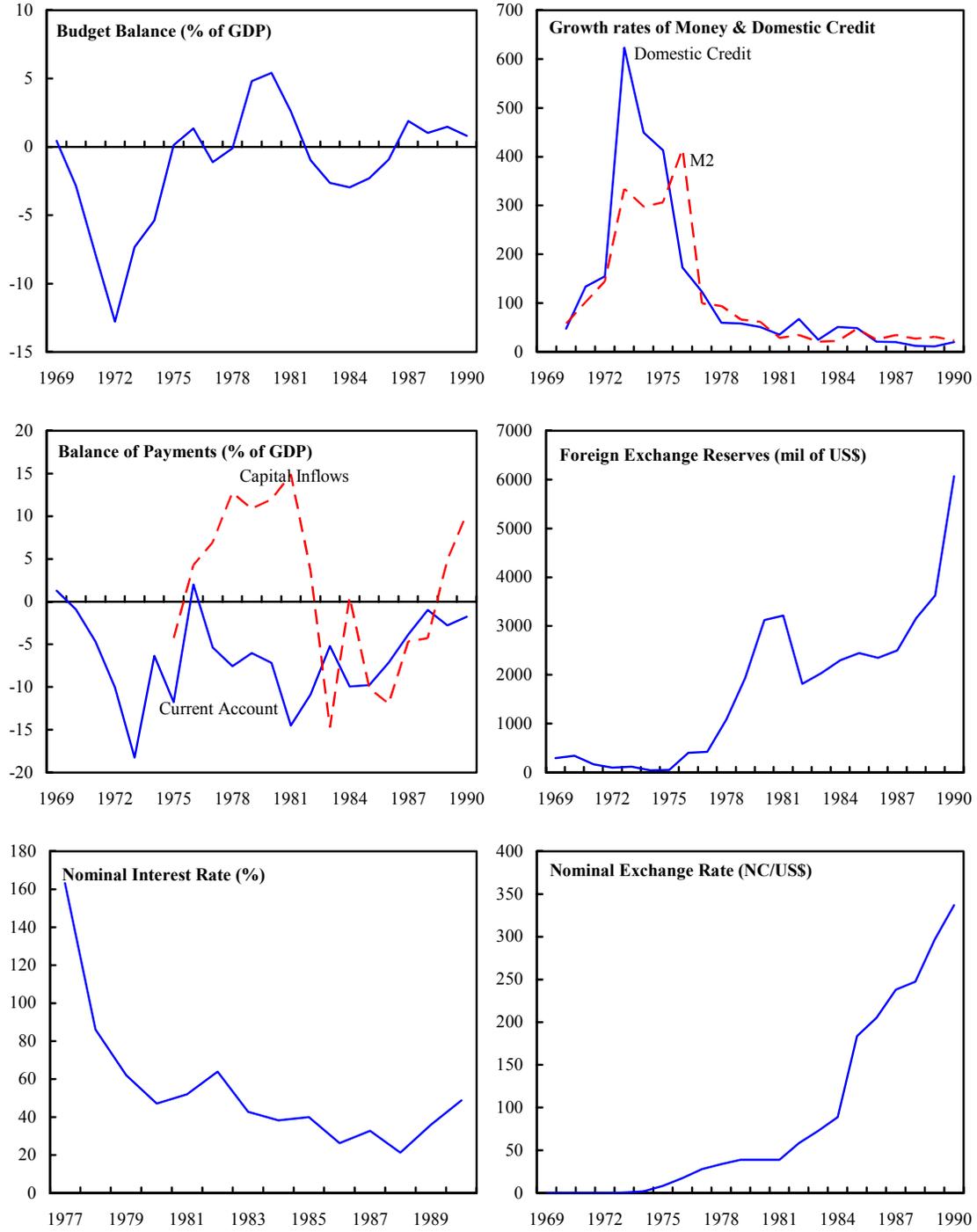
Crisis date: 1982-85

Figure 2. Brazil: First Generation Model



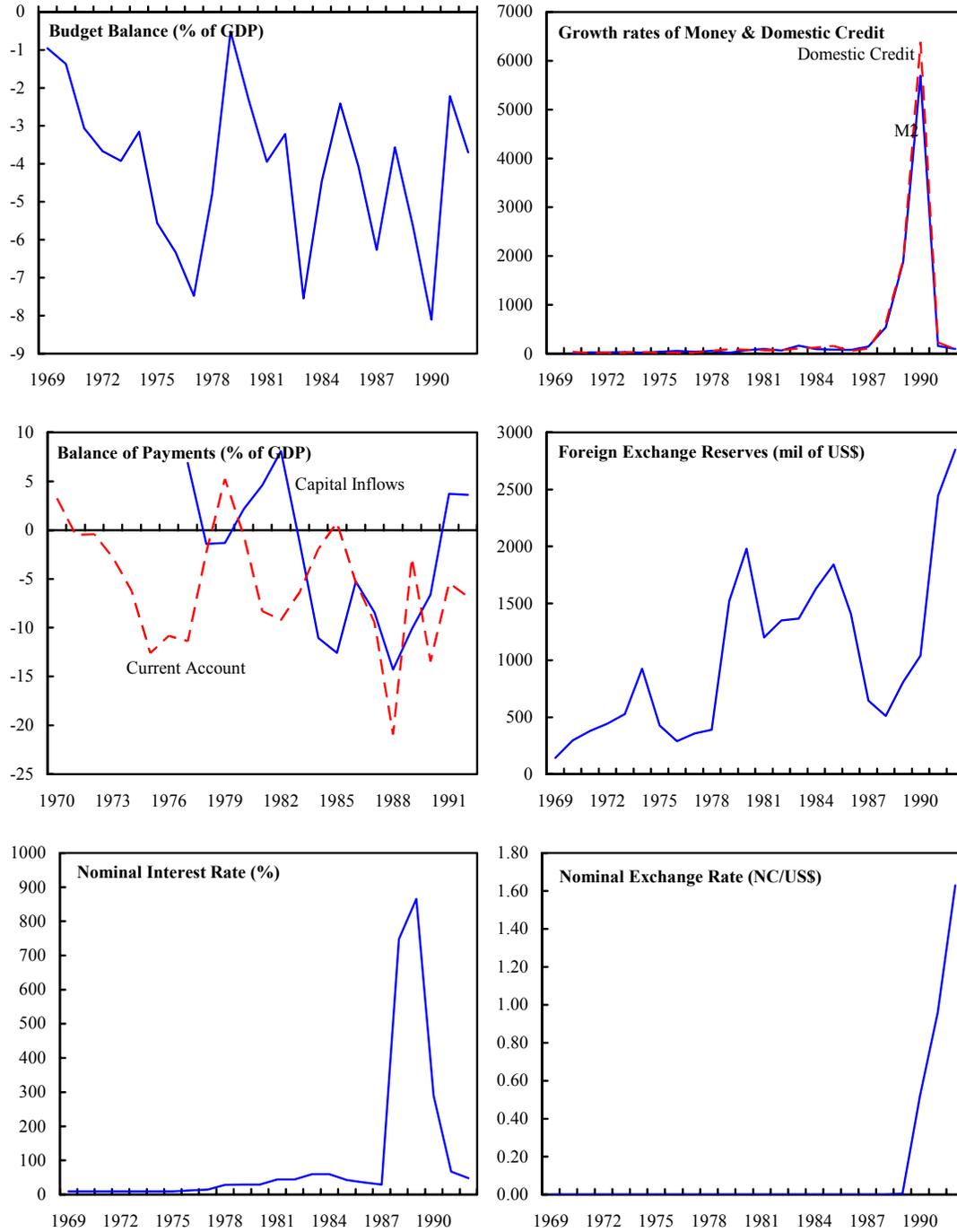
Crises dates: 1983; 1986; 1989-90

Figure 3. Chile: First Generation Model



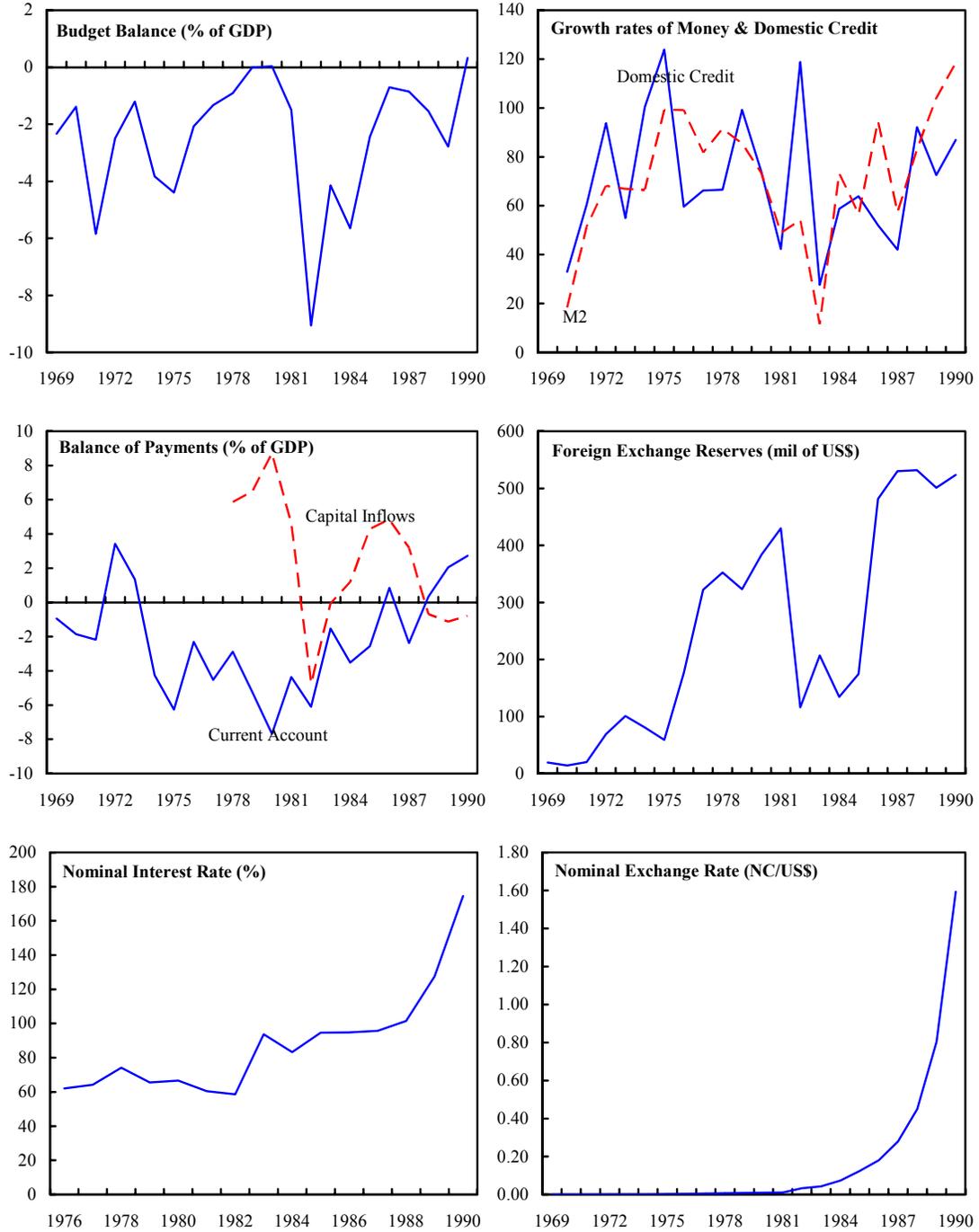
Crisis date: 1971-74

Figure 4. Peru: First Generation Model



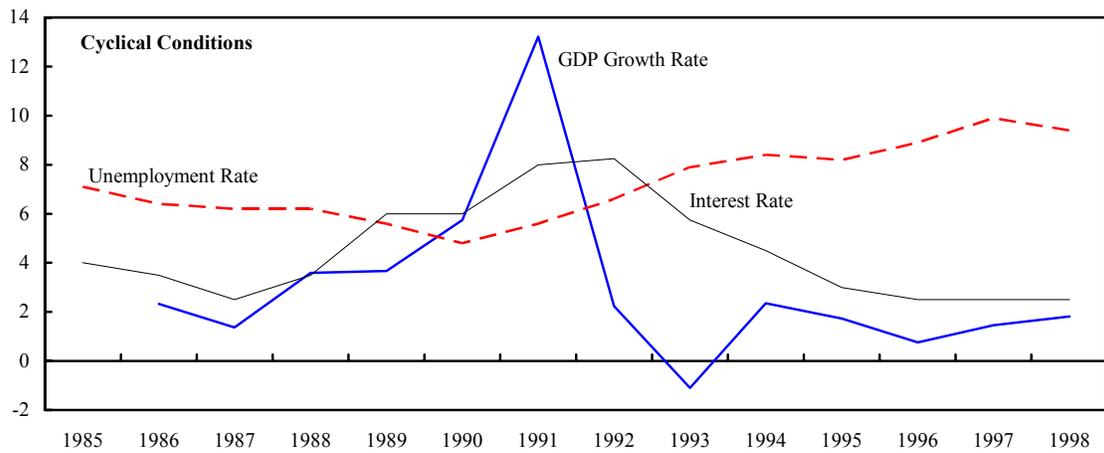
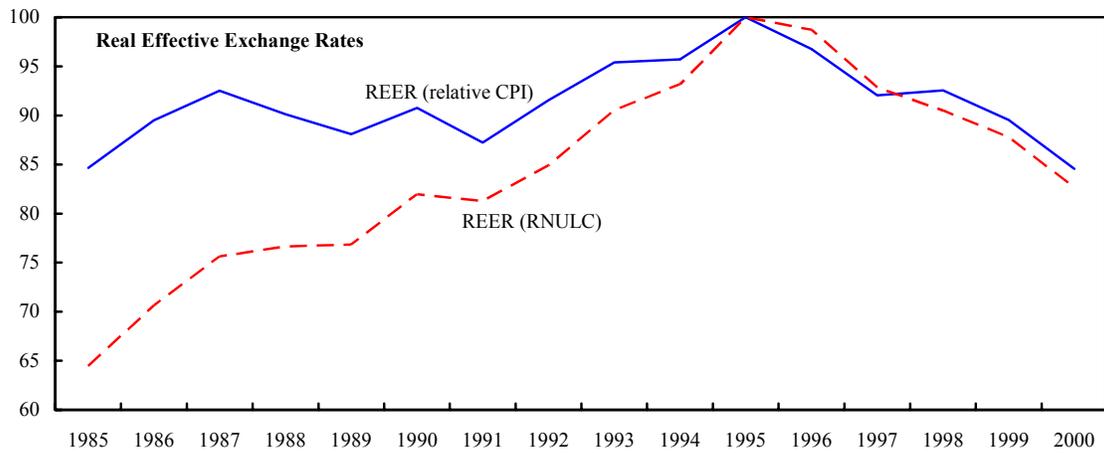
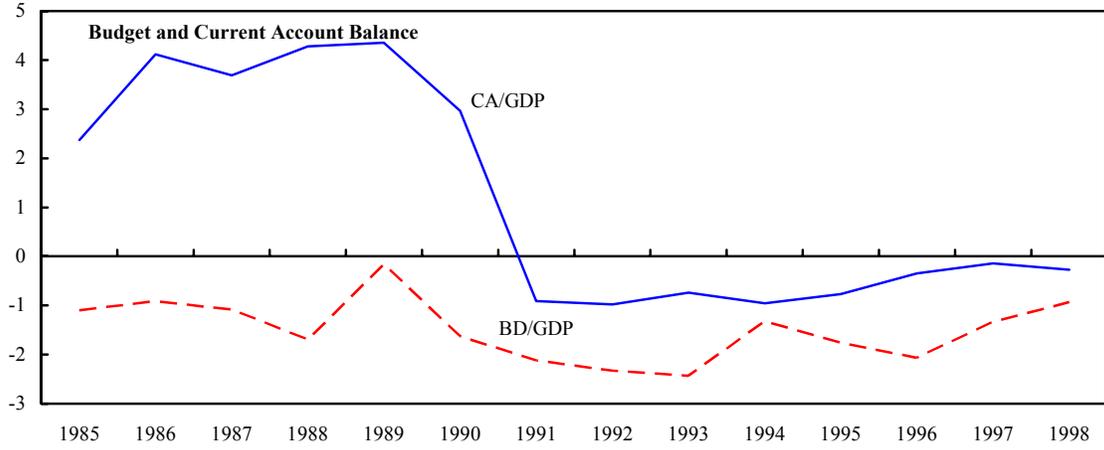
Crises dates: 1976; 1987

Figure 5. Uruguay: First Generation Model



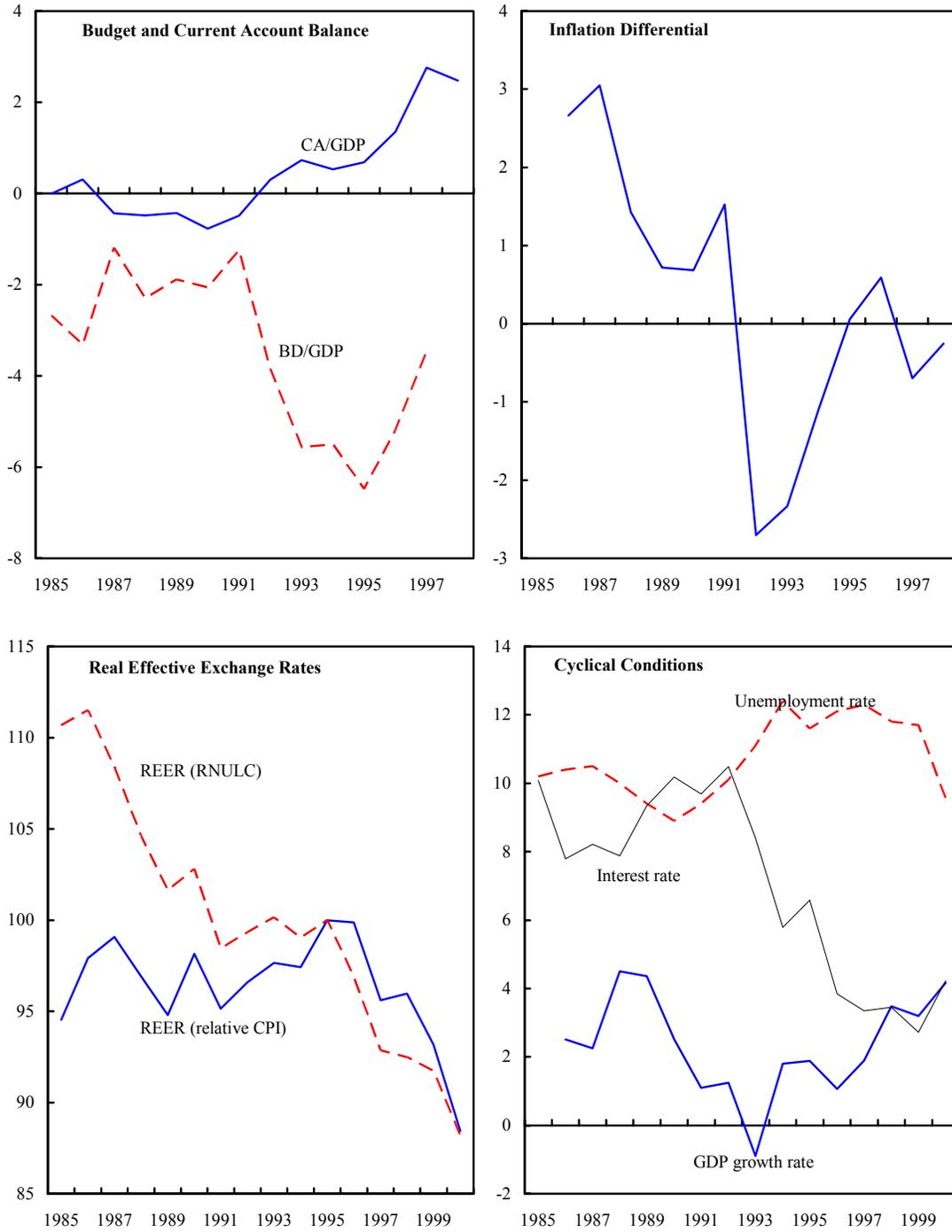
Crisis date: 1982

Figure 6. Germany: Second Generation Model



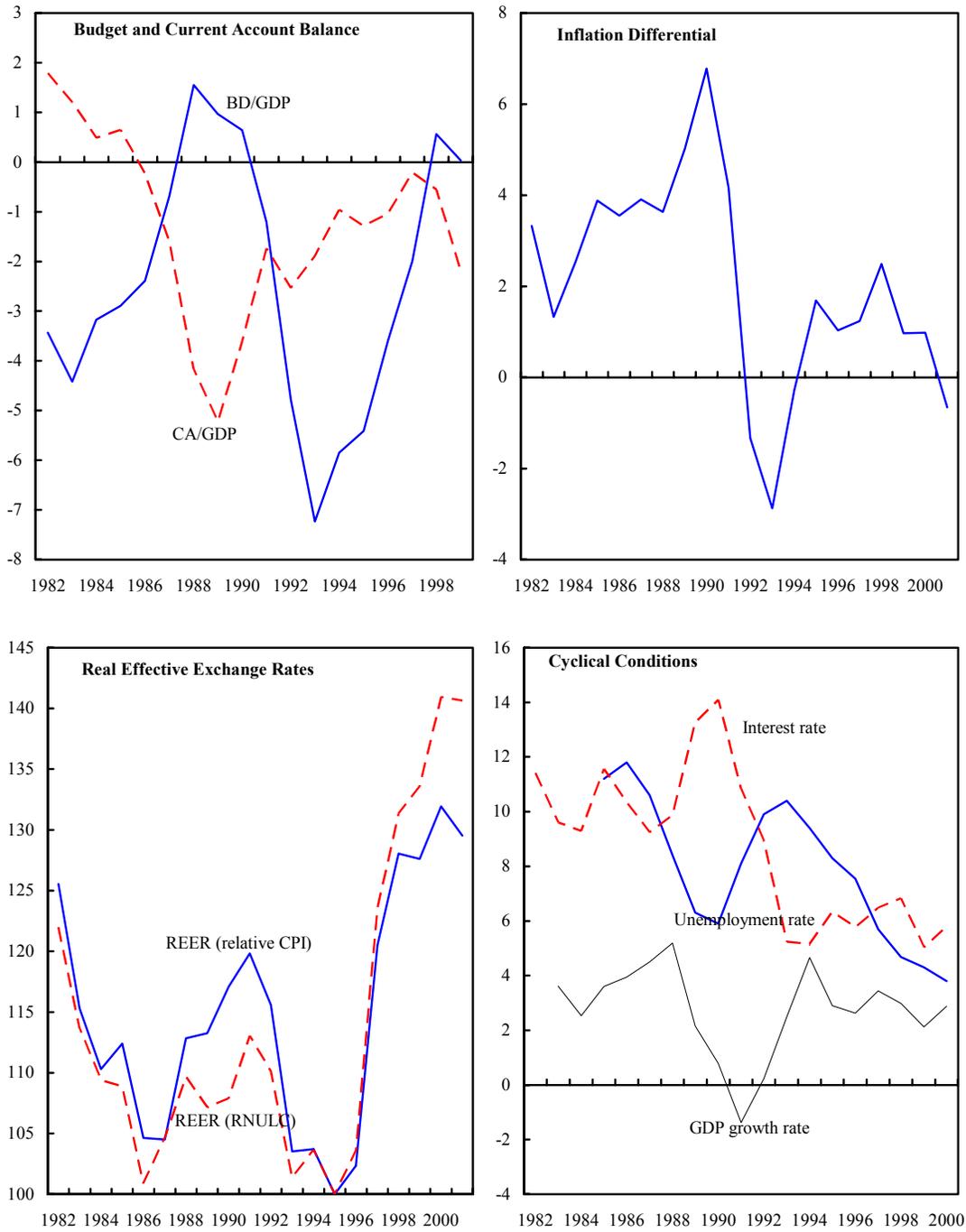
Crisis data: 1992-93

Figure 7. France: Second Generation Model



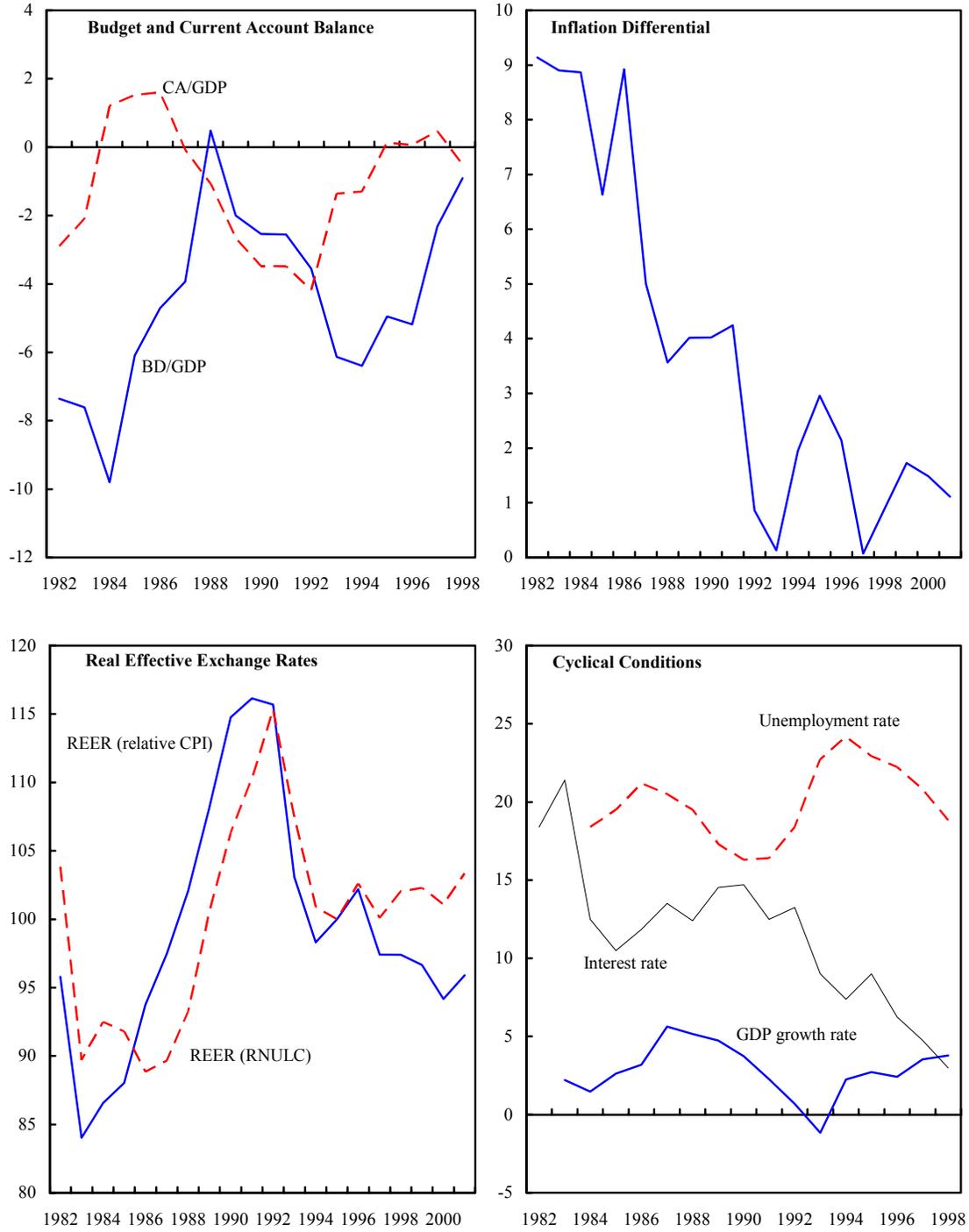
Crisis date: 1992-93

Figure 8. United Kingdom: Second Generation Model



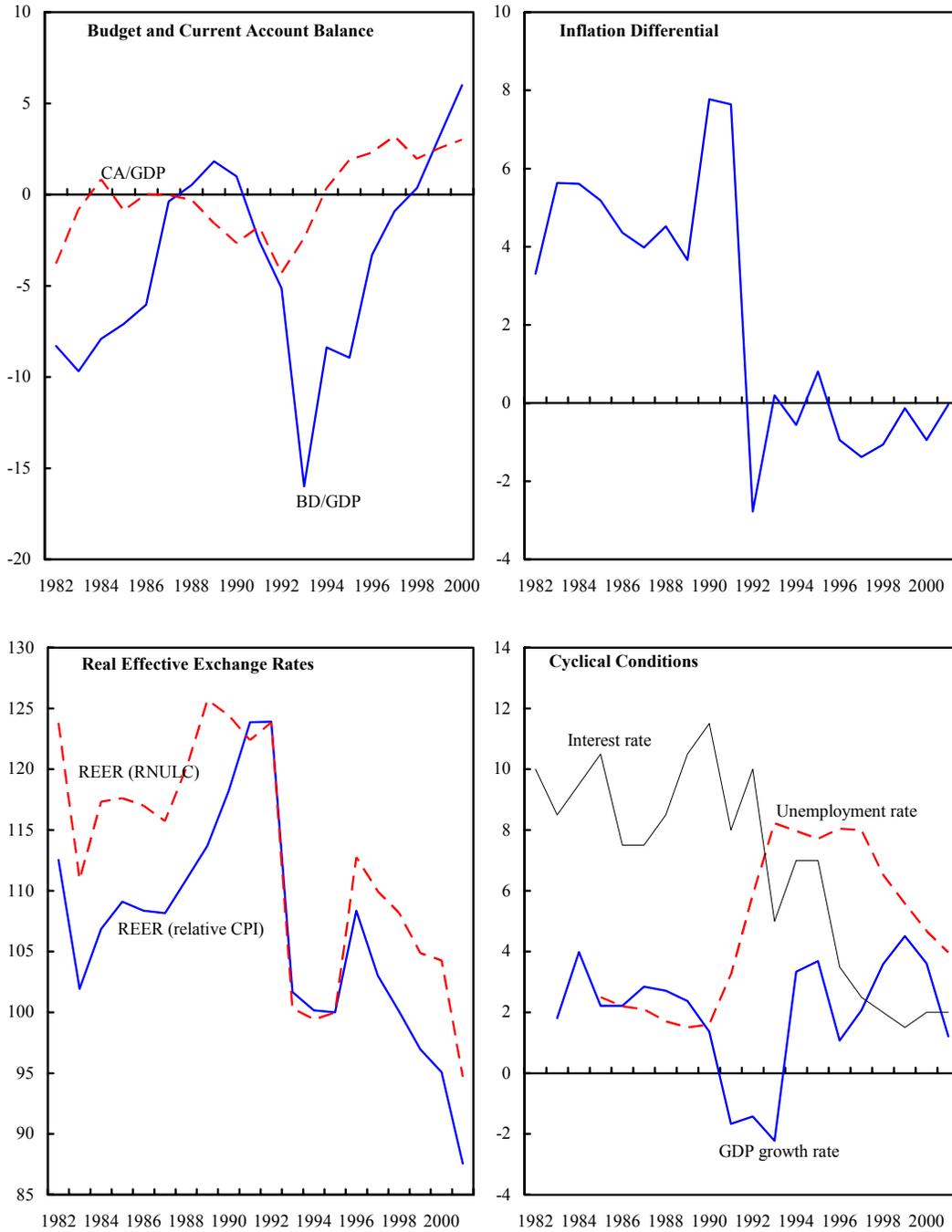
Crisis date: 1992-93

Figure 9. Spain: Second Generation Model



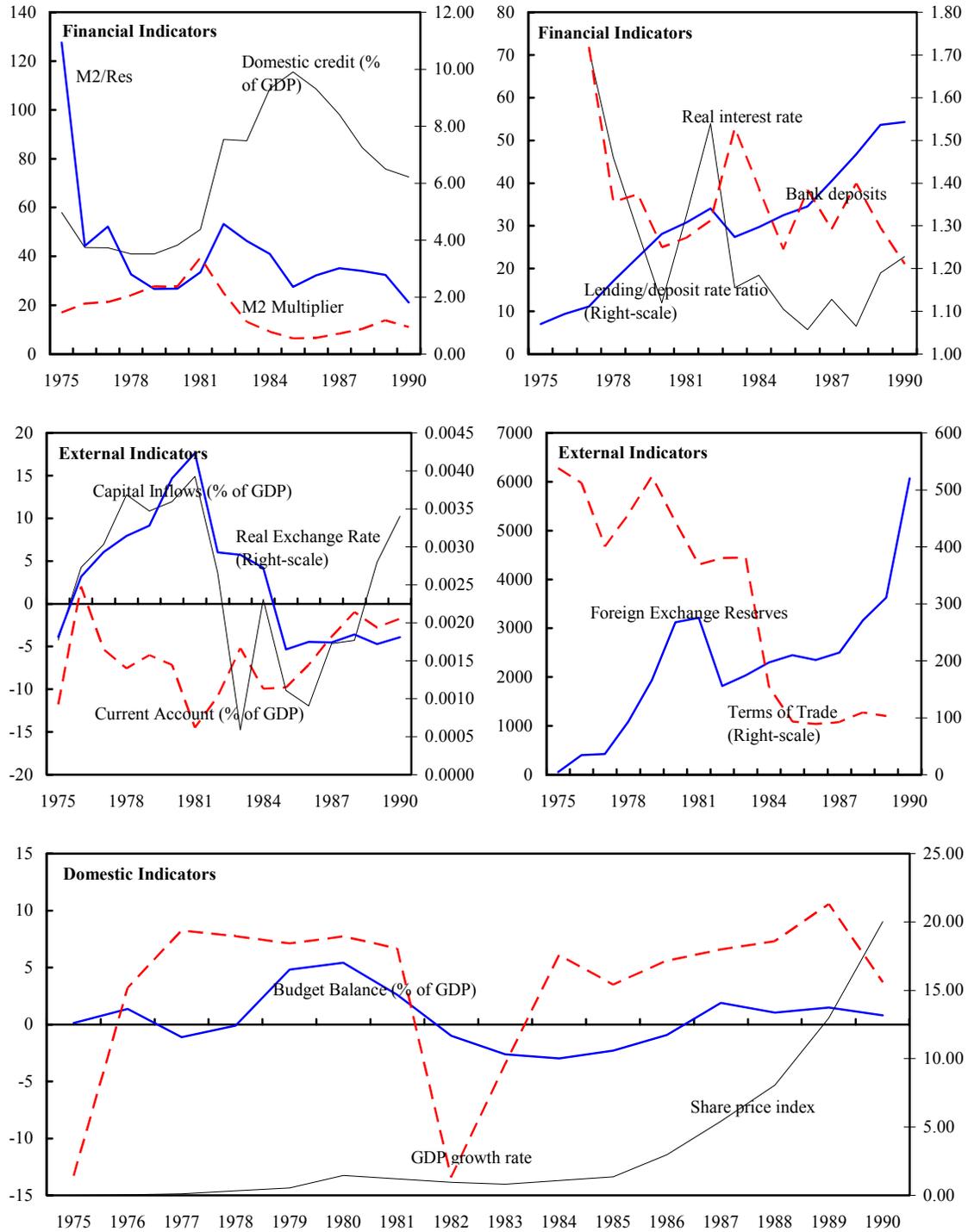
Crisis date: 1992-93

Figure 10. Sweden: Second Generation Model



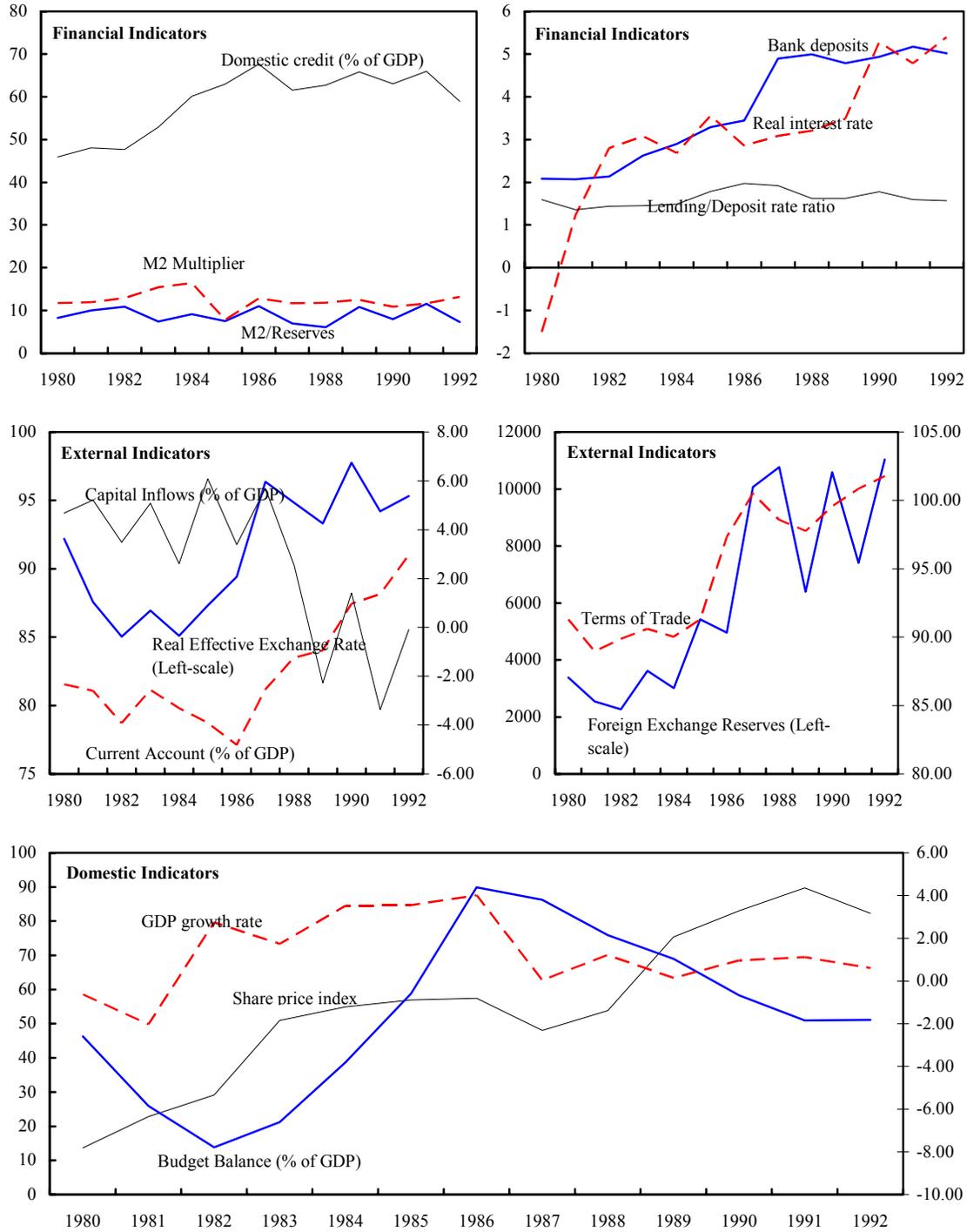
Crisis date: 1992-3

Figure 11. Chile: Twin Crises



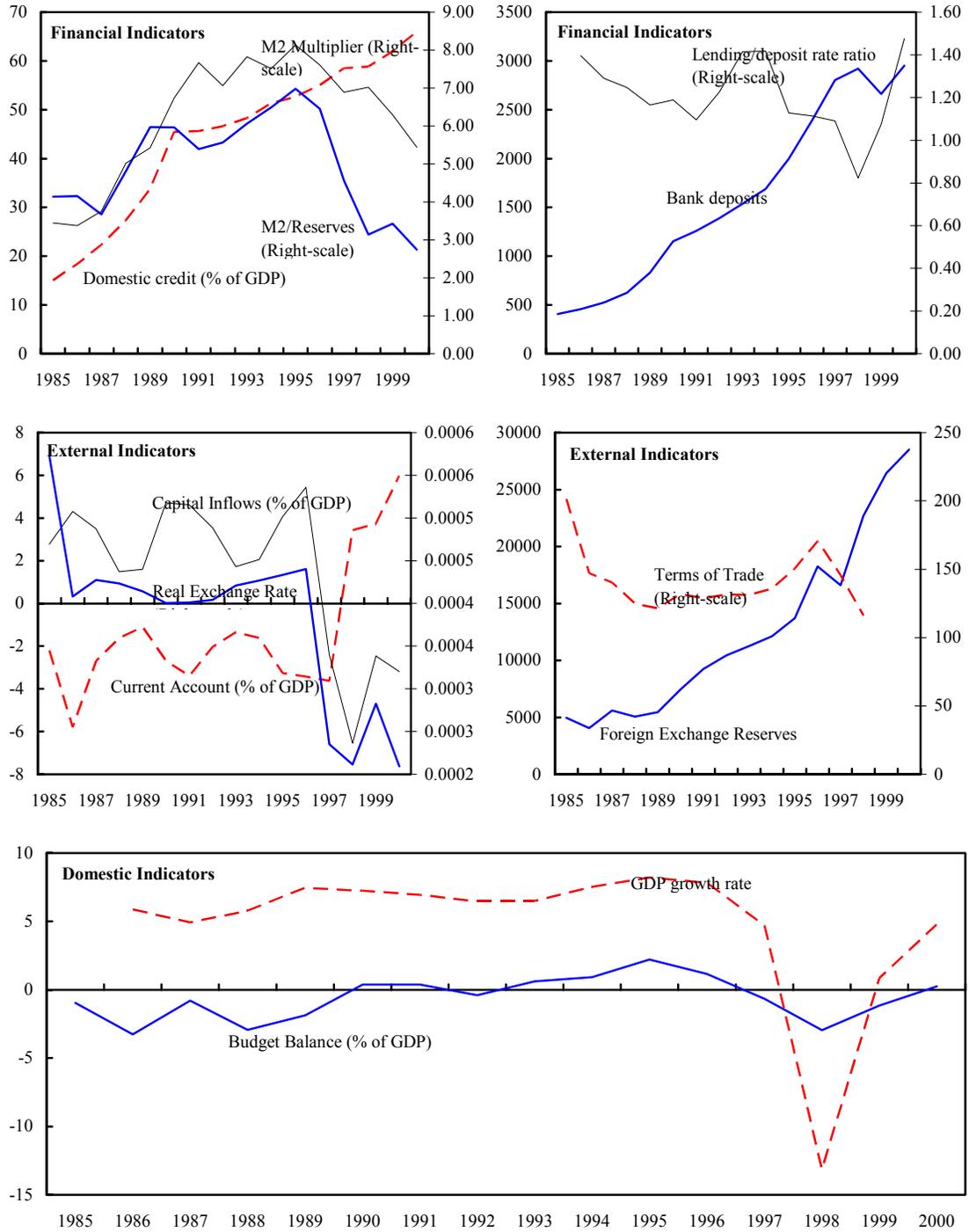
Crisis date: 1981-83

Figure 12. Denmark: Twin Crises



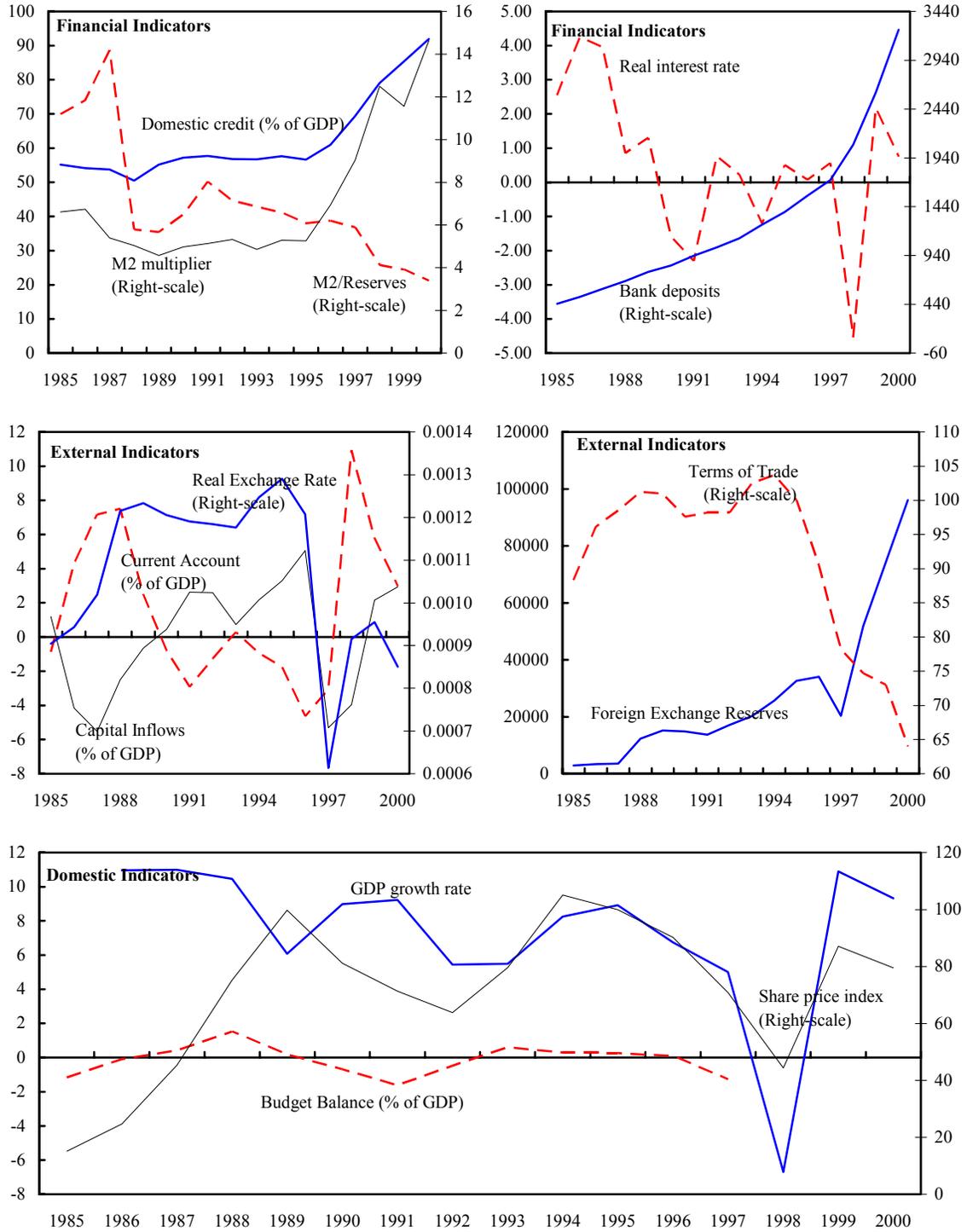
Crisis date: 1987-90

Figure 13. Indonesia: Twin Crises



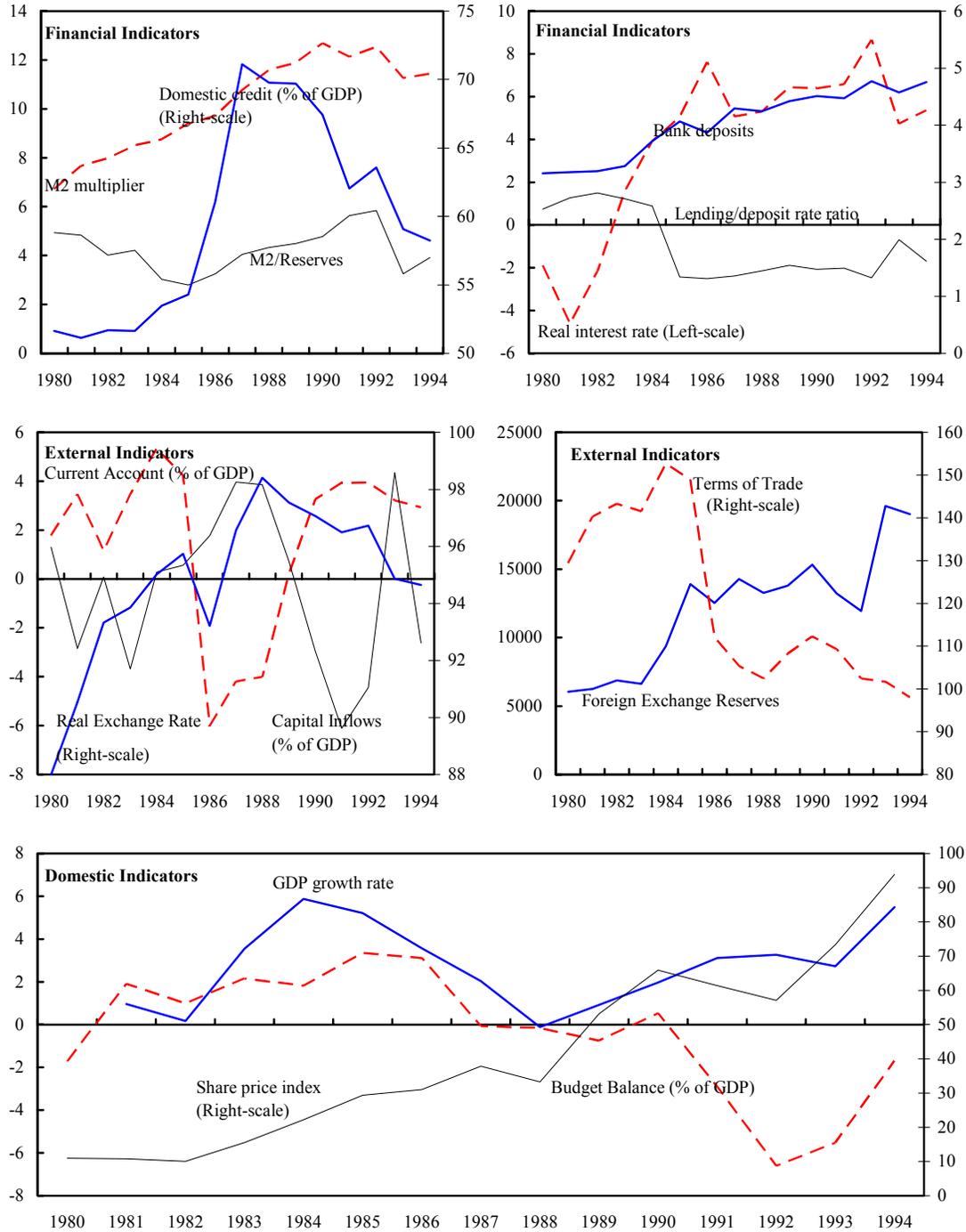
Crisis date: 1997

Figure 14. Korea: Twin Crises



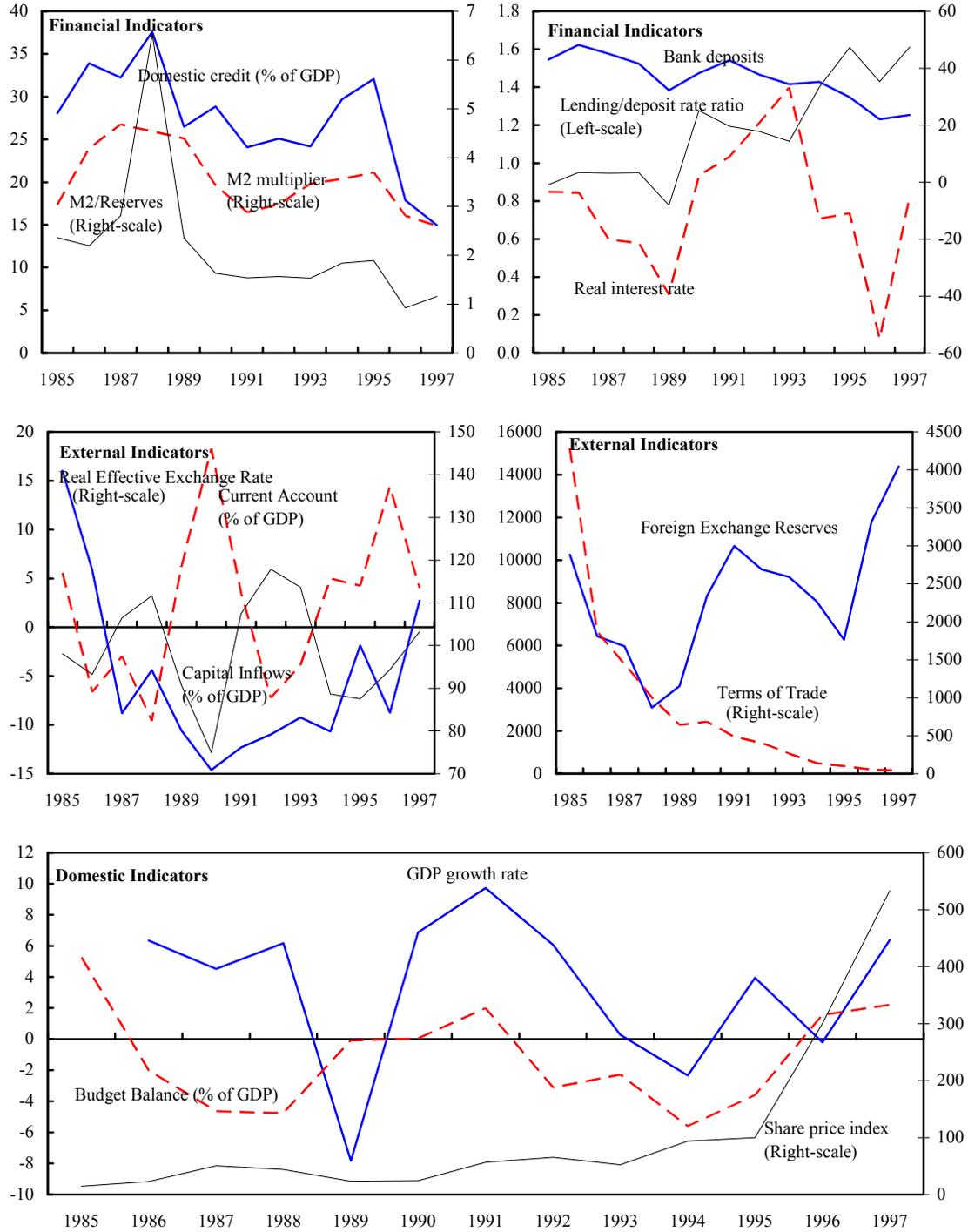
Crisis date: 1997

Figure 15. Norway: Twin Crises



Crisis date: 1988-91

Figure 16. Venezuela: Twin Crises



Crises dates: 1989; 1993-94

Balance sheet recession is the reason for secular stagnation

Richard C Koo

Nomura Research Institute

The Great Recession is often compared to Japan's stagnation since 1990 and the Great Depression of the 1930s. This chapter argues that the key feature of these episodes is the bursting of a debt-financed asset bubble, and that such 'balance sheet recessions' take a long time to recover from. There is no need to suffer secular stagnation if the government offsets private sector deleveraging with fiscal stimulus. However, until the general public understands the fallacy of composition, democracies will struggle to implement such policies during balance sheet recessions.

With the developed economies failing to regain forward momentum after six years of zero interest rates, people are beginning to worry that they may be facing secular stagnation. Although this is an understandable reaction, a large part of the stagnation may be due to the balance sheet recession that these economies are all facing after the bursting of their debt-financed asset price bubbles. And this type of recession takes a long time to overcome, for both economic and political reasons.

The mechanics of balance sheet recession

On the economic front, when a debt-financed bubble bursts, a large number of businesses and households realise that the liabilities they incurred during the bubble days are still on their books, while the assets they bought with borrowed funds have collapsed in value, leaving their balance sheets deep underwater. In order to climb out of their negative equity territory, they have no choice but to pay down debt with their

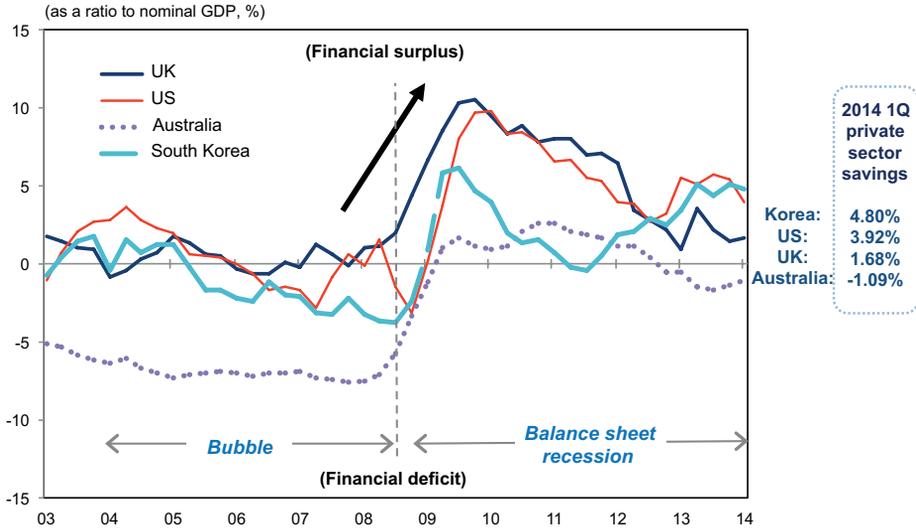
cash flow as quickly and quietly as possible. In other words, they are minimising debt instead of maximising profits.

Although this is the right thing to do for individual businesses and households, when everybody does it at the same time the economy falls into a massive fallacy of composition problems. This is because in a national economy, if someone is saving money or paying down debt, someone else must be borrowing and spending the same amount for the economy to move forward.

In the usual economy that task is borne by the financial sector, which has the incentive to lend or invest all the funds entrusted to it in order to maximise profits. And the usual mechanism to make sure that all saved funds are borrowed and spent is the interest rate; when there are too many borrowers, interest rates are raised and when there are too few rates are lowered.

But after the bursting of a nationwide asset price bubble, those with balance sheets under water are not interested in increasing borrowing at any interest rate. There will not be many lenders either, especially when the lenders themselves have balance sheet problems. The lack of borrowers means a significant portion of the newly saved and deleveraged funds that are entrusted to the financial sector are unable to re-enter the real economy. This in turn means that those unborrowed savings become a leakage in the income stream and a deflationary gap for the economy. If left unattended, this deflationary gap will push the economy ever deeper into balance sheet recession, a highly unusual recession that happens only after the bursting of a nationwide asset price bubble.

Figure 1 The US, UK, South Korean & Australian private sectors are deleveraging after the bubble^{1,2}

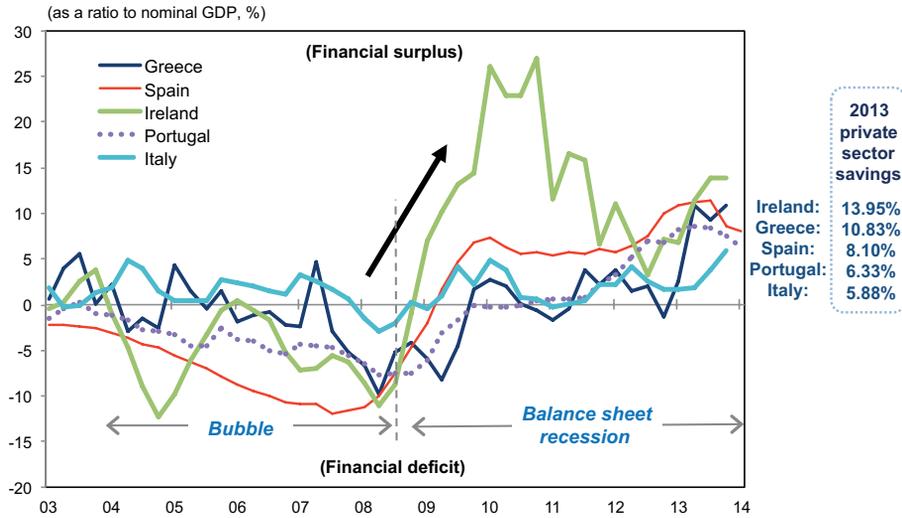


Notes: 1. Private Sector = Household Sector + Non-Financial Corporate Sector + Financial Sector. 2. All entries are four-quarter moving averages. For the latest figures, four-quarter averages ending in 1Q/14 are used.

Sources: Office for National Statistics, UK, FRB, Australian Bureau of Statistics and Bank of Korea.

The flow of funds data for the developed countries that experienced housing bubbles until 2008 indicate that, except for Australia, their private sectors are all in financial surplus, i.e. they are either saving money or paying down debt (Figures 1 and 2). The fact that they are saving money or paying down debt instead of borrowing at zero interest rates means the private sectors of all of these countries are facing severe balance sheet challenges. The same pattern is observed in the private sectors in Japan after the bursting of its massive real estate and equity bubbles in 1990 and in Germany after the bursting of its telecom bubble in 2000 (Figure 3).

Figure 2 Eurozone private sectors are deleveraging massively after the bubble^{1,2}



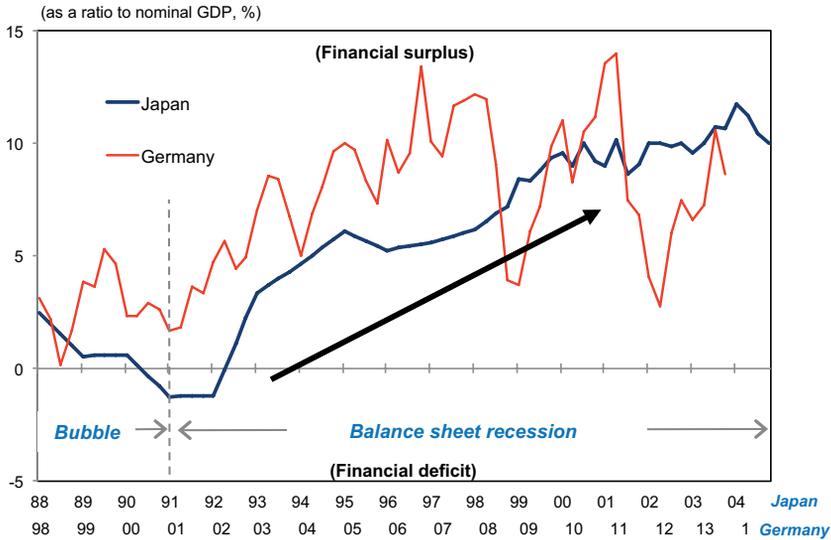
Notes: 1. Private Sector = Household Sector + Non-Financial Corporate Sector + Financial Sector. 2. All entries are four-quarter moving averages. For the latest figures, four-quarter averages ending in 4Q/13 are used for Ireland, Greece and Italy, and those ending in 1Q/14 are used for Spain and Portugal.

Sources: Bank of Greece, Banco de España, National Statistics Institute, Spain, The Central Bank of Ireland, Central Statistics Office Ireland, Banco de Portugal, Banca d'Italia and Italian National Institute of Statistics.

During this type of recession, monetary policy is largely ineffective because, as stated earlier, those with balance sheets under water will not increase borrowing at any interest rate, and financial institutions are also not allowed to lend to those borrowers with balance sheets under water. In addition, the government cannot tell the private sector not to repair its balance sheets because the private sector has no choice but to put its financial houses in order.

This means the only thing the government can do to offset the deflationary forces coming from private sector deleveraging is to do the opposite of the private sector, i.e. borrow and spend the unborrowed savings in the private sector. In other words, fiscal stimulus becomes absolutely essential during this type of recession.

Figure 3 The Japanese and German private sectors also deleveraged after their bubbles^{1,2}



Notes: 1. Private Sector = Household Sector + Non-Financial Corporate Sector + Financial Sector. 2. All entries are four-quarter moving averages. The latest figure for Germany is four-quarter moving average ending in 4Q/13.

Sources: Bank of Japan, Cabinet Office, Japan, Bundesbank and Eurostat.

If the government promptly borrowed and spent the unborrowed savings in the private sector, there would be no leakage in the income stream and the GDP level will be maintained. If the GDP level is maintained, the private sector will have the income to pay down debt. Since asset prices will not fall below zero, as long as the private sector has the income to pay down debt, the balance sheet problem will eventually be resolved.

This also means the government must *sustain* the fiscal stimulus for years until the private sector has finished repairing its balance sheets and has become ready to borrow again. Any premature withdrawal of fiscal stimulus would unleash the deflationary forces, as unborrowed savings are allowed to become a leakage in the economy's income stream. Indeed, the US in 1937, Japan in 1997 and the UK and Eurozone in 2010 all experienced serious double-dip recessions when their governments pursued fiscal consolidation while their private sectors were still in the process of repairing balance sheets.

The forward momentum of economies during this period is necessarily weak because a large part of corporate cash flow is directed towards paying down debt instead of towards research and new product development. Even if corporate research departments are coming up with new ideas and products, their management may be unable to put them into production because of the need to first repair their balance sheets. Many Japanese companies lost their lead to foreign competitors during the last 20 years due to this reason.

Many households will also be rebuilding the savings they thought they had prior to the bubble bursting. That means they will be cutting down on purchases of all kinds, but especially those on credit. The fact that the household sectors of virtually all developed countries have become huge net savers after 2008, in spite of record low interest rates, made even those businesses with clean balance sheets extremely cautious to invest in new capacity.

Recovery from balance sheet recession takes time

When the economy is confronting a fallacy of composition problem affecting a large part of society, the burden cannot be easily shifted to another group. If the government decides to waive all debt for insolvent businesses and households, for example, the problem merely shifts to the entities that lent them the money, i.e. banks and depositors. This means the only option is to wait for the whole of society to get better, a process that takes time.

In a balance sheet recession, the affected businesses and households must use fresh *flows* of savings to slowly repair their balance sheets burdened by the *stock* of excessive debt. The greater the damage to balance sheets, the more time it takes to clean them up. For example, if a company has a \$10 million hole in its balance sheet and can generate \$2 million a year in cash flow that can be used to pay down debt, the repair process will take five years.

But as more firms embark on this process and start to use a large part of their free cash flows to pay down debt, the recession worsens, squeezing cash flow and leading to further declines in the asset prices that triggered the recession in the first place. That is why the government – which is outside the fallacy-of-composition problems – has to proactively take the other side of the bet, so to speak, from the private sector and prevent a vicious cycle. If the government makes the mistake of opting for fiscal consolidation too soon, a recession that people expected would end in two or three years – like Japan’s in 1997 – may persist for seven years, or even ten.

Even after the balance sheets are repaired, people who were forced to deleverage for an extended period of time tend to experience a kind of debt-related trauma that acts as a psychological block to borrowing, even after they have cleaned up their balance sheets. The Americans who had to pay down debt during the Great Depression – the balance sheet recession par excellence – never borrowed money until they died. Even after US private sector balance sheets were repaired thanks to the astronomical government spending of World War II, it still took until 1959 (i.e. full three decades) for US interest rates to return to the average level of 1920s.

The Japanese finished repairing their corporate balance sheets by 2005, but there is no sign that they are resuming their borrowing in spite of the lowest interest rates in human history and the most willing bankers. And that is true even after one full year of Abenomics, which included massive monetary easing.

Democracies are ill-equipped to deal with balance sheet recessions

On the political front, the unfortunate fact is that democracies are ill-equipped to handle such recessions. For a democracy to function properly, people must act based on a strong sense of personal responsibility and self-reliance. But this principle runs counter to the use of fiscal stimulus, which involves depending on ‘big government’ and waiting for a recovery. During a balance sheet recession, people with good incomes and sound

balance sheets will vociferously object to fiscal stimulus and with it the implications of big government, especially once they learn that the stimulus will help rescue people and institutions that participated in the bubble.

Moreover, most people are not aware that this kind of recession is triggered by fallacy-of-composition problems that occur when individual businesses and households begin doing the right and responsible thing by repairing their balance sheets. When the government tries to administer fiscal stimulus, the media, pundits and ordinary citizens who do not understand balance sheet recessions are quick to argue that politicians are wasting taxpayers' money on useless projects to win re-election.

For the past 20 years, the Japanese media and orthodox academics have self-righteously and almost reflexively equated fiscal stimulus with pork-barrel politics. In the US, members of the Tea Party, the Republican Party splinter group that has become so influential, have effectively staked their political careers on preventing the federal government from undertaking fiscal stimulus. German Chancellor Angela Merkel's decision to force through a fiscal compact calling on all Eurozone countries to follow Germany's example and pursue fiscal consolidation was based on a similar philosophy. Since these people were never exposed to the concept of balance sheet recession at university, it is difficult to convince them of the need for fiscal stimulus to cure a disease they have never heard of.

The point is that it is almost impossible to maintain fiscal stimulus in a democracy during peacetime. It is difficult in a democracy because such policies cannot be implemented and maintained during peacetime unless millions of people are persuaded of the need for fiscal stimulus. In contrast, in an autocratic state, only one person – the dictator – needs to be persuaded in order to both administer and maintain fiscal stimulus.

It is difficult in peacetime because during war, when a nation's survival is at stake, no one complains about government spending on armaments or air-raid shelters. There is no danger of getting bogged down in endless debates over how to spend the money, because the answer to that question during wartime is clear to all involved.

Adolf Hitler and Franklin Roosevelt were both elected in 1933 when Germany and the US were in severe balance sheet recessions. The German unemployment rate reached 28% that year and US rate was not far behind at 25%. Although both started with fiscal stimulus, Roosevelt, worried about the criticisms from deficit hawks, reversed course in 1937, resulting in a serious double-dip recession and the unemployment rate increasing to nearly 20% again. Hitler, on the other hand, stayed the course and by 1938, German unemployment had fallen to 2%. And nothing is worse than a dictator with a wrong agenda having the right economic policy, especially when the democracies around him are held hostage to the orthodoxy and remain unable to adopt correct policies.

More recently, the Chinese government implemented a 4 trillion renminbi fiscal stimulus in November 2008 when it was facing a sharp fall in both domestic asset prices and exports. As a percentage of GDP, the stimulus was more than double the size of President Obama's \$787 billion package unleashed three months later. At the time, western observers were laughing when the Chinese government announced that they are going to maintain 8% growth. China's growth soon reached 12%, and nobody was laughing.

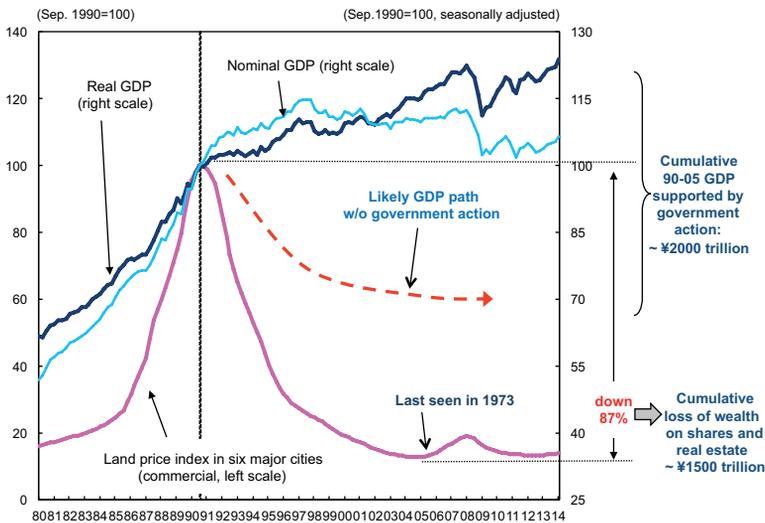
The US government, on the other hand, was extremely cautious with its fiscal stimulus because of the fear that the stimulus package might be criticised for wasting money. As a result, it could not offer the kind of positive jolt its designers have hoped for. The Obama Administration's inability to renew the fiscal stimulus package due to Republican opposition slowed down the subsequent US recovery in no small way.

It is actually not difficult to implement fiscal stimulus when a country experiences a major shock (like the Lehman failure and the Global Crisis). The challenge is whether it can be kept in place long enough for the private sector to finish repairing balance sheets.

At the emergency G20 meeting held in Washington two months after Lehman Brothers collapsed, the 20 nations agreed to administer a dose of fiscal stimulus – a decision attributable in no small part to the efforts of Japanese Prime Minister Taro Aso. Formerly

a corporate executive, Aso was one of the few Japanese politicians who understood that fiscal stimulus was the key to maintaining Japanese GDP when the private sector was saving 8% of GDP at zero interest rates. And at the G20 meeting, he used Figure 4 to tell the leaders of the other 19 countries that Japan was able to maintain its GDP at above the bubble peak for the entire post-bubble period with fiscal stimulus, in spite of commercial real estate prices falling 87% from the peak to the level of 1973.

Figure 4 Japan’s GDP grew despite major loss of wealth and private sector deleveraging



Sources: Cabinet Office, Japan Real Estate Institute

The G20 ultimately agreed to administer fiscal stimulus in 2009, and the global economy staged a V-shaped recovery instead of falling into a depression, as had been feared. But as soon as the economy started to show signs of life, deficit hawks took over the policy debate.

Those who prevent crises never become heroes

It is often said that people who prevent crises never become heroes. Hollywood teaches us that for there to be a hero there must first be a crisis, and the experience of Prime Ministers Taro Aso and Gordon Brown bears that out.

The Japanese media, for example, completely missed the significance of Aso's contribution at G20 in November 2009. Instead, they tried to portray his administration as a care-taker government ahead of the general election scheduled for 2009, and devoted a great deal of coverage to the prime minister's misreading of a single Chinese character in one speech. Partly as a result of such publicity, the LDP was defeated in the election held in August 2009. The UK Prime Minister Gordon Brown, another leader who understood what a balance sheet recession was and used fiscal stimulus to address it, was also defeated in his quest for re-election.

At the Toronto Summit in 2010 – with both Aso and Brown, who had *prevented* crises, out of the picture – the G20 leaders agreed on a plan to halve their fiscal deficits in three years. This in spite of the fact that the private sectors in these countries continued to save massively despite near zero interest rates.

The resulting fiscal retrenchment sent the developed economies into reverse, with the UK and many parts of the Eurozone falling into double-dip recessions. Japan, under the new DPJ government that understood nothing of balance sheet recessions, also stagnated.

In the US, however, Federal Reserve Chairman Ben Bernanke and others soon realised that the Toronto agreement had been a mistake. They kept the US from pursuing premature fiscal consolidation by issuing the warning with the expression 'fiscal cliff', thereby making it the first country to renege on the agreement. Consequently, the US – alone among the developed economies – continued to post modest economic growth, while Japan, the UK, and continental Europe faced severe economic weakness.

Partly because of subsequent reflection on this error, the pendulum had swung back towards the recognition of the importance of fiscal stimulus by the time the St. Petersburg G20 Summit was held in 2013. Although the three years following the Toronto Summit were completely wasted from a global economic perspective, at least there has been some recognition among policymakers that fiscal stimulus is important in this type of recessions. The risk remains, however, that this will turn out to be just another phase in an on/off cycle of fiscal stimulus in democracies during peace time.

The above examples show that there is no need to suffer 'secular stagnation' if proper policies are put in place, but that democracies are very bad at implementing such policies during balance sheet recessions. This predicament will stay with democracies until the general public (the millions) is made aware of the disease called balance sheet recession and how to cure it. Until then, the far-from-ideal on/off cycle of fiscal stimulus and the resultant delayed recovery will make people feel like they are in secular stagnation.

Buttonwood

Firefighting

The sovereign-debt crisis has echoes of the ERM debacle

Jul 14th 2011 | [From the print edition](#)



AS ANYONE who has watched “The Towering Inferno” will know, what starts as a small electrical fire in a cupboard can eventually consume an entire skyscraper. When Greece's debt problems emerged last year, some dismissed them as trivial, given the country's 2.5% weight in the euro area's GDP. Similarly, problems in the subprime-mortgage market were once regarded as too obscure to affect the entire American economy.

The failure to extinguish the Greek flames has allowed the conflagration to spread to Ireland, Portugal and, intermittently, to Spain. Now smoke has been seen in the Italian bond market. On

June 12th Italian ten-year bond yields briefly rose above 6%; the spread (excess interest rate) over German Bunds was more than three percentage points.

The sell-off may have been so rapid because investors previously assumed that Italy was behind the firewall. They were hungry for yield but too cautious to buy the risky bonds of Greece, Ireland and Portugal. So they loaded up on Italian debt as a higher-yielding alternative to that of the core European countries of Germany and the Netherlands.

Although the markets stabilised later in the week, it is too early to assume the crisis is over. A key problem has not been solved: how will the most indebted European nations reduce their debt burdens without faster economic growth?

In many ways, the crisis resembles the demise of the Exchange Rate Mechanism (ERM), a system of currency pegs, in the early 1990s. That episode also illustrated the problems involved when European countries try to tie their currencies to the German powerhouse. In both cases, there has been a domino effect as problems in one country spread to others. In the early 1990s Italy was the first to fall; after it toppled, Britain, Spain, Portugal and eventually France were forced to follow.

Back then European politicians blamed declines in currencies on speculators who failed to understand that countries were committed to maintaining their exchange-rate parities. They have taken a similar line in the current crisis. In May 2010 Angela Merkel, the German chancellor, said: "We must re-establish the primacy of politics over the markets." But the markets are often sending useful signals about the unsustainability of economic policies. If you don't want to worry about the views of bond investors, don't borrow a lot of money from them.

In the early 1990s governments had to choose between defending their exchange rates and the health of their economies. Now they have to choose between disappointing the bond markets and imposing austerity on their voters.

For Britain, the defining moment of the crisis came at lunchtime on September 16th 1992 (also known as Black Wednesday) when an interest-rate increase of three percentage points, the second rise of the day, failed to prop up the pound. Stockmarkets began rallying within minutes as traders worked out that the then Conservative government could never sustain such high rates. The price in unemployment and distress in the housing market would be too great. Sure enough, by the evening Norman Lamont, the then finance minister, announced that Britain would leave the ERM.

The effects were almost entirely benign. The British economy recovered and there was no sign of the surge in inflation that had been predicted as a consequence of devaluation. But European

politicians concluded from this example that the answer was to remove the influence of foreign-exchange markets within Europe by pushing ahead with the single currency.

As a result of this, the choices this time round are much starker. Instead of tightening monetary policy to defend the currency, governments have to tighten fiscal policy to try to keep the lid on their borrowing costs. This will involve many years of painful austerity.



Explore our [interactive guide](#) to Europe's troubled economies

The consequences of default, which could include a banking collapse and inaccessible capital markets, would be far more painful than those of devaluation. That is why peripheral governments in the euro zone have so far swallowed their medicine, although not without a lot of grumbling and plenty of financial help from their neighbours.

None of this is popular, however, in the countries that give and receive financial aid. There may yet come a Black Wednesday moment when politicians cannot ask any more of their voters. Perhaps it will be the parliamentary defeat of an austerity budget or the election of an extreme populist party. Then the markets really will be set ablaze.

Economist.com/blogs/buttonwood

[From the print edition: Finance and economics](#)

Irving Fisher

Out of Keynes's shadow

Today's crisis has given new relevance to the ideas of another great economist of the Depression era

Feb 12th 2009 | WASHINGTON, DC | From the print edition

SHORTLY after he was elected president, Barack Obama sounded a warning: "We are facing an economic crisis of historic proportions...We now risk falling into a deflationary spiral that could increase our massive debt even further." The address evoked not just the horror of the Depression, but one of the era's most important thinkers: Irving Fisher.

Sue Vago



Though once America's most famous economist, Fisher is now almost forgotten by the public. If he is remembered, it is usually for perhaps the worst stockmarket call in history. In October 1929 he declared that stocks had reached a "permanently high plateau". Today it is John Maynard Keynes, his British contemporary, who is cited, debated and followed. Yet Fisher laid the foundation for much of modern monetary economics; Keynes called Fisher the "great-grandparent" of his own theories on how monetary forces influenced the real economy. (They first met in London in 1912 and reportedly got along well.)

As parallels to the 1930s multiply, Fisher is relevant again. As it was then, the United States is now awash in debt. No matter that it is mostly "inside" or "internal" debt—owed by Americans to other Americans. As the underlying collateral declines in value and incomes shrink, the real burden of debt rises. Debts go bad, weakening banks, forcing asset sales and driving prices down further. Fisher showed how such a spiral could turn mere busts into depressions. In 1933 he wrote:

Over investment and over speculation are often important; but they would have far less serious results were they not conducted with borrowed money. The very effort of individuals to lessen their burden of debts increases it, because of the mass effect of the stampede to liquidate...the more debtors pay, the more they owe. The more the economic boat tips, the more it tends to tip.

Though they seldom invoke Fisher, policymakers in America are applying his ideas. In academia Ben Bernanke, now the chairman of the Federal Reserve, sought to formalise Fisher's debt-deflation theory. His research has shaped his response to this crisis. He decided to bail out Bear Stearns in March 2008 partly so that a sudden liquidation of the investment bank's positions did not trigger a cycle of falling asset prices and default. Indeed, some say the Fed has learnt Fisher too well: from 2001 to 2004, to contain the deflationary shock waves of the tech-stock collapse, it kept interest rates low and thus helped to inflate a new bubble, in

property.

Were Fisher alive today, “he would tell us we have to avoid deflation, and to worry about all that inside debt,” says Robert Dimand, an economist at Brock University in Canada, who has studied Fisher in depth. “The ideal thing is to avoid these situations. Unfortunately, we are in one.”

Fisher was born in 1867 and earned his PhD from Yale in 1891. In 1898 he nearly died of tuberculosis, an experience that turned him into a lifelong crusader for diet, fresh air, Prohibition and public health. For a while he also promoted eugenics. His causes, both healthy and repugnant, combined with a lack of humour and high self-regard, did not make him popular.

In 1894, on a trip to Switzerland, he saw, in water cascading into mountain pools, a way to “define precisely the relationships among wealth, capital, interest and income,” Robert Loring Allen, a biographer of Fisher, wrote. “The flowing water, moving into the pool at a certain volume per unit of time, was income. The pool, a given volume of water at a particular moment, became capital.” Over the next 30 years he established many of the central concepts of financial economics.

In 1911, in “The Purchasing Power of Money”, Fisher formalised the quantity theory of money, which holds that the supply of money times its velocity—the rate at which a dollar circulates through the market—is equal to output multiplied by the price level. Perhaps more important, he explained how changing velocity and prices could cause real interest rates to deviate from nominal ones. In this way, monetary forces could produce booms and busts, although they had no long-run effect on output. Furthermore, Fisher held that the dollar's value should be maintained relative not to gold but to a basket of commodities, making him the spiritual father of all modern central banks that target price stability.

During the 1920s Fisher became rich from the invention and sale of a card-index system. He used the money to buy stocks on margin, and by 1929 was worth \$10m. He was also a prominent financial guru. Alas, two weeks after he saw the “plateau” the stockmarket crashed.

To his cost, Fisher remained optimistic as the Depression wore on. He lost his fortune and his home and lived out his life on the generosity of his sister-in-law and Yale. But his work continued. He was prominent among the 1,028 economists who in vain petitioned Herbert Hoover to veto the infamous Smoot-Hawley tariff of 1930. And he developed his debt-deflation theory. In 1933 in *Econometrica*, published by the Econometric Society, which he co-founded, he described debt deflation as a sequence of distress-selling, falling asset prices, rising real interest rates, more distress-selling, falling velocity, declining net worth, rising bankruptcies, bank runs, curtailment of credit, dumping of assets by banks, growing distrust and hoarding. Chart 1 is his: it shows how deflation increased the burden of debt.

Fisher was adamant that ending deflation required abandoning the gold standard, and repeatedly implored Franklin Roosevelt to do so. (Keynes was of similar mind.) Roosevelt devalued the dollar soon after becoming president in 1933. The devaluation and a bank holiday marked the bottom of the Depression, though true recovery was still far off. But Fisher had at best a slight influence on Roosevelt's decision. His reputation had fallen so far that even fellow academics ignored him.

Contemporary critics did poke a hole in his debt-deflation hypothesis: rising real debt makes debtors worse

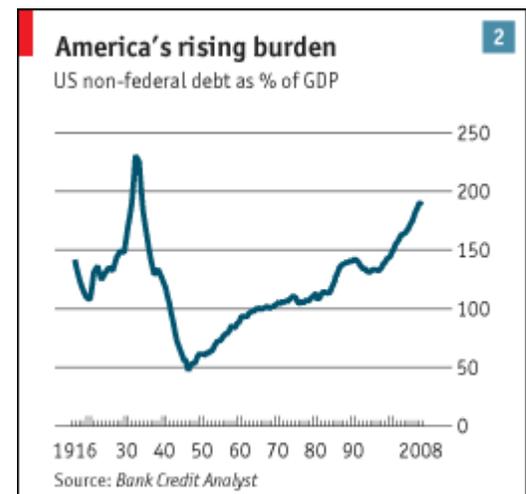
off but creditors better off, so the net effect should be nil. Mr Bernanke plugged this in the 1980s. “Collateral facilitates credit extension,” he said in June 2007, just before the crisis began in earnest. “However, in the 1930s, declining output and falling prices (which increased real debt burdens) led to widespread financial distress among borrowers, lessening their capacity to pledge collateral...Borrowers' cash flows and liquidity were also impaired, which likewise increased the risks to lenders.” Mr Bernanke and Mark Gertler of New York University dubbed this “the financial accelerator”.



The downward spiral can start even when inflation remains positive—for example, when it drops unexpectedly. Consider a borrower who expects inflation of 2% and takes out a loan with a 5% interest rate. If instead inflation falls to 1%, the real interest rate rises from 3% to 4%, increasing the burden of repayment.

Asset deflation can do much the same thing. If house prices are expected to rise by 10% a year, a buyer willingly borrows the whole purchase price, because his home will soon be worth more than the loan. A lender is happy to make the loan for the same reason. But if prices fall by 10% instead, the house will soon be worth less than the loan. Both homeowner and lender face a greater risk of bankruptcy.

Today, debt in America excluding that of financial institutions and the federal government is about 190% of GDP, the highest since the 1930s, according to the *Bank Credit Analyst*, a financial-research journal (see chart 2). There are important differences between then and now. Debt was lower at the start of the Depression, at 164% of GDP. Mortgage debt was modest relative to home values, and prices were not notably bloated: they fell by 24% between 1929 and 1933, says Edward Pinto, a consultant, so were roughly flat in real terms. Debt burdens shot up because of deflation and shrinking output; nominal GDP fell by 46% between 1929 and 1933.



Debt burdens are high today mostly because so much was borrowed in the recent past. This began as a logical response to declining real interest rates, low inflation, rising asset prices and less frequent recessions, all of which made leverage less dangerous. But rising leverage eventually bred easy credit and overvalued homes.

Even without recession, falling home prices would have impaired enough mortgage debt to destabilise the financial system (see chart 3). Recession makes those dynamics more virulent; deflation could do similar damage. Broad price indices fell in late 2008. Granted, that was caused in part by a one-off fall in petrol costs; but America's core inflation rate, which excludes food and energy, has fallen from 2.5% in September to 1.8%. Goldman Sachs sees it falling to 0.25% in the next two years.

That is low enough to mean falling wages for many households and falling prices for many firms. More

widespread and deeper deflation would mean that property prices would have to fall even further to restore equilibrium with household incomes, creating another round of delinquencies, defaults and foreclosures.

What is the solution? Fisher wrote that it was “always economically possible to stop or prevent such a depression simply by reflatting the price level up to the average level at which outstanding debts were contracted.” Alas, reflation is not so simple. Although stabilising nominal home prices would help short-circuit the debt-deflation dynamics now under way, any effort to maintain them at unrealistically high levels (where they still are in many cities) is likely to fail. Higher inflation could help bring down real home prices while allowing nominal home prices to stabilise, and reduce real debt burdens. But creating inflation is easier said than done: it requires boosting aggregate demand enough to consume existing economic slack, through either monetary or fiscal policy.



Though the Fed does not expect deflation, last month it did say that “inflation could persist for a time below” optimal levels. It is mulling a formal inflation target which, by encouraging people to expect positive inflation, would make deflation less likely. But its practical tools for preventing deflation are limited. In December its short-term interest-rate target in effect hit zero. The Taylor rule, a popular rule of thumb, suggests it should be six percentage points below. The Fed is now trying to push down long-term interest rates by buying mortgage-backed and perhaps Treasury securities. With conventional monetary ammunition spent, fiscal policy has become more important.

In 2002 Mr Bernanke argued the government could ultimately always generate inflation by having the Fed finance large increases in government spending directly, by purchasing Treasury debt. Martin Barnes of the *Bank Credit Analyst* thinks this highly unlikely: “You'd have capital flight out of the dollar. The only way it works is if every country is doing it, or with capital controls.”

Fisher died in 1947, a year after Keynes, and remains in his shadow. Mr Dimand notes that Fisher never pulled the many strands of his thought together into a grand synthesis as Keynes did in “The General Theory of Employment, Interest and Money”. More important, Keynes's advocacy of aggressive fiscal policy overcame the limitations of Fisher's purely monetary remedies for the Depression.

Yet Fisher's insights remain vital. They have filtered, perhaps unconsciously, into the thinking of today's policymakers. On February 8th Lawrence Summers, Mr Obama's principal economic adviser, called for the rapid passage of a fiscal stimulus “to contain what is a very damaging and potentially deflationary spiral.” His advice bridges Fisher and Keynes.

From the print edition: Briefing

Free exchange
Economics

Currency crises
Bleeding out

Jun 9th 2010, 13:32 by R.A. | WASHINGTON

THE debt crisis in Europe, it should be clear by now, is as much about exchange rates as it is borrowing. Countries which needed to devalue for years did not, but remained locked in a currency union. This allowed those nations to build up unhealthy capital surpluses amid cheap borrowing, all while loss of domestic competitiveness undermined their productive economic sectors. Finally, as the recession hit, the unsustainable could no longer be sustained, and a breaking point approached. Markets reacted, attacking the troubled nations and forcing a crisis. For now, government interventions have slowed the crisis while the nations in question undertake internal adjustments. But many suspect that it's only a matter of time until the dam breaks.

We've seen this script before. Events played out this way, more or less, in 1992. At the time, the Bundesbank responded to rising German deficits (in the wake of re-unification) by raising interest rates. This put pressure on other countries in the European Exchange Rate Mechanism (a euro pre-cursor). Nations with weak economies were forced to choose between painful domestic interest rate increases—unpalatable given broader economic conditions—and devaluation. Markets betting that governments wouldn't accept domestic economic pain put pressure on currencies. Italy was forced to negotiate a devaluation within the ERM. And then markets turned toward Britain.

In a new book on the history of hedge funds, Sebastian Mallaby dedicates a chapter to the European crisis (full disclosure: Mr Mallaby is married to *The Economist's* Economics Editor). The *Atlantic* has [posted an excerpt](http://www.theatlantic.com/business/archive/2010/06/go-for-the-jugular/57696/1/) (http://www.theatlantic.com/business/archive/2010/06/go-for-the-jugular/57696/1/) from that chapter. It begins:

On Tuesday, September 15, the pound took another beating. Spain's finance minister telephoned Norman Lamont, his British counterpart, to ask him how things were. "Awful," Lamont answered.

That evening Lamont convened a meeting with Robin Leigh-Pemberton, the governor of the Bank of England. The two men agreed that the central bank should buy the pound aggressively the next morning. As the meeting wound down, Leigh-Pemberton read out a message from his press office. Helmut Schlesinger, the president of the German Bundesbank, had given an

interview to the Wall Street Journal and a German financial newspaper, Handelsblatt. According to a news agency report on his remarks, Schlesinger believed there would have to be a broad realignment of Europe's currencies.

Lamont was stunned. Schlesinger's remark was tantamount to calling for the pound to devalue. Already his public statements had triggered an assault on Italy's lira. Now the German central banker was attacking Britain. Lamont asked Leigh-Pemberton to call Schlesinger immediately, overruling Leigh-Pemberton's concern that the punctilious Bundesbanker did not like to have his dinner interrupted...

Oh, those German central bankers.

Druckenmiller walked into Soros's office and told him it was time to move. He had held a \$1.5 billion bet against the pound since August, but now the endgame was coming and he would build on the position steadily.

Soros listened and looked puzzled. "That doesn't make sense," he objected.

"What do you mean?" Druckenmiller asked.

Well, Soros responded, if the Schlesinger quotes were accurate, why just build steadily? "Go for the jugular," Soros advised him.

And the rest is history. Mr Mallaby presents the trading strategy in an interesting way. Essentially, the key was to find bets that basically can't go against you. In that case, it's relatively riskless to bet big, and you can therefore win big. So in this case, it wasn't clear that sterling would fall out of the ERM-approved trading band, but it was almost certain that sterling would *NOT* appreciate against the mark. And so there was little reason not to pile on.

Along these lines, consider [this](http://www.economist.com/blogs/buttonwood/2010/06/emerging_markets_and_sovereign_debt) (http://www.economist.com/blogs/buttonwood/2010/06/emerging_markets_and_sovereign_debt) Buttonwood post from earlier today:

I have just been talking with Alexander Kozhemiakin, a Russian-born but Boston-based bond manager at Standish. He is a great enthusiast for emerging market local currency bonds. And he mentioned in passing that Thai 10-year government bonds, denominated in baht, yield a little over 3%; I just checked on [Bloomberg](http://www.bloomberg.com/apps/news?pid=newsarchive&sid=afSduqh1y9_g) (http://www.bloomberg.com/apps/news?pid=newsarchive&sid=afSduqh1y9_g) and the latest figure I could find is 3.22%. US 10-year yields are, at the time of writing, 3.18%. In other words, the markets are willing to accept virtually the same yield from the Thai government as they do from the American.

Now you might say that this is comparing apples with oranges, as the bonds are in different currencies. But investors used to demand a higher yield for owning an emerging market currency because of the depreciation risk. And this is Thailand, the country which started the Asian crisis of the late 1990s and where demonstrators were besieging the government not so long ago.

Fair enough, but Thailand, like most of the large, emerging, export-oriented, Asian economies has an

overvalued undervalued currency. It's also in a situation, like other export-oriented Asian economies, in which domestic demand must assume a greater role thanks to slower growth in developed nations. The baht might not appreciate in stunning fashion, but it almost certainly won't fall against the dollar. And the risk of the dollar appreciating extensively against the baht is slim; the greenback has nowhere to go but down. So American bond yields are low *despite* the depreciation risk. Indeed, the market pressure is all pushing Asian currencies up. That's one reason, as my colleague [noted today](http://www.economist.com/blogs/freeexchange/2010/06/capital_controls) (http://www.economist.com/blogs/freeexchange/2010/06/capital_controls), that China is restricting certain market activity on currency forwards. Everyone all but *knows* that the renminbi is a one-way bet.

(And interestingly, a subsequent chapter in Mr Mallaby's book addresses the 1998 Thai currency crisis, in which the opposite was true. After years sticking to an unhealthy dollar peg, and suffering persistent current account deficits, Thailand finally had to face up to the inevitable.)

The question is, how does this all apply to the situation in Europe? Let me quote something from a [post](http://www.eurointelligence.com/index.php?id=581&tx_ttnews[tt_news]=2811&tx_ttnews[backPid]=901&cHash=9247ee6091) ([http://www.eurointelligence.com/index.php?id=581&tx_ttnews\[tt_news\]=2811&tx_ttnews\[backPid\]=901&cHash=9247ee6091](http://www.eurointelligence.com/index.php?id=581&tx_ttnews[tt_news]=2811&tx_ttnews[backPid]=901&cHash=9247ee6091)) by Kevin O'Rourke, a professor of economics at Trinity College Dublin:

The lesson of the EMS crisis is that low inflation, low deficits and low government debt are not, it turns out, enough on their own. Low unemployment and economic growth are part of the fundamentals which have to be right, if government policies are to be credible in the eyes of the markets. Speculators bet that governments would not, in the long run, be able to sustain policies which led to rising unemployment: far from enhancing credibility, the 'responsible' and deflationary policies which governments thought markets wanted fatally undermined it. And thus it was that the market forced governments in the UK and elsewhere to adopt policies that were softer, and more growth-oriented, than what orthodoxy had been demanding.

The European crisis hasn't played out as previous currency crises have because there are no individual currencies to bet against. Rather than trying to pry a currency out of a peg, market pressure has primarily come in bond markets, where yields on the debt of troubled nations have risen. But after earlier spikes, a European aid package and the ECB's commitment to buy government debt brought yields back down.

For a while. Now they're creeping back up again. The question I'd ask is whether there's a one-way bet lurking in here somewhere. It's harder to say in this case because of the element of political risk; so far, the combined efforts of European governments and the IMF have been able to influence debt yields. So there is a risk in betting that yields will only rise (for now—markets are no doubt looking to see whether commitment to the rescue might waver).

Wherever the bet may be, you can be sure that many traders are looking at the imbalances involved, the adjustments needed, and the domestic pain necessary to achieve those adjustments, and they're concluding that something has to give. The only task is to find the breaking point.

Chap 5: Before the crisis, the accumulation of internal imbalances

Key issue: the mainstream view of the euro crisis is that it was a balance of payment crisis.

- J. Pisani-Ferry. The euro-crisis and its aftermath. Oxford University Press, chapter 6-8.
- L.A.V Catao. "Why Real Exchange Rates?" Finance and Development, September 2007.
- Baldwin, R. and F. Giavazzi. 'Introduction' in "The Eurozone Crisis, A Consensus View of the Causes and a Few Possible Remedies." A VoxEU.org eBook, Sep. 2015.
- S. Merler and J. Pisani-Ferry. "Sudden Stops in the euro area." Bruegel Policy Contribution, March 2012, companion VOXEU column 2 April 2012.
- P. Lane. "The European Sovereign Debt Crisis." The Journal of Economic Perspectives, Vol. 26, No. 3 (Summer 2012), pp. 49-67.

CHAPTER 6



Germany's Long Penance

Can Germany Still Be Saved? is the surprising title that Hans-Werner Sinn, Germany's most famous and outspoken economist, gave in 2003 to his analysis of his country's economic predicaments.¹ Given Germany's current prosperity, it is all too easy to forget that it was in much poorer shape when the euro was introduced, and that it had to suffer a decade of penance before it emerged from its difficulties.

The figures are striking. In 1999, the euro's first year of existence, France posted the highest current account surplus in its history—€43 billion—whereas Germany had a €26 billion deficit. This was the exact opposite of the standard situation. Twelve years later, things were back to normal: in 2011 France posted a €35 billion deficit on its current account and Germany an impressive €161 billion surplus. This contrast speaks volumes about the euro's first decade of existence.

At the end of the 1990s, Germany had not yet fully recovered from the shock of reunification. In 1989, the fall of the Berlin Wall led, far quicker than anyone expected, to the creation of an "economic, monetary and social union" between West and East Germany. The western half was prosperous and hypercompetitive, whereas East Germany's economy virtually collapsed the moment it opened up to international competition. Germans living in the eastern *Länder* were instantly eager to buy products from the capitalist economy, while the region's infrastructure and capital stock had to be replaced entirely. Transfers from the country's western half financed this boom in demand, which soon resulted in a budgetary deficit. Moreover,

1. *Ist Deutschland noch zu retten?* (2003) Econ Press, 2003. The English translation was published in 2007 by MIT Press under the title *Can Germany Be Saved?*

demand was also directed towards products from West Germany, where the productive capacity was put under considerable pressure.

Inflation followed. Although it remained moderate (it peaked at 4% in 1992), price rises were, year after year, higher than those in France and other countries with currencies pegged to the *Deutsche Mark*. This was enough for Germany to lose competitiveness and start falling behind. In the 10 years between the fall of the Berlin Wall and the 1999 changeover to the euro, its production costs increased by 10% more than France's. Furthermore, Italy and Spain, which had devalued their currencies in 1992 and 1993, had entered the euro at favourable conversion rates. As a result, German producers were struggling.

The conditions were ideal for Germany's European economic partners, whose greatest fear was being steamrolled by Germany's powerful industry. For France, Italy, Spain, and others, the first years of the euro journey were remarkably easy. Germany, however, came into it with a significant handicap that took almost 10 years to remedy.

These were the 10 years that shook German industry. Comparing the evolution of German and French foreign trade over the decade highlights the extent of the country's economic revolution. In 1998—the final year before the euro—resemblances were still strong. Exports and imports of goods, respectively, amounted to 29% and 27% of German GDP. In France, the corresponding figures were 26% and 24%. Exports in both cases were worth roughly a quarter of the country's GDP. Germany was slightly more open, especially taking into account its economic size, but qualitatively the differences were minor. By 2008, however, the situation was entirely different, with German exports reaching 47% of GDP and imports 41%, that is, 14- and 18-point increases during the 10-year period. In France, the export ratio rose by only one percentage point over the same period, reaching 27% of GDP, whereas, on the import side, the rise was 5 percentage points, to 29% of GDP.²

Much of the effort was made by businesses and their employees. Companies were quick to grasp the magnitude of the challenge. They had lost competitiveness vis-à-vis their euro-area peers at the same time that eastern Europe, having completed the first stage of its transition, was becoming an attractive region for industrial investment. China's productive and technological capacities were also catching up.

2. The reader may wonder why German exports exceeded German imports in 1998, while the country posted a deficit. This is because current-account transactions include other items than trade in goods, such as trade in services, factors income, and transfers.

German firms retaliated with unprecedented efforts to boost productivity, restructure, relocate outside Germany segments for which the country had no comparative advantage, and focus their efforts on high-technology segments that required skilled labour. German industry thus reinvented itself. The reason why exports and imports reached such high levels is that companies fully embraced globalisation and the segmentation of value chains. This resulted in frequent transactions with subsidiaries and subcontractors in central Europe and elsewhere. Instead of producing in Germany and exporting to the world, in line with the traditional model, companies increasingly exported components and reimported intermediate products. The fear at one point was that Germany would become a "bazaar economy" (another provocative expression coined by Hans-Werner Sinn), a sort of import/export business with no production of its own. In fact, the exact opposite happened. Offshoring gave Germany's economy the kick start it needed, whereas France's productive structure remained frozen in the 1990s. Jobs were destroyed in the process—with some low-skill activities moved elsewhere—but overall, industrial employment declined much less than in France.

The process weighed heavily on German workers. Nominal wages rose by only 18% between 1998 and 2008, and purchasing power stagnated, whereas the share of profits in value added rose by almost 4 percentage points. This result was partly achieved through mutual consent, with trade unions accepting the stagnation in purchasing power in return for the preservation of jobs in the industrial sector. It was also the result of deliberate policies put in place to foster and support the economic transition. To add momentum to the process Chancellor Gerhard Schröder, whose first term in office, between 1998 and 2002, had been dedicated to stimulating the economy, launched a series of labour market reforms in 2003, known as Agenda 2010. These reforms created a new category of "minijobs" with reduced social security contribution rates, restructured public employment agencies, and reformed social safety nets, reducing the maximum length of unemployment benefits. These were harsh, controversial and socially costly reforms, but they succeeded in reducing unemployment and cutting labour costs. When Angela Merkel's turn came in 2005, she introduced a tax reform that reduced employee contributions and further strengthened German competitiveness. But most of the work had already been done.

The overall climate was not conducive to wage increases, real-estate bubbles, or consumption booms. At a time when most Europeans—including the United Kingdom and central Europe—were spending, Germans saved. Consumption per capita grew by only 9% during the 2000s, compared to 19% in France, 22% in Spain, and 29% in Greece. Initially, this asymmetry

reflected a necessary correction following Germany's loss of competitiveness during the preceding decade. Its persistence, however, turned out to be destabilising, for two reasons.

The first was that the divergence of relative demand—with demand growing in southern Europe and stagnating in Germany—mechanically led to external deficits in high-consumption countries and external surpluses in high-savings countries. The situation in Germany and other northern countries seemed to be the mirror image (or the consequence) of the situation in the South. As Christine Lagarde, then French Minister of Finance, said in May 2010, "It takes two to tango." Deficit countries are not necessarily financed by surplus countries in the euro area, since the area as a whole is not a closed economy, but her observation was true nonetheless.

This interdependence did not come about by chance. One of the objectives of the monetary union was to encourage the flow of savings within Europe, just as savings flow between regions within countries. In the case of the United States, no one is interested in knowing whether investments in Texas are financed by savings from Florida or Texas itself. The only relevant questions are if the investments are sound, if they are adequately financed through, for example, debt or equity and, of course, if the U.S. economy as a whole generates enough savings to finance investment (which it barely does). The euro's architects had very much this example in mind and thought of monetary unification as a way to eliminate internal barriers to savings flows.

The elimination of foreign-exchange risk did have the effect of promoting the circulation of savings within the euro area, but this was certainly not a two-way process. Rather, the euro quickly led to a sustained and ultimately massive widening of external deficits and surpluses. Through portfolio investment and through the banking system, savings in the North took the road to the South, where they financed investment, especially in real estate and consumption. In this sense, the stagnation of demand in Germany can be said to have fuelled the housing boom in Spain.

The second destabilising factor was built into the ECB's interest rate policy. The ECB announced from the outset (and rightly so) that monetary policy decisions would be based on the situation in the euro area as a whole, and that targeting the needs of individual countries was out of the question. Actually this was the only possible stance. Had the ECB wanted to take national situations into account, it would not anyway have had the means to do so via the interest rate channel; the interbank lending rate, which is based on the ECB's policy rates, is supposed to be the same throughout the euro area. However, average inflation, on which the ECB's monetary policy is based, is naturally influenced by the size of countries. Germany accounts, in economic terms, for 28% of the euro area. Very low inflation in Germany,

therefore, had a significant impact on the euro-area average, meaning that monetary policy became too expansionary for booming countries. With low interest rates, credit in the South was not held in check.

Frenzy in the South and lethargy in the North; this was a powerful and self-perpetuating dynamic, which was allowed to go on for far too long. The German government was reluctant to end the country's penance, and the Spanish government did not want to crash the party. Both pushed on for as long as they could—until the crisis broke out.

What happened during the euro's first decade was, therefore, the exact opposite of what many uninformed critics had predicted. German discipline did not create a straitjacket for other countries. Instead, Germany's penance led the ECB to keep interest rates too low for Spain, Greece, and others, thereby fuelling inflation in these countries.

Of course, "the South" was not a uniform block. Though there was a common factor behind the evolution of their situations, the countries that enjoyed the good years of the 2000s and now find themselves in varying degrees of trouble are very much like Tolstoy's families: they are all unhappy in their own ways.

In a first group, the damage was mostly caused by fiscal misbehaviour. Greece until 2009 consistently violated European budgetary rules (and kept it hidden). Fiscal management was also imprudent in Portugal, although it is difficult to pretend that budgetary slippages were the sole cause of problems: unlike Greece, Portugal's budget deficit had, since the country adopted the euro, never been greater than 2.7% of GDP.

In a second group of countries, troubles mainly came from a massive expansion of private credit. This is what happened in Spain. Easy borrowing led to a decrease in private savings and a widening of the external deficit. Ultimately, imprudent lending translated into losses in the banking sector and, after the collapse of tax receipts, a severe deterioration in the public finances. The situation was similar in Ireland, where the cost of bailing out the banking sector reached astronomical proportions, currently estimated at more than 40% of GDP.

In a third and final group of countries, there was no frenzy but instead a gradual deterioration of competitiveness. As indicated, France lost ground significantly compared to Germany, and saw its external deficit widen. Italy also gradually lost market share and its problems were compounded by an extremely disappointing growth performance.

In the context of panic and contagion in sovereign debt markets, individual afflictions tend to be forgotten or turned into variations of a single euro-area epidemic. This is what happened after 2010. However, they must be remembered to avoid hasty generalisations and one-size-fits-all remedies.

CHAPTER 7



The Perfect Culprit

The Irish are famous for their sense of humour. Following the crisis that brought the Icelandic banking system (and with it the whole country) to its knees in autumn 2008, one joke making the rounds was “What is the difference between Ireland and Iceland?” The answer: “1 letter, and 6 months.”

In the end, it took 11 months. The 5-month delay meant that the first country to seek European and IMF financial assistance was Greece, not Ireland. For this reason the endless discussions on financial assistance that occupied Europe during the first half of 2010—on whether to provide it, up to what limit and under what terms—focused on Greece. For public opinion in Europe, and especially in Germany, the crisis was perceived as coming from Greece. Many northern Europeans continue to look at the situation in southern Europe through the Greek prism. They see the financial illnesses of Portugal, Spain, and even Italy as less acute variants of the same disease.

It will never be known how things would have been if Brian Cowen, the Irish prime minister, had requested help first, before his Greek colleague Georges Papandreou. The focus of the discussion on European crises would have been completely different. Ireland would have been a rather embarrassing culprit for Europe, because the country’s real estate and financial follies offered a mirror image of the other side of the Atlantic. All the flaws that brought down Wall Street in 2008 were present, save for financial sophistication. The German regional banks, the *Landesbanken*, had subsidiaries in Dublin, and could not pretend that they were not part of what was going on. The Irish are not known for spending too much of their time on the beaches. And as far as budgetary discipline was concerned, Ireland

was beyond reproach: its debt-to-GDP ratio, which stood at 54% of GDP in 1998, had been brought down to 25% of GDP by 2007. It was among the best in the class. So the discussion would have focused on private leverage and balance-of-payment deficits, not public debt and fiscal deficits.

Greece, on the other hand, was the perfect culprit. After having missed the 1999 entry and having narrowly qualified for joining the euro in 2001, it quickly started to exceed the budgetary deficit limits it had committed to. To avoid sanctions, it started covering up the true state of its public finances. Later, data would show that from 2001 to 2009, when the crisis erupted, the budgetary deficit was never below 4.5% of GDP (against a European ceiling of 3%).¹ By that time, public debt had reached 129% of GDP, a figure much higher than the 60% benchmark set in the Treaty. This was outright deception and it was taken as such by northern Europe.

Furthermore Greece was in a sorry state, as the Troika of IMF, European Commission, and ECB officials dispatched to Athens in spring 2010 to negotiate the request for assistance would soon discover. Unlike Portugal, which had failed to modernize its economy, but had still succeeded in upgrading its infrastructure, Greece had squandered the generous EU grants supposed to help it catch up to the standards of richer EU member states. Clientelism had destroyed an already feeble public machinery, tax evasion was standard practice, and the government had made a habit of suspending tax inspections before elections. Heavy regulation held back on productivity and perpetuated the privileges enjoyed by closed professions. Few companies met international standards, and they could only survive because they were sheltered from international competition. On top of this, the retirement age was lower, pensions were higher, and sector-specific benefits were more generous than in most European countries.

For a country such as Greece to ask the rest of the euro area for help seemed outrageous to governments and public opinion in northern Europe. Academics enumerated the occasions on which Greece had flouted the rules, and began to ponder the meaning of having rules if infringement went unpunished. Newspapers helped turn northern public opinion against their allegedly lazy partners. *Bild*, a German tabloid with a daily circulation that reaches 4 million, turned Greek pensioners into symbols of misguided solidarity, popularizing the idea that Greece first needed to get its act together and follow the footsteps of Germany, which had made major sacrifices in the preceding decade. It was indeed difficult to build a case for

1. To be fair, the gap between the current budget deficit data and the numbers reported in real time is, in part, attributable to a change in the accounting treatment of military spending, which affects the early part of the period.

providing generous assistance to a country in which restaurant owners and doctors continued to conceal their income and pay ridiculously low taxes.

For Germany, the Greek crisis also seemed to offer all the features of a perfect classical tragedy with a script written 20 years ago. In the early 1990s opponents of the euro had argued that mustering the necessary discipline would be beyond the abilities of Mediterranean countries. They had predicted that these countries would pretend their finances were under control, embark on deficit spending, and ask other member states to foot the bill when the burden became too much to bear. Only the villain's name was missing.

To stave off its nightmare scenario, Germany had asked for guarantees and safeguards, which came in the form of the no-bailout clause, the prohibition of monetary financing of public deficits, the obligation to avoid "excessive" public deficits, quantitative benchmarks, the surveillance framework, monetary sanctions, and the Stability Pact. On this basis, the euro advocates had claimed that each country would abide by the rules or face sanctions.

However, as in all authentic tragedies, neither prescience of the fate nor attempts to prevent its occurrence were sufficient to avert disaster. Ironically enough, Germany played a key role in its own misfortune. The rules of the Stability Pact should have led to France and Germany receiving a formal warning in 2003, as their deficits had ballooned as a consequence of the recession and were higher than the authorized threshold of 3% of GDP. Berlin had never really imagined that its own performance would ever attract sanctions, and sided with France to block the recommendations. Both countries requested—and obtained—a reform of the Stability Pact, which Romano Prodi, then president of the European Commission, had anyway deemed "stupid." The reform in itself was an improvement, since it introduced cyclical adjustment in the monitoring of the state of public finances, but it also damaged the Stability Pact's credibility. For many, especially among the fiscal conservatives, this decision amounted to a sort of original sin.

Germany made another mistake in 2005 when the European statistical office, which had doubts about the quality of Greece's reporting of its public accounts, asked for investigative powers. Eurostat wanted to be authorized to conduct on-site inspections instead of only taking note of member states' notifications and asking to have some operations reclassified. However, member states, with France and Germany in the lead, opposed these demands, fearing that such powers might one day be used against them. Doubts about the Greek data were widespread, but the quantitative impact of misreporting was believed to be much smaller than it

was. Europe, therefore, had to wait until Georges Papandreou's confession in 2009 for the truth about Greece to come out.

Of course, doubts about the quality of Greece's accounts were widespread among European and international civil servants. The IMF had made its doubts plain in one of its reports. The extent of the disaster was unknown, however, and few ministers or European officials were brave enough to persistently air their concerns and upset Greek colleagues for perhaps a few tenths of percentage points.

The dramatic failure of surveillance by European authorities was only acknowledged after the fact, but the cost of failure was huge. One of the first decisions taken after the Greek crisis was, therefore, to revise the Stability Pact. The framework created by the new legislation (and enshrined in a new Treaty on Stability, Coordination and Governance signed in March 2012) introduced numerical rules for reducing any debts above 60% of GDP, created new sanctions even before the 3% deficit threshold was breached, and made the sanctions recommended by the European Commission quasi-automatic.² Instead of requiring member states' approval, they will automatically be implemented unless a majority of countries decide to vote against them.

Blaming the euro area's difficulties only on a failure to implement an otherwise adequate surveillance framework would, however, be wrong. At least as serious were the problems arising from the way surveillance had been designed. The crisis revealed two serious flaws in the crisis-prevention framework.

The first mistake was the implicit assumption that threats to stability—if not the only risk—would come from the public sector. Ireland, Spain, and, more generally speaking, the financial crisis of 2007–2008 have demonstrated that the private sector is also a major source of instability. With hindsight, the obsession of the architects of the Maastricht Treaty with government failures made them blind to market failures. The premise was that the private sector was inherently stable. Events, especially in Spain, have shown how flawed this assumption was.

The second, related mistake was to assume that monitoring deficits year after year was enough to prevent public finance disasters from occurring. But this was forgetting that fiscal disasters often materialize in a very sudden way. Spanish public finances were in clear surplus in 2007, but two years later the deficit exceeded 10% of GDP. Until 2007, Ireland complied with the deficit criterion each and every year, but it did not monitor the

2. The treaty was signed by 25 of 27 EU countries. The Czech Republic and the UK did not sign.

contingent liabilities related to the unfettered development of its banking sector. When the latter collapsed and depositors began to panic, the state issued a blanket guarantee to cover deposits to prevent massive withdrawals. In so doing, the Irish state deteriorated massively its own balance sheet. Since the state is the insurer of last resort, its finances are vulnerable to all sorts of economic and financial shocks. Acknowledging this potential fragility and designing an adequate surveillance framework is more difficult than stigmatizing a small Mediterranean country for its (very real) turpitudes.

CHAPTER 8



The Golden Decade

On a warm spring day in May 2006, Miguel Sebastián was beaming with confidence. "It's all very simple," said the chief economic advisor to Spanish prime minister José Luis Rodríguez Zapatero. "Until recently, we thought that Spain would have 35 million inhabitants by 2030. It now seems that this number will exceed 45 million. We have become a country of immigrants, and this changes everything."¹

He was right. To Spain, in the middle of the 2000s, came farmworkers from North Africa, construction workers from Poland or Romania, and waiters from Colombia or Ecuador. More and more pensioners from northern Europe relied on low-cost airlines to spend their last winters under the Spanish sun.

A country whose population is rapidly increasing needs infrastructure, schools, hospitals, and houses. Most of all, houses. So the price of homes skyrocketed: it would increase threefold between 1998 and 2008. Construction boomed: in the middle of the 2000s, more than 700,000 homes were built every year, compared to only 400,000 in France (the population of which is 50% greater). The construction sector employed 13% of the labour force, compared to 7% in France. The urban sprawl sprawled, and the Spanish coastline disappeared beneath dull and uninspired holiday homes.

The demographic argument seemed credible. Except that the story was virtually identical to the one told by U.S. economist John Kenneth Galbraith about Florida in the 1920s: "The Florida boom contained all of the elements of the classic speculative bubble," he wrote, including "the indispensable element

1. Conversation with the author, May 2006.

of substance," on which men and women "had built a world inhabited not by people who have to be persuaded to believe," but by people "who want an excuse to believe." Buyers "did not expect to live on the terrain they bought," nor was it easy to suppose "that anyone ever would," but the reality was that these assets were "gaining value by the day and could be sold for a handsome profit in a fortnight."²

This was Spain in the 2000s. All of those who could take part in the bonanza did so, buying and selling property with the sole aim of making capital gains. Banks and, especially, regional saving banks (the now-infamous *cajas*, most of which took heavy losses and ended up being restructured and recapitalized with public funds after 2007) lent money to anyone who asked. Countless small entrepreneurs borrowed at will to build and sell. Thanks to cheap credit, poor households bought houses they could hardly afford—an Iberian version of the subprime saga. The more affluent bought to resell with a capital gain. Furthermore, regions subsidized the construction of airports that would close down a few years later, having failed to attract any significant traffic, and cities lobbied for high-speed train links of no economic value. This all came to a halt in 2007. The real-estate market began to slow down, before collapsing after the financial crisis of 2008.

The picture was broadly similar in Ireland (but for the sun factor). In both cases the bursting of the housing bubble had tragic consequences. Half-finished houses were abandoned, entrepreneurs were ruined, construction workers were laid off, and more and more of the banks' claims became worthless. GDP plunged and tax revenues collapsed. Households, businesses, and banks realized the extent of their excess leverage and started saving to slowly repay, but it would soon become evident that this was bound to be a protracted process. Worse still, the legacy of the boom years was a very unbalanced economy, in which the nontradable sectors had expanded beyond reason and tradable-goods sectors had shrunk. Reallocating capital and labour from the former and towards the latter could not take place overnight.

The role of the euro in this whole process may not be immediately apparent. After all, there was no housing market boom and bust in Germany. Also the subprime credit crisis in the United States is proof enough that such developments are not the sole preserve of euro-area countries. But monetary union nevertheless played a crucial role.

The main reason was the mechanics of interest rates. In Europe, everything had been done to promote capital mobility and encourage the flow of savings to wherever they would be the most productive. Spanish banks had

easy access to euro loans from counterparts in other countries, and interest rates on similar loans were the same in Barcelona and Frankfurt, as long as the debtors were deemed equally solvent.

Prices were different, however, especially in the nontradable goods and services sectors. A square metre of living space can be more expensive in Paris than in Berlin without Parisians being prompted to migrate in search of better living conditions. By the same token, a restaurant meal can be more affordable at one end of a currency zone than at the other without people moving around in search of the best deal. Because a large part of the economy produces nontradable goods and services, price divergences are frequent and can be long-lasting. As a consequence, inflation rates can also differ for long periods.

But if inflation rates can be different, so are real interest rates on loans. With an identical nominal interest rate—say, 4%—the cost actually borne by borrowers will be 1% in countries where inflation is 3%, and 3% in countries where inflation is 1%. In the first case, household income is likely to increase at a more rapid pace, and the interest burden will decrease accordingly. As the real cost of credit is lower in countries in which the inflation rate is higher, demand for credit is stronger and inflation persists or even increases. Unlike what happens in a country equipped with its own monetary policy, where the central bank can let the interbank interest rate rise and stem the demand for loans, credit will continue to increase as long as bank funding is cheap and that lending is—or seems to be—profitable.

The process needs a trigger to set itself into motion, but once it has started, it quickly becomes a vicious circle. In southern Europe, the trigger came in the form of a drop in interest rates resulting from accelerated convergence. In Spain, the level of long-term interest rates was in excess of 11% in 1995. At the time, markets doubted that the country would be able to join the euro (assuming the latter would come into existence). By the time Spain adopted the euro in 1999, interest rates had dropped to below 5%, close to the German level. The consequences of such a dramatic decrease were an avalanche of credit and massive real-estate investment.

In theory, these destabilising processes are kept in check by automatic correction mechanisms. If inflation persists and prices increase, national products become increasingly hard to sell and consumers turn to foreign products instead. Sectors exposed to international competition—mainly manufacturing, tourism and agriculture—suffer from lower demand and scale back their payrolls. This mechanism eventually ends the boom, and reverses the process. Correction of imbalances is, however, excruciatingly long because for every manufacturing job destroyed, at least one job is created in the sectors that benefit directly or indirectly from lower interest

2. John Kenneth Galbraith (1929) *The Great Crash*. New York: Houghton Mifflin.

rates—construction and services. For long, destabilising forces proved to be stronger than the stabilising ones. In Spain, the external deficit widened from 4% to 10% of GDP between 2000 and 2007, while at the same time unemployment decreased from 14% to 8%. The share of industry in employment dropped from 18% to 15%, whereas the share of the construction and the service sectors rose from 47% to 52%.

There was nothing surprising in this process. The potential for vicious circles was pointed out in the 1980s by Sir Alan Walters, Margaret Thatcher's economic advisor.³ The sluggishness of equilibrating mechanisms was demonstrated in the 1980s and the 1990s by the slow correction of competitiveness gaps among countries participating in the European exchange-rate system. Nor were economists and policymakers unaware of the problem. By 2005–2006 already, divergence within the euro area had become the topic for conferences and academic papers. It was clear that something was going wrong.

There was not much the Bank of Spain could do about interest rates, because the setting of them is within the ECB's remit, but it did have control over bank supervision and the regulation of credit. Like all national central banks, it was still in charge of financial stability, and could, therefore, encourage prudent lending practices. Actually it did: on many occasions the central bank came out with warnings and asked banks to make provision for potential losses on mortgage assets (it was famous in Europe for having promoted what is called dynamic provisioning, the objective of which is precisely to limit the effects of cyclical boom and bust). However, the Bank of Spain's request that banks set aside part of their profit to cover potential losses had only a limited effect on the expansion of credit.

The other way to slow things down would have been to deal with the problem head-on. On the macroeconomic front, the government was actually doing quite well. It posted a 2% budget surplus in 2007, well in excess of European requirements, and as a consequence, the country's debt-to-GDP ratio decreased rapidly during the boom years. It would have been difficult, truth be told, to aim for a much larger surplus. The government in Madrid could, however, have acted directly to prick the bubble, for example by reducing the duration of mortgages (Spain had made it possible to borrow over 40 or 50 years), by capping debt-to-income or loan-to-value ratios, or by eliminating tax incentives for real-estate investment. It *could* have chosen any of these options, but each carried significant political costs and would have disadvantaged specific social and demographic

3. Alan Walters originally used the argument to oppose the UK taking part in the European Exchange-Rate Mechanism, but the reasoning applies equally to monetary union.

groups. Applying them would inevitably have made it more difficult for the middle classes and young workers to become homeowners. In spite of the technicians' warnings, no politician wanted to be held responsible for having taken the punch bowl away just as the party was getting livelier (to paraphrase a famous sentence by Arthur Burns, the Fed chairman in the 1970s). So José Luis Rodríguez Zapatero, the president of the government, decided that inertia was the easiest course.

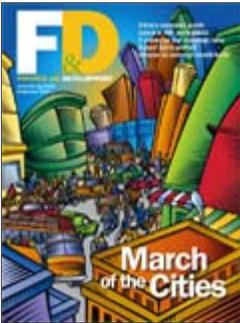
As for Europe's institutions (which were also in charge of supervision) and Spain's European partners (which were supposed to apply peer pressure), they only paid lip service to their obligations. To be fair, the call was not that easy, because inflation rates can also differ between countries for good reasons. For a start, euro-area countries differ in terms of economic development. GDP per capita in Austria was 60% higher than in Portugal in 1999, and price levels were, accordingly, also very different. The process of catching-up by poorer countries is supposed to lead to a catching-up of price levels, and, therefore, also lead to higher inflation. There are no good economic reasons to interfere in this equilibrating process. Price evolutions can also differ across countries. When a country's products climb the quality ladder, for example, demand on international markets increases and so do prices. As a consequence, inflation is higher than in the less-successful neighbouring countries. Inflation differentials can, therefore, be a good thing, just as they can be catastrophic when prices spiral out of control in low-growth and low-competitiveness economies. During the euro's first decade, however, evidence was pointing to problems looming: the average annual inflation rate was 3.2% in Spain compared to 1.7% in Germany. This should have been enough to set off the alarms.

Why did policymakers fail to act? To start with, it was believed that external deficits, even lasting ones, were not important in a monetary union, in the same way they are irrelevant for regions within a country. The somewhat ideological premise was that they are a cause for concern only if they originate in lack of discipline in the public sector. However, from 2001 to 2007, Spain's fiscal balance was, in each consecutive year, stronger than Germany's and the country was never singled out by the EU for having failed to comply with the requirement of the budgetary pact, whereas Germany was for more than four years in an "excessive deficit procedure."

Furthermore, as member states, not the EU, were formally in charge of credit regulation and taxation, no one really wanted to bother the Spanish authorities and take the risk of creating an avenue for more direct EU intervention. So the economics and the politics of surveillance both suggested that unlike budgetary imbalances, external imbalances could be overlooked. For this failure, Europe would pay a high price.



Share This  6



September 2007, Volume 44, Number 3

Search Finance &
Development

Search

[Advanced Search](#)

[About F&D](#)

[Subscribe](#)

[Back Issues](#)

[Write Us](#)

[Copyright Information](#)



**Free Email
Notification**

Receive emails
when we post new
items of interest to
you.

[Subscribe](#) or
[Modify](#) your profile

Back to Basics

Why Real Exchange Rates?

Luis A. V. Catão

How does one determine whether a currency is fundamentally undervalued or overvalued? This question lies at the core of international economics, many trade disputes, and the new IMF surveillance effort.

George Soros had the answer once—in 1992—when he successfully bet \$1 billion against the pound sterling, in what turned out to be the beginning of a new era in large-scale currency speculation. Under assault by Soros and other speculators, who believed that the pound was overvalued, the British currency crashed, in turn forcing the United Kingdom's dramatic exit from the European Exchange Rate Mechanism (ERM), the precursor to the common European currency, the euro, to which it never returned.

But in the ensuing years, neither Soros nor fellow speculators have repeated the feat consistently, and the economics profession itself lacks a foolproof method of establishing when a currency is properly valued. This failure is striking given that the exchange rate is a central price in economics and that there is a measure potentially capable of delivering the answer and for which plenty of data exist: *the real exchange rate* (RER).

What things really cost

Most people are familiar with the *nominal exchange rate*, the price of one currency in terms of another. It's usually expressed as the domestic price of the foreign currency. So if it costs a U.S. dollar holder \$1.36 to buy one euro, from a euro holder's perspective the nominal rate is 0.735. But the nominal exchange rate isn't the whole

story. The person, or firm, who buys another currency is interested in what can be bought with it. Are they better off with dollars or euros? That's where the RER comes in. It seeks to measure the value of a country's goods against those of another country, a group of countries, or the rest of the world, at the prevailing nominal exchange rate.

One can measure the real exchange rate between two countries in terms of a single representative good—say, the Big Mac, the McDonald's sandwich of which a virtually identical version is sold in many countries. If the real exchange rate is 1, the burger would cost the same in the United States as in, say, Germany, when the price is expressed in a common currency. That would be the case if the Big Mac costs \$1.36 in the United States and 1 euro in Germany. In this one-product world (in which the prices equal the exchange rates), the purchasing power parity (PPP) of the dollar and the euro is the same and the RER is 1 (see box). In this case, economists say that absolute PPP holds.

What is the real exchange rate?

The RER between two currencies is the product of the nominal exchange rate (the dollar cost of a euro, for example) and the ratio of prices between the two countries. The core equation is $RER = eP^*/P$, where, in our example, e is the nominal dollar-euro exchange rate, P^* is the average price of a good in the euro area, and P is the average price of the good in the United States.

In the Big Mac example, $e = 1.36$. If the German price is 2.5 euros and the U.S. price is \$3.40, then $(1.36) \times (2.5) \div 3.40$ yields an RER of 1. But if the German price were 3 euros and the U.S. price \$3.40, then the RER would be $1.36 \times 3 \div 3.40 = 1.2$.

But suppose the burger sells for 1.2 euros in Germany. That would mean it costs 20 percent more in the euro area, suggesting that the euro is 20 percent overvalued relative to the dollar. If the real exchange rate is out of whack, as it is when the cost is 1.2, there will be pressure on the nominal exchange rate to adjust, because the same good can be purchased more cheaply in one country than in the other. It would make economic sense to buy dollars, use them to buy Big Macs in the United States at the equivalent of 1 euro, and sell them in Germany for 1.2 euros. Taking advantage of such price differentials is called arbitrage. As arbitrageurs buy dollars to purchase Big Macs to sell in Germany, demand for dollars will rise, as will its nominal exchange rate, until the price in Germany and the United States is the same—the RER returns to 1.

In the real world, there are many costs that get in the way of a straight price comparison—such as transportation costs and trade barriers. But the fundamental notion is that when RERs diverge, the currencies

face pressure to change. For overvalued currencies, the pressure is to depreciate; for undervalued currencies, to appreciate. It can get more complicated if factors such as government policies hinder normal equilibration of exchange rates, often an issue in trade disputes.

Overvalued or undervalued?

How about comparing purchasing power when countries sell more than one product? To do this, economists usually measure the real exchange rate in terms of a broad basket of goods. Because the price of such a basket normally takes the form of an index number—such as the consumer price index (CPI), which includes both goods and services—the RER is also typically expressed as an index that can be benchmarked to any chosen time period. Going back to the dollar-euro example, if an RER index is 1.2, the average consumer prices in Europe are 20 percent higher than in the United States, relative to the chosen benchmark. Indexes don't measure absolute prices (such as the price of the Big Mac), but changes in overall prices relative to a base year (if, say, the index is 100 in the year 2000 and 120 in 2007, average prices are 20 percent higher than in 2000). In this case, if RER indexes between countries don't change over time, we say that relative PPP holds.

RER indexes between two countries can be important. The massive U.S. trade deficit with China has become a political and economic issue, and whether its roots are in a fundamentally misaligned exchange rate is a point of contention.

But, for the most part, economists and policymakers are more interested in the *real effective exchange rate* (REER) when measuring a currency's overall alignment. The REER is an average of the bilateral RERs between the country and each of its trading partners, weighted by the respective trade shares of each partner. Being an average, a country's REER may be in "equilibrium" (display no overall misalignment) when its currency is overvalued relative to that of one or more trading partners so long as it is undervalued relative to others.

To establish when a currency is misvalued, and, if so, by how much, a rough assessment can be obtained by the REER series over time. Under either absolute or relative PPP, there should be no change in REERs over time if currencies are in equilibrium. But because consumption patterns can change faster than the market baskets statisticians construct—as can trade policies and tariffs and transportation costs—deviations in REERs don't necessarily indicate fundamental misalignment.

Yet, even though transportation costs and tariffs have declined sharply over the past century and national consumption baskets have grown more uniform, fluctuations of REERs have intensified. A century ago, among advanced economies, REER fluctuations were within a 30 percent band. In the 1980s, the United States experienced swings in

its REER as wide as 80 percent! Other countries have had similar experiences.

Tough calls

But not all large REER fluctuations should be interpreted as indications of misalignment. Some large REER adjustments are remarkably smooth, suggesting that there may be factors besides transportation costs, tastes, and tariffs that play a key role in moving about the REER of a currency that is not misaligned.

Technological progress leading to productivity increases in goods commonly traded, called tradables, is thought to be one of those factors. Higher productivity lowers production costs, thus lowering prices of such tradable goods in the higher-productivity country, which then translates into lower tradable goods prices elsewhere through international competition. But not all goods are tradables. Nontradable sectors, such as housing and many personal services, face minimal international price competition. So the prices of tradable goods will tend to fall relative to those of nontradable goods. To the extent that nontradable goods have a large weight in the country's consumption basket, the country's consumer price index will rise relative to the international consumer basket; hence, its REER will tend to appreciate. This mechanism is often referred to as the "Balassa-Samuelson effect." Both theory and data support that much of the REER variations across countries are accounted for by fluctuations in the prices of nontradables relative to those of tradables, and particularly so among developing countries.

Persistent changes in terms of trade (such as oil producers usually experience) and differences in fiscal policies, tariffs, and even financial development can also help explain why REERs can differ across countries. The IMF and economic analysts take such real exchange rate fundamentals into account in estimating the "*equilibrium*" REER, around which the actual REER should hover if there is no misalignment.

Estimating equilibrium REERs can be difficult because prices are somewhat sticky in the short run and the nominal exchange rate is not (at least in countries where exchange rates are market determined). So REERs typically display considerable short-run volatility in response to news and noise trading, and it's not surprising that many market participants and policymakers get things wrong—sometimes very wrong. That can lead to massive realignments with devastating consequences—such as the 1992 ERM crisis. Imperfect though they may be, REERs have signaled large exchange rate overvaluations in the run-up to many financial crises, making it important for the IMF and others to monitor bilateral RERs and multilateral REERs.

Luis A.V. Catão is a Senior Economist in the IMF's Research Department.

Introduction

Richard Baldwin and Francesco Giavazzi

The Graduate Institute and CEPR; Bocconi University and CEPR

The Eurozone crisis which broke out in May 2010 is a long way from finished. Charles Wyplosz puts it bluntly in his chapter: “Five years later, growth is miserable and is forecasted to remain miserable as far as the forecasters’ can see. Governance is in disarray as the tragic Summit of July 13 – the last of an incredible series of official meetings – showed.” Worse yet, there is widespread belief that the fragilities and imbalances that primed the monetary union for this crisis are still present.

As a first step to finding a broad consensus on what more needs to be done, this eBook gathers the views of a dozen world-renown economists on a simple question: “What caused the Eurozone Crisis?” The focus was to be on thinking about the causes as a prelude to developing remedies.

Although the essays were largely uncoordinated – and the authors hark from diverse backgrounds – a remarkably coherent message emerges from this collection. In this introduction, we have two goals.

- The first is to pull together what we feel is the mainstream view of the crisis narrative.

In the world of policymaking, narratives are incredibly important since if we cannot agree on what happened – or more precisely, on what were the most important things that happened – then we cannot agree on how to remedy the situation.

- The second is to identify a consensus view of the causes of the EZ crisis.

Again, we cannot know what needs fixing until we agree broadly on what was broken.

While the eBook's focus is on the causes of the crisis, some of the authors also discussed possible remedies. These ideas will lead us in the next step of this project: an eBook on ways forward for the EZ.

Thinking systematically about the crisis

The core reality behind virtual every crisis is the rapid unwinding of economic imbalances. The size and duration of the crisis typically depends upon: (i) the size of the initial imbalances; (ii) how the initial shock gets magnified by a variety of 'amplifiers'; and (iii) how rapidly and effectively policy responds.

In the case of the EZ crisis, the imbalances were extremely unoriginal. They were the standard culprits that have been responsible for economic crises since time immemorial – namely, too much public and private debt borrowed from abroad. Too much, that is to say, in relation to the productive investment financed through the borrowing.

From the euro's launch and up until the crisis, there were big capital flows from EZ core nations like Germany, France, and the Netherland to EZ periphery nations like Ireland, Portugal, Spain and Greece. A major slice of these were invested in non-traded sectors – housing and government services/consumption. This meant assets were not being created to help pay off in the investment. It also tended to drive up wages and costs in a way that harmed the competitiveness of the receivers' export earnings, thus encouraging further worsening of their current accounts.

When the EZ crisis began – triggered ultimately by the Global Crisis – cross-border capital inflows stopped. This 'sudden stop' in investment financing raised concerns about the viability of banks and, in the case of Greece, even governments themselves. The close links between EZ banks and national governments provided the multiplier that made the crisis systemic.

Importantly, the EZ crisis should not be thought of as a sovereign debt crisis. The nations that ended up with bailouts were not those with the highest debt-to-GDP ratios. Belgium and Italy sailed into the crisis with public debts of about 100% of GDP and yet did not end up with IMF programmes, while Ireland and Spain, with ratios of just 40%, (admittedly kept artificially low by large tax revenues associated with the real estate bubble) needed bailouts. The key was foreign borrowing. Many of the nations that ran current account deficits – and thus were relying of foreign lending – suffered; none of those running current account surpluses were hit.

The initial shock – the rapid loss of EZ investors’ trust in the deficit nations– was amplified in several ways. Given the EZ design, governments who got in trouble had no lender of last resort. Which meant their euro denominated borrowing was akin to foreign currency debt in traditional sudden stop crises. The natural lender of last resort, the ECB, was explicitly forbidden from playing the role. This ruled out one of the classic ways out of avoiding government default – having the central bank print the money needed to service the debt.

The predominance of bank financing was another amplifier of problems. European banks were thinly capitalised and extremely large relative to the countries’ GDP. They were so large that they had to be saved, but their size also created a ‘double drowning’ scenario. This is exactly what happened in Ireland. In what might be called a tragic double-drowning scenario, Ireland’s banking system went down first, and the government of Ireland went down trying to save it. Spain and Belgium flirted with, but ultimately avoided the same fate.

A third amplifier is the so-called doom-loop – the potential for a vicious feedback cycle between banks and their government. Fear about the solvency of a nation’s government fans fears about the solvency the nation’s banks, which in turn weakens the economy, thus worsening the sustainability outlook for the nation. A deadly helix of rising risk premiums and deteriorating budget deficits can suck nations into a debt default vortex. This happened to Portugal and came close to happening to Italy, Spain and Belgium.

Even France and Austria flirted with the shadow of doom-loops. The key element in this mechanism is that EZ banks are heavily invested in the debt of their own government.

The final amplifier was the rigidity of factor and product markets in nations that could not restore competitiveness via a currency devaluation. Indeed, five years down the road, few of the EZ nations have recovered their pre-crisis growth or employment rates, although Spain surprises observers with the strength of its recovery.

The third determinant of crisis severity – appropriateness of the policy response – was clearly a big problem in the Eurozone. Nothing in the EZ institutional infrastructure was prepared for a crisis on this scale. The possibility had simply not been considered when setting up the euro’s architecture. Many mistakes were made. Indeed, judging from market reactions, each policy intervention made things worse. The corner was only turned in the summer of 2012 with the ‘whatever it takes’ assertion by ECB President Mario Draghi.

Outline of the introduction

The next section presents what we believe is a consensus narrative of the EZ crisis. It has not been approved by all the contributors to this eBook, but we believe it is in line with the essence of their analyses.

The subsequent section gathers what we view as a consensus on the causes of the crisis – both the proximate causes and, if you will, the causes of the causes.

While the eBook’s focus is on the causes of the crisis, some of the authors presented remedies and the next to last section summarises the most specific of these.

The final section presents a summary and our concluding remarks.

EZ Crisis: A consensus narrative

The first step to repairing the Eurozone is to answer the questions: ‘what is broken?’, and ‘what broke it?’ This section presents what we believe is a consensus view on the main elements of the crisis’s evolution as a prelude to answering the two questions.

Before detailing the damage, “it is worth recalling that the Eurozone did not fare badly in the first years of the Global Crisis,” as Giancarlo Corsetti writes in his chapter. “Participating in the European monetary union appeared to shelter countries from the early difficulties experienced by countries with a large financial sector relative to their tax base, such as the UK and Switzerland.” That was not to last.

Building up problems

The 1990s was a time of exchange rate turbulence in Europe with major crises in 1992 and 1993. In the face of large differences in inflation rates and the occasional devaluations, markets demanded very large risk premiums. Nations like Italy, Spain and Portugal paid far higher rates on their debt than Germany and other DM-bloc countries (Netherlands, Austria, France, and Belgium). That all changed with the move towards monetary union.

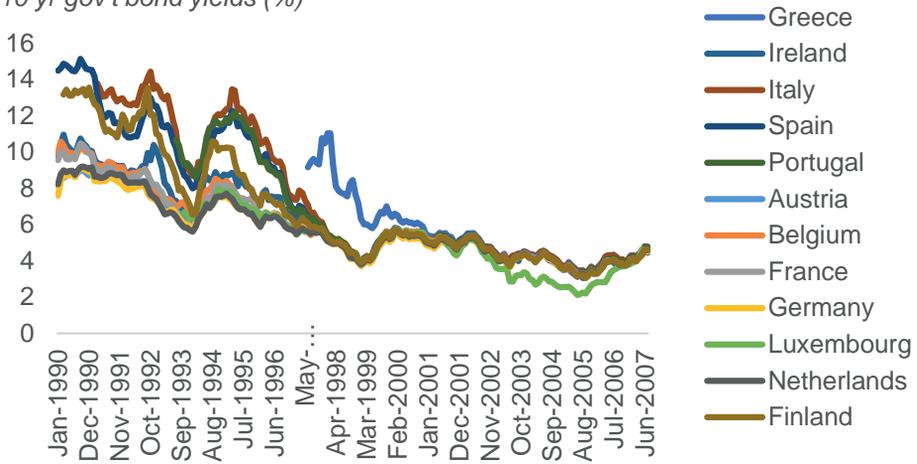
Starting from 1995, Eurozone interest rates converged in anticipation of the single currency.

- Risk premiums evaporated as markets believed the Maastricht Treaty promises of ‘no devaluation’, ‘no default’, and ‘low inflation forever for all’.

As Figure 1 shows, government bond yields for all prospective Eurozone members converged to the German rate which itself fell substantially and remained low until the start of the financial crisis. As Feld et al. write in their chapter: “Despite substantial differences in macroeconomic and fiscal policies, member countries were able to access financial markets at almost identical yields between 2001 and 2007”.

The convergence, however, was as Jeff Frankel writes in his chapter, “viewed as a good thing rather than a bad thing.” A sign of real convergence between North and South EZ countries.

Figure 1. Risk premiums disappeared in the run up to the crisis.
 10 yr gov't bond yields (%)



Source: OECD online database with authors' elaboration.

The magnitude of the changes were astounding. Italy saw the nominal cost of borrowing fall from 13% to 3% in a decade. The cheapening of credit was not quite as stark for other Eurozone members, but even Germany's government bond yields dropped from 7% to 3%. Indeed all around the world, the cost of capital fell steadily in what Ben Bernanke has been called a 'global savings glut'.

This had consequences:

- Cheaper credit encouraged borrowing throughout the monetary union, some public, some private and some foreign.

Each type of borrowing played a role in setting up the pre-crisis imbalances.

The critical imbalance: Intra-EZ lending and borrowing

The EZ crisis was not, at its roots, a sovereign debt crisis. The culprit was the large intra-EZ capital flows that emerged before the crisis.

Daniel Gros puts it succinctly: “The euro crisis started as a classic ‘sudden stop’ to cross border capital inflows. As boom turned into bust, government lost their tax base and had to assume private debt, thus creating a public debt crisis. The highly leveraged banking system of the Eurozone, tightly linked to national governments, provided a multiplier, which made the crisis systemic.”

The problem, as Paul de Grauwe notes in his chapter was that “the European monetary union lacked a mechanism that could stop divergent economic developments ... which were crystallised in the fact that some countries built up external deficits and other external surpluses.”

These imbalances baked problems into the EZ ‘cake’ since, as Guido Tabellini writes in his chapter, “if a sudden stop occurs, the sovereign most likely will lack the fiscal resources to cope with it. The size of the financial sector has grown just too large.”

Recalling that a nation’s current account is its net borrowing from abroad, large increases in foreign indebtedness shows up as a negative current account. A positive current account indicates that the nation is, on net, lending to foreigner nations.

To interpret the individual current accounts, we must depart from an essential fact:

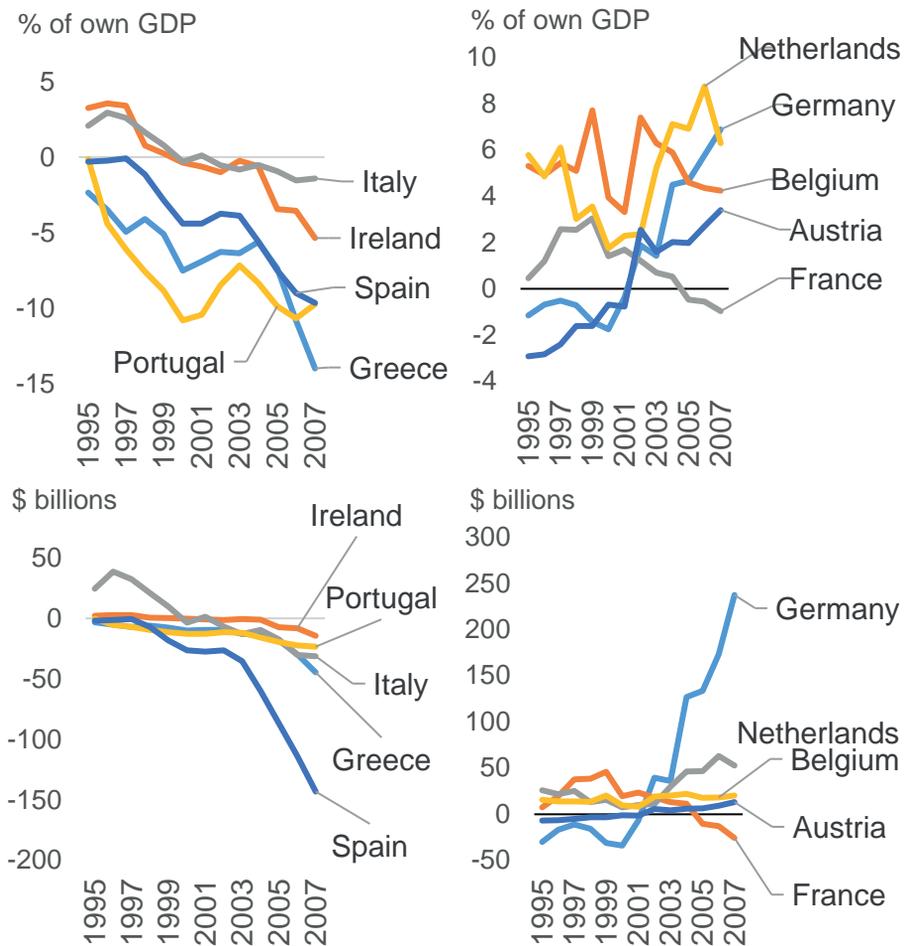
- The Eurozone’s current account as a whole was in balance before the crisis and remained close to balance throughout.

Thus there was very little net lending from the rest of the world to EZ countries. Unlike in the US and UK, the global savings glut was not the main source of foreign borrowing – it was lending and borrowing among members of the Eurozone. For example, Germany’s large current account surpluses and the crisis countries deficits

mean that German investors were, on net, lending to the crisis-hit nations – Greece, Ireland, Portugal and Spain (GIPS).

Nevertheless, as Philip Lane points out, even if the savings wasn't coming from China, the atmosphere mattered. “The evidence indicates that the 2003-2007 period can be characterised as a ‘credit supply’ shock,” he writes. “The global financial system was more willing to tolerate large net debt flows to these advanced economies.” He points to the deep causes as: “low policy rates adopted by advanced countries’ central banks, financial innovations (such as new types of securitisation) and shifting beliefs about risk levels and risk absorption capacity combined to foster an extraordinary boom in international capital flows.”

Figure 2. Current accounts: The core lent to the GIIPS from 2000 to 2007.



Source: WEO online database with authors' elaboration. Note: Current account deficit (-) and surplus (+).

With these points in mind, Figure 2 shows the evolution of the current account surplus for the turnkey nations. The top charts shows the figures for the pre-crisis years as a share of each nation's own GDP. This gives an idea of how important the flows were to the macroeconomy of each nation. As the top left panel of Figure 2 shows:

- All the nations that eventually ended in bailout programmes – Greece, Ireland, Portugal and Spain, or the GIPS for short – ran substantial and increasing current account deficits.

Italy, which almost got in trouble, was also increasingly relying, though to a much lesser extent than, say Spain, on foreign lenders during the decade preceding the crisis.

The top right panel shows that:

- The nations who had to contribute to the bailouts were all significant net lenders to other EZ nations.

France is the exception that tests the rule. It started out with a positive current account but saw its position deteriorate in the late 2000s. Unlike the other core nations, France started to have problems in the height of the crisis (its debt rating was cut in 2012).

These balance-of-payment figures as a share of GDP illustrate the importance of foreign capital flows from the perspective of the individual nations. But given the gigantic size differences among EZ members, the ratios hide important information. The second row of charts shows the current account figures in billions of dollars, not as a proportion of national GDP. Here we see that there were really two outliers:

- By 2007, Germany was, on net, lending almost \$250 billion per year to other EZ nations.

The figure for the next biggest net lending, Netherlands, was \$50 billion or less.

- Spain was by far the largest net borrower, with its capital inflows reaching \$150 in the year before the crisis.

No other nation was even close in terms on the lending or borrowing sides.

Investment versus savings

The current account is, by construction, the difference between the amount a nation invests and saves. Looking at savings and investment thus provides hints as to the drivers of current account imbalances.

As Figure 3 makes it clear:

- Broadly speaking – the core nations had above-average savings, while the GIIPS had below-average savings (Ireland being the exception).

The Dutch, Belgian and (at the end) German cases stand out in particular. Italy, Spain and Portugal, by contrast, stand out on the low side.

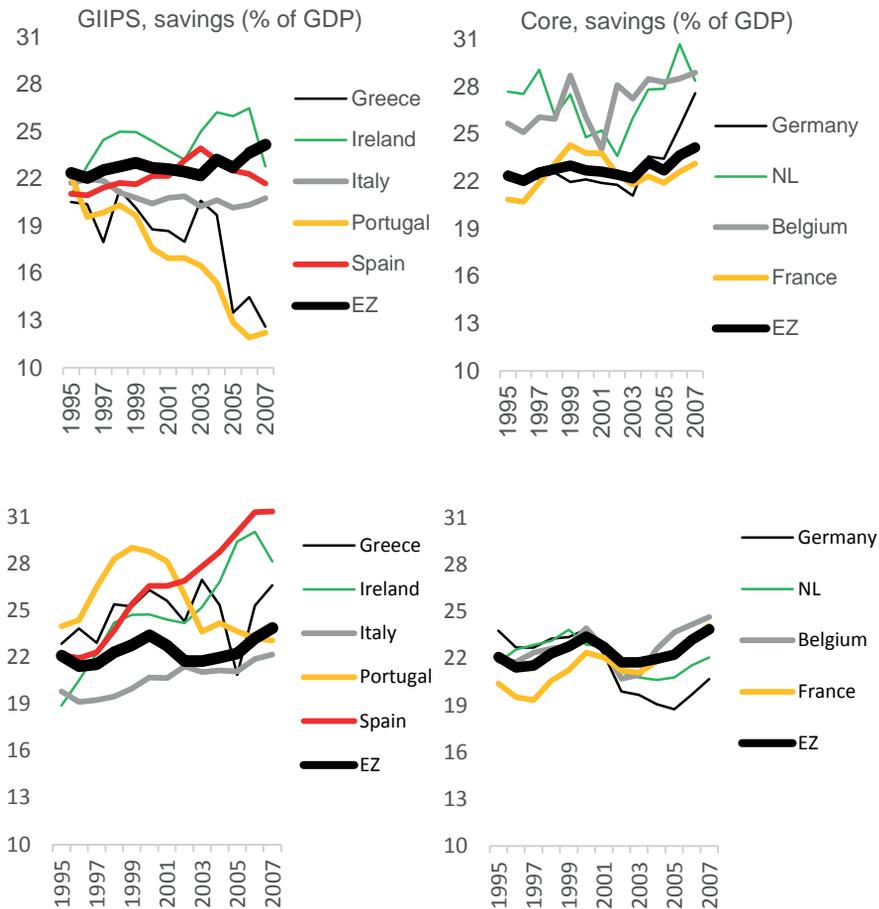
- When it comes to investment, the core nations tended to underinvest at home (compared to the EZ average), while the GIIPS tended to overinvest.

Spain and Ireland are the outliers here; they invested far more than the EZ average during the pre-crisis years.

At the time, this basic pattern was not viewed as source of EZ problems, but rather as a badge of success. It was widely believed that these sorts of private capital flows were part of the natural real convergence within a monetary union. The poorer nations, which had abundant investment opportunities, were attracting investors from richer nations where capital faced diminishing returns. Or at least that was the contemporaneous thinking (see for instance Blanchard and Giavazzi in *The Brookings Papers on Economic Activity* in 2002).

A big problem was that much of the investment headed towards non-traded sectors like government consumption and housing.

Figure 3. Core and GIIPS savings and investment patterns.



Source: WEO online database with authors' elaboration. Note: Current account deficit (-) and surplus (+).

Competitiveness imbalances

As Agnès Bénassy-Quéré put it: “capital flows tended to feed non-tradable sectors in the periphery of the Eurozone. In receiving countries, the increase in liabilities was not sustainable since it did not correspond to the building up of export capacities. Worse, it contributed to house price bubbles that would inevitably burst at some point.”

The inflows also tended to drive up wages and costs that resulted in competitiveness losses that validated the current account deficits. All four nations that eventually signed bailout packages – Greece, Ireland, Portugal and Spain – had inflation well in excess of the average.

As Feld et al write, these nations “suffered a considerable loss in price competitiveness during the debt-financed booms, due to major wage increases and high inflation. Consequently, domestic export companies were put at a competitive disadvantage and lost trade shares. The loss of price competitiveness combined with the debt-financed increase in domestic demand and the associated imports resulted in high current account deficits.”

By contrast, all the core nations except Netherlands and Luxembourg had inflation below the norm, especially Germany.

It was not supposed to work this way. In their chapter, Thorsten Beck and José-Luis Peydró note: “In fact, a key argument in favour of peripheral countries to adopt the euro was that the only way to go out of a crisis would be with a more flexible, competitive economy through structural reforms. Sadly, the substantial lower risk premiums ... that came with the euro implied strong booms but too little economic reforms.”

Public debt build up

In her chapter, Beatrice Weder di Mauro crystallises the thinking of all authors in writing: “There can hardly be any disagreement on the diagnosis that there was too much debt accumulation over the course of the first eight years of the existence of the euro.”

Figure 4 shows that the evolution of public debt does not line up well with the nations that subsequently got in trouble (Greece, Ireland, Portugal and Spain, and later Cyprus). The Eurozone as a whole lowered its debt-to-GDP ratio from 72% in 1999 to 66%

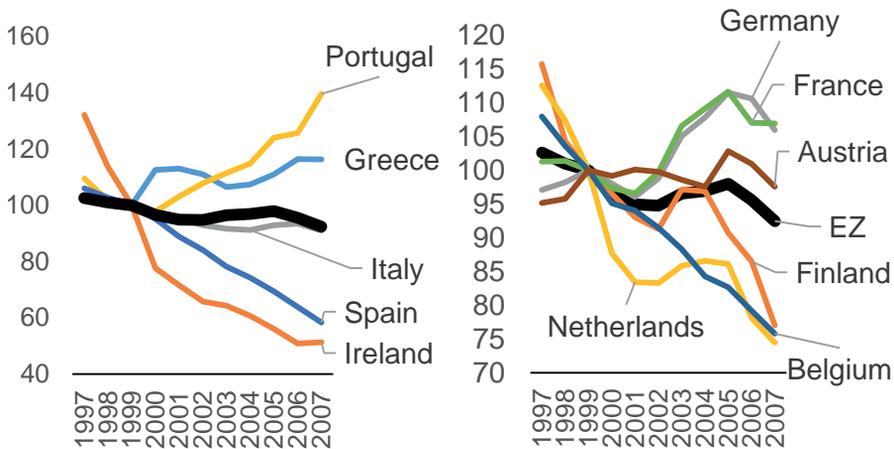
in 2007. Of course, some nations did see higher public debt ratios but others fell substantially.

- For Ireland and Spain, public debt was not a problem before the crisis.

These two were paragons of fiscal rectitude, dramatically lowering their public debt burden to far, far below the Maastricht limit of 60%. In 2007, Ireland's and Spain's ratios were, respectively, 24 and 36 percent of GDP. It should be noted, however, that both countries' government revenues were kept artificially high by tax revenues associated with a real estate boom.

Belgium, Netherlands and Finland likewise slashed their public debt piles.

Figure 4. Government debt improved for most EZ nations (1999 = 100).



Source: IMF WEO online database with authors' elaboration.

Public debt did become an issue for the other two crisis-hit nations.

- Portugal and Greece ran their debt ratios above the Maastricht limit during the calm years.

Greece's burden reached 103% but Portugal went into the crisis with a modest debt ratio of 68%. But this was not a North versus South development.

- Germany and France debt ratios drifted upwards beyond the Maastricht limit.

This was despite the sharp decline in the budgetary burden of interest payments that came with lower borrowing costs compared to the 1990s.

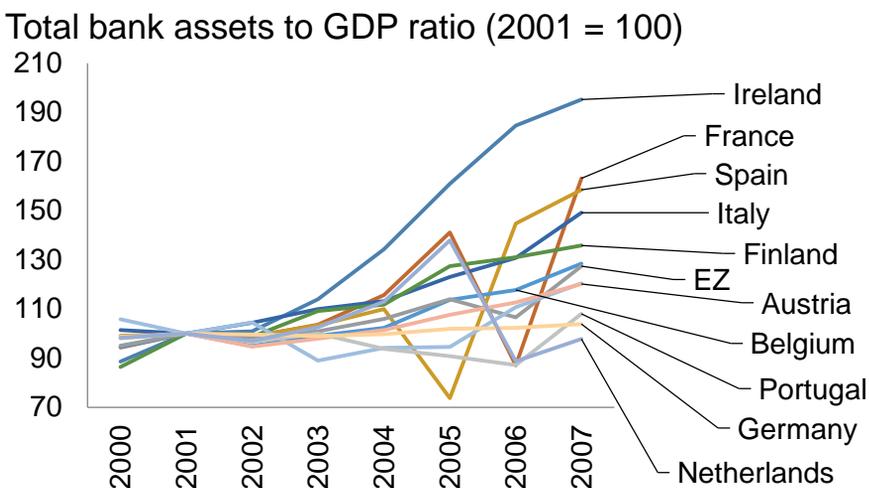
Private debt build ups

Private debt accumulation – which became a huge issue during the crisis – was run up during the heydays of the Eurozone’s first decade in some EZ members.

- Ireland’s total bank assets as a percent of GDP soared from 360% in 2001 to 705% in 2007;
- French bank debt rose from 229% to 373%;
- Italy’s from 148% to 220%, and
- Spain’s rose from 177% to 280%.

Figure 5 shows the evolution. Of course, this was not a Eurozone specific issue. It occurred in the US, the UK and other OECD nations as well.

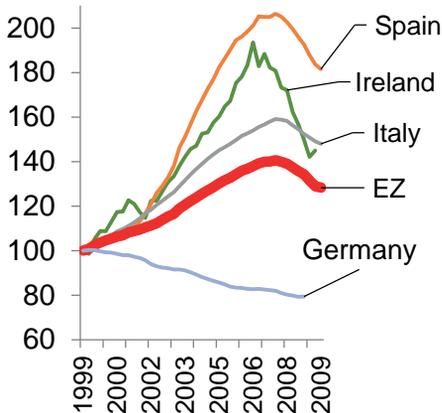
Figure 5. Rapid accumulation of bank debt was a problem.



Source: OECD online database with authors’ elaboration.

Much of the bank lending in Spain and Ireland was going to the housing sector. As Figure 6 shows, prices were rising steadily. The prices in Germany, by contrast, were falling thus creating what might have been a very tempting opportunity for German banks flush with more loanable funds than local investment opportunities.

Figure 6. House prices rose in the GIIPS and fell in Germany, 2000-2007.



Source: OECD. House price indices in real terms. Index value in first quarter of 1999 equals 100. Adapted from Baldwin, Gros and Laeven (2010).

EZ banks' cross-border lending

We can zoom in more precisely for one particularly important form of cross-border private lending/borrowing – that of banks. Table 1: Total lending from core countries' banks into the periphery (billion euros)

shows that banks from the 'core' (Germany, France, Austria, Belgium and the Netherlands) bought very large amounts of debt from the nations that would eventually get in trouble.

Table 1. Total lending from core countries' banks into the periphery (billion euros)

	1999 4th quarter	2009 4th quarter	% change 99-2009
Portugal	26	110	320
Ireland	60	348	481
Italy	259	822	217
Greece	24	141	491
Spain	94	613	554
GIPS	204	1,212	495
Total	463	2,033	340

Note: EZ core is Germany, France, Austria, Belgium and Netherlands.

Source: Adapted from Baldwin, Gros and Laeven (2010), which draws on BIS Consolidated Banking Statistics, June 2010.

This inter-linkage among core-nation banks and periphery-nations came to be a critical piece of the puzzle as the crisis unfolded.

In particular, it meant that the obvious solution of writing down Greek debt would have forced the problem back onto the core-nations leading the bailout. Solving Greece's problem in the time-honoured way might well have created classic bank crises in France and Germany (much as the eventual Greek write-down sparked a banking crisis in Cyprus).

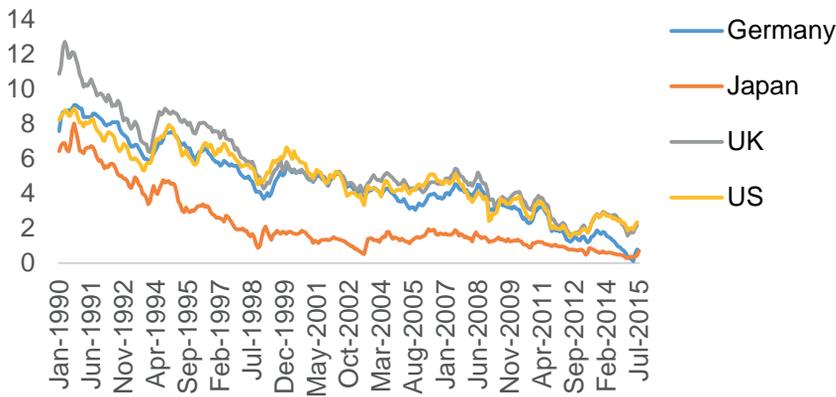
The Eurozone followed global trends

Importantly, nothing about this was unique to the Eurozone's debt accumulation.

The sum of public and private indebtedness skyrocketed in nations ranging from the US and Japan to emerging markets. This has been called the global savings glut, since the reward to holding debt fell even as its level rose. As Figure 7 shows, borrowing costs plummeted even before central banks undertook quantitative easing and other non-standard monetary measures.

Figure 7. German long-run interest rates fell in line with global trends.

% per annum



Source: OECD online database with authors' elaboration. 10-year government bond yields.

Summing up: The imbalances on the eve of the crisis

The linchpin of the EZ vulnerabilities stem from the build-up of large current account imbalances. There is nothing intrinsically wrong with such flows. If nations borrow from foreigners to invest in capacity that helps them pay off the loan, everyone can be better off. To a large extent, this was not the case. In all the GIPS, the funds ended up in various non-traded sectors.

The first column of Table 1 shows the cumulated imbalance from the euro's inception till the last year before Lehman Brothers went down in flames. The numbers for Greece, Cyprus, Portugal and Spain are enormously negative. This meant that these nations were investing far, far more than they were saving and implicitly financing it via foreign borrowing. Of course, all of this was the outcome of open markets. None of the governments (with the exception of Portugal and Greece) were involved systematically in the foreign borrowing or lending.

On the creditor side, the figures are high but not quite as high for the large core nations – Germany, France, and Netherlands. Italy, interestingly, is only modestly negative during these years at -8%.

The second column shows that for some of these nations, the inflow of foreign capital was implicitly financing budget deficits. In Greece and Portugal the large negative numbers in the first column (foreign borrowing) are matched by large negative ones in the second column (government borrowing). The large cumulative current account deficit also stand out for Spain, but it is not matched by corresponding government deficits.

Even Germany and France had cumulative public borrowing on the order of 20 percentage points of GDP over this period. Italy's was on a similar scale, although a bit higher. None of this countries however had large current account imbalances. On the surplus side, Finland and Luxembourg have unusually large numbers.

Table 1. Summary of pre-crisis imbalances

	1999 to 2007 (% of own GDP)		Bank assets (% of GDP)		%	
	Cumulative current account balance	Cumulative budget deficit	2000 to 2008 increase (p.p.)	Bank assets, 2008	Debt-GDP ratio, 2008	Excess inflation (1999- 2007)
Portugal	-96	-36	44%	262%	72	7.5
Greece	-84	-47	36%	173%	109	9.9
Spain	-60	2	121%	296%	39	9.2
Ireland	-21	14	464%	783%	43	11.6
Italy	-8	-26	85%	235%	102	1.8
EZ	-2	-17	94%	335%	69	0.0
France	6	-23	180%	395%	68	-2.9
Austria	16	-19	305%	379%	69	-3.2
Germany	27	-19	18%	316%	65	-4.8
Belgium	47	-5	83%	392%	92	-1.1
Netherlands	48	-5	-9%	375%	55	2.8
Finland	61	33	101%	197%	33	-4.9
Luxembourg	98	23	-577%	2367%	14	5.5

Source: IMF and European Banking Association online data with authors' elaboration.

Plainly, a great deal of public debt was being created during the Eurozone's peaceful years. But as Figure 4 showed, the good growth during this period resulted in falling

debt burdens in most euro nations. For reference, the endpoint government debt-to-GDP ratio is shown in the final column.

The third and fourth columns of Table 2 show the increase from 2000 to 2008 in bank assets as a fraction of GDP, and, respectively, the asset-to-GDP ratio on the cusp of the crisis. The numbers are remarkable.

- Ireland's banks added almost 4 times the nation's GDP;
- Austria's banks added 2.5 times GDP.
- Spanish, Belgian and French bank assets rose by over 100%.

By 2007, many banks were not only too big to fail, they were too big to save (Gros and Micossi 2008). Ireland's banks had assets (and thus loans) worth seven times Irish GDP. The core economies were not much better with their banks holding more than twice their nation's GDP. The figures were over three times for Germany, France and the Netherlands. Luxembourg's number was astronomic.

Unnoticed build-up of imbalances

It is, ex post, surprising that the building fragilities went unnoticed. In a sense, this was the counterpart of US authorities not realising the toxicity of the rising pile of subprime housing loans.

Till 2007, the Eurozone was widely judged as somewhere between a good thing and a great thing. The rose-garden feeling, however, started to disintegrate with the fall of Lehman Brothers in September 2008. Slowing growth and heightened fear of risk soon started to tell on the Eurozone economy as a whole, but especially for those that had built up large stocks of public and private debt, or run up large current account deficits.

- Risk premiums that had been measured in basis points for years jumped up to two or three percentage points for Greece, Ireland, Italy and Portugal.

Yet as it became clear in the Summer of 2009 that the Lehman shock would not create a second Great Depression, EZ spreads declined substantially. This was not to last.

Triggers of the crisis

Every crisis has a trigger. In Europe it was revelation of the Greek deficit deceit.

- In October 2009, the newly elected Greek government announced that the previous governments had masked the true size of the budget deficit.

The true deficit was twice as large as previously announced – a whopping 12.5%.

“The Greek fiscal crisis acted as a detonator in two ways,” write Stefano Micossi, “It alerted the authorities and public opinions in Germany and the other ‘core’ countries to the possibility of large (and hidden) violations of the common fiscal rules; and it alerted financial markets to the risk of a sovereign default in a system where the provision of liquidity to ensure the orderly rollover of distressed sovereigns is not guaranteed.”

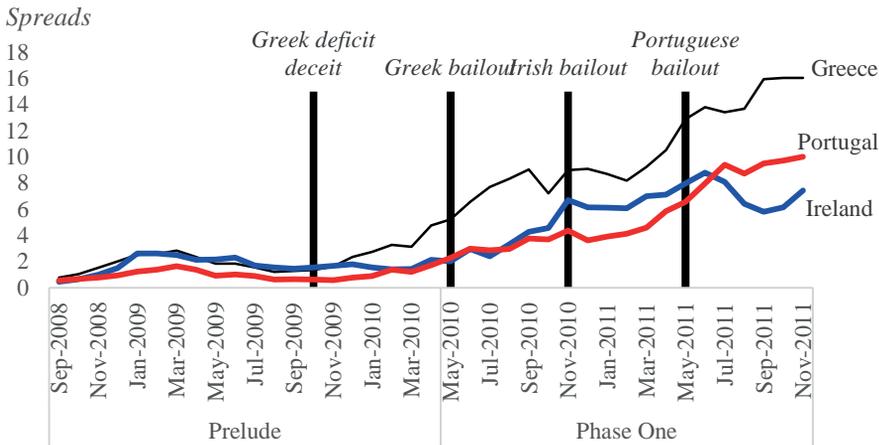
What followed was a six-month Greek effort to save itself. This failed. Greece was caught in a classic public debt vortex.

A nation’s debt is sustainable when the debt burden – commonly measured by the debt-to-GDP ratio – is not rising forever. Higher debt-service costs combined with a plummeting GDP made many investor suspect that Greek debt might be unsustainable. As markets raised their estimate of Greek default risk, they demanded higher interest rates to compensate for the extra risk. Higher rates, however, raised Greece’s debt-service payments, worsening the budget deficit.

In response, Greece slashed spending and raised taxes. But this backfired; it set off an austerity cycle. Since the austerity involved mostly tax rises, the belt tightening fanned a recession which reduce tax revenues and raised social spending, thus dragging the nation ever closer to the precipice of unsustainability. Credit agencies repeatedly

downgraded Greek government debt and its borrowing cost rapidly rose from 1.5% to 5%. Note that all this happened before the first bailout.

Figure 8. Prelude and Phase One of the Crisis in EZ Periphery.



Source: OECD online database with authors' elaboration. Note: The spreads are the difference between national 10-year government bond yields and those of Germany, in percentage points.

If a public debt vortex like this goes for long enough, there are only two ways to stop it: a default or a bailout by a 'lender of last resort'.

Phase One: Failed bailouts and contagion

Europe's leaders decide it was unthinkable for a Eurozone member to default, so Greece had to be bailed out. In the event, the 'lenders of last resort' were the Troika: Eurozone zone governments, the ECB and the IMF.

The rescue did not work; Greece's package was too little too late. Markets did the math and realised that Greece was not on a clear path to debt sustainability. The rushed and politically charged way in which the package was put together did nothing to bolster confidence in EZ leaders' ability to handle fast-moving crises.

- Greece's borrowing cost continued to soar (Figure 8).

But worse was to come.

Contagion in the periphery: Sudden stops

From early 2010, markets wondered whether Greece's inability to save itself might also apply to other nations. These doubts – combined with the remorseless logic of public debt vortices – was enough to drive up the yields in other Eurozone nations.

Importantly, debt levels were not the determinant issue when it came to which nations got in trouble.

- The nations with the highest debt ratios were not the ones hit; current account deficits were what mattered.

Given the worldwide recession, all Eurozone governments were running deficits and thus having to borrow more on the markets. Only nations that relied (implicitly) on foreigner lending (i.e. were running current account deficits) faced contagion. The borrowing costs of Portugal and Ireland rose briskly once the Greek bailout was announced (Figure 8).

This was the beginning of a 'sudden stop'. Nations that relied on foreign capital to cover their savings-investment gap – Ireland, Portugal, Spain and Italy – all nations with substantial current account deficits were affected. As it turned out, Eurozone investors were far more wary about lending to other EZ governments than they were about lending to their own.

The rise in the risk premiums set in train debt vortices that pulled down both Ireland and Portugal although via very different mechanisms. In Ireland's case, the imbalance that mattered lay in the state of its banks.

Bank debt vortices: The 'doom loop', or 'diabolic loop'

Banks, like nations, can be subject to debt vortices. Banks borrow money short-term to lend it out long term. For each euro borrowed short term, the bank makes long-term

loans of a dozen or more euros – this is called leverage. It is as profitable in good times as it is dangerous in bad times.

What puts the ‘bank’ in bankruptcy is the fact that banks go broke any time their short-term funders refuse to rollover the short-term funding. Banks, in other words, operate with a business model that would look financially irresponsible in any other sector. The whole thing only works since people believe that the banks can be rescued by a ‘lender of last resort’ – typically the national government or national central bank.

Yet two critical differences make systemic banking crises extremely pernicious: leverage and maturities.

The average Eurozone nation had a debt stock of about 70% of GDP going into the Crisis. EZ banks were holding vastly larger debts. In 2007, Irish banks held debt equal to 700% of the entire countries GDP. Plainly a systemic banking crisis in Ireland could – and in fact did – bring down the whole nation.

Moreover, the maturity of bank borrowing is typically much shorter than that of nations, so the need for new funding is radically more pressing. A typical Eurozone government may have to seek fresh loans to cover, say, 10% of its outstanding debt per year. A typical Eurozone bank has to seek fresh loans worth 10% or more of its total debt on a daily basis.

This daily need for billions means that the vortex – once it gets going – can accelerate at a frightening pace. During the Lehman Brothers debacle, one bit of bad news – Lehman’s default – brought the entire US credit market to a halt within hours; it spread to the rest of the world within days.

The ‘doom loop’ is closed by the fact that the banks – who view their national government as their lender-of-last resort – are also major lenders to the governments. The rescue, in essence, could require the rescuers to borrow from the rescued. This is how Ireland went down.

Irish banks got in trouble in 2010 and the Irish government bailed them out. In this way, private debt imbalances became a public debt imbalance.

- Despite having a very low debt-to-GDP ratio going into the crisis, this extra dollop of debt – together with the fear in the markets – pushed Ireland over the sustainability edge.

Like a tragic double-drowning, Ireland's banking system went down first, and the government of Ireland went down trying to save it. The Irish bailout was signed in November 2010. This was the Eurozone's first example of the doom-loop linking bad bank debt to national solvency.

Ireland's bailout package did little to calm investors' fears. The borrowing costs of Greece, Portugal and Ireland continued to rise. As with the Greek bailout, the Irish bailout saved the day but worsened the crisis.

By the time of the Portuguese bailout in May 2011, markets were demanding 16% for holding Greek bonds – a ruinous level even for nations in good economic shape. The Greek economy, however, was collapsing. After contracting about 5% in 2009 and 2010, it crashed by almost 9% in 2011.

In July 2011, the second Greek package was agreed in principle, but one of its elements enflamed the overall situation. As part of the EZ leaders' new view that the private sector should bear part of the cost of the bailout, private holders of Greek government debt would see about half the face value of investment disappear in what was called Private Sector Involvement (PSI). This was a wake-up call for investors who still believed the Maastricht Treaty's no-default clause.

Seeing private investors explicitly having to write down EZ government debt, and seeing how EZ leaders seemed unable to put the crisis behind them, markets drew the natural conclusion that holders of the debt of other EZ nations might also be forced into a write down.

Markets, already leery of lending across borders, became even more reluctant. Portugal, who had borrowed (via the current account) 10% of its GDP abroad in 2009 and 2010, was the next to suffer a sudden stop. Its bailout was signed in November 2011.

Once again, the bailout saved the day but worsened the crisis. After a brief respite, Irish and Portuguese rates continued their ascent towards levels that would bankrupt almost any nation. But worse was to come.

The three countries hereto caught in the crisis were small and their debts were insignificant compared to the overall EZ output. Worries started to mount when markets started demanding higher rates for the government bonds of Belgium, Spain and Italy. Italy in particular was a mortal threat to the Eurozone given the size of its economy and its massive debt.

Phase Two: Contagion spreads to the core

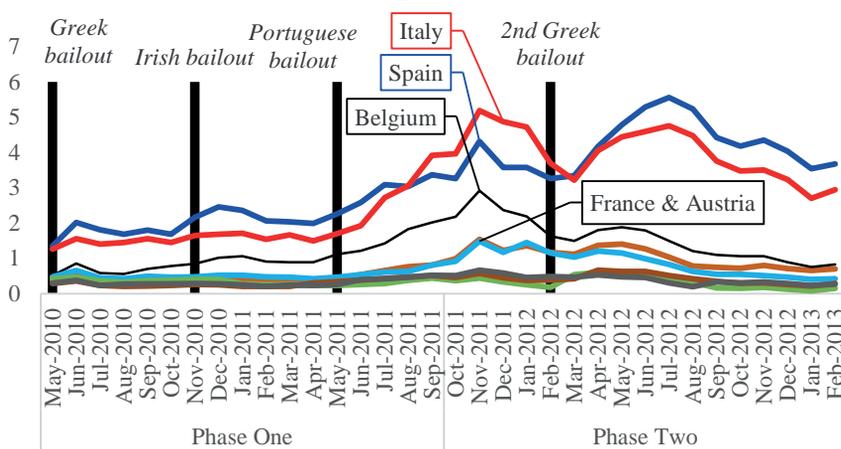
‘Phase Two’ of the Eurozone crisis had started. As IMF Chief Christine Lagarde candidly put it: “Developments this summer have indicated we are in a dangerous new phase” (Lagarde 2011).

Repeated attempts at getting ahead of the curve failed, giving rise to a general suspicion that the Eurozone crisis might spiral out of control.

- Nations that had previously been considered safe investments were now feared to be potential defaulters; borrowing costs started to move in ways that threatened to validate the fears.

This self-feeding aspect is one feature of the crisis that is both essential and elusive as it rests on perceptions.

Figure 9. Phase Two: Contagion spreads to the EZ core



Source: OECD online database with authors' elaboration. Note: The spreads are the difference between national 10-year government bond yields and those of Germany, in percentage points.

Self-feeding panic: Good equilibrium, bad equilibrium

When investors started to lose confidence in Italy – a trend which was not helped by its volatile political situation – they sold Italian government bonds in an effort to avoid future losses. These sales pushed interest rates up, making it harder for Italy to fund the rollover of its debt at reasonable rates. Seeing the funding difficulties, markets demanded higher interest rates and the spiral continued. In this way, a liquidity crisis (i.e. difficulty in rolling over debt) can – all on its own – become a solvency crisis. It's a matter of expectations.

A good way to think of this is as there being two equilibrium situations. In the first, the 'good equilibrium', markets believe Italy is solvent and thus are willing to rollover Italy's debt at reasonable rates, so Italy stays solvent. This was the situation up the beginning of Phase Two. In the second, the 'bad equilibrium', markets suspect Italy is insolvent and thus demand interest rates that make Italian debt unsustainable – thus confirming their suspicions.

Private sector involvement

The botched Portuguese bailout and EZ leaders' inability to get ahead of the curve seems to have switched Europe from the good equilibrium to the bad one. It is impossible to know why markets think what they do, but many point to the 50% 'haircut' that was a pre-condition for Greece's second bailout.

By insisting that private holders of Greek debt lose money, Eurozone leaders transformed fears of losses into real losses. Any lingering belief that default was impossible inside the Eurozone were erased. The thought foundations of the pre-Crisis imbalances in public, private and cross-border debt were shattered. The consequences were not long in coming.

The massive pre-2008 lending across EZ borders had exposed banks in the core to debt in the periphery. In early October 2011, a Franco-Belgian bank, Dexia, was pushed into a bank-debt vortex by worries over its exposure to Greek government debt. It was nationalised by Belgium by month's end. Fearing an Irish-like end to the story, the sudden stop started to drive up Belgian interest rates. Belgium had, by this point of the recession, turned from a net creditor to foreigners into a net borrower (it was in current deficit from 2008).

Given Spain's large bank debt and collapsing property markets, similar worries had begun to spread to Spain since the Portuguese bailout. Events in Belgium accelerated the process.

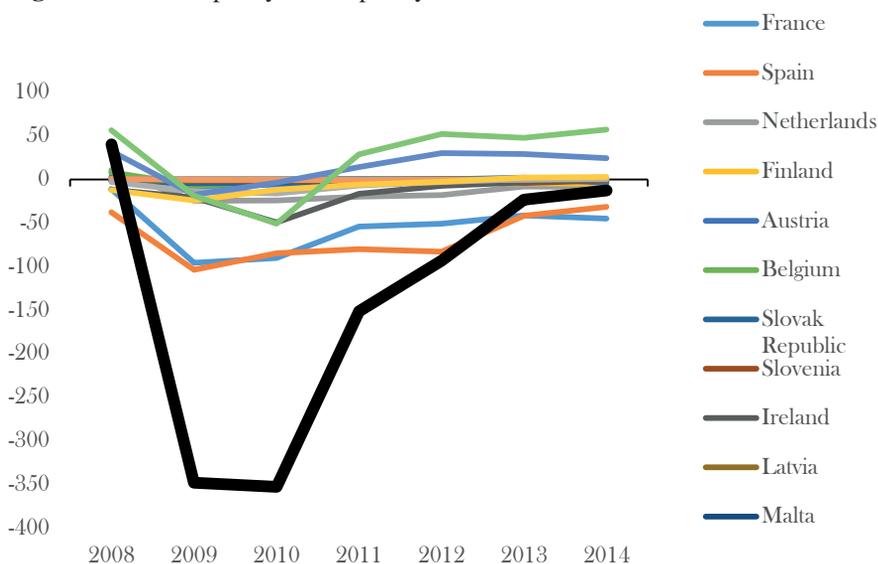
Italian yields also soared, but not due to bank problems. With a debt-to-GDP ratio over 100%, Italy needed both good growth and reasonable borrowing costs to stay afloat. It was vulnerable to a sudden stop since its implicit reliance of foreign investors rose steadily during the crisis with its rising current account deficit. Italy, in short, was coming down with the sudden-stop disease.

Sharp actions by national governments calmed the waters for a few months, but attempts to switch investors’ from the bad-equilibrium expectation to the good-equilibrium expectations failed. It all started up again when after the second Greek bailout once again disappointed markets.

Pro-cyclical fiscal tightening

The Great Recession produced counter-cyclical fiscal policy via the usual automatic stabilisers – falling tax receipts and rising social spending. This surely dampened the shock and prevented the Great Recession from becoming the second Great Depression.

Figure 10 Fiscal policy turned pro-cyclical from 2010.



Note: General government primary net lending/borrowing (billion EUR).

Source: IMF WEO online database with authors’ elaboration.

From 2010, however, the fiscal policy stance flipped from stimulus to contraction, as Figure 10 shows. The Eurozone as a whole saw its 2010 primary deficit move from about minus €350 billion in 2010 to €10 billion in 2014. This was a massive contractionary shock – equal to 4 percentage points of the monetary union’s economy.

The tightening by the GIIPS was unavoidable as they were either in bailout packages that prescribed fiscal tightening, or they were doing the tightening themselves to evade the debt-vortex.

- The GIIPS accounted for 48% of the fiscal swing even though they account for only a third of the EZ GDP (Table 2).

The effects on the economy were amplified by the fact that most countries achieved the tightening mostly by raising taxes; cutting public spending would have been less contractionary (Alesina et al 2015).

Table 2. Pro-cyclical fiscal policy, 2010 to 2014

	% of own potential GDP	bill EUR	%	%
	2010 to 2014 swing	2010 to 2014 swing	Share of EZ swing	Share EZ 2014 GDP
Greece	7%	14	4%	2%
Ireland	28%	49	14%	2%
Italy	2%	28	8%	16%
Portugal	9%	17	5%	2%
Spain	5%	53	16%	11%
EZ	4%	340	100%	100%
Austria	1%	2	1%	3%
Belgium	0%	2	0%	4%
Finland	0%	0	0%	2%
France	2%	46	13%	21%
Germany	4%	108	32%	29%
Luxembourg	1%	1	0%	0%
Netherlands	2%	15	5%	6%

Source: IMF WEO online database with authors' elaboration.

The economic impact of the crisis was also made worse by tightening in the core nations who were not suffering from the sudden stop. As Guido Tabellini puts it: "When hit by a sudden stop, domestic fiscal policy has no option but to become more restrictive, and a credit squeeze cannot be avoided as domestic banks are forced to deleverage. To avoid

a deep and prolonged recession, active aggregate demand management at the level of the Eurozone as a whole is required. But this did not happen.”

- Tightening by Germany accounted for 32% of the EZ’s overall tightening. France’s austerity amounted to 13% of the EZ total.

Denouement

Things were plainly going from bad to worse. Each attempt to end the crisis seemed to make matters worse.

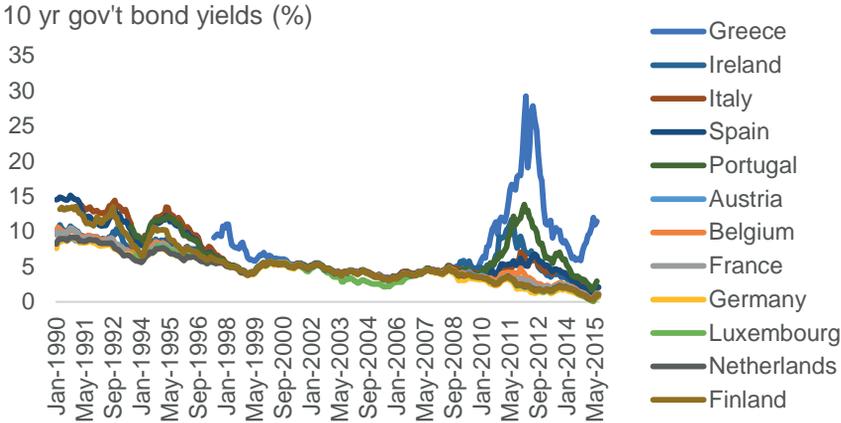
By this time, the contagion spread all the way to France. Its debt was downgraded and market yields rose substantially above those of other ‘core’ EZ nations like Germany and the Netherlands. British Prime Minister Gordon Brown unhelpfully suggested that Italy and France might need a bailout. The Belgian problem – domestic banks in trouble due to Greek lending – spread to Cyprus. Its banks were severely affected by the Greek debt write down, so the nation asked for a bailout in June 2012 (granted in March 2013).

Draghi’s “whatever it takes”

Needless to say, a crisis that threatened Italy and France was a crisis of global dimension. This was no longer an issue of Greece fiddling the books to pay for the Olympics – this had the potential of blowing up the Eurozone and the possibility the EU itself. The world economy was looking at another Lehman-sized shock. With EZ leaders manifestly incapable of mastering events, something had to be done.

Figure 11. Yields converged again after Draghi's intervention.

10 yr gov't bond yields (%)



Source: OECD online database with authors' elaboration.

That something was a forceful intervention by ECB President Mario Draghi in his famous July 2012 speech. He told markets that the ECB would do “whatever it takes” to keep the Eurozone together. That did the trick. It switched expectations from 2011 and 2012’s doom-is-inevitable back to the old we-will-get-through-this-thing expectations of 2009 and 2010. Borrow cost returned to pre-Crisis levels (Figure 11).

The basic switching mechanism that Draghi triggered is a direct corollary of the debt-vortex logic. The rush to unload debt is driven by fear. The fear is driven by the suspicion that everyone else will sell the nation’s debt, thus driving borrowing costs up to the point where the nation actually goes broke. But if there is a debt buyer-of-last-resort – someone who can buy unlimited amounts of debt – the suspicion dissolves and investors are happy to hold the debt. This is what Mario Draghi did in the Summer of 2012. So far it has worked.

Even in Greece, things seemed to be getting better. In April 2014, Greece successfully sold new debt on the open market at reasonable rates and the economy appeared to be recovering slowly. The election of the far left Syriza coalition threw things off track. Whether the Greek government’s stance was justified or not (our authors disagree), the

outcome is indisputable. Borrowing costs soared back to levels that made Greek debt unsustainable.

Causes of the crisis

The proximate cause of the EZ crisis was the rapid unwinding intra-EZ lending/borrowing imbalances that built up in the 2000s. Some of this was to private borrowers (especially in Ireland and Spain) and some of it to public borrowers (especially in Greece and Portugal), but in every case the difficult debt mostly ended up in government hands. As Thorsten Beck and José-Luis Peydró put it: “Often this private over-indebtedness ends up on governments’ balance sheets, so that the rise in public debt is more a consequence than a cause of a financial crisis.”

This ‘sudden stop’ was a crisis rather than a problem since EZ members could not devalue and their central banks could not bail out the government. As Paul de Grauwe writes: “Countries that have their own currency and that are faced with such imbalances can devalue or revalue their currencies.” This was not an option for the GIIPS.

Causes of the causes

The causes of the crisis – imbalances and lack of crisis management mechanisms –tell us that there are really three sorts of underlying causes:

- Policies failures that allowed the imbalances to get so large;
- Lack of institutions to absorb shocks at the EZ level; and
- Crisis mismanagement.

Some of these failures involved unanticipated events. Others were a failure to implement the provisions agreed in the Maastricht Treaty.

At one level, all these causes stem from a fundamental design flaw, emphasised by Feld et al. The key deficiency was the misalignment between accountability and authority, or

as Feld et al put it, the “divergence of liability and control”. If the control and liability had been supranational – as it is in US’s monetary union – the imbalances could surely have been handled without provoking a continent-wide economic crisis. It is much more likely that at least the public debt run up would have been not allowed to go so far in Greece. Likewise, if control and liability had been effectively unified at the national level, nations would have had to deal with their own debt problems, perhaps with the help of the IMF. This might also have prevented or reduced some of the pre-crisis build-ups, as happens among US states (most of whom have balanced budget clauses in their state constitutions).

Paul de Grauwe elaborates on the dire consequences of divorcing authority and accountability. “When the Eurozone was started, a fundamental stabilising force that existed at the level of the member-states was taken away from these countries. This is the lender of last resort function of the central bank.” EZ governments, “could no longer guarantee that the cash would always be available to roll over the government debt.” Unlike stand-alone nations, EZ members did not have “the power to force the central bank to provide liquidity in times of crisis.”

This created a fundamental fragility in the monetary union. Without a buyer-of-last-resort, shocks that provide re-funding difficulties in banks or nations can trigger self-fulfilling liquidity crises that degenerate into solvency problems.

At an even deeper level, Elias Papaioannou stresses the differences in national institutions as the bedrock source of problems. “Somewhat paradoxically, the criteria for joining the euro did not touch upon key institutional issues, related to state capacity (tax collection), property rights protection, investor rights, red tape, and administrative-bureaucratic quality. The high growth during the convergence period came mostly from increased investment and some limited reforms, mostly on banking and monetary policy stability.” The Maastricht criteria failed, he asserts since they focused on nominal targets rather than the fundamental institutions that are source of divergences in the first place.

Allowing the imbalances to get so large

- As discussed, the imbalances involved public and private debt as well as cross-border borrowing and lending.

Failure to bring national debt under 60%

The Maastricht Treaty assigned monetary policy to the ECB. The Bank was made politically independent and instructed to maintain price stability. This worked. Fiscal policy was a different matter.

Many warned that a monetary union without a fiscal union would be problematic as EZ members might run up unsustainable debts that would either lead to pressure on the ECB to inflate it away, or to pressure on members to bailout an insolvent member to avoid the fallout from a sovereign default.

Ultimately, fiscal union was viewed as politically unrealistic so it was left to member governments. Three safeguards were put in place to prevent problems:

- The Stability and Growth Pact was to keep deficits below 3% of GDP in normal times, and debt levels below, or at least heading towards, 60% of GDP;
- The ECB's independence was to protect against political pressures to inflate away debt, and the ECB was explicitly forbidden from financing members' deficits; and
- The no-bailout clause was to protect against the pressure to bailout failing EZ members.

During the pre-crisis years, the SGP failed. As Feld et al write: The SGP "sanctioning mechanisms were barely employed; there were 34 breaches of the 3% threshold for the general government deficit between 1999 and 2007. ... The breaches of the pact by Germany and France set particularly detrimental precedents."

Failure to control excessive bank leverage

Many Eurozone banks were dangerously overleveraged going into 2010 due to regulatory failures before 2007, and half-hearted bank clean-ups after the Global Crisis. The responsibility for bank regulation was, as with fiscal policy, left to EZ member states, but there was no SGP for banking.

Banking supervision was not really a focus of the 1990s discussion leading up to the Maastricht Treaty. Coordination of banking rules was at best mentioned in passing in the 1990s and few thought of moving deposit insurance, supervision, or bank resolution to EZ level.

National central banks in both the surplus and the deficit countries -- which at the time still in charge of local banking supervision -- failed to realise what the huge intra-EZ credit flows were financing.

Lack of institutions to absorb shocks

The public and private debt problems interact in a deadly embrace called the doom loop. This is a cause of the sudden stop and crisis contagion in the sense that it opens the door self-fulfilling crises. It makes the monetary union vulnerable to shocks that get amplified all out of proportion even if the debt imbalances are not extreme.

The problem, as Daniel Gros writes is that “In Europe the banks and the sovereign are usually so closely linked that one cannot survive without the other.” As discussed above in the context of the doom loop, EZ national governments are the ultimate guarantor of their banks, but the banks are key holders of public debt. As a result: “insolvency of a government would also wipe out the capital of the banks and bankrupt them as well. But an insolvent government would no longer be able to save its banks.”

The key element in this doom loop is that banks hold large amounts of the debt of their own government. If banks instead held more diversified portfolios of public debt, there would be much less doom in the doom loop.

The problem this creates, according to Agnes Bénassy-Quéré, is that it makes sovereign default very difficult. “The risk is that a sovereign default could trigger bank insolvency. The latter would then trigger a liquidity crisis since insolvent banks cannot get refinancing within the lender-of-last-resort procedures. To avoid a collapse of the national economy, there would be no other way than to re-introduce a national currency to refinance the banks.” This was a real possibility for Greece during its Summer 2015 crisis.

Beatrice Weder di Mauro illustrates the problems created by the inability to deal with insolvent EZ members. “Although Greece is extreme in many ways, it has repeatedly highlighted the European failure to establish a regime for dealing with cases of unsustainable debt. After the no-bail out clause failed to prevent excess debt accumulation, the Eurozone had to find a quick fix when Greece lost market access in 2010.”

Lack of collective action mechanisms

The risks of credit imbalances can be diminished by surveillance and avoiding the accumulation of excessive imbalances. But it will never disappear. As Paul de Grauwe says, a fundamental feature of a capitalistic system “is that it is characterised by booms and busts; bubbles and crashes.”

If a sudden stop occurs, the sovereign most likely will lack the fiscal resources to cope with it. The size of the financial sector has grown just too large: at the end of 2007, bank assets were several multiple of GDP in most EZ countries (Table 1). Unlike a typical country hit by a sudden stop, an EZ member cannot devalue its currency to cope with

the crisis. Yet, as became crystal clear in Greece but also in other countries of Southern Europe, currency risk is a major concern for market participants.

During a sudden stop, the bank-sovereign loop that we have seen at work during the crisis becomes inevitable. The home bias of bank portfolios aggravates the loop. But in the presence of currency risk, even a bank with well diversified assets would not be able to withstand the flight to safety of its depositors. And the sovereign would typically not be in a position to help, given that it cannot devalue nor print money.

In countries with high public debt the sovereign itself could be the primary source of fragility, and its exposure to debt runs could activate the bank-sovereign loop. Any country with a large public debt, and with no access to monetary financing, could be subject to a run on its debt, even if it was solvent in the long run. In other words, a liquidity crisis triggered by lack of confidence could push into insolvency not only banks, but also sovereigns with high public debts and no access to the printing press.

By the end of the acute phase of the crisis, Spain was able to borrow from an EZ-level institution to fix its banks. This was a major improvement over the situation faced by Ireland at the start of the crisis. But the funds lent by the ESM did not go directly to Spanish banks, increasing their capital. They were borrowed by the Spanish government, which in turned used them to increase bank capital. The outcome was a further increase in sovereign debt.

Failures in crisis management

The EZ crisis was mismanaged on many levels. When it came to the firefighting moment, it is puzzling that, as Beck and Peydro state: “European policy makers decided not to draw on the extensive crisis resolution experience in and outside Europe, a decision which has led to substantial number of policy mistakes.”

Charles Wyplosz provides the response: “The simple answer to these questions is that the European treaties never anticipated that there could be such a crisis.” He continues

by judging that: “politicians have reluctantly been led to micromanage complex technical issues as the result of an amazing accumulation of economic mistakes, which they are unwilling to recognise.”

The lack of clear lines of liability and control caused problems for crisis management. As Giancarlo Corsetti writes; “Instability grew out of a disruptive deadlock between national governments forced to address and correct fundamental weaknesses in their national economies on their own, and the EZ-level policymaking, which could have created the conditions for successful implementation of national policies, but did too little, too late (at best).”

This lacunae was on fully display in Summer 2015. Despite repeated debt restructuring, the economic fallout from this year’s long-lasting negotiations between Greece and her creditors has probably made the nation insolvent. According to the IMF Greek debt is unsustainable without further restructuring.

More specifically, when Greece got into its sovereign debt crisis in 2010, the standard solution would have been for Greece to turn to the IMF for help. But, as Jeff Frankel writes: “the reaction of leaders in both Frankfurt and Brussels was that going to the IMF was unthinkable, that this was a problem to be settled within Europe. They chose to play for time ...”. This turned out to be a critical mistake.

When it became clear that the Irish banks were in trouble, no EZ structures existed to facilitate a collective rescue. In this sense, the policy mistake was not having anticipated the problems that would arise when the responsibility for banks was with nation governments but the capacity to bailout them was only at the EZ level.

While there were important mistakes made in crafting the bailout packages, this is not the place for detailed critiques. Nevertheless common problems came in the sequencing of reforms and a general lack of national ownership of the credit-for-reform deals. In retrospect, it is also clear that it was a mistake to not write down more of Greece’s debt early on when most of the debt was still held by private creditors. As Frankel writes,

private debt holders: “could usefully have taken a ‘haircut,’ in a way that public sector creditors cannot ... But again, leaders in both Frankfurt and Brussels insisted in 2010 and 2011 that writing down the debt was unthinkable.”

Some of the ‘mistakes’ were in fact linked to deeper conflicts. As Philip Lane stresses, the creditors and debtors in the bailout packages share a common currency and are deeply intertwined in terms of economic and political linkages. This inevitably creates conflicts of interest when it comes to the design of the bailout programmes and the potential role of debt restructuring mechanisms.

Failure to understand the monetary union regime

EZ governments systemically failed to understand the regime change implied by adoption of a common currency. A key aspect of this was the implication of financial integration. The thinking in the 1980s and 1990s, when financial integration was introduced before the common currency, was that it would contribute to both convergence and macroeconomic stability. Instead, capital flows tended to feed non-tradable sectors in the periphery of the Eurozone. Second, financial integration did not play as a smoothing device when the crisis hit. Quite the opposite, crisis countries suffered sudden stops.

In this same line is the idea that the Eurozone was as a whole a large but fairly closed economy, while EZ governments continued with the mindsets of small open economies – each ignoring the impact of their own actions on the situation faced by the collective.

Some remedies

The goal of this eBook is to establish a consensus on the causes and a narrative for the EZ crisis. The idea is to agree what happened as a first step towards developing a consensus on what should be done to fix the current problems and to create mechanisms that will make the next crisis less damaging.

Although the main focus of this eBook is the analysis of what happened trying to see if there is a common narrative, some authors did not refrain from suggesting how the monetary union should be reformed. Here proposals vary and understandably is harder to find a common theme.

Guido Tabellini stresses the need to develop shock absorbers at the EZ level. Referring to the classic macroeconomic trilemma (the Eurozone cannot have full financial integration, financial stability and no common fiscal policy), he writes: “This trilemma implies that, in order to preserve financial integration and avoid future crisis, we need adequate common fiscal resources to cope with both systemic banking crisis and sovereign debt runs.”

He also stresses the need for a remedy to the tendency of EZ members to go for pro-cyclical fiscal policy. “This aggregate demand mismanagement was not just the result of human error. It reflects the institutional design of the Eurozone.” His solutions would surely attract disagreement from other authors. “These institutional features that led to this mismanagement ought to be corrected by changing the mandate of the ECB, by removing the constraints on monetary financing in order to facilitate a coordinated monetary and fiscal expansion, and by endowing the Eurozone with the possibility of issuing and servicing its own debt.”

A fresh solution to the difficulties of adjustment with a fixed nominal exchange rate is suggested by Pesenti. The idea is that structural reforms make it possible to substitute relative price changes with shifts in the composition of output. But setting up firms and new production lines is costly and typically requires financial resources. Structural reform cannot succeed without appropriate policies that address tight credit constraints on investment and firms’ activity due to liquidity and balance sheet problems hitting banks. We no longer have a dichotomy between costly reforms and anti-recessionary monetary policy, but rather an integrated and perhaps coordinated vision of monetary and structural policies to restore growth. The alternative, unattractive, option is

continuing reliance on deflationary adjustment in a currency union stuck at the zero lower bound.

Beatrice Weder di Mauro stresses the legacy debt as the core fragility that must be eliminated if any solution is to work. Referring to a report she helped write earlier this year that accepts the no-mutualisation red line as a given, she writes: “Corsetti et al (2005) have proposed an alternative approach for a rapid and concerted debt reduction. The proposal involves an agreement by all Eurozone countries to commit future revenues for the sake of retiring debt. They would bring forward current and future income streams and commit their net present value to buy back the national debt now. Capitalising even small current and future income streams over a long horizon generates in net present value terms a large sum of money to buy back the debt. In addition, elements of solidarity and a debt equity swap could make the debt reduction deal viable and equitable.”

“The aim of the debt reduction deal is to eliminate legacy public debt, bringing debt in countries (except Greece, which is a special case), below 95 percent and thus plausibly into the zone of solvency.”

“The second pillar of the proposal is a regime to deal with cases of unsustainable sovereign debt, which would help prevent countries from becoming too big to fail, again.”

Concluding remarks

When Europe’s Economic and Monetary Union (EMU) was being designed in the late 1980s, there was no clear vision on the standards of political and institutional cohesion among members that would be required to make the project viable. “The consensus view was that the member states of the union would be able to reach agreement and cooperate on how to create the common currency institutions over time,” as Giancarlo Corsetti put it. Euro nations would be able to muddle through any problems.

Although shocks would create fault lines and policy conflicts, it was widely expected that the cooperation would prevail – as it had always done before. The historical importance of tying together Europe would provide sufficient motivation to overcome obstacles, smooth differences over policy, and elicit solidarity.

And the need for further changes was quite clear. For example, CEPR's Monitoring the ECB report (Begg et al. 1998) wrote: "The ECB suffers serious faults in its design that sooner or later will surface. This is likely to happen when large shocks, such as the world financial crisis, hit euroland," where the world crisis referred to here was the 1997 Asian Crisis. "The lack of centralised banking supervision, together with the absence of clear responsibilities in crisis management, risk making the financial system in euroland fragile. No secure mechanism exists for creating liquidity in a crisis, and there remain flaws in proposals for dealing with insolvency during a large banking collapse. ... These design faults are due to a failure to put sufficient decision making power at the centre of the system."

Europe's bad luck exposed the costs of relying on Monnet's view. Shocks the size of the Global Crisis and Great Recession were not really what the Eurozone's architects had in mind when they thought they could rely on muddling-through. From 2010 to 2012, Monnet's logic was turned on its head. Discussion of how to complete the currency union proved divisive and destabilising. It made matters worse and a consensus had to be reached on institutional issues while knitting together agreements on immediate emergency measures.

The consequences were and still are dreadful. Europe's lingering economic malaise is not just a slow recovery. Mainstream forecasts predict that hundreds of millions of Europeans will miss out on the opportunities that past generations took for granted. The crisis-burden falls hardest on Europe's youth whose lifetime earning-profiles have already suffered.

Money, however, is not the main issue. This is no longer just an economic crisis. The economic hardship has fuelled populism and political extremism. In a setting that is

more unstable than any time since the 1930s, nationalistic, anti-European rhetoric is becoming mainstream. Political parties argue for breaking up the Eurozone and the EU. It is not inconceivable that far-right or far-left populist parties could soon hold or share power in several EU nations.

Many influential observers recognise the bind in which Europe finds itself. A broad gamut of useful solutions have been suggested. Yet existing rules, institutions and political bargains prevent effective action. Policymakers seem to have painted themselves into a corner.

This eBook is a first step in a bigger project called “Rebooting Europe”. It seeks to marshal a critical mass of Europe’s best thinkers in developing ways to get Europe working again. To undertake a systematic rethink of today’s European socio-economic-political system. In short, to figure out a way to update Europe’s ‘operating system’ and reboot.

References

Alesina, Alberto, Omar Barbiero, Carlo Favero, Francesco Giavazzi, and Matteo Paradisi, “Austerity in 2009–13”. *Economic Policy*, 30 (83), pp. 383-437.

Baldwin, Richard, Daniel Gros and Luc Laeven (2010). “Introduction” in Baldwin R, D Gros and L Laeven (eds.), *Completing the Eurozone rescue: What more needs to be done?* VoxEU.org eBook, CEPR Press.

Begg, David, Francesco Giavazzi, Paul De Grauwe, Harald Uhlig and Charles Wyplosz (1998), *The ECB: Safe at Any Speed? Monitoring the ECB*, Vol. 1, CEPR Press.

Gros, Daniel and Stefano Micossi (2008). “The beginning of the end game...”, VoxEU.org column, 20 September 2008.

Lagarde, Christine (2011). “Global Risks Are Rising, But There Is a Path to Recovery”: Remarks at Jackson Hole, 27 August 2011.

SUDDEN STOPS IN THE EURO AREA

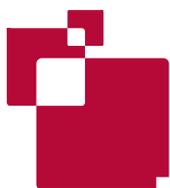
SILVIA MERLER AND JEAN PISANI-FERRY

Highlights

- The single currency was expected to make balance of payments irrelevant between the euro-area member states. This benign view has been challenged by recent developments, especially as imbalances between euro-area central banks have widened within the TARGET2 settlement system.
- Current-account developments can be misleading as indicators of financial-account developments in countries that receive significant official support. Greece, Ireland, Italy, Portugal and Spain experienced significant private-capital inflows from 2002 to 2007-09, followed by unambiguously massive outflows.
- We show that such reversals qualify as '**sudden stops**'. Euro-area sudden-stop episodes were clustered in three periods: the global financial crisis, a period following the agreement of the Greek programme and summer 2011. The timeline suggests contagion effects were present.
- We find evidence of substitution of the private capital flows with public components. In particular, weak banks in distressed countries took up a major share of the central bank refinancing. The steady divergence of intra-Eurosystem net balances mirrors this.
- In the short term, TARGET2 imbalances could be addressed by tightening collateral requirements for central bank liquidity. For the longer term, the evidence that the euro area has been subject to internal balance-of-payment crises should be taken as a strong signal of weakness and as an invitation to reform its structures.

Telephone
+32 2 227 4210
info@bruegel.org

www.bruegel.org



Jean Pisani-Ferry (jean.pisani-ferry@bruegel.org) is Director of Bruegel. Silvia Merler (silvia.merler@bruegel.org) is a Research Assistant at Bruegel.

SUDDEN STOPS IN THE EURO AREA

SILVIA MERLER AND JEAN PISANI-FERRY, MARCH 2012

THERE IS A VIEW that the euro crisis is a balance-of-payments crisis at least as much as a fiscal crisis¹. This claim could have a bearing on the nature of the policy response, which thus far has concentrated on strengthening budgetary discipline and has treated external imbalances as a second-order matter.

The issue has become more relevant with the widening of imbalances between euro-area central banks within the TARGET2 settlement system – the Eurosystem's interbank payment system². The cumulated net position of the northern euro-area central banks reached €800 billion in December 2011, being matched by the southern euro-area central banks' equivalent negative position.

The balance-of-payments discussion lacks clarity, however. First, it seems awkward to speak of balance-of-payments crises within a monetary union that was designed to make such crises impossible. Second, few of the proponents of the balance-of-payments crisis view have substantiated their claims with clear evidence. Unlike a standard balance-of-payments crisis, within the euro area, current-account deficits have adjusted partially and slowly. Third, the relationship between TARGET2 balances and balance-of-payment imbalances remains confused.

The purpose of this paper is to fill these gaps. We start in section 1 with a brief discussion of the possibility of a balance-of-payment crisis within a monetary union and an overview of the evolution of current-account balances. In section 2 we analyse the evolution of private capital flows to southern Europe before and during the euro crisis. In section 3 we proceed to a more formal test and apply standard sudden-stop criteria to the evolution of capital flows. In section 4 we discuss the roles played by central banks and official financing. We return to policy issues in section 5 to discuss the consequences of our findings.

1 CRISIS? WHAT CRISIS?

In one of the earliest papers on European monetary union, Ingram (1973) notes that in such a union “*payments imbalances among member nations can be financed in the short run through the financial markets, without need for interventions by a monetary authority. Intracommunity payments become analogous to interregional payments within a single country*”³. This view was not challenged in the debate of the 1980s and the 1990s on the economics of Economic and Monetary Union (EMU). It quickly became conventional wisdom. The European Commission's *One Market, One Money* report (1990) similarly posits that “*a major effect of EMU is that balance-of-payments constraints will disappear [...]. Private markets will finance all viable borrowers, and savings and investment balances will no longer be constraints at the national level*”⁴. The important words here are “*all viable borrowers*”, meaning that the budget constraint applies to individual borrowers, not to countries as such. In other words a solvent company in Italy or a solvent bank in Spain cannot be cut off from market financing because of the situation of the sovereign or the households. There is no such thing as a specific country-level intertemporal budget constraint – only those of individual agents matter.

This view was so widespread in the early 1990s that the Maastricht negotiators decided to exclude members of the common currency from the benefit of EU balance-of-payments assistance under Article 143 of the Treaty – with the result that the euro area was left without an instrument to provide assistance to Greece and had to rely in a first step on bilateral loans from its member countries, before the European Financial Stability Facility and the European Stability Mechanism were created. As reported in Marzinotto, Pisani-Ferry and Sapir (2010), this exclusion had nothing to do with the no-bail out clause. It was simply assumed that

1. See for example Carney (2012), Giavazzi and Spaventa (2011), Sinn (2012); Martin Wolf in the *Financial Times* has reflected this view on a number of occasions.

2. The Eurosystem is the monetary authority of the euro area, comprising the European Central Bank and the central banks of countries using the euro.

3. Ingram (1973), p10.

4. European Commission (1990), synthesis chapter, p24.

balance-of-payment crises within the euro area would become as unthinkable as they are within countries⁵.

To our knowledge, the only one to challenge this benign view was Peter Garber in a 1998 paper on the role of TARGET in a crisis of monetary union (Garber, 1998). The paper insightfully recognised that the federal structure of the Eurosystem and the corresponding continued existence of national central banks with separate individual balance sheets made it possible to imagine a speculative attack within monetary union. According to Garber, the precondition for an attack “*must be scepticism that a strong currency national central bank will provide through TARGET unlimited credit in euros to the weak national central banks*”. His conclusion is that “*as long as some doubt remains about the permanence of Stage III exchange rates, the existence of the currently proposed structure of the ECB and TARGET does not create additional security against the possibility of an attack. Quite the contrary, it creates a perfect mechanism to make an explosive attack on the system*”.

As said, the benign view prevailed during the first ten years of EMU. It even continues to dominate today. Indeed, casual data observation seems to vindicate it. Figure 1 reports the 2007-11 evolution of current-account balances in the three non-euro area EU countries and the three euro-area countries with the highest deficits in 2007⁶. It is apparent that the two groups of countries have not followed the same path: whereas adjustment has been brutal for the first group, with deficits amounting to 15 to 25 percent of GDP transformed into surpluses over three or four years, it has been

very slow for the second. One may even wonder if Greece and Portugal have adjusted at all.

2 PRIVATE CAPITAL FLOWS

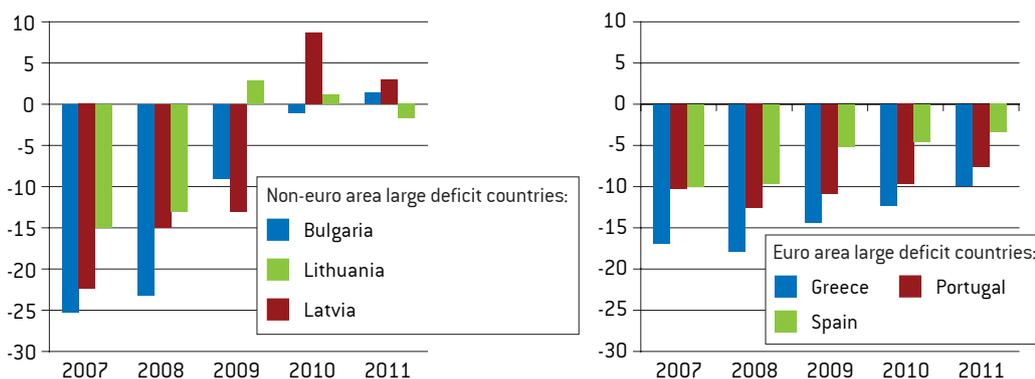
Assessing if there has been a balance-of-payment crisis by looking at the evolution of the current account is however a flawed approach. It is adequate to look at the evolution of current-account balances as long as it offers a mirror image of net private capital flows. In a stand-alone country, this is largely the case except for foreign exchange interventions by the central bank – at least as long as the country is not under an International Monetary Fund programme. This is however not the case for monetary union, because the financial account includes official capital flows. The correct accounting identity (neglecting the balance of the capital account as well as errors and omissions) is:

$$(1) CAB + PCI + T2F + PGM + SMP = 0$$

in which CAB stands for the current-account balance, PCI for private capital inflows, T2F for Eurosystem financing through the TARGET2 system (change in the net liability of the national central bank vis-à-vis the rest of the Eurosystem), PGM for financing through official IMF and European assistance, and SMP (Securities Markets Programme) for European Central Bank purchases of government securities from residents. Of these five flows, four are recorded statistically and only one (SMP) is not known.

In what follows we evaluate private capital inflows to southern Europe from 2002-11 using monthly

Figure 1: A tale of two adjustments: current accounts outside and within the euro area



Source: ECFIN Forecasts November 2011.

5. The literature of the 1990s explored this comparison and showed that the Feldstein-Horioka paradox vanishes entirely when applied to regions within countries. See for example Bayoumi (1999).

6. We have excluded Cyprus because of its small size.

financial account data. Capital flows are taken from national balance-of-payments as published by national central banks, and we deduct from them official inflows resulting from changes in TARGET2 balances (see Box 1) and assistance

under IMF/EU programmes (see Appendices 1 and 2 for details).

As we want to focus on inflows and reversals, not short-term fluctuations, and to compare evolu-

Box 1: TARGET2

TARGET2 (Trans-European Automated Real-time Gross settlement Express Transfer) is the Eurosystem's operational tool through which national central banks of member states provide payment and settlement services for intra-euro area transactions. Intra-Eurosystem claims arise from different types of transactions and they can or cannot have a 'real' counterpart: they might be the result of transfers of goods that require a cross-border payment (ie imports) or the transfer of deposits to a different euro-area country. When capital is transferred (eg a deposit is moved) from an Irish bank to a German bank via TARGET2, the transaction is settled between the Irish central bank and the Bundesbank, with the former incurring a liability to the latter. TARGET2 can be used for all credit transfers in euro and it processes both interbank and customer payments. There are transactions for which TARGET2 must be used⁷ but for all the other payments – interbank and commercial payments in euro – market participants are free to use TARGET2 or any other payment system of their choice. Banks prefer the TARGET2 system because most banks in Europe are reachable through it and payments are settled immediately (immediate finality of the transaction) and in central bank money (allowing credit institutions to transfer money held in accounts with the central bank among themselves)⁸.

The settlement of intra-Eurosystem payments via TARGET2 gives rise to cross-border obligations that are aggregated and netted out at the end of each single business day, leaving national central banks with a certain net TARGET2 balance (positive, negative or zero). There is no a priori limit to the transactions that can be processed by the system – and therefore to the size of TARGET positions. Daily net balances are generally remunerated at the respective interest rate for main refinancing operations⁹.

TARGET2 balances are balances that each central bank accumulates from the operations conducted vis-à-vis other national central banks in the euro area, but the final balance is a claim or a liability against the ECB, the ultimate manager of liquidity. In a way, it is as if the ECB were intermediating all transactions among national central banks¹⁰.

Until 2007, TARGET2 positions remained close to balance. From 2007 (and more so with the intensifying of the sovereign debt crisis in 2010) the balances started to diverge, with Germany becoming the largest creditor and Greece, Spain, Ireland and Portugal being traditional net borrowers and Italy moving into a negative position during the summer of 2011. The huge increase in TARGET2 claims and liabilities has recently drawn attention, triggering a debate on the forces behind this steady divergence (see Sinn and

7. Operations including the Eurosystem monetary policy operations as well as for the settlement of position in large-value net settlement system that effectively operate in euro.

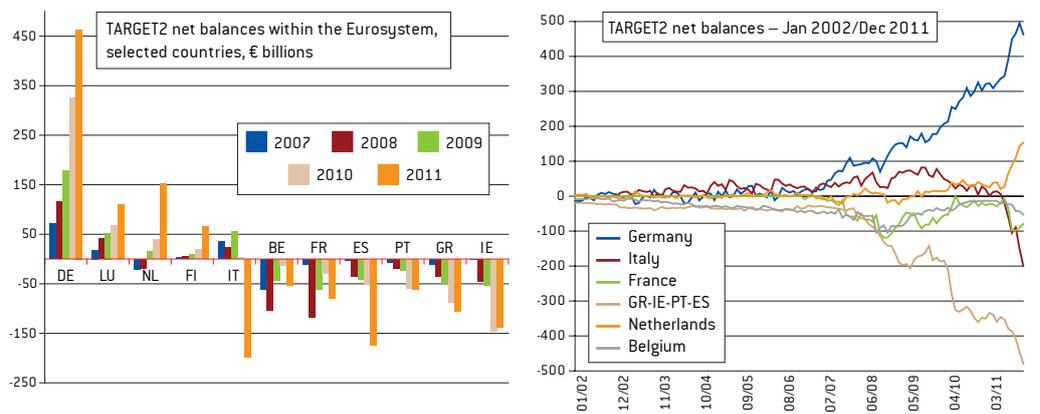
8. See Kokkola (2010).

9. Deutsche Bundesbank Annual Report 2010.

10. The multilateralisation of the claims is an important feature of the system.

It implies that any loss resulting from a central bank's failure to settle its debts would be shared among all the members of the Eurosystem, irrespective of their creditor or debtor positions in the TARGET2 system.

Figure B1: TARGET2 Balances in the euro area



Source: Bruegel, national central banks.

Wollmerhaeuser, 2011; Buitier, Rahbari and Michels, 2011; Bindseil and Koenig, 2012; and Bornhorst and Mody, 2012).

The build-up of such imbalances presupposes that capital does not flow uniformly across countries, meaning that central banks that report a deficit position have been systematically settling more outward payments than inward payments. In other words, some countries have been constantly net borrowers and other countries have been net lenders. This development is closely related to the tensions on the interbank markets and the increase in the perceived country risk in southern Europe. While payments between credit institutions can or cannot be processed via TARGET2, the transfers related to the Eurosystem monetary policy operations are managed through the system, so when the use of central bank liquidity becomes unevenly distributed across countries, TARGET2 balances will reflect it. The steep increase in TARGET2 claims and liabilities from 2008 onwards suggests that tensions in the financial system may have an important role in explaining the divergence. In a period of financial crisis, banks in countries undergoing net payment outflows (eg deposit flights) need liquidity but can find it difficult to refinance on the interbank market (also because of the shortage of valuable collateral), and will therefore resort more to central bank liquidity than banks in countries to which money is flowing.

Germany is a good example of this mechanism: the volume of central bank refinancing attributable to German banks decreased from €250 billion at the start of 2007 to €130 billion in 2010¹¹, signalling that German banks have been reducing on average their reliance on central bank liquidity. Symmetrically, demand for ECB liquidity from banks located in troubled countries increased considerably over the same period. In light of these considerations, TARGET2 imbalances can therefore be interpreted as evidence of a changing distribution in countries' refinancing operations, and as a compensation mechanism that allows sound banks in stressed countries to cover their liquidity needs.

tions across countries, we plot for all countries cumulated capital inflows in proportion to their 2007 GDPs, taking as a starting point the end-2001 net investment position of the country as recorded by Eurostat¹². Figure 2 on the next page presents the results for Greece, Ireland, Italy, Portugal and Spain. In each case the blue line gives total cumulated flows, and the red line total cumulated private flows.

Figure 2 provides evidence that all five countries experienced significant private capital inflows from 2002 to 2007-09, followed by unambiguous and rather sudden outflows. In Greece inflows and outflows each amounted to about 40 percent of 2007 GDP. In Ireland inflows were limited but outflows reached 70 percent of 2007 GDP. In the other three countries outflows were less sizeable and started later, but nevertheless they were of significant size.

It is interesting also to observe the timing of reversals: capital stopped flying into Greece even before the announcement in October 2009 by the Papandreou government that public finance data had misreported deficit and debt. In Portugal there was a noticeable outflow at the time of the first Greek programme in spring 2010, followed by a

second outflow in early 2011. In Ireland, private capital inflows dropped the first time in the early stage of the financial crisis (2008Q3). The outflow then paused temporarily, starting again when the Greek programme was agreed in the second quarter of 2010. In Spain also there was a first, short-lived outflow in spring 2010, followed by a second, in summer 2011, concurrent with the one experience by Italy.

3 EVIDENCE OF SUDDEN STOPS

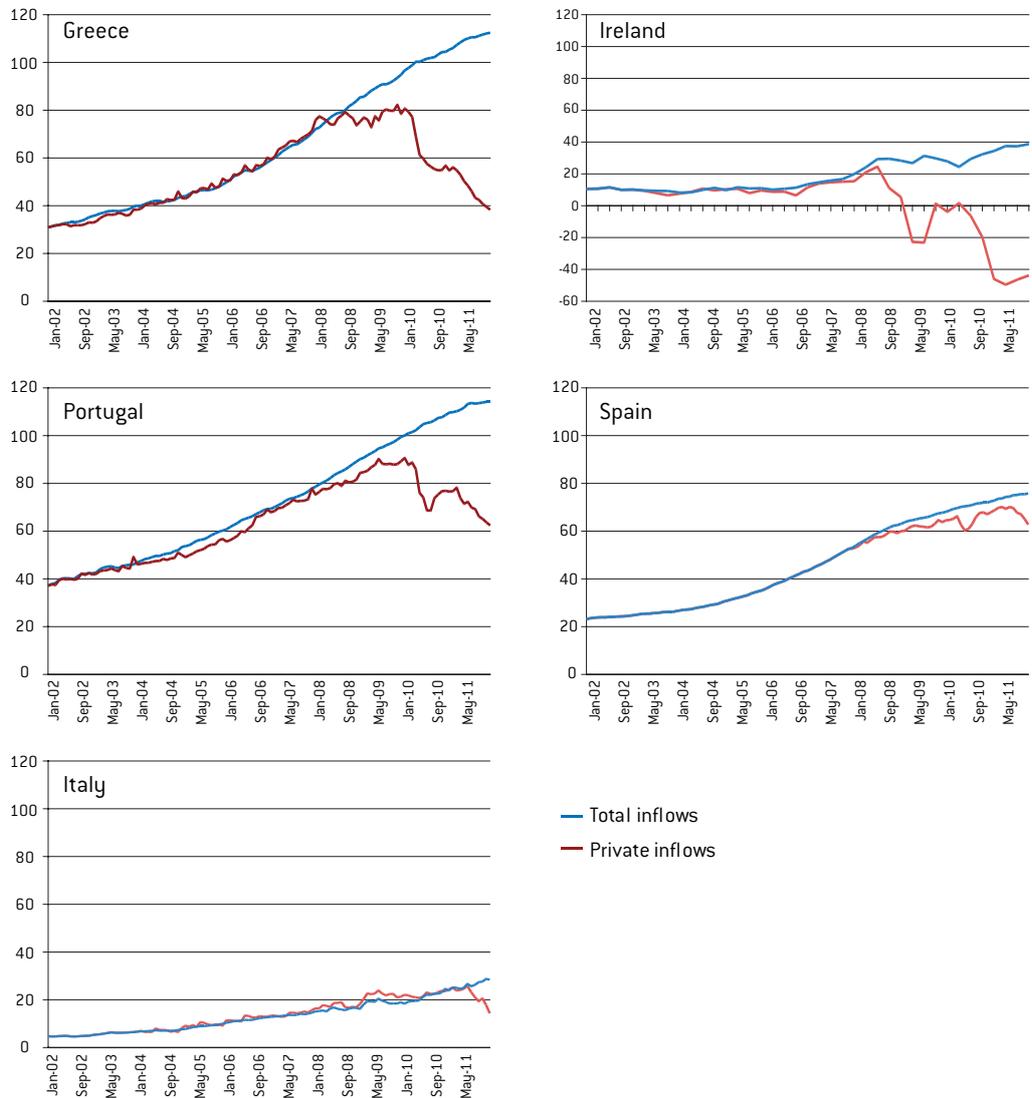
Figure 2 provides *prima facie* evidence of sudden stops of capital inflows. We complement this observation with a more formal test based on the standard methodology introduced by Calvo *et al* (2004). The Calvo methodology is based on monthly data and identifies a sudden stop as an episode in which there is at least one observation with year-on-year capital inflows two standard deviations below the mean. Calvo's methodology has two advantages: it provides a more rigorous and systematic comparison of the experience within the euro area with the experience of emerging countries; and it dates the sudden stop.

After a sudden stop has been identified, it is considered to start with the first observation for which

11. Deutsche Bundesbank Monthly Report, March 2011.

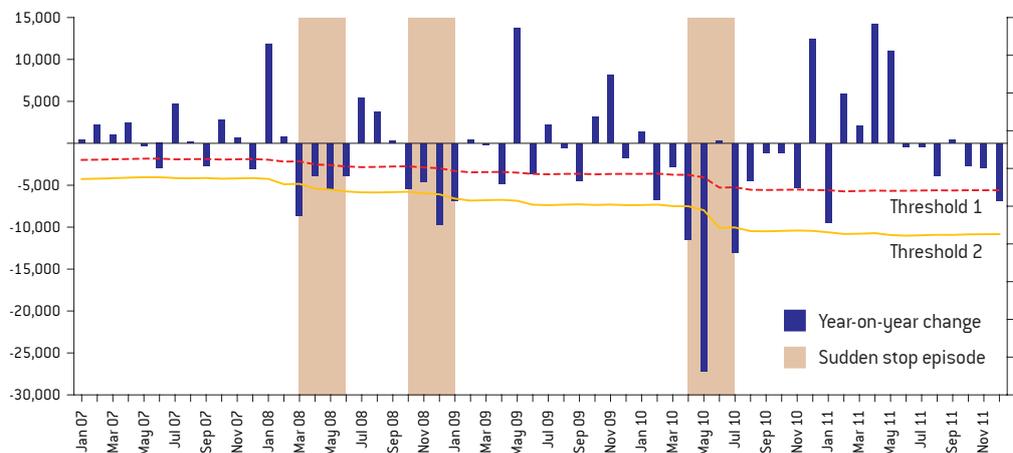
12. We cannot exactly replicate the evolution of the international investment position simply by cumulating financial account flows. This is because the international investment position can be subject to major valuation effects, including the effect of market prices and of exchange rates (European Commission, 2006).

Figure 2: Total and private capital inflows, selected southern euro-area countries, 2002-11 (% 2007 GDP)



Source: Bruegel calculations with national and Eurostat data. Figures show cumulative capital inflows relative to the international investment position debt in 2001.

Figure 3: Identifying sudden stops: Greece (€ millions)



Source: Bruegel.

capital inflows are one standard deviation below the mean, and to end with the first observation for which capital inflows return above one standard deviation below the mean (see Appendix 1 for details). In Figure 3, we present the application of this methodology to the case of Greece. The grey areas correspond to sudden stop episodes.

It is apparent in Figure 3 that the Calvo methodology provides a straightforward way to identify a sudden stop that takes place after a sustained period of capital inflows, but the methodology yields more ambiguous results when it comes to identifying sudden stops that take place during protracted periods of capital outflows. An alternative to the Calvo methodology is to freeze the thresholds after the first episode, instead of de facto toughening the criterion, as apparent in Figure 3. We use both methodologies, and find no significant difference in results except for Ireland, for which the fixed-threshold methodology results in the identification of a series of sudden stop episodes throughout 2010 (see Appendix 1).

The dating of sudden stop episodes helps identify contagion effects, showing how reversals of capital flows spread among crisis countries. Figure 4 shows the number of countries in a sudden stop episode (counting only episodes of at least three months in order to eliminate short-term variations). We find three sudden stop periods:

- **The global financial crisis.** The rise in risk aversion and the clogging of the interbank market

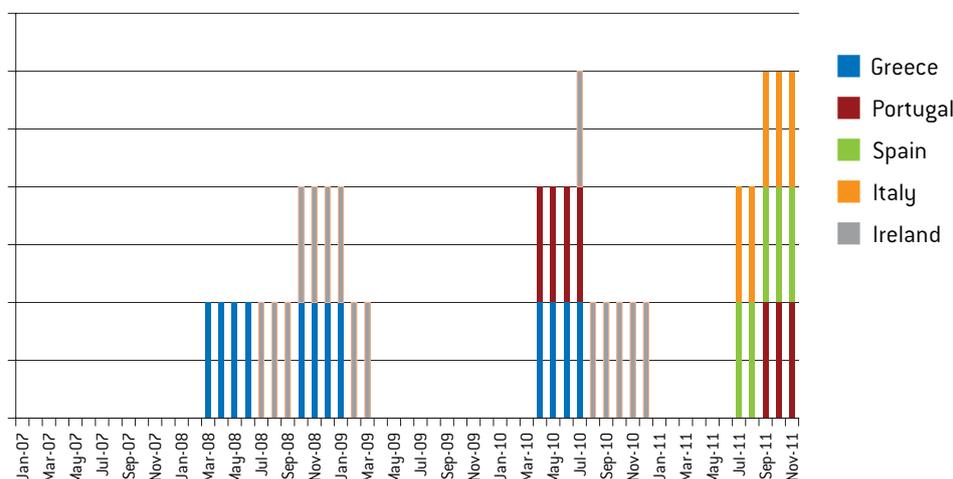
affected both Greece and Ireland. Capital started flowing out of Greece early in 2008 (between March and June), before the Lehman shock and well before the misreporting of fiscal statistics was revealed. This phase was followed by another episode between October 2008 and January 2009, corresponding with the intensification of the financial crisis. At the same time, private capital also started leaving Ireland, which entered a long sudden stop phase (2008Q3 to 2009Q1).

- **Spring 2010.** The agreement of the IMF/EU programme marked the beginning of a third Greek episode (April 2010 to July 2010¹³), which also triggered an impressive contagion effect. Portugal entered a sudden stop immediately but it was relatively short, whereas Ireland experienced a serious and prolonged capital outflow that eventually led the country to ask for support.
- **End 2011.** The third wave of sudden stops involved Italy¹⁴ and Spain – both put under increased scrutiny and pressure by sovereign bond markets during the summer – and Portugal. Contrary to reasonable expectations, we cannot detect (at least using Calvo’s methodology) any episode of sudden stop for Portugal in May 2011, even though the cumulative capital flows continued to fall steadily. In this respect, it is important to recall that we are not taking the Securities Markets Programme (SMP) out of the financial account and this could partly account for the overestimation of capital inflows.

13. June 2010 would not satisfy the requirement of being at least one standard deviation below average. However, given that the year-on-year change in capital inflow was almost zero in June 2010 and it is preceded and followed by two observations falling below the second threshold, we decided to include it in the sudden-stop period.

14. As in the case of Greece, the observation of October 2011 would not satisfy the criterion, but the year-on-year positive change is very small and followed by two observations below the second threshold, so we include it in the sudden stop.

Figure 4: Sudden stop episodes in southern euro-area countries, 2009-11



Source: Bruegel.

An important question is whether capital outflows simply result from sovereign crises, ie from the disposal by non-residents of their portfolios of government securities, or if their impact is broader, also affecting solvent private agents¹⁵. It is only in the second case that it is justified to challenge conventional wisdom and speak of balance-of-payment crises instead of sovereign crises. Lack of detailed comparable data does not make it possible to proceed to a formal test, but discussion can draw on orders of magnitude in the cases of Italy and Spain.

In the Italian case, data holdings of government debt by agents measured at nominal value are available and can be compared to balance-of-payment flows. Outflows during the end-2011 episode were significantly larger than the selling of government bonds by non-residents, which suggests that other agents were also affected by the sudden stop. For Spain, the same can be done but with quarterly data only. Again, the data indicates that the outflows meaningfully exceeded what could be accounted for by the withdrawal of non-residents from the government bond market. These are rough assessments only, and our estimate of capital outflows is admittedly imperfect because we do not take into account the impact of the SMP. But our reading of the evidence is that the data tends to confirm the view that capital outflows exceeded what can be explain by the withdrawal of non-residents from the government bond market.

4 THE ROLE OF OFFICIAL FINANCING AND THE TARGET2 DEBATE

The evidence presented in section 3 shows that the three programme countries – and more recently Italy and Spain – have experienced significant reversals of capital inflows. This was not evident from the official balance-of-payment statistics, because the private capital outflows were compensated for by an equally sizeable increase in public capital inflows. These flows have prevented the official financial account from shrinking.

Public support has taken three forms in the euro area: EU/IMF assistance programmes; provision by the Eurosystem of liquidity to the banking sector (captured by the development of TARGET balances); and ECB purchases of sovereign bonds under the SMP. As previously discussed, we have not been able to build estimates for the third component, so our estimates of private capital inflows tend to err on the optimistic side.

Figure 5 shows the relative size and importance of the two first components in filling the void left by private capital flight. The decomposition is obtained simply by cumulating separately changes in TARGET net *liabilities*, programme flows, and our measure of private capital inflows over the same period (2002-11) for all countries. The sum of these three components has been plotted against the cumulated total inflow (the official financial account data).

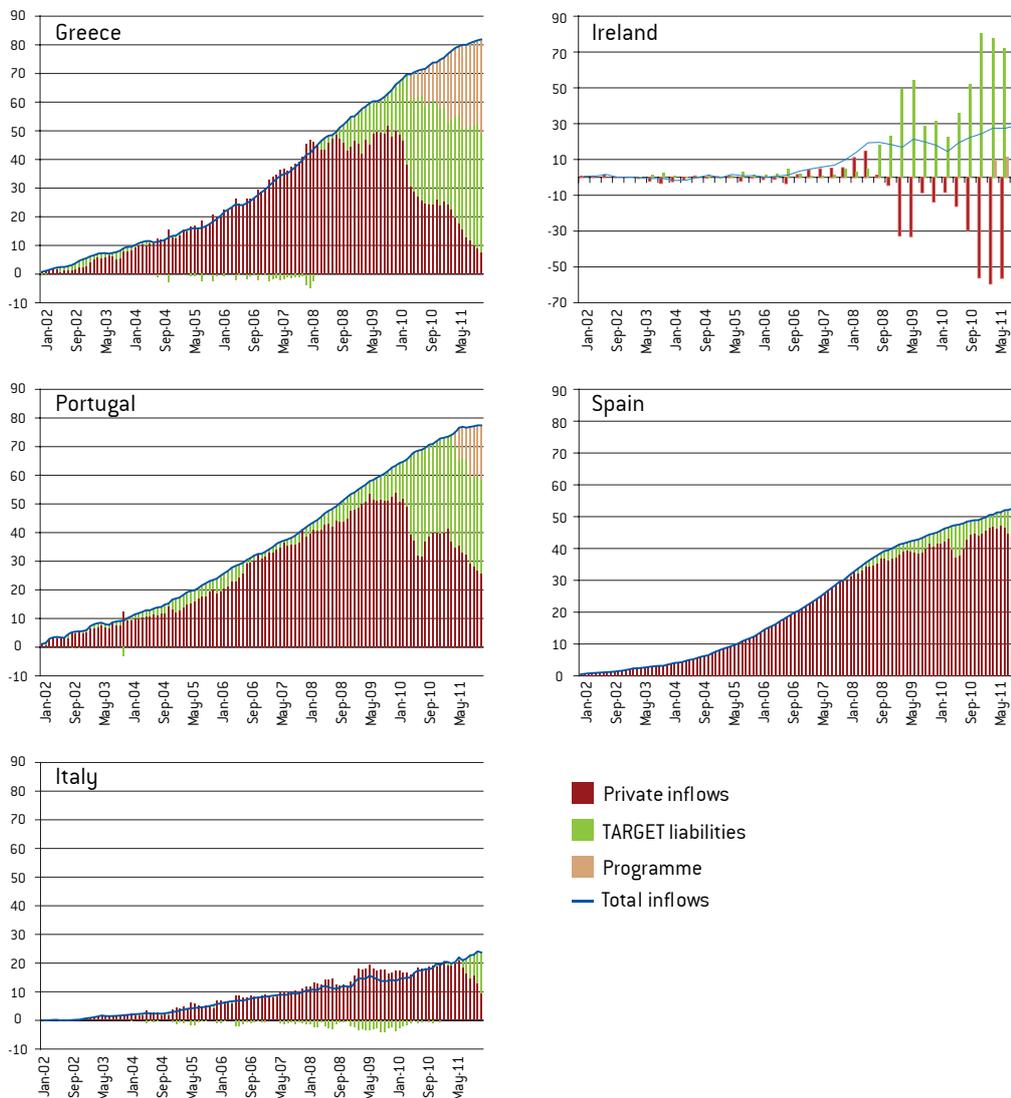
For Greece, at the end of 2011, programme and TARGET liabilities accounted respectively for 44 percent and 56 percent of total official financing. For other countries, however, TARGET financing was by far the largest component. Intra-Eurosystem liabilities amounted to 69 percent of GDP in Ireland (at end-2011Q3) and 32 percent in Portugal (December 2011), against only 14 percent and 19 percent respectively accounted for by the programme in the two countries. ECB financing has also been sizable in Italy and Spain, amounting to 13 percent and 11 percent of GDP respectively, as of November 2011.

These findings help to shed light on the debate on the role of TARGET2 financing. Early contributions focused mostly on the link between TARGET balances and current-account balances, arguing that the former financed the latter to some extent. As we have shown, the pace of current-account adjustment in the euro area was clearly much slower than for non-euro area EU countries. Substitution of private-capital inflows by public inflows, especially Eurosystem financing, helped accommodate persistent current-account deficits

15. See Merler and Pisani-Ferry (2012) for evidence on the withdrawal of non-residents from the government bonds market.

'The three programme countries – and more recently Italy and Spain – have experienced significant reversals of capital inflows. This was not evident from the official balance-of-payment statistics. These flows have prevented the official financial account from shrinking.'

Figure 5: Decomposition of cumulative capital inflows (% of 2007 GDP)



Source: Bruegel based on national central banks, IMF, ECFIN, EFSF.

in a context in which capital markets were no longer willing to accommodate them.

However, large current-account balances *per se* are neither a necessary nor a sufficient condition for incurring significant TARGET liabilities (Bindseil and Winckler, 2012). What was instead crucial was *how* these current-account balances were financed in the euro area before the outbreak of the financial crisis. As stressed by the European Commission already in 2006 (European Commission, 2006), the countries with large current-account deficits (Greece, Portugal and Spain) were mostly financed via portfolio debt securities and bank loans, whereas the contribution of foreign direct investment was very limited. Such a financ-

ing structure, biased towards banks' intermediation, rendered the deficit countries very exposed to the unwinding of capital inflows, especially in a financial crisis. We have shown that a reversal of private inflows indeed took place and that it was sizeable enough to qualify as a sudden stop. The Eurosystem has provided a buffer against the associated drying up of liquidity on the interbank market, and this is reflected in the evolution of intra-Eurosystem claims.

Reliance on Eurosystem financing primarily reflects the distress of euro-area banking systems in the aftermath of the global crisis. The difficulty that banks had to refinance on the interbank market led the Eurosystem to perform this stan-

dard role as a lender of last resort to the banking system through the provision of liquidity in large amounts. From October 2008 onwards, the fixed-rate, full allotment procedure adopted by the ECB made a large part of the euro-area banking system reliant on central bank financing, while weak banks in distressed countries ended up taking up a disproportionately large part of the central bank refinancing (Figure 6). These figures do not include the Emergency Liquidity Assistance (ELA) extended by single national central banks to their banking systems. ELA – the risk being entirely borne at national level – has been extensively used in Ireland and more recently also in Greece, where the government has approved €60 billion in guarantees to facilitate the process (IMF, 2011). The operation is generally recorded in central banks' balance sheets under 'Other assets' (Figure 6), an item that had jumped to €45 billion in Ireland and €58 billion in Greece as of November 2011¹⁶. The rationale for ELA is to ensure that the banking system can access liquidity even when it faces shortages of good collateral to pledge at the ECB. Therefore, any tightening of collateral requirements that makes it more difficult for banks to access ECB refinancing could result in a larger share of the demand for central bank liquidity being covered by national emergency liquidity assistance.

These developments raise an analytical question and a policy question. The analytical question is if the low cost of ECB refinancing and its long maturity (especially but not only since the launch

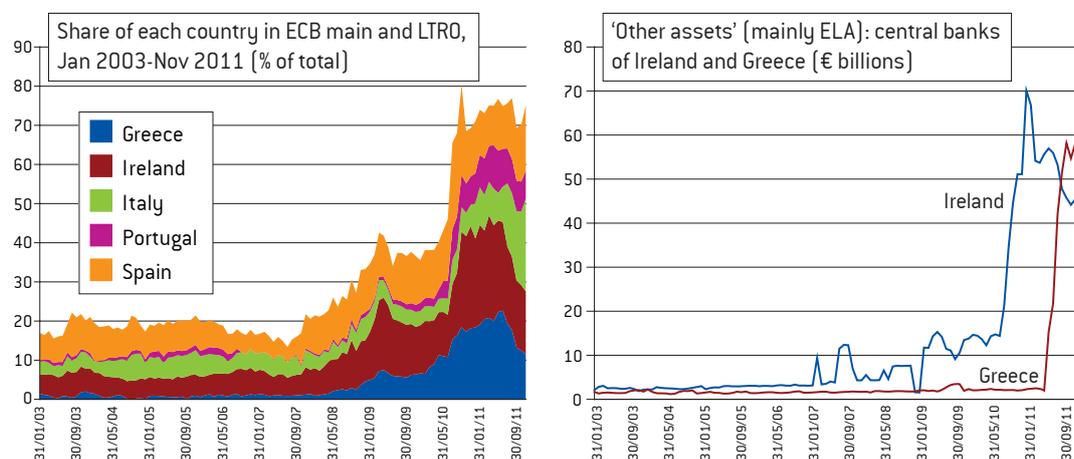
of the three-year Long-Term Refinancing Operations (LTRO) in December 2011) might have contributed to the increase in demand for Eurosystem financing, crowding out private capital flows. The correlation between private capital outflows and increased reliance on Eurosystem financing should be treated with care, because causality could run in both directions. However, for each of our three periods of capital outflows we find hard to reconcile the view that private capital could be crowded out with the sequence of events. The first period started before the adoption by the ECB of its fixed rate, full allotment procedure. In the second period, the coincidence of the drops in private capital flows experienced by Greece, Ireland and Portugal suggests that it was the change in market sentiment rather than the availability of ECB financing that triggered the rise in intra-Eurosystem liabilities. Similarly, capital outflows from Italy and Spain in the second half of 2011 took place before, not after, the extension of the LTRO to three years.

Turning to policy, several proposals have been advanced to shelter national central banks from the perceived risk involved in the accumulation of positive TARGET2 balances. This risk however must be qualified:

- First, as far as TARGET balances reflect the uneven distribution of central bank liquidity within the Eurosystem, they do not entail specific risks for the creditor central banks, over and above the risk from monetary policy oper-

Figure 6: Share of countries affected by sudden stops in take-up of Eurosystem liquidity and ELA

16. There is lack of transparency in both the financing and the amount of ELA, but there is consensus on the fact that the operation is recorded under 'Other assets' (see, for example, Buiter *et al.*, 2011). This is reinforced by the jump observable in this item in crisis periods. For Greece in particular 'Other items' reached €58 billion in November 2011, very close to the €60 billion in guarantees the Greek government approved to back ELA (IMF, 2011).



Source: Bruegel based on national central banks and ECB.

ations. Losses from Eurosystem monetary policy operations could occur in case that there is counterparty failure and the value of collateral posted at the ECB is not sufficient to cover the claim entirely. Such losses would however be shared by national central banks according to the extent of their participation in the Eurosystem's capital. In other words, the possible loss faced by each national central bank would be the same, irrespective of the size of the TARGET claims/liabilities recorded in their own balance sheets. For example the Bundesbank, being the largest shareholder in ECB capital, would bear the greatest loss even if private capital flows from the periphery had been directed massively towards France rather than towards Germany.

- Second, the only scenario in which TARGET would represent an actual additional risk for national central banks would be if one (or more) country decided to leave the euro area. In that case, the net claims against the rest of the system would constitute an additional risk. Any approach that would be interpreted as the introduction of a hedge against the break-up of the euro would involve the risk of sending the message that this break-up is indeed likely.
- Third, any proposal to limit the size of TARGET balances to a fixed threshold underestimates both the importance of a smoothly functioning payment system in a currency union, and the risk of speculative attacks that such limits would imply. The purpose of introducing the single currency was to overcome the weaknesses of fixed-exchange regimes, and this requires all capital flows between members to be treated in the same way. Placing caps on the size of TARGET balances would imply that euros would be entirely fungible across countries only up to a limit (Bindseil and Koenig, 2012), and this would in turn implicitly amount to the creation of two currencies. The threshold would offer a clear target to speculators in the same way that limited reserves offer a target in a fixed exchange-rate regime. Other proposals include the 'collateralising' of the TARGET balances of weaker countries and their disposal for an annual settlement (Sinn and Wollmerhaeuser, 2012). Though more reasonable in principle, such solutions would be very difficult to implement safely at present, given the size

of TARGET balances and the shortage of good collateral. Again, an approach of this sort would give an incentive for speculation against the possibility of the exhaustion of collateral reserves or the inability/unwillingness of countries to mobilise resources for periodic settlements.

TARGET2 balances are the symptom of the uneven distribution of central bank liquidity within the Eurosystem. Those who focus on TARGET2 imbalances as having significance beyond this confuse consequence and causes. Rather than tinkering with the symptom, with the risk of creating doubts about the very viability of the euro, attention should focus on curing the disease, in other words the underlying banking-system problems.

The Eurosystem can tackle the short-term high demand for liquidity by weak banks, against collateral of declining quality, by tightening the quality of the required collateral. This would be likely to reduce TARGET imbalances and is an option the central bank can consider without hampering the functioning of the euro area. Naturally, however, it can only be contemplated if banks are adequately recapitalised and if the threat of a vicious circle of bank and sovereign insolvency is removed. The introduction of a three-year LTRO at the end of 2011, and the extension of the range of eligible collateral, resulted from the Eurosystem's assessment that the risk of a funding crisis in major countries was significant enough for a massive provision of liquidity to be necessary, even though it implied almost by definition a widening of the TARGET imbalances. Only if the situation normalises further will the Eurosystem be able to mop up liquidity, reinstate its collateral policy and thereby contribute to the gradual unwinding of these imbalances.

This, in turn, requires underlying factors that contribute to bank weakness to be addressed: bad loans on bank balance sheets must be provisioned, and recapitalisation must take place wherever needed; public finances must be made convincingly sustainable; and on the macro front, persistent current account deficits can also be tackled through the 'Excessive Imbalances Procedure' recently adopted as part of the so-called Six-Pack legislation (European Commission, 2011).

Private capital flows will only return after the disease has been addressed.

5 CONCLUSIONS

European monetary union involved from the outset many 'known unknowns' and a few 'unknown unknowns'. The possibility that countries within the monetary union would experience balance-of-payment crises belonged to the latter category: conventional wisdom in research and policy was that among euro-area countries, balance-of-payments would become as irrelevant as among regions within a country. Yet developments since 2009 have challenged the wisdom of this view.

In this Policy Contribution, we have examined in detail the financial account of five southern European countries, and have provided evidence of a dramatic reversal in their private components. Considering only the private capital flows, we find that all countries have undergone episodes of sudden stops, more usually seen in emerging markets. These episodes were clustered in three phases (the outbreak of the global financial crisis; spring 2010 at the time of the launch of the Greek programme; and the second half of 2011), which suggests that there has been contagion across countries.

Countries within the euro area can experience such crises because they do not exhibit the same degree of market and policy integration as regions within a country. Regions rarely rely on their own banking systems, implying that the bursting of a regional credit bubble will not translate into a banking crisis. Should a banking crisis nevertheless develop, it does not affect the regional state because responsibility for bank rescue and restructuring is generally a federal competence. Regions therefore can hardly be subject to confidence crises of the sort that affected euro-area countries.

A striking feature of the euro-area crisis is that whereas capital outflows have been dramatic, the current accounts of deficit countries have adjusted only partially. Decomposition of capital inflows highlights the crucial role of Eurosystem financing in mitigating the effect of private capital

outflows (with a contribution of international financial assistance of a comparable order of magnitude in the case of Greece). The injection of liquidity has helped accommodate persistent current-account adjustments in the southern part of the euro area, but most importantly it has protected countries that could no longer rely on adjusting their exchange rates from the full negative impact of a sudden stop. Given the level of integration of euro-area financial markets, the effects of unmitigated sudden stops in southern Europe would have endangered the entire system and put at risk the survival of the single currency.

The smooth functioning of a payment system is essential for maintaining the stability of the financial system, preserving confidence in the common currency and allowing the implementation of a single monetary policy. Introducing constraints on the operations of the payment system would suggest an unwillingness to provide unlimited liquidity across the euro area and open a window for speculation. The more important question is how to address the underlying disease. Together with a gradual mopping up of exceptional liquidity provision and the tightening of collateral requirements, the cure is likely to require interventions to foster the sustainability of public finances, the resilience of the financial system and the reduction of the remaining external imbalance. However confidence cannot be regained overnight and in the meantime, the Eurosystem should not be blamed for playing fully its role.

For the longer run, the evidence that the euro area went through internal balance-of-payment crises should be taken as a clear signal of weakness and as an invitation to reform its structures. Contrary to common belief, a monetary union of this sort is closer to a fixed exchange-rate system among independent countries than to a fully integrated economy. Financial-market participants have realised this and certainly will not forget it. In response, the fostering of a pan-European banking industry and the creation of a banking union with centralised supervision and access to resources to recapitalise weak financial institutions should feature high on the policy agenda. Only a closer integration of markets and policies will preserve the euro area from the risk of further attacks.

REFERENCES

- Bayoumi, T. (1999) 'Is there a world capital market?' in Siebert, H. (ed.) *Globalisation and labour*, Tübingen, J.C.B. Mohr
- Bindseil, Ulrich and P.J. Koenig (2011) 'The economics of TARGET2 balances', *SFB 649 Discussion Paper* 2011-035
- Bindseil, Ulrich and P.J. Koenig (2012) 'TARGET2 and the European sovereign debt crisis', mimeo
- Bindseil, Ulrich, and A. Winckler (2012) 'Dual liquidity crises under alternative monetary frameworks – a financial accounts perspective', mimeo
- Buiter, W., J. Michels and E. Rahbari (2011) 'ELA: an emperor without clothes?', CITI Global Economics
- BIS (2003) *Payment systems in the euro area*
- Bornhorst, Fabian, and Ashoka Mody (2012) 'TARGET imbalances: Financing the capital-account reversal in Europe', *Vox-EU*, 7 March
- Buiter, Willem H., Ebrahim Rahbari and Jürgen Michels (2011) 'The implications of intra-eurozone imbalances in credit flows', *Policy Insight* No. 57, CEPR
- Calvo, Guillermo A., A. Izquierdo and L. F. Mejia (2004) 'On the empirics of sudden stops: the relevance of balance-sheet effects', *Working Paper* 10520, NBER
- Carney, Mark (2012), *Remarks at the World Economic Forum*, 28 January
- Deutsche Bundesbank (2011) *Monthly Report – The German balance of payments*, March
- European Central Bank (2011) *Monthly Bulletin*, October
- European Commission (1990) 'One Market, One Money: An evaluation of the potential benefits and costs of forming a monetary union', *European Economy* no 44
- European Commission (2006) *Quarterly report on the euro area*
- European Commission (2011) Regulations 1173-77/2011 and Directive 85/2011
- Eichengreen Barry, P. Gupta and A. Mody (2006) 'Sudden stops and IMF-supported programmes', *Working Paper* 06/101, International Monetary Fund
- Garber, P. M. (1998) 'Notes on the role of TARGET in a Stage III crisis', *Working Paper* 6619, NBER
- Garber, P. M. (2010) 'The mechanics of intra-euro capital flight', *Research Special Report*, December, Deutsche Bank
- Giavazzi, F. and L. Spaventa (2010) 'Why the current account may matter in a monetary union: lessons from the financial crisis in the euro area', *Discussion Paper* 8008, CEPR
- IMF (2011) *Greece: fifth review under the stand-by arrangement*, December
- Ingram, J. C. (1973) 'The case for European monetary integration', *Essays in International Finance* 98, Princeton University
- Kokkola, T. (2010) *The payment system*, European Central Bank
- Marzinotto, Benedicta, Jean Pisani-Ferry and André Sapir (2010) 'Two crises, two responses', *Policy Brief* 2010/1, Bruegel
- Merler, Silvia and Jean Pisani-Ferry (2012) 'Who's afraid of sovereign bonds?', *Policy Contribution* 2012/02, Bruegel
- Sinn, Hans-Werner and Timo Wollmershaeuser (2011) 'Target loans, current-account balances and capital flows: the ECB's rescue facility', *Working Paper* No. 17626, NBER
- Sinn, Hans-Werner (ed) (2012) 'The European balance-of-payment crisis', *CEifo Forum Volume* 13
- Whittaker, J. (2011) *Intra-Eurosystem debts*, Lancaster University Management School Monetary Research

APPENDIX 1: METHODOLOGY

Data

Following Eichengreen *et al* (2006) we focus on the *financial-account balance*, a comprehensive variable that includes Net Foreign Direct Investments, Net Portfolio Investment and Net Other Investment¹⁷. To maximise the chances of detecting a sudden stop episode, we work with monthly data from national central banks or statistical offices. Only for Ireland we have to use quarterly data and adjust the computations accordingly¹⁸.

From the financial account we derive a measure of *private capital flows*, constructed as the official financial account net of the changes in TARGET2 balances and of the inflow associated to disbursements under the IMF/EU programmes. Both these components are classified in balance-of-payments statistics under 'Other investment' (respectively of monetary authorities and of general government) where they can be clearly identified, provided that the balance of payments is sufficiently disaggregated (see Table A1 for an example). Data on TARGET2 balances is not available for all countries over the same time span¹⁹, but we include them from the earliest date we have.

Table A1: Greece – detecting TARGET2 in the balance of payments (figures in € millions)

	May 11	June 11	July 11	Aug 11	Sept 11	Oct 11	Nov 11	Dec 11
Financial Account								
– Other Investment	8312	5453	-5600	6248	3304	4934	3627	-4565
Liabilities – Monetary Authorities								
Change in TARGET2 liabilities	8313	5452	-5600	6248	3304	4934	3627	-4565

Source: Central Bank of Greece.

17. Calvo *et al* (2004) deals with a panel of many countries (including also emerging ones), which makes it difficult to have consistent financial account data at monthly frequency. Therefore he uses a proxy constructed as the Trade balance net of change in foreign reserves. We do not have such problem because balance of payment data for Euro Area countries are generally published by Central Banks at monthly frequency.

18. In particular: when dealing with monthly data all the computations are done on a minimum of 24 months of observations, whereas the equivalent with quarterly data is a minimum period of 8 quarters.

19. For Portugal and Greece we have data since 2002; for Ireland since 2003; for Italy we would have data since 2004 but due to some inconsistencies between yearly and monthly data before 2004, we consider TARGET balances only starting from this date; for Spain we only have data since 2007.

Unfortunately there is no fully accurate way to account for the impact of the ECB's Securities Market Programme (SMP). First, the ECB only publishes the aggregate outstanding portfolio without any country decomposition, neither of the stock nor of the purchases. Estimates of the composition exist, but they would anyway tell us nothing about the nationality of the agents the ECB bought the bonds from. For example, if the ECB bought Greek bonds from non-resident holders (as seems reasonable given the decline in non-residents' bond holdings observable over the same period), this would not immediately affect the Greek total financial-account balance, as bonds would only pass from one non-resident entity to another. But at the same time, the capital inflows represented by the foreign ownership of those bonds would change from private to public. Given the impossibility of making any assumption that allows the SMP to be taken into account, our measure of private capital inflows is likely to overestimate to some extent the actual private capital inflows.

Identification of sudden stops

Using this measure of private capital flows, we assess if a country has experienced a sudden stop. Following the methodology proposed by Calvo *et al* (2004) we identified a sudden stop as an episode with the following characteristics:

- At least one month in which capital flows fall (year-on-year) two standard deviation below the sample mean
- The start of a sudden stop coincides with the first months in which year-on-year change in capital flows drops one standard deviation below the mean (obviously a fall by two standard deviations below the mean would also qualify as the trigger of a sudden stop, provided that it is not an extemporaneous one).

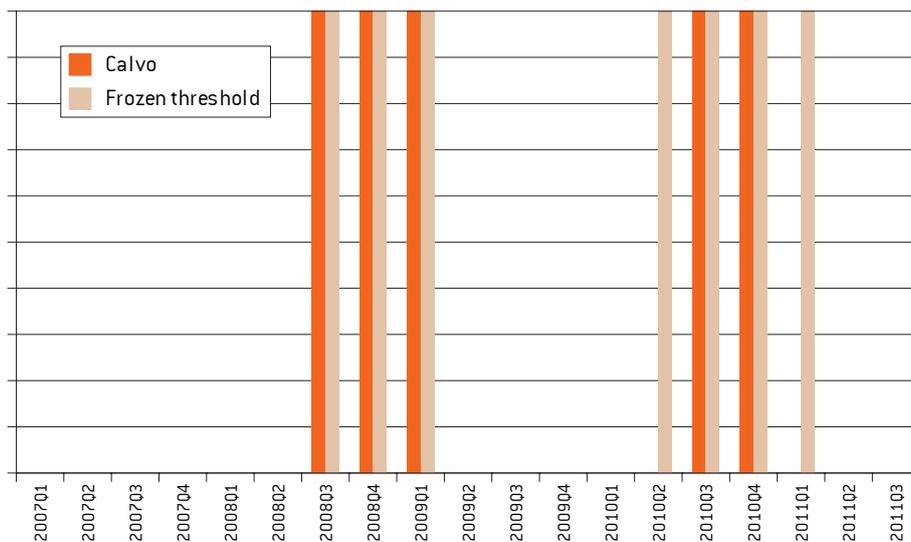
- The end of a sudden stop coincides with change in capital flows reverting to the mean, namely above average minus one standard deviation.

Again following Calvo *et al* (2004), both average and standard deviations are computed in each month over an expanding window with starting date fixed at the earliest data available and a minimum width of 24 months. Moments and threshold are computed in each month t considering only data up to $(t-1)$, so excluding the potential crisis year. In this way we obtain 'adaptive' thresholds that keep track of the past evolution of capital flows but at the same time incorporate the increase in the volatility of capital flows recorded towards the end of the time series and toughen the requirements accordingly. However, thresholds take some time to adapt and therefore we risk detecting too many episodes of sudden stops especially in periods of high volatility (eg during the financial crisis). Therefore we decide to complement the Calvo *et al* criteria with an additional requirement and consider only episodes of sudden stops that last for at least three months. The time series of financial accounts have a different length for all countries, but for the purpose of identifying sudden stops we restricted the sample to the same period for all (2002-11). We did this for the sake of consistency, but we also replicated the analysis considering the whole (different) periods, and results are unaffected.

The Calvo methodology results in toughening the criterion for sudden stops in the case of repeated episodes. For this reason we have explored an alternative methodology to identify the months of sudden stop.

We 'freeze' the thresholds at the value observed the last month before a significant capital drop²⁰ and compared post-sudden stop observations with the pre-sudden stop threshold. This variation does not change anything relevant for Greece, Ireland, Portugal and Spain²¹ whereas it makes a difference for Ireland, stretching the second Irish episode over two more quarters. This is probably due to the fact that quarterly data miss most of the information given by monthly data and they are more sensitive to changes in the threshold.

Figure A1: Alternative dating of sudden stops in the case of Ireland



Source: Bruegel.

20. We identified the huge capital drop by looking at the evolution of monthly financial account flows compared to their long-term average (the same drops are also evident in the cumulative capital inflows graphs).

21. Only the third episode for Greece lasts one month longer, until September 2011.

APPENDIX 2: DISBURSEMENTS UNDER IMF/EU PROGRAMMES

Greece (Source: ECFIN)

Disbursement (€bn)	Euro area	IMF	Total
May 2010	14.5	5.5	20
September 2010	6.5	2.6	9.1
December 2010		2.5	2.5
January 2011	6.5		6.5
March 2011	10.9	4.1	15
July 2011	8.7	3.2	11.9
December 2011	5.8	20.1	8

Portugal (ECFIN, IMF, EFSF)

Disbursement (€bn)	EFSF	EFSM	IMF	Total
May 2011		1.75	6.1	7.85
June 2011	5.8	4.75		10.55
September 2011		7	3.98	10.98
October 2011		0.6		0.6
December 2011			2.9	2.9

Ireland (ECFIN, IMF, EFSF)*

Disbursement (€bn)	EFSF	EFSM	IMF	UK	Total
January 2011		5	5.8		10.8
February 2011	3.3				3.3
March 2011		3.4			3.5
May 2011		3	1.58		4.58
September 2011		2	1.48		3.48
October 2011		0.5		0.5	1
November 2011	3				3
December 2011			3.9		3.9

* data has been aggregated at quarterly level for the sudden stop exercise

The European Sovereign Debt Crisis[†]

Philip R. Lane

The capacity of the euro-member countries to withstand negative macroeconomic and financial shocks was identified as a major challenge for the success of the euro from the beginning (in this journal, for example, see Feldstein 1997; Wyplosz 1997; Lane 2006). By switching off the option for national currency devaluations, a traditional adjustment mechanism between national economies was eliminated. Moreover, the euro area did not match the design of the “dollar union” of the United States in key respects, since the monetary union was not accompanied by a significant degree of banking union or fiscal union. Rather, it was deemed feasible to retain national responsibility for financial regulation and fiscal policy.

On the one side, the ability of national governments to borrow in a common currency poses obvious free-rider problems if there are strong incentives to bail out a country that borrows excessively (Buiter, Corsetti, and Roubini 1993; Beetsma and Uhlig 1999). The original design of the euro sought to address the over-borrowing incentive problem in two ways. First, the Stability and Growth Pact set (somewhat arbitrary) limits on the size of annual budget deficits at 3 percent of GDP and the stock of public debt of 60 percent of GDP. Second, the rules included a “no bailout” clause, with the implication that a sovereign default would occur if a national government failed to meet its debt obligations.

On the other side, the elimination of national currencies meant that national fiscal policies took on additional importance as a tool for countercyclical macroeconomic policy (Wyplosz 1997; Gali and Monacelli 2008; Gali 2010). Moreover, since

■ *Philip R. Lane is Whately Professor of Political Economy at Trinity College Dublin, Dublin, Ireland, and Research Fellow, Centre for Economic Policy Research, London, United Kingdom. His email address is <plane@tcd.ie>.*

[†]To access the Appendix, visit <http://dx.doi.org/10.1257/jep.26.3.49>.

banking regulation remained a national responsibility, individual governments continued to carry the risks of a banking crisis: both the direct fiscal costs (if governments end up recapitalizing banks or providing other forms of fiscal support) and also the indirect fiscal costs since GDP and tax revenues tend to remain low for a sustained period in the aftermath of a banking crisis (Honohan and Klingebiel 2003; Reinhart and Rogoff 2009).

There are three phases in the relationship between the euro and the European sovereign debt crisis. First, the initial institutional design of the euro plausibly increased fiscal risks during the pre-crisis period. Second, once the crisis occurred, these design flaws amplified the fiscal impact of the crisis dynamics through multiple channels. Third, the restrictions imposed by monetary union also shape the duration and tempo of the anticipated post-crisis recovery period, along with Europe's chaotic political response and failure to have institutions in place for crisis management. We take up these three phases in the next three major sections of this article, and then turn to reforms that might improve the resilience of the euro area to future fiscal shocks.

As will be clear from the analysis below, the sovereign debt crisis is deeply intertwined with the banking crisis and macroeconomic imbalances that afflict the euro area. Shambaugh (2012) provides an accessible overview of the euro's broader economic crisis. Even if the crisis was not originally fiscal in nature, it is now a full-blown sovereign debt crisis and our focus here is on understanding the fiscal dimensions of the euro crisis.

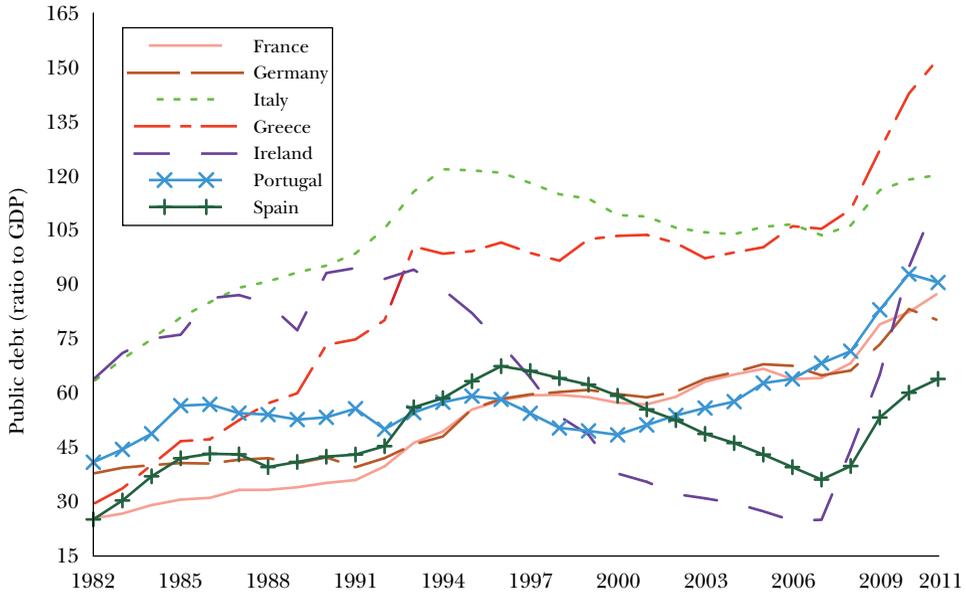
Pre-Crisis Risk Factors

Public debt for the aggregate euro area did not, at least at first glance, appear to be a looming problem in the mid 2000s. During the previous decade, the euro area and the United States shared broadly similar debt dynamics. For example, the ratio of gross public debt to GDP in 1995 was about 60 percent for the United States and 70 percent for the set of countries that would later form the euro area, based on my calculations with data from the IMF Public Debt Database. In both the United States and the euro area, the debt/GDP ratios declined in the late 1990s, but had returned to mid 1990s levels by 2007. The debt/GDP ratios then climbed during the crisis, growing more quickly for the United States than for the euro area.¹

However, the aggregate European data mask considerable variation at the individual country level. Figure 1 shows the evolution of public debt ratios for seven key euro area countries over 1982–2011. These countries were chosen because Germany, France, Italy, and Spain are the four largest member economies, while the fiscal crisis so far has been most severe in Greece, Ireland, and Portugal (of course, Italy

¹ For a detailed country-by-country breakdown of the evolution of public sector debt across these seven countries from 1992–2011, see the Appendix available online with this paper at (<http://ejep.org>).

Figure 1
The Evolution of Public Debt, 1982–2011



Source: Data from IMF Public Debt Database.

and Spain have also been flagged as fiscally vulnerable countries during the crisis). Clearly, these countries have quite different debt histories.

In one group, both Italy and Greece had debt/GDP ratios above 90 percent since the early 1990s; these countries never achieved the 60 percent debt/GDP limit specified in the European fiscal rules. Ireland, Portugal, and Spain each achieved significant declines in debt ratios in the second half of the 1990s, dipping below the 60 percent ceiling. While the Portuguese debt ratio began to climb from 2000 onwards, rapid output growth in Ireland and Spain contributed to sizable reductions in debt–output ratios up to 2007. Finally, France and Germany had stable debt/GDP ratios at around 60 percent in the decade prior to the onset of the crisis; indeed, their debt ratios were far above the corresponding values for Ireland and Spain during 2002–2007. Thus, circa 2007, sovereign debt levels were elevated for Greece and Italy, and the trend for Portugal was also worrisome, but the fiscal positions of Ireland and Spain looked relatively healthy. Moreover, the low spreads on sovereign debt also indicated that markets did not expect substantial default risk and certainly not a fiscal crisis of the scale that could engulf the euro system as a whole.

However, with the benefit of hindsight, 1999–2007 looks like a period in which good growth performance and a benign financial environment masked the accumulation of an array of macroeconomic, financial, and fiscal vulnerabilities (Wyplosz 2006; Caruana and Avdjiev 2012).

Table 1
Private Credit Dynamics

	<i>Loans to private sector from domestic banks and other credit institutions (percent of GDP)</i>		
	<i>1998</i>	<i>2002</i>	<i>2007</i>
Greece	31.8	56.5	84.4
Ireland	81.2	104.4	184.3
Portugal	92.1	136.5	159.8
Spain	80.8	100.1	168.5
Italy	55.7	77.3	96.5
Germany	112.2	116.7	105.1
France	81.0	85.6	99.3

Source: World Bank Financial Database.

Financial Imbalances and External Imbalances

A key predictor of a banking crisis is the scale of the preceding domestic credit boom (Gourinchas and Obstfeld 2012). Table 1 shows the evolution of credit/GDP ratios for the seven euro area countries. The European periphery experienced strong credit booms, in part because joining the euro zone meant that their banks could raise funds from international sources in their own currency—the euro—rather than their previous situation of borrowing in a currency not their own (say, U.S. dollars or German marks or British pounds) and then hoping that exchange rates would not move against them. In related fashion, lower interest rates and easier availability of credit stimulated consumption-related and property-related borrowing (Fagan and Gaspar 2007).

A related phenomenon was the increase in the dispersion and persistence of current account imbalances across the euro area. Table 2 shows that current account imbalances were quite small in the pre-euro 1993–1997 period. But, by the 2003–2007 period, Portugal (–9.2 percent of GDP), Greece (–9.1 percent), and Spain (–7.0 percent) were all running very large external deficits. Conversely, Germany ran very large external surpluses averaging 5.1 percent of GDP, while the overall euro area current account balance was close to zero.

To the extent that current account imbalances accelerated income convergence by reallocating resources from capital-abundant high-income countries to capital-scarce low-income countries, this would be a positive gain from monetary union (Blanchard and Giavazzi 2002). Similarly, current account deficits might have facilitated consumption smoothing by the catch-up countries to the extent that current income levels were perceived to be below future income levels. However, if capital inflows rather fueled investment in capital that had little effect on future productivity growth (such as real estate) and delayed adjustment to structural shocks (such as increasing competition from Central and Eastern Europe and emerging Asia in the production of low-margin goods),

Table 2
Current Account Balances
(percent of GDP)

	1993–1997	1998–2002	2003–2007	2008–2011
Greece	-2.0	-5.9	-9.1	-11.1
Ireland	3.4	-0.2	-2.6	-1.6
Italy	2.1	0.2	-1.8	-2.9
Portugal	-2.4	-9.0	-9.2	-10.5
Spain	-0.6	-3.1	-7.0	-5.8
France	1.1	2.0	-0.2	-1.9
Germany	-0.9	-0.3	5.1	5.7

Source: International Monetary Fund's World Economic Outlook database.

then the accumulation of external imbalances posed significant macroeconomic risks (Blanchard 2007; Giavazzi and Spaventa 2011; Chen, Milesi-Ferretti, and Tressel forthcoming).

For countries running large and sustained external deficits, Blanchard (2007) identifies several risk factors. In terms of medium-term growth performance, a current account deficit can be harmful if increased expenditure on nontradables squeezes the tradables sector by bidding up wages and drawing resources away from industries that have more scope for productivity growth. This is especially risky inside a currency union, because nominal rigidities mean that the downward wage adjustment required once the deficit episode is over can only be gradually attained through a persistent increase in unemployment.

In addition, a large current account deficit poses short-term risks, if there is a sudden stop in funding markets such that the deficit must be narrowed quickly. Large and sudden capital flow reversals have often proven costly in terms of output contractions, rising unemployment, and asset price declines (Freund and Warnock 2007). A reversal in capital flows is also associated with a greater risk of a banking crisis, especially if capital flows have been intermediated through the domestic banking system.

The 2003–2007 Boom

The most intense phase of the dispersion in credit growth and current account imbalances did not occur at the onset of the euro in 1999. Rather, there was a discrete increase during 2003–2007 (Lane and Pels 2012; Lane and McQuade 2012). A complete explanation for the timing of this second, more intense phase of current account deficits and credit booms is still lacking, but the simultaneous timing with the securitization boom in international financial markets, the U.S. subprime episode, and the decline in financial risk indices suggest that the answer may be found in the underlying dynamics of the global financial system and the unusually low long-term interest rates prevailing during this period.

The credit boom in this period was not primarily due to government borrowing. For Ireland and Spain, the government was not a net borrower during 2003–2007. Rather, households were the primary borrowers in Ireland and corporations in Spain, with the property boom fueling debt accumulation in both countries. In Portugal and Greece, the government and corporations were both significant borrowers, but these negative flows were partly offset during this period by significant net accumulation of financial assets by the household sector in these countries.

Failure to Tighten Fiscal Policy

Looking back, the failure of national governments to tighten fiscal policy substantially during the 2003–2007 was a missed opportunity, especially during a period in which the private sector was taking on more risk. In some countries (Ireland and Spain), the credit and housing booms directly generated extra tax revenues, since rising asset prices, high construction activity, and capital inflows boosted the take from capital gains taxes, asset transaction taxes, and expenditure taxes. Faster-growing euro member countries also had inflation rates above the euro area average, which also boosted tax revenues through the non-indexation of many tax categories. Finally, low interest rates meant that debt servicing costs were below historical averages. However, these large-scale revenue windfalls were only partially used to improve fiscal positions, with the balance paid out in terms of extra public spending or tax cuts. Overall, fiscal policy became less countercyclical after the creation of the euro, undoing an improvement in cyclical performance that had been evident in the 1990s (Benetrix and Lane 2012).

A contributory factor in the failure to tighten fiscal policy was the poor performance of the analytical frameworks used to assess the sustainability of fiscal positions. In evaluating the cyclical conduct of fiscal policy from 2002–2007, domestic authorities and international organizations such as the IMF, OECD, and European Commission primarily focused on point estimates of the output gap in order to estimate the “cyclically adjusted” budget balance, without taking into account the distribution of macroeconomic, financial, and fiscal risks associated with the expansion in external imbalances, credit growth, sectoral debt levels, and housing prices. A more prudential and forward-looking approach to risk management would have suggested more aggressive actions to accumulate buffers that might help if or when the boom ended in a sudden and disruptive fashion (Lane 2010).

For the euro periphery, the 2008 global financial crisis triggered a major reassessment among investors of the sustainability of rapid credit growth and large external deficits. In turn, this took the form of significant private sector capital outflows, the tightening of credit conditions, and a shuddering halt in construction activity, with national banking systems grappling with the twin problems of rising estimates of loan losses and a liquidity squeeze in funding markets. In turn, the combined impact of domestic recessions, banking-sector distress, and the decline in risk appetite among international investors would fuel the conditions for a sovereign debt crisis.

The Financial Crisis and the Sovereign Debt Crisis

August 2007 marked the first phase of the global financial crisis, with the initiation of liquidity operations by the European Central Bank. The high exposure of major European banks to losses in the U.S. market in asset-backed securities has been well documented, as has the dependence of these banks on U.S. money markets as a source of dollar finance (McGuire and von Peter 2009; Acharya and Schnabl 2010; Shin 2012). The global crisis entered a more acute phase in September 2008 with the collapse of Lehman Brothers. The severe global financial crisis in late 2008 and early 2009 shook Europe as much as the United States.

From Financial Shock to Sovereign Debt Crisis

Through 2008 and 2009, there was relatively little concern about European sovereign debt. Instead, the focus was on the actions of the European Central Bank to address the global financial shock. In tandem with the other major central banks, it slashed short-term interest rates, provided extensive euro-denominated liquidity, and entered into currency swap arrangements to facilitate access by European banks to dollar-denominated liquidity.

But the global financial shock had asymmetric effects across the euro area. Cross-border financial flows dried up in late 2008, with investors repatriating funds to home markets and reassessing their international exposure levels (Milesi-Ferretti and Tille 2011). This process disproportionately affected countries with the greatest reliance on external funding, especially international short-term debt markets. Inside the euro area, Ireland was the most striking example: the high dependence of Ireland's banking system on international short-term funding prompted its government at the end of September 2008 to provide an extensive two-year liability guarantee to its banks (Honohan 2010; Lane 2011).

More generally, the global financial crisis prompted a reassessment of asset prices and growth prospects, especially for those countries that displayed macro-economic imbalances. For instance, Lane and Milesi-Ferretti (2011) show that the pre-crisis current account deficit and rate of domestic credit expansion are significant correlates of the scale of the decline in output and expenditure between 2007 and 2009, while Lane and Milesi-Ferretti (forthcoming) show that "above-normal" current account deficits during 2005–2008 were associated with sharp current account reversals and expenditure reductions between 2008–2010. The cessation of the credit boom was especially troubling for Ireland and Spain, since the construction sectors in these countries had grown rapidly. The decline in construction was a major shock to domestic economic activity, while abandoned projects and falling property prices indicated large prospective losses for banks that had made too many property-backed loans.

Still, euro area sovereign debt markets remained relatively calm during 2008 and most of 2009. During this period, the main focus was on stability of the area-wide banking system, with country-specific fiscal risks remaining in the background. Furthermore, the relatively low pre-crisis public debt ratios of Ireland and Spain

gave some comfort that these countries could absorb the likely fiscal costs associated with a medium-size banking crisis. Demand for sovereign debt of euro area countries was also propped up by banks that valued government bonds as highly rated collateral in obtaining short-term loans from the European Central Bank (Buiter and Sibert 2006).

In late 2009, the European sovereign debt crisis entered a new phase. Late that year, a number of countries reported larger-than-expected increases in deficit/GDP ratios. For example, fiscal revenues in Ireland and Spain fell much more quickly than GDP, as a result of the high sensitivity of tax revenues to declines in construction activity and asset prices. In addition, the scale of the recession and rising estimates of prospective banking-sector losses on bad loans in a number of countries also had a negative indirect impact on sovereign bond values, since investors recognized that a deteriorating banking sector posed fiscal risks (Mody and Sandri 2012).

However, the most shocking news originated in Greece. After the general election in October 2009, the new government announced a revised 2009 budget deficit forecast of 12.7 percent of GDP—more than double the previous estimate of 6.0 percent.² In addition, the Greek fiscal accounts for previous years were also revised to show significantly larger deficits. This revelation of extreme violation of the euro's fiscal rules on the part of Greece also shaped an influential political narrative of the crisis, which laid the primary blame on the fiscal irresponsibility of the peripheral nations, even though the underlying financial and macroeconomic imbalances were more important factors.

These adverse developments were reflected in rising spreads on sovereign bonds. For example, the annual spread on ten-year sovereign bond yields between Germany and countries such as Greece, Ireland, Portugal, Spain, and Italy was close to zero before the crisis. Remember that sovereign debts from these countries are all denominated in a common currency, the euro, so differences in expected yield mainly represent perceived credit risks and differences in volatility.

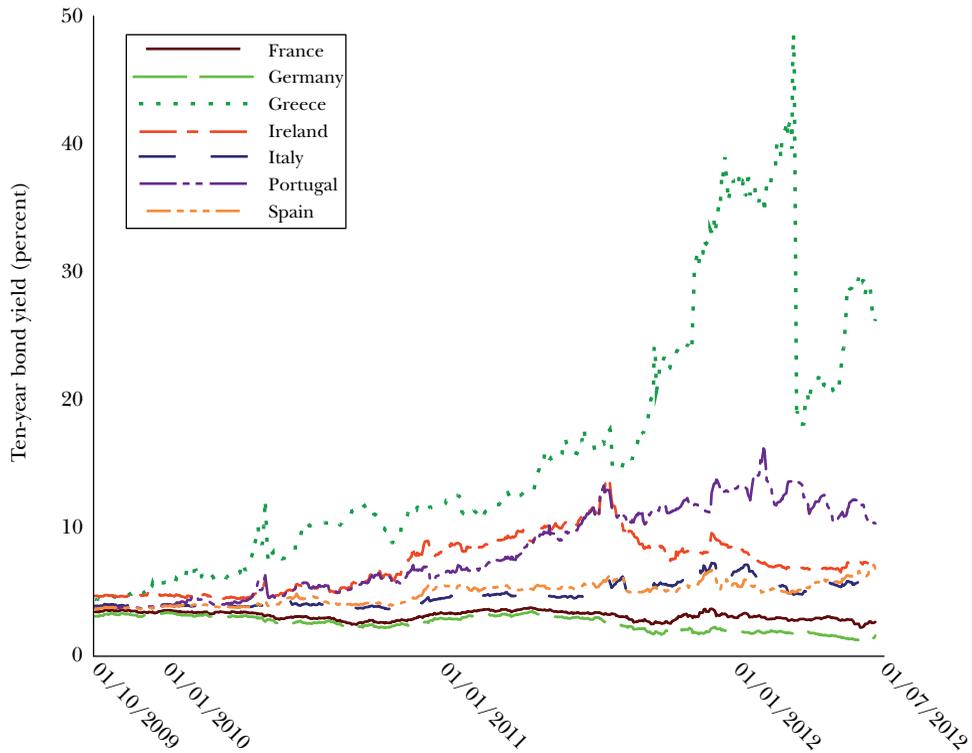
Figure 2 shows the behavior of country-level ten-year bond yields for seven euro area countries from October 2009 through June 2012. Three particularly problematic periods stand out. First, the Greek yield began to diverge from the group in early 2010, with Greece requiring official assistance in May 2010. Second, there was strong comovement between the Irish and Portuguese yields during 2010 and the first half of 2011 (Ireland was next to require a bailout in November 2010, with Portugal following in May 2011). Third, the yields on Italy and Spain have moved together, with these spreads at an intermediate level between the bailed out countries and the core countries of Germany and France. For Italy and Spain, the spread against Germany rose above 300 basis points in July 2011 and remained at elevated levels thereafter. In 2011, a visible spread also emerges between the French and

² See also Gibson, Hall, and Tavlás (2012). These authors also point out that the Greek announcement was coincidentally soon followed by the surprise request from Dubai World for a debt moratorium, such that the climate in international debt markets markedly deteriorated in October/November 2009.

Figure 2

Yields on Ten-Year Sovereign Bonds, October 2009 to June 2012

(percent)



Source: Author's calculations based on data from Datastream.

German yields, although the greater relative vulnerability of France is not pursued in this paper.

Cobbling Together a Response to the Sovereign Debt Crisis

Greece was the first country to be shut out of the bond market in May 2010, with Ireland following in November 2010, and Portugal in April 2011. (In June 2012, Spain and Cyprus also sought official funding. At the time of writing, it is unclear whether Spain will require only a limited form of official funding to help it recapitalize its banking system or a larger-scale bailout.)

In each of the three bailouts, joint European Union/IMF programs were established under which three-year funding would be provided on condition that the recipient countries implemented fiscal austerity packages and structural reforms to boost growth (especially important in Greece and Portugal) and recapitalized and deleveraged overextended banking systems (especially important in Ireland). The scale of required funding far exceeded normal IMF lending levels, so the European

Union was the major provider of funding. At that time, it was also decided to set up a temporary European Financial Stability Facility that could issue bonds on the basis of guarantees from the member states in order to provide official funding in any future crises. In addition, the pre-existing European Stability Mechanism, which had previously only been used for balance-of-payments foreign currency support for non-euro member countries, was adapted to also provide funding for euro member countries.

In principle, a temporary period of official funding can benefit all parties. For the borrower, it can provide an opportunity for a government to take the typically unpopular measures necessary to put public finances on a trajectory that converges on a sustainable medium-term path, while also implementing structural reforms that can boost the level of potential output. For the lender, avoiding default can benefit their creditor institutions (especially banks), while guarding against possible negative international spillovers from a default.

The details of the funding plans for Greece, Ireland, and Portugal largely copied standard IMF practices, but they faced a number of potential problems. Here are six issues, in no particular order.

First, given the scale of macroeconomic, financial, and fiscal imbalances, the plausible time scale for macroeconomic adjustment was longer than the standard three-year term of such deals. In particular, fiscal austerity by individual member countries cannot be counterbalanced by a currency devaluation or an easing in monetary conditions, which is especially costly if a country has to simultaneously close both fiscal and external deficits. By June 2011, it was clear that Greece would need a second package, while it is also likely that Ireland and Portugal will not be able to obtain full market funding after the expiry of their current deals. The slow pace of adjustment was also recognized in Summer 2011 through the extension of the repayment period on the official debt from 7.5 years to 15–30 years.

Second, in related fashion, excessively rapid fiscal consolidation can exacerbate weaknesses in the banking system. Falling output and a rising tax burden shrinks household disposable income and corporate profits, increasing private sector default risk. This was identified as an especially strong risk in the Irish program in view of the scale of household debt.

Third, the fiscal targets were not conditional on the state of the wider European economy. As growth projections for the wider European economy declined throughout 2011, the country-specific targets looked unobtainable for external reasons.

Fourth, the original bailouts included a sizable penalty premium of 300 basis points built into the interest rate, which is standard IMF practice. A penalty rate discourages countries from the moral hazard of taking such loans when not really needed and also compensates the funders for the nontrivial default risk. However, it also makes repaying the loans harder and gives an appearance that the creditor EU countries are profiteering at the expense of the bailed-out countries. This penalty premium on the European component of the official loans was eliminated in July 2011, although the interest rate on the IMF-sourced component of the funds continued to include a penalty premium.

Fifth, the bailout funds have been used to recapitalize banking systems, in addition to covering the “regular” fiscal deficits. So far, this element has been most important in the Irish bailout, but it was also a feature of the Greek and Portuguese bailouts; it is also the primary element in the official funding requested by Spain in June 2012. While publicly funded recapitalization of troubled banks can ameliorate a banking crisis, this strategy is problematic if it raises public debt and sovereign risk to an excessive level (Acharya, Drechsler, and Schnabl 2010). Moreover, excessive levels of sovereign debt can amplify a banking crisis for several reasons: domestic banks typically hold domestic sovereign bonds; a sovereign debt crisis portends additional private sector loan losses for banks; and a highly indebted government is likely to lean on banks to provide additional funding (Reinhart and Sbrancia 2011). Furthermore, the generally poor health of major European banks and the cross-border nature of financial stability inside a monetary union means that national governments are under international pressure to rescue failing banks in order to avoid the cross-border contagion risks from imposing losses on bank creditors.³ Despite these international externalities, at least until mid 2012, the only type of European funding for bank rescues was plain-vanilla official loans to the national sovereign, with fixed repayment terms. Under this approach, the fates of national sovereigns and national banking systems remain closely intertwined.

The sixth issue involves a standard IMF principle that funding is only provided if the sovereign debt level is considered to be sustainable. If it is not sustainable, the traditional IMF practice has been to require private sector creditors to agree to a reduction in the present value of the debt owed to them. Under the joint EU–IMF programs, such “private sector involvement” was not initially deemed necessary in the three bailouts of 2010 and 2011.

The argument against requiring private sector involvement is that it can spook an already nervous sovereign debt market. For example, when the prospect of requiring private sector involvement was broached in October 2010 (in the Franco-German “Deauville Declaration”), interest rate spreads immediately increased, especially for Greece, Ireland, and Portugal. Ireland’s efforts to avoid a bailout came to a halt soon thereafter in early November 2010. European banks also had increased difficulties in raising funds, especially the local banks in the troubled periphery, in line with the increase in the perceived riskiness of their home governments.

The March 2012 agreement to provide Greece with a second bailout package did require that private sector creditors accept a haircut, which eventually turned out to be about 50 percent of value, which is equal to 47 percent of Greek GDP.⁴ But

³ The poor design of European bank resolution regimes has also increased the fiscal cost of rescuing banks, since it is difficult to shut down failing banks and impose losses on holders of the senior bonds issued by banks.

⁴ Although the plausibility of this projection has been disputed by many commentators, the second bailout package is officially projected to deliver a Greek debt/GDP ratio of 120 percent by 2020, which is a shade above the debt ratios of some of the other troubled euro member countries. See also Ardagna and Caselli (2012) for an account of the Greek crisis.

as this requirement was discussed during the course of 2011, it contributed to the sharp widening of the spreads on Spanish and Italian debt.

Listing some of the difficulties in this way may make the European response to its sovereign debt issues appear more coherent than it has actually been. Instead, it may be fair to characterize Europe's efforts to address its sovereign debt problem as makeshift and chaotic, at least through the middle of 2012.

Risks of Multiple Equilibria when Sovereign Debt is High

A significant factor during the crisis has been the increased volatility in euro area sovereign debt markets. A country with a high level of sovereign debt is vulnerable to increases in the interest rate it pays on its debt (Calvo 1988; Corsetti and Dedola 2011). This risk can give rise to self-fulfilling speculative attacks: an increase in perceptions of default risk induces investors to demand higher yields, which in turn makes default more likely. In contrast, if default risk is perceived to be low, interest rates remain low, and default does not occur. This multiple equilibria problem may have greater force in the context of a multicountry currency union, since a small adverse shift in the fundamentals of one individual country can trigger a large decline in demand for the sovereign debt of that country as investors "run for the exit" and switch to sovereign debt of other safer euro area countries.

What policies might encourage the "good" equilibrium? One option is to create a firewall through the availability of an official safety net. This would reduce the risk of the "bad" equilibrium arising because investors would not need to fear that a country will be pushed into involuntary default by an inability to rollover its debt. As of mid 2012, the available funding through the European Financial Stability Facility and its successor, the European Stability Mechanism, was only enough to address the bailouts of Greece, Ireland, and Portugal—and thus not nearly sufficient to offer substantial support to Spain and/or Italy. Proposals to create a large firewall fund are politically unpopular in creditor countries for many reasons, including fear of taking losses, and concerns that such a fund would tempt politicians in at-risk countries to postpone or avoid tough fiscal and structural reform decisions.

The European Central Bank's program to purchase sovereign bonds can also be viewed as a way to reduce the risk of the "bad" equilibrium. Between May 2010 and October 2010, about 65 billion euro of bonds were bought by the ECB; a further 125 billion euro were committed during the market turmoil between August 2011 and November 2011 such that the cumulative bond holdings grew to over 200 billion euros (about 2 percent of euro area GDP). The ECB has taken pains to emphasize that these purchases are not monetizing debt because liquidity created is canceled out through offsetting sterilization operations. Instead, the program seeks to provide liquidity and depth when certain sovereign debt markets are troubled. A useful analogy here is to the modern argument for currency market interventions. Such interventions do not try to fix asset values; instead, limited intervention by a central bank can be temporarily stabilizing by breaking momentum dynamics.

There have also been calls for the European Central Bank to take further steps to stabilize the sovereign debt market (for example, De Grauwe 2012). At

one level, it could increase the firepower of the European Stability Mechanism by allowing it to borrow from the ECB. Going further, the ECB could announce a ceiling to the interest rate it would tolerate on the sovereign debt of countries that meet certain fiscal criteria (such as taking credible steps to ensure debt declines to a safe level over the medium term), and guarantee to buy the debt at that price if needed.

Even more controversially, outright debt monetization might be viewed in some quarters as preferable to outright default by large member countries if it becomes clear that solvency concerns are so great that market funding will not be available for an extended period. While debt monetization exceeds the current legal mandate of the European Central Bank, debate over these proposals might heat up if a more acute and severe phase of the crisis were to take hold. At least for now, it is hard to envisage that such a change would be supported by all member countries of the euro area. However, it is also important to appreciate that the reserve capacity to monetize debt is commonly cited as the reason why highly indebted governments such as Japan, the United Kingdom, and the United States are still able to borrow at low interest rates.

Prospects for Post-Crisis Reduction in Sovereign Debt

The legacy of the euro area sovereign debt crisis is that a number of countries will have dangerously elevated public debt ratios, while others will have debt levels that are lower by comparison but still high relative to long-term normal values. Even if current austerity programs are sufficient to stabilize debt ratios, there remains the post-crisis adjustment challenge of gradually reducing government debt to safer levels. This medium-term challenge is viewed with trepidation in European circles. Consider four reasons why the underlying fundamentals for reducing the debt/GDP ratio are not promising.

First, growth in nominal GDP is likely to be low. Debt/GDP ratios are stickier in high-income countries than in emerging economies in part because there is less scope for rapid output growth in the former group of countries. There is nothing to suggest that real growth rates for advanced economies should exceed a long-term annual average of about 2 percent. Indeed, real annual growth of 2 percent may be optimistic given several factors: the erosion of human capital from the prolonged unemployment of the last few years (DeLong and Summers 2012); the likelihood of tax increases and reduced public investment; and the historical pattern that output growth can be compromised for a decade in the aftermath of a banking crisis (Reinhart and Rogoff 2010). For the most-indebted countries, nominal GDP is unlikely to grow much faster than real GDP. The European Central Bank has a 2 percent aggregate inflation target (approximately), and the most indebted member countries are likely to have average inflation substantially below that level in view of the correlation between domestic demand and the price level of nontradables (Lane and Milesi-Ferretti 2004).

Second, the political economy environment is likely to be challenging. The highly indebted countries will need to be led by governments that must impose spending cuts and tax increases with no short-term prospect of fiscal relaxation. Adjustment fatigue can set in, making it difficult to sustain long-term fiscal austerity.

Third, the possibilities for financing at least some of the sovereign debt through “financial repression” are limited. This approach uses tight regulations on domestic financial institutions—including banks, pension funds, and others—so that these institutions are pressured to put a greater portion of their assets than they would otherwise choose into sovereign debt (Reinhart and Sbrancia 2011). However, the principle of open capital markets across the European Union means that countries have fairly limited scope for financial repression in comparison to what was historically possible.

Fourth, risk premia will likely remain nontrivial for most indebted member countries. The large losses experienced by private sector investors in Greek sovereign debt underline that the sovereign debt of euro area member countries can no longer be categorized as risk-free investments. Indeed, the historical evidence suggests that further rounds of debt restructuring will form part of the adjustment process (see also the discussion by Reinhart, Reinhart, and Rogoff in this issue).

Accordingly, the medium-term outlook suggests that sovereign debt is likely to pose significant policy challenges for the euro area over the next few years. The next section outlines some possible reforms that could help to alleviate the situation and avoid a similar disaster in the future.

Reforms to Address Sovereign Debt Concerns

The high outstanding sovereign debt levels and the importance of avoiding future fiscal crises in the euro area have induced reforms to the fiscal rules for the euro area, with a new Fiscal Compact Treaty that is scheduled to go into effect at the start of 2013 (if it is ratified by 12 members of the euro area by then). The Fiscal Compact requires that the new fiscal principles be embedded in each country’s national legislation. These fiscal governance reforms are based on two principles: first, high public debt levels pose a threat to fiscal stability; and, second, the fiscal balance should be close to zero “over the cycle.”

The operation of the pre-crisis fiscal rules focused on the overall budget balance, with a maximum annual budget deficit set at 3 percent of GDP, while there was no strong pressure on highly indebted countries (such as Greece and Italy) to reduce debt levels below the specified 60 percent ceiling. Even on its own terms, this approach had two main defects: it did not adequately allow for cyclical variation in budget positions, and it did not provide much discipline for countries inside the limit.

In contrast, the new system focuses on the structural budget balance, thus stripping out cyclical effects and one-off items. A structural budget balance

target encourages a government to bank cyclical revenue gains during upturns in exchange for a greater slippage in the overall budget balance during recessions. Under the new system, there is a specified time frame for reducing public debt below the ceiling of 60 percent of GDP, with the excess above the ceiling eliminated at an average rate of “one twentieth” each year.

This new approach faces several implementation problems. For example, a fiscal framework based on structural budget balance faces knotty measurement problems because it requires macroeconomic forecasters to differentiate between cyclical fluctuations and trend fluctuations in output almost in real time. For this reason, the Fiscal Compact requires that governments enact a mechanism that requires adjustments if the forecast errors for the structural budget balance cumulate over several years to a significant level. In the German fiscal law, for example, a cumulative overshoot above 1.5 percent of GDP requires a gradual correction by running tighter structural budgets until the excess is eliminated (Bundesbank 2011).

Another potential issue is that, in contrast to the original Stability and Growth Pact, the primary source of fiscal discipline is intended to be national. The Fiscal Compact requires that the fiscal rules are written into domestic legislation and that national independent fiscal councils be created to monitor the compliance with the specified fiscal rules. The hope is that national-level discipline will be more effective, since it should have greater political legitimacy than external surveillance. However, external surveillance and the threat of external sanctions remain as a “second line of defense” against fiscal misbehavior.

In recognition that fiscal stability can be quickly undone by financial and macroeconomic shocks, the Fiscal Compact is accompanied by new European regulations that go beyond narrow fiscal governance in monitoring “excessive imbalances.” A wide range of risk indicators will be tracked, including credit growth, house price indices, and external imbalances. The intention is that a country experiencing severe imbalances should respond with policy interventions to mitigate crisis risks and improve resilience. However, it remains unclear whether national governments have the capacity to identify excessive imbalances accurately or to deploy policy instruments that can be effective in managing such risk factors.

Given the limited nature of these initiatives, more extensive reforms are also under discussion. A partial list of such proposals includes the following. First and foremost is the creation of a banking union, since the diabolic loop between national banking systems and national sovereigns has been central to the fiscal crisis. The ingredients of banking union are well-known and include European-level regulatory responsibility, deposit insurance, bank resolution policies, and a joint fiscal backstop in the event that fiscal resources were deemed necessary to stabilize the banking system (Allen, Beck, Carletti, Lane, Schoenmaker, and Wagner 2011; Brunnermeier et al. 2011; Marzinotto, Sapir, and Wolff 2011). A partial move in this direction was announced at the June 2012 European Council meeting, which also opened up the possibility of the European Stability Mechanism making direct equity injections into troubled banks. However, the details of these new plans have yet to be ironed out.

A second step is the introduction of common areawide “eurobonds,” with the goal of avoiding the disruptive impact of destabilizing speculative attacks on national sovereign debt markets inside the euro area (Favero and Missale 2012). Fiscally stronger member states might support eurobonds if it is cheaper than the alternatives for reducing default risk, for instance with bigger bailout funds. To prevent fiscally weaker member states from using eurobonds to overborrow, these could be restricted in various ways. One option is to limit eurobonds to short maturities, so that ill-disciplined countries could quickly be denied access to funding (Philippon and Hellwig 2011); another option is to limit eurobond funding only for sovereign debt up to 60 percent of GDP, with the excess still requiring funding through the issuance of national bonds (Delpla and Von Weizsäcker 2011); or eurobonds could be limited to countries that satisfy certain criteria for good macroeconomic and fiscal fundamentals (Muellbauer 2011).⁵

Alternatively, Brunnermeier et al. (2011) point out that many of the advantages of eurobonds can be obtained even if sovereign debt remains a national responsibility. In particular, a European Debt Agency could be established that would buy up large quantities of national sovereign bonds (up to a limit of 60 percent of GDP in each case). This agency would be funded by the issuance of two tranches of bonds—European Safe Bonds and European Junior Bonds—with the latter having the primary exposure in the event of defaults on the underlying portfolio of national sovereign bonds. Accordingly, the senior European Safe Bonds should be safe assets, which in turn should make them preferred collateral for central bank liquidity operations. Since this proposal does not require joint backing of sovereign debt issues, it avoids the moral hazard problems that plague the eurobond proposals.

Third, Europe might seek *a deeper level of fiscal union*, agreeing to share certain tax streams or spending programs in a way that would be delinked from fluctuations in national-level output. In related fashion, *enhanced coordination of national fiscal policies* would also be helpful, thereby enabling the collective fiscal position of the euro area to be appropriately calibrated in relation to the prevailing macroeconomic conditions.

Many of these policy proposals would require changes in the treaties governing the European Union and imply a transformative increase in the level of political integration. Paradoxically, the European crisis has generated severe political tensions across the member states, while at the same time prompting much discussion of the desirability of more extensive types of political union. In this debate, the parallels with the historical development of fiscal federalism in the United States have been well-flagged (Henning and Kessler 2012; Sargent 2012).

⁵ A temporary type of eurobond has been suggested by the German Council of Economic Experts (Bofinger, Feld, Franz, Schmidt, and Weder di Mauro 2011). Under this proposal, a jointly-backed Debt Redemption Fund would refinance the excess debt above 60 percent of GDP, thereby relieving the roll-over pressures facing highly-indebted countries. Once debt levels fall back to the 60 percent ceiling, the Debt Redemption Fund would be wound up.

In conclusion, the origin and propagation of the European sovereign debt crisis can be attributed to the flawed original design of the euro. In particular, there was an incomplete understanding of the fragility of a monetary union under crisis conditions, especially in the absence of banking union and other European-level buffer mechanisms. Moreover, the inherent messiness involved in proposing and implementing incremental multicountry crisis management responses on the fly has been an important destabilizing factor throughout the crisis.

The most benign perspective on the European sovereign debt crisis is that it provides an opportunity to implement reforms that are necessary for a stable monetary union but that would not have been politically feasible in its absence. A more modest hope is that the unfolding reform process will deliver a monetary union that can survive, even if it remains vulnerable to recurring crises. However, the alternative scenario in which the single European currency implodes is no longer unthinkable, even if it would unleash the “mother of all financial crises” (Eichengreen 2010). The stakes are high.

■ *I thank Michael Curran, Michael O’Grady, and Clemens Struck for diligent research assistance. I am grateful to my fellow members of the euro-nomics group for many insightful discussions of the euro crisis and Claudio Borio, Giancarlo Corsetti, Jordi Gali, Paolo Mauro, Ashoka Mody, Maury Obstfeld, and the JEP editors for helpful suggestions.*

References

- Acharya, Viral V., and Philipp Schnabl.** 2010. “Do Global Banks Spread Global Imbalances? The Case of Asset-Backed Commercial Paper during the Financial Crisis of 2007–09.” *IMF Economic Review* 58(1): 37–73.
- Acharya, Viral V., Itamar Drechsler, and Philipp Schnabl.** 2010. “A Pyrrhic Victory? Bank Bailouts and Sovereign Credit Risk.” *CEPR Discussion Paper* 8679.
- Allen, Franklin, Thorsten Beck, Elena Carletti, Philip R. Lane, Dirk Schoenmaker, and Wolf Wagner.** 2011. *Cross-Border Banking in Europe: Implications for Financial Stability and Macroeconomic Policies*. CEPR Report. Center for Economic Policy Research.
- Ardagna, Silvia, and Francesco Caselli.** 2012. “The Political Economy of the Greek Debt Crisis: A Tale of Two Bailouts.” Centre for Economic Performance Special Paper No. 25, London School of Economics.
- Beetsma, Roel, and Harald Uhlig.** 1999. “An Analysis of the Stability and Growth Pact.” *Economic Journal* 109(458): 546–71.
- Benetrix, Agustín, and Philip R. Lane.** 2012. “Fiscal Cyclicalities and EMU.” Unpublished paper, Trinity College Dublin.
- Blanchard, Olivier.** 2007. “Current Account Deficits in Rich Countries.” *IMF Staff Papers* 54(2): 191–219.
- Blanchard, Olivier, and Francesco Giavazzi.** 2002. “Current Account Deficits in the Euro Area: The End of the Feldstein–Horioka Puzzle?” *Brookings Papers on Economic Activity*, no. 2, pp. 147–86.
- Bofinger, Peter, Lars P. Feld, Wolfgang Franz, Christoph M. Schmidt, and Beatrice Weder di**

- Mauro.** 2011. "A European Redemption Pact." VoxEU, November 9.
- Brunnermeier, Markus, Luis Garicano, Philip R. Lane, Marco Pagano, Ricardo Reis, Tano Santos, David Thesmar, Stijn Van Neuwelburgh, and Dimitri Vayanos.** 2011. "European Safe Bonds (ESBies)." <http://euro-nomics.com/wp-content/uploads/2011/09/ESBiesWEBsept262011.pdf>.
- Buiter, Willem, Giancarlo Corsetti, and Nouriel Roubini.** 1993. "Excessive Deficits: Sense and Nonsense in the Treaty of Maastricht." *Economic Policy* 8(16): 57–100.
- Buiter, Willem, and Anne Sibert.** 2006. "How the Eurosystem's Treatment of Collateral in its Open Market Operations Weakens Fiscal Discipline in the Eurozone (and What to Do about It)." In *Fiscal Policy and the Road to the Euro*, 29–60. National Bank of Poland.
- Bundesbank.** 2011. "The Debt Brake in Germany: Key Aspects and Implementation." *Monthly Bulletin*, October, pp. 15–39.
- Caruana, Jaime, and Stefan Adjiev.** 2012. "Sovereign Creditworthiness and Financial Stability: An International Perspective." *Banque de France Financial Stability Review*, no. 16, April, pp. 71–85.
- Calvo, Guillermo.** 1988. "Servicing the Public Debt: The Role of Expectations." *American Economic Review* 78(4): 647–61.
- Chen, Ruo, Gian Maria Milesi-Ferretti, and Thierry Tresselt.** Forthcoming. "Euro Area Debtor Countries: External Imbalances in the Euro Area." *Economic Policy*.
- Corsetti, Giancarlo, and Luca Dedola.** 2011. "Fiscal Crises, Confidence and Default: A Bare-Bones Model with Lessons for the Euro Area." <https://sites.google.com/site/giancarlocorsetti/main/calvo-code.pdf?attredirects=0>.
- De Grauwe, Paul.** 2012. "Fragile Eurozone in Search of a Better Governance." *Economic and Social Review* 43(1): 1–30.
- DeLong, J. Bradford, and Lawrence H. Summers.** 2012. "Fiscal Policy in a Depressed Economy." *Brooking Papers on Economic Activity*, forthcoming.
- Delpla, Jacques, and Jakob Von Weizsäcker.** "Eurobonds: The Blue Bond Concept and Its Implications." Bruegel Policy Contribution 2011/02.
- Eichengreen, Barry.** 2010. "The Breakup of the Euro Area." In *Europe and the Euro*, edited by Alberto Alesina and Francesco Giavazzi, 11–56. University of Chicago Press.
- Fagan, Gabriel, and Vitor Gaspar.** 2007. "Adjusting to the Euro." *ECB Working Paper* 716.
- Favero, Carlo, and Alessandro Missale.** 2012. "Sovereign Spreads in the Euro Area: Which Prospects for a Eurobond?" *Economic Policy* 27(70): 231–73.
- Feldstein, Martin.** 1997. "The Political Economy of an European Economic and Monetary Union: Political Sources of an Economic Liability." *Journal of Economic Perspectives* 11(4): 23–42.
- Freund, Caroline, and Frank Warnock.** 2007. "Current Account Deficits in Industrial Countries: The Bigger They Are, The Harder They Fall?" In *G7 Current Account Imbalances: Sustainability and Adjustment*, edited by Richard Clarida, 133–68. University of Chicago Press.
- Gali, Jordi.** 2010. "Notes on the Euro Debt Crisis." <http://www.crei.cat/people/gali/debt%20crisis%2003%2006.pdf>.
- Gali, Jordi, and Tommaso Monacelli.** 2008. "Optimal Monetary and Fiscal Policy in a Currency Union." *Journal of International Economics* 76(1): 116–32.
- Giavazzi, Francesco, and Luigi Spaventa.** 2011. "Why the Current Account Matters in a Monetary Union." In *The Euro Area and The Financial Crisis*, edited by Miroslav Beblavy, David Cobham, and L'udovit Odor, 59–80. Cambridge University Press.
- Gibson, Heather D., Stephen G. Hall, and George S. Tavas.** 2012. "The Greek Financial Crisis: Growing Imbalances and Sovereign Spreads." *Journal of International Money and Finance* 31(3): 498–516.
- Gourinchas, Pierre-Olivier, and Maurice Obstfeld.** 2012. "Stories of the Twentieth Century for the Twenty-First." *American Economic Journal: Macroeconomics* 4(1): 226–65.
- Henning, C. Randall, and Martin Kessler.** 2012. "Fiscal Federalism: US History for Architects of Europe's Fiscal Union." *Bruegel Essays and Lectures Series*, Brussels: Bruegel.
- Honohan, Patrick.** 2010. "The Irish Banking Crisis: Regulatory and Financial Stability Policy 2003–2008." Report for the Commission of Investigation into the Banking Sector in Ireland. Available at: http://www.bankinginquiry.gov.ie/Preliminary_Reports.aspx.
- Honohan, Patrick, and Daniela Klingebiel.** 2003. "The Fiscal Cost Implications of an Accommodating Approach to Banking Crises." *Journal of Banking and Finance* 27(8): 1539–60.
- Lane, Philip R.** 2006. "The Real Effects of European Monetary Union." *Journal of Economic Perspectives* 20(4): 47–66.
- Lane, Philip R.** 2010. "Some Lessons for Fiscal Policy from the Financial Crisis." *Nordic Economic Policy Review* 1(1): 13–34.
- Lane, Philip R.** 2011. "The Irish Crisis." In *The Euro Area and The Financial Crisis*, edited by Miroslav Beblavy, David Cobham, and L'udovit Odor, 59–80. Cambridge University Press.

- Lane, Philip R., and Gian Maria Milesi-Ferretti.** 2004. "The Transfer Problem Revisited: Real Exchange Rates and Net Foreign Assets." *Review of Economics and Statistics* 86(4): 841–57.
- Lane, Philip R., and Gian Maria Milesi-Ferretti.** 2007. "Europe and Global Imbalances." *Economic Policy* 22(51): 519–73.
- Lane, Philip R., and Gian Maria Milesi-Ferretti.** 2011. "The Cross-Country Incidence of the Global Crisis." *IMF Economic Review* 39(1): 77–110.
- Lane, Philip R., and Gian Maria Milesi-Ferretti.** Forthcoming. "External Adjustment and the Global Crisis." *Journal of International Economics*.
- Lane, Philip R., and Barbara Pels.** 2012. "Current Account Imbalances in Europe." *Moneda y Credito*, forthcoming.
- Lane, Philip R., and Peter McQuade.** 2012. "Domestic Credit Growth and International Capital Flows." Unpublished paper, Trinity College Dublin.
- Marzinotto, Benedicta, Andre Sapir, and Guntram B. Wolff.** 2011. "What Kind of Fiscal Union?" *Bruegel Policy Brief No. 2011/06*.
- McGuire, Patrick, and Goetz von Peter.** 2009. "The US Dollar Shortage in Global Banking and the International Policy Response." *BIS Working Paper No. 291*.
- Milesi-Ferretti, Gian Maria, and Cedric Tille.** 2011. "The Great Retrenchment: International Capital Flows during the Global Financial Crisis." *Economic Policy* 26(66): 285–342.
- Mody, Ashoka, and Damiano Sandri.** 2012. "The Eurozone Crisis: How Banks and Sovereigns Came to be Joined at the Hip." *Economic Policy* 27(70): 199–230.
- Muellbauer, John.** 2011. "Resolving the Eurozone Crisis: Time for Conditional Eurobonds." *CEPR Policy Insight No. 59*.
- Philippson, Thomas, and Christian Hellwig.** 2011/ "Eurobills, Not Eurobonds." VoxEU, December 2. <http://www.voxeu.org/article/eurobills-not-euro-bonds>.
- Reinhart, Carmen M., and Kenneth S. Rogoff.** 2009. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press.
- Reinhart, Carmen M., and Kenneth S. Rogoff.** 2010. "Growth in a Time of Debt." *American Economic Review* 100(2): 573–78.
- Reinhart, Carmen M., and M. Belen Sbrancia.** 2011. "The Liquidation of Government Debt." *NBER Working Paper 16893*.
- Sargent, Thomas J.** 2012. "United States Then, Europe Now." *Journal of Political Economy* 120(1): 1–40.
- Shambaugh, Jay.** 2012. "The Euro's Three Crises." *Brookings Papers on Economic Activity*, forthcoming.
- Shin, Hyun Song.** 2012. "Global Banking Glut and Loan Risk Premium." *IMF Economic Review*, forthcoming.
- Wyplosz, Charles.** 1997. "EMU: Why and How It Might Happen." *Journal of Economic Perspectives* 11(4): 3–22.
- Wyplosz, Charles.** 2006. "European Monetary Union: The Dark Sides of a Major Success." *Economic Policy* 21(46): 207–261.