What is the main limitation of much modern macroeconomic theory, among the failings pointed out by William R. White at the 2010 Mayekawa Lecture? We argue that the main deficiency is a failure to incorporate the possibility of default, including that of banks, into the core of the analysis. With default assumed away, there can be no role for financial intermediaries, for financial disturbances, or even for money. Models incorporating defaults are, however, harder to construct, in part because the representative agent fiction must be abandoned. Moreover, financial crises are hard to predict and to resolve. All of the previously available alternatives for handling failing systemically important financial institutions (SIFIs) are problematical. We end by discussing a variety of current proposals for improving the resolution of failed SIFIs.

Keywords: Default; Transversality; Money; Bankruptcy cost; Asset bubbles; Resolution mechanisms

JEL Classification: B40, E12, E30, E40, E44, G18, G20, G28, P10
I. Introduction

Last year, our friend and colleague, Bill White, with whom one of us has worked now for more than 40 years, first at the Bank of England and subsequently, gave the Mayekawa Lecture on the subject of “Some Alternative Perspectives on Macroeconomic Theory and Some Policy Implications” (White [2010]). The lecture this year is complementary to that earlier lecture and tries to extend and to build on it. Bill pointed out several of the main deficiencies of modern theory.

Let us quote from the opening section of his paper:

[T]he prevailing macroeconomic frameworks simply allowed no room for crises of the sort we are currently experiencing.... Absent an analytical framework that included the possibility of crises and deep economic slumps, it is not surprising that the crisis was not commonly anticipated. Nor is it surprising that no policy efforts were made to prevent the crisis from happening.

Moreover, absent any fears of crisis, few ex ante preparations were made to help improve crisis management (e.g., adequate deposit insurance, special legislation for the insolvency of financial institutions, and so on). Further, ex post crisis management was also inadequate in that each stage of the downturn was treated as the last, and recovery was constantly said to be imminent. By way of example, problems in the banking sector were initially treated as having to do with liquidity rather than solvency, and it was generally assumed that traditional Keynesian policy responses would suffice to restore full employment. . . .

It will be contended in this paper that the two workhorses of post-World War II macroeconomics have serious practical deficiencies. These workhorses are referred to here as modern macroeconomics (made up of the New Classical and New Keynesian models favored by academics) and applied Keynesian models (generally empirically estimated IS/LM models of the type still favored by policymakers and other applied economists). The former models rule out crises and deep slumps by assumption. The latter set of models underestimates the contributions made to deep slumps by developments occurring in the upswing. Thus, they overestimate the capacity of Keynesian policies to moderate deep slumps when they do occur. In effect, they also rule out deep slumps, but on the basis of the assumption that policy will always work effectively to moderate them. Taken together, these points also imply a greater need to lean against the upswings of credit cycles rather than to simply try to clean up afterward.

To remedy these deficiencies, it will be argued here that a new analytical synthesis is required. The building blocks of such a synthesis would be an increased focus on credit, stocks rather than flows (balance sheets), the possibility of stock “imbalances” (in particular excessive levels of debt), and the process of transition into crisis. In effect, the work of Keynes needs to be complemented by additional insights from the Austrian School of Economics, and still others from the work of Hyman Minsky.
In our view, Bill White’s assessment of the defects of current macroeconomic theory is pertinent, but we would not follow his approbation for Austrian theory, at least not in general. While we do agree with the praise for Hyman Minsky, we would note that neither Hy Minsky nor Bill himself have appeared able to express their insights in the form of a well-defined mathematical model. There were no equations, no algebra, not even any charts or tables in Bill’s lecture last year. While this may have come as a blessed relief to many of his listeners, and we shall follow his example in this respect, such a lack of formalization remains a considerable handicap to having one’s ideas accepted by a profession that has pretensions to being a proper science, and gives greater weight to mathematical precision than to empirical realism.

But above all, we do not think that Bill quite put his finger on what we see as the central, crucial deficiency of the standard forms, and mechanisms, of modern macro theory. This is their failure to incorporate the concept of default, that is, that borrowers of credit may fail fully to repay what they had previously promised, to renege on their debt commitments, into their core models.

As soon becomes obvious, once one tries to include the potentiality for default into a macro model, such modeling becomes much harder. Partly as a result, default is assumed away in the standard Walrasian models, from which the dynamic stochastic general equilibrium (DSGE) model is directly descended. In the technical jargon, the transversality condition implies that all debtors repay all their due debts in full by the terminal date of the model.

But this assumption requires two conditions to hold, both of which are patently false. The first is that no debtor fails to repay, even when it is in his own self-interest so to do (i.e., no strategic defaults); that is, that we are maximizing something beyond our own utility. The second is that, whatever happens in the future, the debtor will still be able to repay. This must logically require complete financial markets, wherein all eventualities, including Donald Rumsfeld’s famous “unknown unknowns,” can be hedged. How can you price and hedge the unknown? Since the number of potential future outcomes is infinite, any transaction cost, however minute, would make the whole exercise infeasible.

Nevertheless, Walrasian models, and their modern descendants, are based on this fiction of perfect frictionless financial markets (plus debtors who place honest repayment above self-interest, a condition perhaps more nearly realized in Tokyo than in London or Athens!). Apart from its obvious lack of realism, such perfect, frictionless financial markets have strong consequential implications for the structure of the financial system.

First, there is no need for financial intermediaries, whereby firms interpose their own credit standing or credit analysis between the ultimate saver and ultimate borrower, since there is no credit risk. All borrowers—whether the government or a noodle-shop owner, Governor Shirakawa, or yourself—are known with certainty to have no credit risk at all. All can borrow, or lend, in a frictionless world without transaction cost at the risk-less interest rate. So there is no room in such a world for banks or money market mutual funds. Moreover, in such a perfect financial system where the probability of all future eventualities is assumed to be known, and can be hedged without transaction cost, there would seem to be no room for specialist insurance companies or pension
funds. The complete perfect market will do it all. Our core models assume away the whole structure of financial intermediation. As Bill White correctly emphasized, this is not the mark of an evidence-based science.

Indeed, in Michael Woodford’s iconic book, *Interest and Prices* (Woodford [2003]), this implication is accepted and financial intermediaries play no role at all. However, whereas standard macro modelers have been willing to embrace the logic of their assumptions in respect of the absence of any role for financial intermediation (at best, a veil), they gag at the further but equally logical step of eliminating money—and hence inflation—from their models, perhaps because such models have a nominal interest rate, a Taylor reaction function and inflation (expectations) as key working parts.

It has long been known at the theoretical level that Walrasian (Arrow, Debreu, Hahn) models with perfect and complete financial markets have no logical role for money. When everyone is risk-less, then anyone’s IOU can and would be immediately and fully acceptable in payment for goods or services. There would still be a need for an accounting system (the divine accountant marking to a perfect market) and possibly for a numeraire, but not for a special asset category called money. Why need anyone hold “money,” when his or her own IOU is entirely acceptable in any exchange? To include “money” in a system with perfect financial markets in someone’s utility function is just a logical error. It is, indeed, the concept of default—that not all debts are fully repaid—which gives substance and meaning to the human instruments of money, liquidity, banks, and the whole panoply of finance.

If there ever could be such a frictionless, perfect financial system, could we still give meaning to the rate of interest? Yes, of course, but it would be a real rate, affected by real factors such as time preference, the marginal efficiency of investment, and so on. Could the government then influence it, if other frictions, such as wage and price stickiness, or other shocks such as war or national disasters strike, causing unemployment?

The answer is yes. By its powers of taxation, the government can reduce the private sector’s current access to consumption, and by increasing its expenditures, it could raise present consumption relative to the future. So the government can vary the trade-off between present and future goods. Of course, if there were no such frictions (or shocks), the private sector, assuming also that the government never defaults, could and would offset the government’s actions (as Barro [1974] showed long ago). But so long as there are real frictions, there can be a fiscal theory of (real) interest rates.¹

Without a clear appreciation of the essential nexus between default and money (or liquidity), monetary theory tends to get into a mess. This, for example, is exemplified by the attempt to base theories of the evolution, and role, for money on the need to minimize transaction costs in markets. But the appropriate counterfactual, at least for most purposes, to a monetary economy is not a barter economy, but a pure credit economy. Moreover, money—and credit—developed over time not so much in the context of markets for goods and services, but in the context of social inter-relationships. But one of us has already written on this subject at some length in his paper on “The Two

---

¹ There can also be real automatic stabilizers. In Egypt, which had not invented money, taxation was related to the height of the Nile flood. The higher the flood, the better the harvest. So the tax level for the year, payable in grain or labor, was adjusted according to the flood gauge.
The Mayekawa Lecture: The Role of Default in Macroeconomics

Concepts of Money” (Goodhart [1998]). And today’s paper is to focus on default, rather than money.

So, in Section II, we shall discuss ways to model default. It has been, we would suggest, the inherent difficulty of doing so that has led most economists to assume it away. Then, in Section III, we shall move on to outline various ways of trying to forecast an economy in which systemic financial failure is a possibility, though perhaps (hopefully) only as a tail risk. In the final section (Section IV), we shall consider a variety of methods for handling, and minimizing the cost of, financial failure (especially of banking failure).

II. Including Default in Macroeconomic Models

Default is hard to model formally, partly because it is a discontinuous variable. A company (or other economic agent) is either in default of a commitment, or not. So, prior to Martin Shubik’s original work (Shubik and Wilson [1977]), there was relatively little attempt to include default in formal macroeconomic models, although the probability of default has always played a central role in finance.

Shubik described money as an “institutionalized symbol of trust,” and Kiyotaki and Moore (2002) coined a nice phrase, “Evil is the root of all money.” And they are correct in this. If everyone always repaid all their debts with certainty, all that would be needed to complete a transaction would be a handshake and the acknowledgment that the buyer is indebted to the seller. Of course, the good that the seller would like to receive at some future date would not necessarily be what the buyer could offer, but this discrepancy could easily be resolved in complete financial markets.

It is, indeed, the possibility of default on contractual obligations that implies the necessity for the cash-in-advance constraints that we employ in our model. The interplay of liquidity and default justifies fiat money, based on the (tax) powers of government, as the stipulated mean of exchange. Otherwise, the mere presence of a monetary sector without the possibility of endogenous default or any other friction in equilibrium will become a veil without affecting real trade and, eventually, final equilibrium allocations.

The closest precursor to our present suite of models is the work of Shapley and Shubik (1977) and Shubik (1973, 1999), who introduced a central bank with exogenously specified stocks of money and cash-in-advance constraints in a strategic market game. The commercial banking sector of our models follows closely Shubik and Tsomocos (1992), who used, however, gold-backed money and modeled a mutual bank with fractional reserves. Finally, the modeling of money and default in an incomplete markets framework is akin to the models developed by Dubey and Geanakoplos (1992) and Dubey, Geanakoplos, and Shubik (2005). However, Dubey and Geanakoplos (1992, 2003) is a one-period model with money and default, Dubey and Geanakoplos (2003) include incomplete asset markets and money, and Dubey, Geanakoplos, and Shubik (2005) have incomplete asset markets, default, and no money. None of the previous papers combined all these three ingredients, incorporates a competitive commercial banking sector, and focuses on financial fragility. However, Goodhart, Sunirand, and Tsomocos (2006a) and Tsomocos (2003a, b) combine all three characteristics. Default
is modeled as in Shubik (1973) and Shubik and Wilson (1977), namely, by subtracting a linear term from the objective function of the defaulter proportional to the debt outstanding.

In these papers, Shubik developed a method whereby default can be incorporated into formal models. It runs as follows: the advantage of default is that it enables the borrower not to repay all that he owes. The disadvantage is that society imposes certain costs on defaulters, whether pecuniary (lawyers’ fees, no access to further borrowing, etc.) or non-pecuniary (reputational damage, debtors’ jail, etc.). If the costs of bankruptcy are zero, everyone would default totally and no one would lend. If the cost were infinite, no one would borrow. So there must be an interior optimum at which the benefits of default just balance its costs. The actual incidence of default then for any agent depends on the interaction of the agent’s preferred strategy interacting with the state of the economy, so that when output is low and consumption temporarily depressed, the marginal utility of extra consumption (from default) reduces the share of the debts that the debtor will pay back, the repayment rate. In other words, default (the inverse of the repayment rate) is always and simultaneously both strategic, dependent on each agent’s character, and endogenous, dependent on the state of the world.

An implication of this approach is that the mean expected default rate for all agents, however low, remains non-zero in almost all states of the world. The expected, and ex post actual, repayment rate is rarely 100 percent. Of course, in cases where bankruptcy costs are high, economic conditions are good, and the agent is both risk averse and honorable, the repayment rate may rise close to 100 percent. In reverse circumstances, the repayment rate may fall sharply.

Formally, this relationship can be described as a constrained maximization problem; and the specification of default can include the idea (first introduced by Shubik and Wilson [1977]) that utility decreases monotonically in the level of default because it becomes more expensive to default. In equilibrium, agents equalize the marginal utility of defaulting (i.e., increased consumption) with the marginal disutility of the bankruptcy penalty. Then, under rational expectations, expected rates of delivery of repayments of all kinds of debt in all periods and states of nature are equal to actual rates of delivery in equilibrium. This concept is the crucial ingredient of the model, since it allows default as an equilibrium outcome without necessarily collapsing the orderly functioning of the financial system.

There are some who find the concept of a partial repayment rate objectionable, because agents either default, or not (a binary division). But in a sector consisting of many small individual units—for example, companies or households—one can think of the average repayment rate of the sector as a whole. And where the sector consists of one, or a few, agents—for example, government, oligopolistic banks or car companies—even when the agent defaults, most creditors get some partial repayment. Thus, the bondholders with claims on Argentina or the Irish banks usually get some partial repayment.

So the Shubik approach, with its assessment of ex ante expected, and its post actual, repayment rates is, we claim, realistic. Alternatively, we can model default on collateralized loans. In the initial period, agents finance their investment through collateralized borrowing. When they borrow from the collateralized loan markets, they pledge the
assets purchased as collateral. In the second period, the borrowers either deliver in full
the amount of the collateralized loan, or default. In the case of default, the collateral
pledged is foreclosed and is put up for sale in the secondary capital market. The receipts
are transferred to the banks and determine the effective return on the collateralized loan.
More involved collateral requirements can be introduced. Due to our general equilib-
rium framework, every contract is priced in equilibrium. When equilibrium prices are
such that the value of the collateral is higher than the nominal value of the loan, the
agent will repay fully. When equilibrium prices are such that the value of the collateral
in the future is less than the amount the agent must repay, he would rather default,
purchase, for instance, the same amount of capital from the secondary market, and
be better off. Default is an endogenous decision stemming from utility optimization.
This process of default, and subsequent fire sale of assets can then generate the debt
deflation channel whereby monetary policy and money supply matter for the deter-
mination of asset prices, such as the interest rate on the collateralized loan, and they
affect the decision to default and the level of aggregate output (see Lin, Tsomocos, and
Vardoulakis [2010]).

The other main problem that this approach engenders is that of dimensionality. At
the very least models of this kind involve, as an essential component, an additional
banking sector. If one is concerned with interactions, within the banking sector for
example, then the requisite number of agents in the model rises much faster. More-
over, unless all the agents have prior knowledge of the long-run equilibrium properties
of the system, which seems implausible, the progress of the system will be path de-
pendent. Thus, we show in Bhattacharya et al. (2011) that financial collapses will be
more extreme after a sequence of good outcomes, since agents will have revised their
expectations in such a way as to assume more risk.

The opposite side of the coin, however, is that we can make models of this kind
quite feasible to explore whichever problem is at hand. Thus, we can use this approach
to develop a DSGE model in which the only friction is financial, or in which there are
both wage/price and financial frictions. The initial endeavors, by Curdia and Woodford
(2009), Leao and Leao (2007), de Walque, Pierrard, and Rouabah (2010), and Iacoviello
and Neri (2010), which introduced some aspects of financial friction into the DSGE
framework, have not taken into account simultaneously liquidity, agent heterogeneity,
money, and default risk. The majority of these models attempt to model default as an
out-of-equilibrium phenomenon that never occurs in equilibrium. Nevertheless, these
models are valuable efforts toward a plausible explanation of the phenomena observed
during and after the credit crisis. One paper that attempts to achieve this synthesis is

One of the crucial elements remaining to be introduced into the DSGE framework
is the liquidity constraint the agents face, because goods are not fully readily tradable
for other goods. Acharya and Pedersen (2005), Acharya, Shin, and Yorulmazer (2011)
and Acharya and Viswanathan (2011), Brunnermeier and Pedersen (2009), and Vayanos
(2004) have all studied liquidity within partial equilibrium models. In our models, li-
quidity will be modeled following Espinoza and Tsomocos (2008). Such an extension
to the dynamic framework will be a direct and useful tool to assess the impact of the
financial and real shocks, since it has two important advantages. The first is the ability
to monitor the impact of a liquidity shock in the short, but also medium, run. The second is that the dynamic setting allows us to parameterize different liquidity environments (i.e., steady-state values) and examine how shocks impact the economic variables in each case.

Perhaps the main purpose of our program of work has been to develop models that might be used to explore financial stability. Thus, our earlier models focused on interactions within the banking sector. In particular, Goodhart, Sunirand, and Tsomocos (2006b) and Tsomocos (2003a, b) set out a tractable model that illuminates problems relating to individual bank behavior, to possible contagious inter-relationships between banks, and to the appropriate design of prudential requirements and incentives to limit “excessive” risk-taking. Our model is rich enough to include heterogeneous agents, endogenous default, and multiple commodity, and credit and deposit markets. Yet it is simple enough to be effectively computable and can therefore be used as a practical framework to analyze financial fragility. Financial fragility in our model emerges naturally as an equilibrium phenomenon. Among other results, a nontrivial quantity theory of money is derived, liquidity and default premia co-determine interest rates, and both regulatory and monetary policies have nonneutral effects. The model also indicates how monetary policy may affect financial fragility, thus highlighting the trade-off between financial stability and economic efficiency. Other papers, using this—or a closely similar—model, mostly developed within central banks, include Saade, Osorio, and Estrada (2007) and Goodhart, Sunirand, and Tsomocos (2006b).

Given the importance of the boom/bust cycle in the property market, both residential and commercial, we have developed models to explore and to simulate such crises. Goodhart, Tsomocos, and Vardoulakis (2010) build on our previous model of a system in which default plays a central role for both borrowers and banks, and in which financial intermediation and money, therefore, have a necessary real function, to include both an additional good—housing—in addition to the prior composite basket of other goods and services, and an additional agent—a new entrant to the housing market. So our model is general enough to allow for the examination of a wide variety of shocks, which can lead to financial instability.

III. Default and Forecasting

A. Extrapolating the Past

Our past history is all that we (partially)2 know. From this we can extract trends, for example, of output, productivity, monetary growth, velocity, inflation, and so on. Given such trends, the actual time series exhibit cycles, of varying periodicity, around the trend. One of the strongest maintained assumptions in economics is that the system has an equilibrium (rate of growth), to which the economy would revert in the longer run, in the absence of future shocks (including defaults and financial shocks). DSGE models

---

2. The downgrading of the teaching of economic history from its prior role as part of the core curriculum of an economics degree to a lesser status as an inessential specialization has meant that most economists know even less of their past history than previously.
are in general constructed around a strong assumption of reversion to an equilibrium (Brunnermeier and Sannikov [2011]).

At any time, however, the economic system is out of equilibrium, and is at some stage of a cycle, with its own autoregressive momentum. Thus, stripped of its detail, most economic forecasting represents a balance between the centrifugal forces of ongoing cyclical momentum and the centripetal forces of reversion to the long-run equilibrium. In practice, most forecasters are poor at predicting cyclical turning points. Hence they tend to predict some fairly constant combination of autoregressive momentum and reversion to the mean. As a result, so long as the up phase of a cycle continues, forecasters tend to underestimate outcomes, and vice versa in the down phase. This gives rise to a common pattern of outcomes and predictions looking like Figure 1; see Goodhart and Lim (2011).

While many of the shocks giving rise to this pattern are real—for example, technology, productivity, and supply shocks—some are financial. The most serious financial shocks are those that occur after periods of strong and stable growth, with increasing optimism about declining volatility and lower risks. Explanations of this syndrome are given by Minsky (e.g., Minsky [1982, 1992]) and modeled by Bhattacharya et al. (2011). Examples are the United States (1929), Japan (1990), Southeast Asia (1997), and the developed economies (2007–08). After such a financial shock, there is generally a longish period of stagnant credit expansion and slow growth (Reinhart and Rogoff [2009]).

The financial crisis leads to a regulatory response, to tighten up, and more government intervention in the operation of the financial system. Some, for example, Congdon (2009, 2010, 2011), argue that such intervention and tightened regulatory
requirements are themselves a cause of the resultant slower growth of money, credit, and output. Be that as it may, the new and reinforced regulatory controls, following the Great Depression of 1929–33, then ushered in a long period from the 1930s to the 1960s that was characterized by constrained, repressed banking, but which was free of banking/financial crises.

Although the crises in the early 1970s were in some countries associated with financial liberalization—for example, the United Kingdom—the main blame was placed on poor government policies (e.g., incomes policies) and (oil) supply shocks. So despite a number of other episodes of financial stress (the less-developed-country crisis of 1981–82, involving Mexico, Argentina, and Brazil; the housing/European exchange rate mechanism [ERM] crisis of 1991–92; and the Southeast Asian/Long-Term Capital Management crisis of 1997–98), the process whereby further and extended financial liberalization and globalization coincided with increasing official success in achieving low and stable inflation and steady growth, the “Great Moderation,” continued until punctuated by the financial crisis of 2007–08.

If one calls this period from the 1929 crisis until the 2008 crisis a single super-cycle, then we have really only had one such episode. This is far too small to extrapolate into the future. But perhaps the example of Japan, whose own idiosyncratic crisis occurred much earlier in the 1990s, gives a warning that developed economies more generally may now be prone to a period of much slower growth both of the monetary aggregates and of output. How far such slower growth in money and credit is responsible for, and how far it is simply caused by, the slower growth in demand and output is not easy to discern.

The adoption of unconventional monetary policy measures—notably, quantitative easing (QE)—has been an attempt to reinvigorate monetary expansion. Insofar as this was expected to run via injections of base money, through the traditional money multiplier to the broader monetary aggregates, it was, however, a dramatic example of the effect of Goodhart’s law. Insofar, however, as it was expected to run via enhanced liquidity (and money balances) into asset prices more generally and into a recovery of certain dysfunctional financial markets, it has been quite successful, though exactly how successful remains a contentious and contested subject.

Clearly the virulence of the current boom/bust cycle, both in Japan and more recently in developed economies, owes much to the interaction of the housing and financial cycles. Standard DSGE models incorporate neither sector. We have tried to incorporate both, notably in our 2010 paper with Vardoulakis, already cited. Be that as it may, a policy response that focuses almost entirely on the monetary side and fails to deal with the continuing downward pressures on housing and land prices, for example, from foreclosures, is likely to be lopsided and only partially effective. Surely it is difficult to design a satisfactory policy to cope with mortgage default and consequential foreclosures, but the failure to do so has been a millstone holding down recovery in the United States and Japan. Looking again at the set of policies for housing, including importantly the bankruptcy provisions and “skin in the game” loan-to-value (LTV) requirements, remains a key element of unfinished business.

Equally the nexus between additional financial regulation and the pace of monetary and credit expansion remains a highly contentious issue. Many economists (Admati
et al. [2011], Miles, Yang, and Marcheggiano [2011], and Barrell et al. [2010]) and regulators (Basel Committee on Banking Supervision [2010a, b]) believe that tougher capital and liquidity requirements, so long as these are introduced quite gradually, will only have minor effects on interest rate spreads, and hence on credit and monetary aggregates. A few economists (Congdon again) and many in the industry contest this (Institute of International Finance [2010]) claiming that the effect already has been, is, and will be significant and damaging to recovery.

In our view, a crucial issue is the way that such tougher requirements are introduced. If banks are given a target for the absolute value of core tier 1 equity, based on their current total value of assets/liabilities, to prevent them from achieving a required ratio by delevering further, and if that target value of capital is enforced by dividend restrictions—and maybe bonus restrictions—until it is reached, then it could be met without adverse effects on the real economy. The continued proclivity of banks to pay out dividends and bonuses after the crisis struck in August 2007 and the premature relaxation on U.S. dividend payments in March 2011 should not have been allowed to occur. On this, see Acharya et al. and Goodhart et al. (2010).

Nevertheless, what is clear is that now that we in developed economies have entered the post-crisis (bust) stage, we still do not have much ability to forecast how the financial sector will respond to tougher regulation, and/or how developments in the financial sector will influence the real economy. We simply do not have the models to do so. Of course, a few diehard adherents of the previously mainstream DSGE models, without housing or finance sectors, may claim that none of this matters, that the only shocks are real, but it is hard to take such protestations seriously.

But policymakers, and most outside commentators, are not so much concerned with forecasting in the depressed, recovery stage (though they should be, since that is where we are now), as with the more high-colored (sexier) question of whether economists can provide an “early warning system” (EWS) to predict, and so perhaps to forestall, the advent of both systemic crises and of idiosyncratic individual bank failures. It is to this that we now turn.

**B. Can Crises Be Forecast?**

The first international banking problems since the interwar period surfaced in 1974 (Herstatt, Franklin National Bank, and fears about the recycling of petrodollars). Immediately thereafter, the Group of Ten (G10) central bank governors at the Bank of International Settlements (BIS) meetings and finance ministers called for an EWS to anticipate such crises in the future. Indeed, it was largely for the purpose of providing such an EWS that the Basel Committee on Banking Supervision (BCBS) was set up initially in 1974–75 (Goodhart [2011]). Fortunately, since attempts to estimate such an EWS have been of strictly limited success throughout, the BCBS soon turned its attention instead to the somewhat more amenable subject of international, cross-border, bank regulation and supervision.

It is not perhaps surprising that the Lehman failure caught everyone by surprise. After the handling of Bear Stearns and of Fannie Mae and Freddie Mac, by the Federal Reserve and the Treasury, the general assumption had been that these authorities would find some way to keep Lehman Brothers, and the other broker/dealer investment houses,
in continuing operation. Indeed, it was partly the shock to such prior expectations, and the resultant uncertainty about how the U.S. authorities might act from then on, that caused the crisis to become so sharp and abrupt (see Taylor [2009]).

Nevertheless, there are intrinsic reasons to doubt whether a really successful EWS can ever be developed. As implied by the Lucas critique (and Goodhart’s law), if such a warning system ever appeared to have predictive power, then agents’ behavior would change, and the crisis would be averted, one way or another. The best that we can hope to achieve is a greater appreciation of the conditions of stress that may foster a crisis. Several of these are known already. They include credit expansion, leverage, housing and property price inflation, and output growth, which are both rapid and above trend. The crisis will also probably be worse when it is preceded by a current account deficit (Barrell et al. [2010] and Reinhart and Rogoff [2009]). There is also evidence that periods in which the risk of a collapse (recovery)—as assessed by tail risk in asset returns—is seen as higher than the opposite, but asset prices continue rising, herald a sharp reversal (Cascon, Shadwick, and Shadwick [2011]).

As is fairly obvious, almost all the signs of potential stress are equivalently signs of an asset price boom. In a boom, almost everyone is making money and is optimistic. Only the Cassandras and the doomsayers will have lost money. The precept of “buy on the dips” will have led to fortune. Nobody can ever regularly predict either a turning point or a change in trend. In this respect, macroeconomics has something in common with seismology. There are always many who predict a continuing faster expansion: “it is different this time;” “the end of boom and bust;” “Dow 40,000;” and so on. Indeed, during the course of such a boom, the most accurate forecasters will be those who have been generally most optimistic. It is difficult not to get caught up in the general enthusiasm; indeed as Frankel and Froot (1986, 1987) showed, the boom (bust) only collapses once all the uncertain investors have finally jumped on the bandwagon.

Asset price booms and busts are market phenomena. The market itself by definition cannot, and does not, predict its own crisis and reversal. Credit default swap (CDS) spreads were never lower, and bank equity prices never much higher, than in June 2007. Within a cyclical phase the market can predict quite well which company/bank will do best/worst, but the market gives little or no guidance about future macro-developments, or of its own sharp reversals.

Even if central bankers should manage to maintain their scepticism about asset price bubbles (“irrational exuberance”), they would face intense (political) opposition if they were to try to take measures to counter such a boom. “Taking away the punch bowl just when the party gets going” will not endear one to the guests. Given all the other difficulties of using countercyclical measures to halt an asset price boom, leaving this just to the discretion of central banks is likely to result in them being underused. We have, however, made this point in other papers.

Besides the repeated call, never effectively answered or answerable, for an EWS, the other persistent nostrum of regulators seeking to forestall crisis is for stress tests. These also are of very limited use. We are confident that had regulators applied a stress test in June 2007 on the assumption of a decline in U.S. housing prices by June 2009 of 20 percent (five times the maximum expected by econometric exercises and which would then have been dismissed as totally implausible), not a single U.S. or European
bank would have appeared in any serious danger. Recall how the Irish banks passed the European stress tests.

There are several inherent problems with stress tests. The first is that the regulators may be unwilling even to contemplate the most serious risk events, for instance, the default of a sovereign debtor. A second is that the regulators running the stress test need to have a convincing backstop policy in place to restore any bank, or other financial intermediary, to reasonable health should it fail the test before the stress test is applied. Otherwise, there would be growing doubts whether either a bank (or other financial institution) or the regulators themselves would be prepared to announce such a test failure. Such a backstop (provided by the Troubled Asset Relief Program [TARP]) was in place in the successful U.S. stress test, but not in the unsuccessful European test.

But the most important deficiency of any stress test is that it explores the effect of some exogenous, macroeconomic shock on an individual bank. It does not, and in practice cannot, examine the endogenous risk to the system as a whole, as banks themselves react to the worsened situation by deleveraging, restricting (interbank) loans, requiring more collateral from counterparties, selling assets, and hoarding liquidity. There are multiple sets of self-reinforcing spiral effects that drive financial bubbles and busts (Brunnermeier et al. [2009]), and stress tests on individual banks will not reveal these.

This is not to suggest that such stress tests are a waste of time. They may be able to indicate comparative weakness, that is, which banks would be seen as most at risk from some exogenous event. But they are not of much value as a guide to the probability of a systemic crisis. It is our view that the individual bank stress tests need to be complemented by top-down, model-based simulations of the banking (financial) system as a whole. One of the purposes of our work in building models in which banks and default play a central role is precisely to provide a basis for carrying out such simulations. We would not expect such models to be able to predict—to forecast—the actual occasion and initial cause of the crisis itself, any more than any other (reduced-form) EWS, but they might be able to estimate both the fragility of the system to such a crisis, for example, as risk-aversion coefficients could be assumed to alter and also perhaps to chart the likely further direction of the crisis once it was underway.

This latter consideration leads to the question of how to assess the likely interaction of the financial system with the real economy, once a financial crisis has begun.

**C. Forecasting Procedures after the Crisis Has Begun**

A cyclical downturn that has been triggered by a financial crisis is likely to differ, in several respects, from a downturn caused by other factors, for example, official action to counter inflation (Reinhart and Rogoff [2009, 2011] and Reinhart [2011]). The supply of credit to marginal borrowers at least, and monetary growth, will be cut back. Spreads between such government debt as is perceived to remain riskless and on riskier assets are likely to widen.

Indeed, the simplest and most straightforward way to incorporate financial difficulties into a forecast for the real economy is just to substitute a risk-spread adjusted interest rate, in place of the official rate, into the model’s domestic expenditure function(s),
and then proceed as before. This is in effect the proposal of Curdia and Woodford (2009), and it has much to recommend it (also see Gilchrist and Zakrajeski [2011]).

There are, as always, problems. Exactly which risk spread should be used? Is the effect on expenditures of a rise in official rates of 1 percent more or less equal to a rise in the risk spread of 1 percent (official rates remaining constant)? We can think of arguments why the effect of the latter (risk spread rise) might be greater (additional non-price rationing of credit) or less (not so pervasive; less effect on confidence and expectations perhaps) than a rise in the official rate. We are not currently aware of any conclusive econometric tests on this issue.

But the main problem with this approach is that the risk spread is treated as an exogenous variable, rather than being derived endogenously from the optimizing behavior of agents. Moreover, the forecast will have to contain some estimate of the future path of the risk spread itself. Of course, this too can be provided by some auxiliary model. It could be univariate; thus, after an (unforeseeable) shock, risk spreads tend to revert asymptotically to the low, long-run equilibrium, unless distorted by some further shock. Or one could run a reduced-form equation, relating the spread to factors such as failure rates, expected bank profitability and capital ratios, the increase in base money, and so on. Some might ask why not relate risk spreads to market variables such as option and CDS prices; but such market pricing variables are in a sense just another facet of the same market factor.

Ideally, we would like to be able to forecast risk spreads as one of the outputs of our more general equilibrium financial model, in which default plays a central role. But for the time being, this is beyond our capabilities. Nevertheless, this can only be achieved within a model of liquidity and endogenous default whereby spreads emerge in equilibrium and, hence, can be calibrated and used for policy determination and regulatory purposes. At the same time, financial stability measures should be constructed that can be easily implemented and used contemporaneously with inflation targeting to assess monetary and financial stability policy.

IV. How to Handle the Default of a Global SIFI

Few events have more destructive power than the default of a global systemically important financial institution (G-SIFI), especially if badly handled, as in the case of the failure of Lehman Brothers. Not only is there a major loss of value, following the event of default, to creditors and/or taxpayers, but much more important, the blow to confidence and the subsequent effect on market prices can have an impact on the real economy that is a large multiple of the direct loss from the initial failure. The failures of Lehman Brothers, of Credit-Anstalt, and of the Knickerbocker Trust Company in 1907 are examples. The extension of credit, and the working of the financial system, are based in some large part on trust and confidence, and—should that trust suddenly evaporate—the inner fragility of the capitalist system is revealed. This can be terrifying.

And for the creditors of the failing bank, the occasion of bankruptcy will tend of itself to destroy value. We understand (from personal conversation) that, at the weekend which settled the fate of Lehman Brothers, the assessed shortfall of value, once equity
had been wiped out, was somewhere in the region of US$25 billion. But after the bankruptcy had occurred, the eventual shortfall was around three times as large (or more). What makes default such an expensive exercise? First, the generalized shock of the event and the expectation of forced fire sales will cause the value of the assets held by the bank to decline; with no ongoing reputational ties to maintain, debtors to the bank will seek ways to delay or to diminish their repaying commitments; the specialized personal knowledge of the bank employees suddenly loses value (goodwill disappears), and the bankruptcy procedure itself is extremely expensive and time consuming.

So creditors of SIFIs, especially bondholders, often would have a common interest in providing more equity capital up front to rescue the failing SIFI rather than allow it to default. There is, however, a major coordination problem to overcome in arranging this. It is the purpose of contingent convertible (CoCo) bonds and of bail-in procedures for bondholders to try to overcome this coordination problem. But this is to run ahead of the currently unfolding story, since neither CoCos nor bail-in arrangements are yet in general widespread use.

Instead at present, when faced with the default of a bank or SIFI, the authorities have only had three alternatives:

1. Try to arrange a merger with a stronger financial institution, sometimes sweetened by some (overt or covert) subsidy.
2. Rescue the bank using taxpayers’ money.
3. Allow it to go bankrupt and be liquidated.

All these alternatives have major disadvantages. The encouraged merger route increases concentration, reduces competition, may involve the outlay of taxpayers’ money (and/or can involve risky central bank loans), and may create a much larger weak bank, where formerly the absorbing bank was much healthier. The encouraged mergers of Bear Stearns (into JPMorgan Chase), HBOS (into Lloyds TSB), and Merrill Lynch (into Bank of America) each illustrate some of these disadvantages. Nevertheless, in the throes of a crisis this is often the preferred solution for the authorities.

When this solution became unavailable, the choice lies between propping up the failing bank with taxpayers’ money or liquidating it. The experience of the failure of Lehman Brothers was so appalling that most governments thereafter decided that liquidation of an SIFI could not be tolerated. So they moved rapidly to the use of taxpayers’ money to recapitalize failing banks. The problem with this was that it has often turned out to be too expensive for the public sector to sustain, “too big to save” as in Ireland, and the political consequences of imposing austerity on taxpayers and public-sector employees to save banks and bankers—who are perceived to have behaved badly—have become insupportable. So this route has, in many respects, also reached its limits.

Since none of the current alternatives for handling failing SIFIs are acceptable, or are capable of much further utilization, the search is on for other methods for handling their default. We shall discuss three here; though they are separable, they are not necessarily mutually exclusive:

1. Improve the liquidation process.
2. Make the bondholder liable, via CoCos and bail-ins.
3. Reform Prompt Corrective Action (PCA).
The first alternative, as largely adopted in the Dodd-Frank Act, involves four main elements.
(a) Enactment of a Special Resolution Regime enabling the authorities to take over and manage the running of a bank prior to actual bankruptcy, to split it into a good bank/bad bank, and so on, as soon as its supervisors come to the view that its normal operations are no longer sustainable. This can be well before it is formally bankrupt.
(b) The requirement for all SIFIs that they complete “living wills” to the satisfaction of their supervisors. This should include both a crisis recovery component and a funeral plan, should the recovery fail (Huertas [2010a, b]). The funeral plan should enable the authorities to have the capacity to know in advance how to close down this SIFI expeditiously and efficiently.
(c) An orderly liquidation process, ready to be put in place.
(d) A fund, to be provided ex ante, or ex post, by the remaining banks to meet the residual cost of the liquidation.

Much of this program, especially (a) and (b), is admirable, but we doubt whether it will work effectively for the following reasons. On (a), assessing when a bank has gotten into an unsustainable state is extremely difficult, and will be contested by the bank’s owners and equity holders. Given the threat of a lawsuit, if such closure is to be left to the discretion of the supervisors, it will come too late. On (b), the procedure for completing and updating “living wills” is, we expect, likely to be so time-consuming and expensive for both the regulator and the regulated that it will be done only occasionally and less thoroughly and completely than would in principle be ideal. On (c), the main problem is that G-SIFIs are by definition cross-border (indeed most SIFIs are), and the institutional structure for organizing an orderly cross-border liquidation is nonexistent (Avgouleas, Goodhart, and Schoenmaker [2010], Goodhart and Schoenmaker [2009], and Claessens, Herring, and Schoenmaker [2010]). Much of the disaster attending the Lehman failure arose not from the treatment of the U.S. part of the firm, but from the chaos that occurred with the liquidation of Lehman Brothers (Europe) in London and Lehman Brothers (Asia) in Tokyo. On (d), no one knows the correct premium (tax rate) to charge ex ante, and the banks will fight the proposal. On the other hand, charging the better-run banks, which did not fail, ex post just when the banking system as a whole is much weaker would be misguided (though this is proposed by Dodd-Frank!). For additional criticisms of the Dodd-Frank Act, see Acharya et al. (2010). This is not to say that such a process of orderly liquidation cannot work, nor that answers to these objections cannot be attempted; rather, the claim made here is that the concept of an “orderly” liquidation process remains highly problematical.

In contrast to the Americans, the Europeans now appear much keener on option (2), making the bondholder pay, whereas the Americans seem somewhat more sceptical. There are two versions, CoCos and bail-ins. With the CoCo, a bond is issued that automatically transforms into equity when a preordained trigger is met, with a conversion factor that is also settled in advance. The optimal form of both settings is quite complex to decide. In our view, the best approach would be a market price of equity (maintained over some 20 working days to avoid flash crash and sudden manipulation problems) that was quite high above the failure level and converted at a rate which diluted existing
equity sufficiently to encourage existing equity holders to recapitalize themselves via a new issue, rather than allow the trigger to come into action.

Bail-ins are mechanisms to require the bondholder to recapitalize the bank when it approaches the point of failure. They can come in two versions, contractual or statutory. With the contractual version, the extent of requirement for the bondholder to meet the residual costs of the failure (after the bank reaches the near point of failure) is set out in advance in the prospectus and documentation (and is thereby limited). In this respect, it is akin to a CoCo with a very low (zero equity price) trigger.

With a statutory bail-in, the government assesses the residual loss to be met, plus the need for recapitalization, and then allocates the required burden across the various categories of bondholder in order from the most junior to the most senior. To avoid uncertainty, the principles to be followed would have to be clearly and publicly established in advance, probably incorporated into legislation since it involves the authorities encroaching on private property rights.

The newfound enthusiasm for imposing (some significant part of) the cost of failure on the bondholder extends beyond bank bondholders to holders of government bonds. All member countries of the eurozone are to include collective action clauses (CACs) into their sovereign bonds from 2013 onward. The idea is that CAC bonds will be considerably easier to restructure, that is, to impose a partial default, than bonds without such clauses.

In view of the disadvantages of the other methods for handling failing banks, one can see why this proposed alternative has been greeted so warmly at least in Europe. But it too has numerous problems. For example, poorly chosen settings for CoCos could make fragility worse rather than improve it. Next, if the authorities were to allow other banks or levered financial intermediaries to hold CoCos or bank bonds subject to bail-in, the potentiality for contagion and falling rows of dominoes is obvious. So if the holders are to be restricted to pension, hedge, and sovereign wealth funds or insurance companies, then the market could be rather thin. Moreover, the ultimate beneficiaries of such intermediated funds are much the same set of people as taxpayers. It is not necessarily clear that switching the burden of preventing liquidation from the taxpayer to the fund beneficiary will be politically much more palatable. It may seem so in advance when few such imposed losses of value have yet been suffered, but it may look quite different after the event.

But the most serious problem, in our view, is that of contagion, and the danger of making a crisis worse rather than better. Too often, the pros and cons of schemes like CoCos and bail-ins are discussed in the context of a single failing SIFI rather from the standpoint of the system as a whole. In general, a well-designed CoCo or bail-in mechanism does have much to recommend it in the context of a single failing bank (or eurozone government).

But the systemic problems could be acute. The failure of a single financial intermediary, even quite a large one (e.g., the Bank of Credit and Commerce International or Barings), is not that serious a problem if its cause was clearly idiosyncratic, *sui generis*. Most severe crises, however, are caused by *common* shocks. So if a CoCo or a bail-in gets triggered for one bank, it will raise the likelihood of it taking place at other banks. But no one will be sure of what the probability might be. So hedging on CDS and
short sales would increase. Moreover, the cost of issuing new CoCos and bail-in bonds for other banks subject to the common shock would rise sharply. So long as senior bondholders thought that they would not be subject to loss in the event of a default, then banks (and eurozone governments) could refinance themselves in the markets, even if they were under some suspicion. In the aftermath of the 2007–08 crisis, many banks resorted to debt issues to meet their funding needs. Would they have been able to do so, and at what price, if the bonds to be issued were subject to bail-in?

Markets can dry up and become dysfunctional when faced with the prospect of uncertain loss. Do we really want that to happen for the bank (and eurozone government) bond markets? The authorities could perhaps respond in the face of a generalized crisis by making the bonds issued during the duration of the crisis guaranteed against loss. But unless the world were lucky, the proportion of bail-in bonds then might be so small that the loss imposed on them for recapitalization would have to be large, perhaps so large as to sully the reputation of the asset category for a long time.

A besetting weakness of our regulatory approach has always been that it has focused on resolving the problems of the individual bank, rather than those of the wider financial system. We are concerned that the move toward requiring banks to issue CoCos and bail-in bonds may have the same flaw.

In our view, the conceptually best approach to dealing with banking fragility that has been tried was to impose a regime of PCA, as set out under the Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991. This not only set out a carefully considered and graduated ladder of sanctions as tier 1 capital fell below a fully adequate leverage ratio, but also required banks either to recapitalize or be closed when their leverage ratio fell below 2 percent, with the aim of never leaving a loss that other parties, for example, taxpayers or bondholders, would have to bear.

Yet this failed in 2008. Banks that failed generally had accounting capital, at their final audit, not only over the 2 percent limit but even on average higher than the comparable banks that did not fail (Haldane [2011]). Losses were recorded by the FDIC in resolving these failing banks. The approach was right, but the practice was deficient. As has now been frequently pointed out (Calomiris [2011]), market equity prices gave a much better guide to future failure than tier 1 accounting data. The latter involve some considerable delay (lags) and can be manipulated.

Simply switching from an accounting measure to a market measure for a leverage ratio would, however, lead to other problems. Market prices can to some extent be manipulated, and the use of a market measure would feed the hysteria about short sales. Then there is the problem of a flash crash; while if one responded to it by making the market price trigger depend on averaging over a certain number of days, how would the averaging be done? It could lead to a “sword of Damocles” effect on the bank involved, and attempts at end-period manipulation.

Perhaps the best option would be to use a combination of both accounting and market price triggers, with a considerably higher accounting ratio, say 4 percent for a tangible core equity leverage ratio from the accounting measure and a 2 percent tangible common equity (TCE) leverage ratio using market prices, averaged over two days. Both would have to be broken simultaneously to force PCA.
Under the FDICIA, a bank that fell below 2 percent would either be recapitalized or liquidated. In view of the costs of liquidation, we would propose that the bank either be recapitalized, or taken over by the government. One of the errors of handling the recent crisis was that political ideology prevented the government in either the United States or the United Kingdom from taking failing banks into temporary public ownership. Fear of the word “nationalization” was too great. But when banks get into such straits, not only should the management team and the board of directors be removed, but also existing shareholders should be expropriated. As soon as practically possible, the bank should be sold back to the private sector. If the receipts from such sales exceed the government’s costs and a reasonable rate of return on the interim investment of public-sector money, then the excess could be distributed to the prior shareholders.

A reformed and improved PCA mechanism should be able to cope with most banking problems, and a combination of living wills and special resolution regimes should enable the authorities to better handle the few that are so sudden and catastrophic that they are not picked up by the reformed PCA. Yet this is not the direction in which the regulatory authorities appear to be moving. In our view, bank resolution procedures will remain problematical both in the United States and Europe, though in different ways.
References


———, “Has Bernanke Broken His Promise to Friedman?” Central Banking, 20 (3), 2010.
The Mayekawa Lecture: The Role of Default in Macroeconomics


