RETHINKING MACROECONOMICS: WHAT FAILED, AND HOW TO REPAIR IT

Joseph E. Stiglitz
Columbia University

Abstract
The standard macroeconomic models have failed, by all the most important tests of scientific theory. They did not predict that the financial crisis would happen; and when it did, they understated its effects. Monetary authorities allowed bubbles to grow and focused on keeping inflation low, partly because the standard models suggested that low inflation was necessary and almost sufficient for efficiency and growth. After the crisis broke, policymakers relying on the models floundered. Notwithstanding the diversity of macroeconomics, the sum of these failures points to the need for a fundamental re-examination of the models—and a reassertion of the lessons of modern general equilibrium theory that were seemingly forgotten in the years leading up to the crisis. This paper first describes the failures of the standard models in broad terms, and then develops the economics of deep downturns, and shows that such downturns are endogenous. Further, the paper argues that there have been systemic changes to the structure of the economy that made the economy more vulnerable to crisis, contrary to what the standard models argued. Finally, the paper contrasts the policy implications of our framework with those of the standard models.

1. Introduction

Those who claim to be disciples of Adam Smith should be unhappy with what has happened in the last few years. The pursuit of self-interest (sometimes called greed) on the part of bank executives did not lead, as if by an invisible hand, to the well-being of all; in fact it was disastrous for the banks, workers, taxpayers, homeowners, and the economy more broadly. Only the bankers seemed to have fared well.1

1. Of course, Smith himself took a broader perspective on self-interest than his modern-day disciples, one which recognized some sensitivity to the effects of one’s actions on others. See, for instance, Nick Phillipson (2010). Indeed, Rothschild (2001) and Kennedy (2009) argue that Smith used the term “invisible hand” with some irony—with markedly different views about market perfection than those held by Smith’s latter-day descendants.
Modern general equilibrium theory has explained why markets are almost never (constrained) Pareto efficient whenever there is imperfect and asymmetric information or when risk markets are incomplete—which is always the case (see Greenwald and Stiglitz 1986, 1988). Both before and after that paper, there have been a large number of studies showing that even with rational expectations, markets are not in general constrained Pareto efficient. Much of modern macroeconomics forgot these insights, and constructed models centering around special cases where market inefficiencies do not arise, and where the scope for welfare-enhancing government intervention, either to prevent a crisis or to accelerate a recovery, is accordingly limited. This has made the models of limited relevance either for prediction, explanation, or policy—at least in times of severe downturns, when markets evidently are working so poorly.

Prediction is the test of a scientific theory. But when subject to the most important test—the one whose results we really cared about—the standard macroeconomic models failed miserably. Those relying on the Standard Model did not predict the crisis; and even after the bubble broke, the Fed Chairman argued that its effects would be contained. They were not. In the months that followed, policymakers floundered—and the Standard Model provided little guidance as to what they should do, for example the best way to recapitalize the banks. Many of the critical policies before, during, and after the crisis were based on analyses of modern macroeconomics. Monetary authorities allowed bubbles to grow, partly because the Standard Models said there couldn’t be bubbles. They focused on keeping inflation low, partly because the Standard Model suggested that low inflation was necessary and almost sufficient for efficiency and growth. They focused on nth-order distortions arising from price misalignments that might result from inflation, ignoring the far larger losses that result (and have repeatedly resulted) from financial crises. Belief in the efficiency of the market discouraged the use of the full panoply of instruments (for example, restrictions on mortgage lending) at the disposal of central banks and regulators; these would at least have dampened the bubble and mitigated its consequences. Instead, it was repeatedly claimed that it would be cheaper to clean up the aftermath of any bubble that might exist than to interfere with the wonders of the market. Thus, while financial markets and regulators have been widely blamed for the crisis, some of the blame clearly rests with the economic doctrines on which they came to rely (Stiglitz 2010a).

There are some, such as Ben Bernanke (2010), who take a markedly different view, arguing that economic science did not do a bad job. The fault, he argued, lay not with economic science, but with economic management. I believe he is wrong—if by economic science we mean the central macroeconomic models that have played key roles in the formulation of economic policy and thinking in recent years. He is right, of course, that there were many mistakes in the application of economic science. Economic policymakers should have been aware, for instance, of the consequences of the perverse incentive structures that had become prevalent within the financial sector.

2. On 28 March 2007 in testimony before the US Congress, after the bubble had already broken, the Fed Chairman asserted: “the impact on the broader economy and financial markets of the problems in the subprime market seems likely to be contained.”
But the standard macroeconomic models neither incorporated them nor provided an explanation for why such incentive structures would become prevalent—and these failures are failures of economic science.  

Of course, there is enormous diversity within economics, and even within the subfield of macroeconomics. Some macroeconomists warned of the looming bubble and the consequences that would follow upon its bursting. Economists like Minsky, who warned of the dangers of credit cycles (1992) or Kindleberger (1978), who described repeated patterns of manias, panics, and crashes, have come back into fashion (see also Reinhart and Rogoff 2009). Still, there was a single model, albeit with many variations, that came to dominate, sometimes referred to as the DSGE (dynamic, stochastic, general equilibrium) model. At the risk of considerable oversimplification, we refer to that model, and the standard policy prescriptions that were associated with it, as the Standard Model, or the Conventional Wisdom (CW). As in other areas, such simplifications help to clarify what is at issue.

Some advocates of that model recognize its limitations, arguing that it is, however, just the beginning of a research strategy that will, over time, bring in more and more of the relevant complexities of the world. Anything left out—agency problems, financial constraints, and so forth—will eventually be incorporated. I will argue, to the contrary, that that model is not a good starting point. Such Ptolemaic exercises in economics will be no more successful than they were in astronomy in dealing with the facts of the Copernican revolution.

Section 2 lays out in broad terms the failures of the Standard Model, while Section 3 develops the economics of deep downturns, arguing that the major disturbances giving rise to such downturns are endogenous, not exogenous; this crisis is not just an accident, the result of an unusually large epsilon, but is man-made. I explain why economic systems often amplify shocks and why recoveries are sometimes so slow. Section 4 argues that there have been systemic changes to the structure of the economy—changes that, within the Standard Model, should have led to enhanced stability, but which in fact made the economy more vulnerable to precisely the kind of crisis that has occurred. Section 5 contrasts the policy implications of our framework with that of the Standard Model.

---

3. There is a long list of flaws in the incentive structures, discussed and documented well before the crisis. See, for example, Stiglitz (1982a, 1987a, 2003, 2010a) and Nalebuff and Stiglitz (1983a, 1983b). These include: (i) incentive structures should have been based on relative performance—not, for example, stock market value which could increase due to an industry shock or to an increase in equity prices; (ii) incentive structures should have attempted to differentiate between increases in profitability due to increases in \( \alpha \) (hard to achieve) and to \( \beta \) (anyone can get higher average returns, simply by taking more risk). As designed, the incentive structures encouraged excessive risk taking and bad accounting. The compensation schemes were also not tax efficient. In practice, there was simply a weak relationship between pay and performance.

4. For a textbook treatment of both the basic classical and new Keynesian DSGE model, see Galí 2008. For a detailed analysis of the use of the New Keynesian model to evaluate monetary policy, see Woodford (2003), and Clarida et al. (1999), and Galí and Gertler (2007).
2. The Failure of Economic Science and Alternative Approaches

Any model is an idealization, an abstraction. The central challenge of macroeconomics is to identify the salient aspects of the economy that help us explain what it is that we want to explain. And that, of course, is where macro-economists begin to differ. What is it that they seek to explain? And to what use do they want to put the model? Because models can be used for different purposes, it makes little sense to strive for a single model. Yet, to a large extent, the single model upon which much of macroeconomics focuses is ill-suited for most of the purposes for which one might hope that such a model might be used. It was of limited usefulness either for short-run prediction, ex post interpretation, or the design of policies to prevent fluctuations, to minimize their scale, or to respond once they occurred. In this paper, I am especially concerned with deep downturns, such as the Great Recession or the Great Depression.

Economic Theory as Blinders. Models by their nature are like blinders. In leaving out certain things, they focus our attention on other things. They provide a frame through which we see the world. Psychologists have explained how we discount information that is contrary to our cognitive frame. The result is that there can be equilibrium fictions—given the information that individuals actually process, the world as they see it supports their beliefs. For those believing in perfect markets, even repeated crises are seen as rare events, accidents that don’t really need to be explained (see, for example, Greif and Tabellini (2010) and Hoff and Stiglitz (2010). Shiller (2008) talks about social contagion.)

The neoclassical investment function provides example of how theory can lead modeling in the wrong direction. (As is typically the case, the problem lies not with “theory” but with a specific theory.) For a long time, economists dismissed incorporating cash flow effects into investment functions, even though empirical studies (such as Kuh and Meyer 1957) suggested that they should be, because, it was said, economic theory said that they such effects shouldn’t exist. But, only a little later, economic theories that took into account capital market constraints arising from imperfect information explained why such effects should be important, at least at certain times. A vast subsequent literature established empirically the importance of these constraints (see Gilchrist and Himmelberg 1995).

Trade-offs in Modeling. Because any model is a simplification, an idealization, of reality, it is not a criticism to suggest that some aspect of reality has been left out. But it is a criticism if what is left out is essential to understanding the problem at hand, including the policy responses. If one is interested, for instance, in understanding unemployment, it makes little sense to begin with a model that assumes that the labor market clears.

In illuminating some questions, a model of two or three periods may have to suffice—not because we believe that the world only lasts for three periods, but because the complexities resulting from the infinite extension preclude incorporating more important complexities. There are trade-offs in modeling just as there are in economics.

For instance, it has become acceptable, even fashionable, to use particular parameterizations, for example, constant elasticity utility functions, often of the
Dixit–Stiglitz (1977) variety, and Cobb–Douglas production functions. In using them, we should be aware not only of their special nature, but that they have empirical predictions that can be (and typically are) refuted. For some purposes (such as the analysis of behavior towards risk), these utility functions provide a bad description, and one should use such models with extreme caution. When Dixit and I used the particular utility function that has become fashionable, we chose it because it provided the benchmark case where markets traded off optimal diversity and firm scale. The diversity/quantity tradeoff was, we thought, the fundamental tradeoff in the theory of monopolistic competition, and the partial equilibrium models that had been at the center of the theory of monopolistic competition until then simply could not even address this issue. We would never have thought to have concluded using that model that markets were on average or in general efficient. Rather, a better interpretation was that the market was almost surely inefficient; the direction of bias was a subject of some complexity, though our model provided a framework within which one could address that issue.

By the same token, if the distribution of income (say between labor and capital) matters, for example, for aggregate demand and therefore for employment and output, then using an aggregate Cobb–Douglas production function which, with competition, implies that the share of labor is fixed, is not going to be helpful. The large changes in the share of labor imply, of course, that such a model does not provide a good description of what has happened.

If economics is the science of scarcity, economic modeling is the art and science of selecting which among the many economic complexities to incorporate. The question, then, is have the Standard Models focused on what is of critical importance, e.g. for purposes of predicting the length and depth of the current downturn in employment or output or the design of the policy responses?

2.1. The Representative Agent Model

While modern macroeconomics has gone well beyond the representative agent model, that model has helped shape the direction of research. It is important to understand its major limitations, and to assess the extent to which more recent developments, for example in New Keynesian DSGE models, have failed to come to terms with these.

Methodological Missteps. The Standard Model takes as its methodological foundation that macroeconomic behavior has to be derivable from underlying microeconomic foundations. That proposition seems on the face of it uncontroversial. But it was important that macroeconomics be based on the right microeconomic assumptions, those consistent with actual behavior, taking into account information asymmetries and market imperfections. Yet, in carrying out that research agenda, particular microeconomic foundations (competitive equilibrium, rational expectations, and so forth) were employed, and, to make the analysis tractable, particular parameterizations, which in fact are inconsistent with microeconomic evidence, were used (see Greenwald and Stiglitz 1987).
The timing of the new classical revolution (which in turn led to the Standard Models currently in use) was unfortunate. The well-established micro-foundations of the standard competitive equilibrium model were just being undermined by advances in the economics of information and game theory, but it was to the “perfect markets” models that they turned to provide their micro-foundations. The problem is that with perfectly functioning markets we would not expect to see the kinds of fluctuations that we see—and that we seek to explain.

With information asymmetries, markets behave markedly differently than they do with perfect information: markets may not clear; there can be credit and equity rationing, or unemployment (for a survey, see Stiglitz 2002). Imperfect information leads to imperfect risk markets, and the two together alter the behavior of product, labor, and capital markets, and firms and other agents operating in those markets in fundamental ways that have macroeconomic implications.

Ironically, the standard paradigm always claimed more “virtue” than it deserved. For instance, in the absence of the representative agent assumption (all individuals are identical) virtually any aggregate function can be consistent with the standard competitive model (Sonnenschein 1972; Mantel 1974; Debreu 1974; Kirman 1992). And when it comes to a key piece of the macro-model—money and finance—the analysis is ad hoc and hard to justify in terms of reasonable first principles. Money, for instance, is not needed for most transactions; credit can be used. With fluctuations in the supply of credit being at the center of many economic fluctuations, a theory that has little to say about credit and its determinants is obviously of limited use. Ad hocery was introduced in other ways as well: the shocks to the economy (typically modeled as productivity shocks) were simply assumed exogenous. Even if disturbances to productivity were the major causes of economic fluctuations, surely they are related to investments in R&D, which should be modeled as endogenous. But more fundamentally, most of the important shocks to the economic system—including those leading up to the current crisis—are endogenous. The subprime mortgage crisis was man-made. To assume that it was exogenous is both wrong, and obviates one of the major objectives of economic and policy analysis—to prevent the occurrence of such shocks.

The standard paradigm claimed more virtue than it deserved in another respect: one advantage of the rational expectations hypothesis is that it helps immunize the models against the Lucas critique, which emphasized that behavior itself was endogenous to policy. But behavior is also sensitive to expectations about policy change. Typically, one looks at whether behavior is consistent with a given policy regime. But policies change—indeed, one of the points of macroeconomic analysis is to consider the consequences of policy changes; and one of the arguments for democracy is that citizens have the right to change governments, leading to policy changes. To varying degrees, individuals anticipate these policy changes. Observed behavior, thus, is not just a function of current policies but about beliefs about what future policies might look like. A full rational expectations model would have to embrace some kind of objective probabilities of those changes; conceptually, it is not clear from where these would come; practically, it is obvious that individuals differ in these judgments.
Moreover, the Standard Models treat the government as outside the system: it too is modeled as if it were exogenous. But in fact, political actors are agents, who may use policies to pursue their objectives. They are engaged in a stochastic game with market participants, and this too should have been formally modeled (see Korinek and Stiglitz 2008, 2009). Such stochastic games provide a framework in which we can formally model market agents’ expectations about policy changes. What they make clear is that behavior at time $t$ can be highly dependent not just on policies at time $t$, but on beliefs about policies which might exist at all later dates. A full rational expectations equilibrium ought to incorporate these; there is considerable ad hocery in deciding which aspects of the broader socio-political-economic system to embrace within the rational expectations equilibrium.

The prevailing methodology was, moreover, dominated by *as if* modeling: True, most individuals may not be able to solve complex intertemporal optimization problems of the kind that they are assumed to solve in the Standard Model, but they behave *as if* they do, and that is all that counts. Yet, it is *all* the predictions of the model that need to be tested. And many of the predictions of the model—such as those concerning the microeconomic behavior of the constituents—are inconsistent with the empirical evidence. Attention was centered on certain facts that seemed consistent with the theory, but those that were not were conveniently ignored.6 In the end, the macroeconomics is evidently *not* truly ground on microeconomics. For instance, the large variations in employment with little changes in real wages suggest a highly elastic labor supply (if one assumes that there is no unemployment and that individuals are therefore on their labor supply function), yet micro-studies suggest highly inelastic labor supply functions.7 Some of these inconsistencies arise from the use of parameterizations designed to make the models tractable, but they are not only implausible—there is

---

5. There is a large and burgeoning literature on macro-political economy. See, for example, Besley (2004) and Besley and Persson (2009). Many of the reduced-form econometric estimates, for example of what happens if the government engages in expansionary fiscal policy, are predicated on predictable policy responses elsewhere in the system, for instance, from monetary authorities. For a critique, see Section 5.

6. As Korinek (2010c) points out, “If DSGE models abstract from certain features of reality or, even more, if they need to employ fundamental parameter values that are at odds with empirical estimates at the micro level in order to replicate certain aggregate summary statistics of the economy, then the model is not actually capturing the true microeconomic incentives faced by economic agents, but is ‘bent’ to fit the data, as was the case with 1970s-style macroeconomic models.”

7. The problem has been recognized by some of those in the DSGE tradition. The high elasticity of labor motivated alternative models in which the extensive margin of employment is more central, for example Hansen (1985). Defenders of the Standard Paradigm might rightly point out that any piece of the model (here, that describing the labor market) could be replaced with a more reasonable specification. One might, for instance, incorporate search or efficiency wage considerations in the labor market. As I point out in what follows, no one can object to models that appropriately model the general equilibrium properties of dynamic economies facing endogenous and exogenous shocks, so at one level DSGE is unobjectionable. Part of the concern is with the particular simplifying assumptions used to make the models tractable. As Korinek (2010c) points out, tractability often biases results towards very special specifications in which markets are stable and efficient; in many cases, key questions of interest are, in effect, answered by assumption. As he also points out, tractability leads to a focus on the ergodic steady state (which, given the model structure, is usually unique). Many real-world processes are not ergodic—and under quite plausible specifications the equilibrium is not unique.
ample evidence that they generate behavior that is inconsistent with what is observed. For example, many models employ constant-elasticity separable utility functions, implying that all individuals buy the same portfolio of assets; richer individuals buy the same portfolio that poorer individuals do.

**Why the Representative Agent Models Had to Fail.** The major deficiency in the representative agent model is that there can be no (meaningful) information asymmetries (at least without the representative agent facing acute schizophrenia, which was inconsistent with the assumption of unparalleled rationality); no financial markets (who is lending to whom?); no scope accordingly for excess indebtedness (who owes money to whom?) or for deleveraging (who is reducing their indebtedness to whom?); no problem of debt restructuring; no meaningful capital structures (since the single individual is bearing all the risk, it is obvious that nothing can depend on whether finance is provided in the form of debt or equity); no role for bankruptcy; no agency problems; no externalities. Because there are no agency problems, there is no scope for problems of corporate governance. Because there can be no externalities, there is no role for government intervention to align social and private interests. Because there are no distributive issues, there is no scope for exploitation—for example by the banks of uninformed borrowers. Changes in wages and interest rates can have large distributive effects, and therefore large macroeconomic consequences; but not in the representative agent model: for instance, what the worker loses through lower wages, he gets back in his role as “owner” through higher profits. In short, the assumptions underlying the representative agent model bias the results: there is little scope for the kinds of market failures that require government action; and redistributive policies that might affect aggregate demand can’t in these models.

### 2.2. Limitations of the Representative Agent Model and its Descendants

Repeatedly in history, there have been booms and busts, bubbles that broke, of the kind that occurred in 2008. They involved breakdowns in financial markets. Booms typically were marked by excess leverage. Resolving the crises entailed deleveraging. The inability of the representative agent model to incorporate meaningful information asymmetries and financial constraints makes the model of particularly limited use in understanding such fluctuations. By the same token, the issues that are center stage in this crisis are of no moment. The major complaint about bank bailouts, that they redistribute money from taxpayers to bank bondholders and shareholders, is of little concern, because the representative agent gains as owner of the bank what he loses as a taxpayer; there would be no impediments to restructuring mortgages, because what the bank loses in the restructuring the homeowner-cum-bank owner gets back. Not even unemployment is of much concern: the representative agent may change the number of hours he works, but only by a small amount, and, with perfect capital markets, there is little effect on consumption, since the impact can be easily smoothed out over time (Lucas 1987). Variations in the demand for labor are of such concern because of how the burden is distributed—a relatively small percentage are unable to sell all the labor that they would like, and this can impose enormous hardship on them. Private
insurance markets do not spread this risk. In the Standard Models (with no information asymmetries) there is no explanation for the absence of these key markets—like so many other aspects of those models, the absence is just ad hoc. The models are immune from the Lucas critique only by assumption—that there is no impact of policy on risk sharing because there is no one with whom to share risks.

In the following sections, I want to elaborate on the major deficiencies—assumptions that were included that shouldn’t have been and simplifications that make the model of limited use in studying these fluctuations.

*Distribution Matters.* As we have noted, it is inequality in the distribution of work—the fact that some individuals are fully employed while many cannot get the work they seek—that we seek to understand. As the economy expands and contracts and the demand for labor increases and decreases, we want to know (and be able to predict) how those changes are reflected in hours worked per worker versus the number of workers. In downturns and recessions, the focus rightly is on aggregate demand, which can be affected in fundamental ways by the distribution of income. With individuals differing in their marginal propensities to consume, aggregate savings (consumption) rates depend not just on income, but on its distribution.

Many interpretations of the current crisis have emphasized the importance of distributional concerns. The growing inequality (which itself should be explained within the model) would have led to lower consumption but for the effects of loose monetary policy and lax regulations, which led to a housing bubble and a consumption boom. It was, in short, only growing debt that allowed consumption to be sustained. But with the breaking of the bubble, even if banks were fully functional, the level of indebtedness—and the levels of consumption—that prevailed before the crisis can’t be sustained.

Prior to the crisis, some analysts looked at the average equity of homeowners in their home. Even a marked decline in housing prices would, in these calculations, leave significantly positive average home equity. But again, distribution matters: what was of concern were the large numbers of homeowners who had sufficiently little equity that, say, a 20% or 30% decline in home prices would leave them underwater, and therefore at risk of foreclosure.

Distributional conflicts are at the center of the impasse in dealing with the mortgage crisis. As we note later, the critical question is, who bears the losses? Distributional impacts are at the center of many other policy choices. Lifecycle is central to behavior; yet models with infinitely lived individuals have no lifecycle. When interest rates are lowered to near zero, policymakers should worry about the plight of risk-averse elderly, who have much of their savings in short term T-bills. But even if one were to focus only on aggregate demand, these distributional effects can be of first-order importance: If the interest elasticity of investment is low, the increased investment may be lower than the decreased consumption of the elderly. Lowering interest rates may then actually weaken aggregate demand. When the impact of savings of those approaching retirement is considered, matters could be even worse: these individuals might increase their savings rate, to make up for the low return.
Price changes always have distributional effects (outside of the representative agent model), and it is only under highly special cases that the effects of the winners just offset those of the losers. Indeed, unexpected interest rate or price changes lead to unexpected gains in some firms’ net worth, matched by losses to others. But firm investment and supply is in general a concave (non-linear) function of net worth (taking into account financial constraints that arise endogenously from information imperfections), so that such changes can, in general, lower aggregate demand and supply. Indeed, there can accordingly be adverse effects in the short run both from increases and decreases in prices (evidenced, for instance, in the adverse effects observed when the price of oil increased in the 1970s, and decreased in the 1980s; see Greenwald and Stiglitz 1993).

Interestingly, if we formulated a representative agent model of the world, changes in exchange rate should make no difference: gains to some are just offset by losses to others. But almost all economists recognize that exchange rate changes do matter. The Standard Models assume, implicitly, that differences between countries matter, differences within countries don’t. But that is, of course, both ad hoc and wrong.

*Markets Are Not Fully Rational.* Several critical aspects of the behavior of the economy seem so patently inconsistent with any model of rationality that attempting to construct a model predicated on rationality that explains such behavior is almost doomed from the start. Most neoclassical investment models entail an analysis of the cost of capital, taking into account given tax structures. But with full tax deductibility of interest, if marginal investment is financed by debt, the corporation tax leaves the effective cost of capital unchanged. Most investment models use instead of a marginal cost of capital a variable more appropriately interpreted as an average cost, taking into account how investment on average is financed. But it is hard to reconcile overall corporate financial policy with any model of rationality: there are ways of distributing funds from the corporate sector to the household sector that entail the payment of lower taxes (the dividend paradox).8

In the run-up to the crisis, investors, consumers, banks, and regulators all exhibited behavior that is hard to reconcile with the hypothesis of rationality as it is incorporated in most Standard Models. Alan Greenspan’s mea culpa put the matter forcefully:

“[T]hose of us who have looked to the self-interest of lending institutions to protect shareholders [sic] equity, myself especially, are in a state of shocked disbelief.”9

---

8. See Stiglitz 1973. For a discussion of other tax paradoxes—and other aspects of firm behavior that are hard to reconcile with the Standard Models, see Stiglitz 1982a. Since then, there have been innumerable attempts to explain the paradox, none of which I find convincing. Most telling, as the appreciation of the point has grown, a smaller fraction of funds distributed from the corporate sector to the household sector have been in the form of dividends. The market seems to have “learned”. But the process has been slow, and the learning incomplete. Details of the tax code matter: the earlier observation that with debt finance, the effective marginal cost of capital is unchanged is true if the tax laws provided for Samuelsonian “true economic depreciation”. Since virtually all tax systems have depreciation allowances which are accelerated relative to true economic depreciation, at the margin, debt-financed investment is effectively encouraged.

9. See Committee on Oversight and Government Reform 2008. Greenspan also said: “I made a mistake in presuming that the self-interest of organizations, specifically banks and others, were such . . . that they were best capable of protecting their own shareholders and their equity in the firms.”
Some of the “deviant” behavior can be explained by the incentive structures of bank managers. They (the decision makers) might have been managing risk in a way that maximized their own welfare—even if it didn’t maximize the welfare of shareholders, let alone society as a whole. Indeed, given the incentive structures, we should have been surprised if banks had not undertaken excessive risk taking. Textbooks would have had to have been rewritten: it would have meant that, after all, incentives did not matter. Thus, I was surprised at Greenspan’s surprise that banks had not managed their risks better. Some of the deviant behavior can also be explained by distorted organizational incentives arising from too-big-to-fail financial institutions. But the Standard Models didn’t incorporate these incentive distortions—and therefore, like Greenspan, didn’t anticipate the problems to which they would give rise.

But the failures run deeper: standard theory argues that markets should be efficient in their choice of incentive structures. Modern microeconomics (of the kind not incorporated into modern macroeconomics) explains this failure. Macroeconomics based on modern microeconomic foundations would have gone beyond the simplistic models of firms and finance of the past. Admittedly, this would have been difficult (though this is the thrust of much of the alternative macro-developments described later in this paper). But it should be obvious that we can place little reliance on macroeconomic models based on flawed micro-foundations.

Regulators and investors should have recognized (i) the pervasiveness of the agency problems and the risks to which that exposed them and the economy; (ii) the peculiarities (and risks) associated with commonly employed incentive structures—including the incentives for non-transparency and the provision of distorted information by firms and banks; (iii) the risk associated with increased leverage, unmatched with any (social) benefits. Their failure to do so and to take appropriate actions is itself evidence of market irrationality.

While it is amply clear that the neoclassical assumptions underlying the standard model cannot explain widespread behavior, it is not always evident which assumption fails; for example, whose irrationality was pivotal. For instance, under the standard assumptions, the Modigliani–Miller theorem would hold; indeed the high risk and costs of bankruptcy would exert a strong force limiting leverage. The Modigliani–Miller theorem ceases to hold, of course, even with rational market participants, in the presence of important agency problems and other information asymmetries. But it is not clear

10. Besides the dividend paradox noted above, there are many other instances of market irrationality (see e.g. Shiller 2000; Stiglitz 1982a, 1982b). Economic theory during the last 30 years has “explained” the persistence of many of these seeming anomalies. For instance, take-over mechanisms (see, for instance, Stiglitz 1972a, 1982b, 1985a; Grossman and Hart 1980; Edlin and Stiglitz 1995) and evolutionary processes (see Stiglitz 1975, 2010a; Nelson and Winter 2002) often don’t work—at least in the naïve way that market advocates claim, and in the relevant time frame. Years earlier, Berle and Means (1932) had called attention to problems of corporate governance that arise with a separation of ownership and control.

11. Stiglitz (1969) explains how bankruptcy costs modify the standard Modigliani–Miller theorem that leverage has no costs or benefits. Nor can taxation explain the drive for leverage: a close examination of America’s overall tax system, taking into account corporate and individual taxation, including preferential treatment given to capital gains, suggests that if individuals were rational, the tax benefits are likely small in comparison to the costs of bankruptcy (ignoring for the moment the benefits associated with bailouts). See Stiglitz 1973. Signaling and screening models provide further explanations for limiting leverage.
that most market participants (including bank management) fully understood the risks associated with their high leverage—unless, of course, we assume that they were in fact counting on some form of bailout, a hidden or open transfer to them from the taxpayer. In retrospect, such expectations appear to be rational, and it was the regulators who ignored this that were irrational. By the same token, many of the banks *might* have been rational in exploiting uninformed borrowers; but it is hard to believe that many of those taking out some of the worst forms of subprime mortgages were rational. Many of those in the financial sector had an irrational optimism about the ability of poor borrowers to repay.\footnote{That this is so retrospectively is obvious; but Shiller’s work (2008) makes clear that it should have been obvious ex ante. The models used for forecasting default rates on securities *irrationally* ignored correlations and the chance of price declines—again, something that is clear retrospectively, but was argued on the basis of theory and historical experience well before the crisis (Stiglitz 1992c). Greenspan’s argument in favor of variable rate mortgages (see Chapter 5 of Stiglitz 2010a) suggests a deep lack of understanding of risk sharing in the market.} In short, even were we to pin blame for the crisis—and for the failure in our models—on irrationalities, there is a question of whose irrationality?

Recent years have seen the development of behavioral economics, and its application to macroeconomics (Akerlof 2002; for an interpretation of the crisis, see Fuster et al. 2010). There are many important and systematic aspects of behavior that simply can’t be reconciled with the standard utility-maximizing model, and increasingly, these have come to play a role even in policy (for example in the form in which taxes were reduced in the Obama stimulus package.)

The Limitations of Rational Expectations. The hypothesis of rational expectations which has played such an important role in modern macroeconomics is questionable in general\footnote{See, for instance, Shiller 2000. Grossman and Stiglitz (1980) explain why markets cannot be informationally efficient—a view that, in the aftermath of the crisis, has come to be widely accepted. Much of the observed behavior can only be explained on the hypothesis of differences in beliefs. See Stiglitz (1972a), Allen et al. (1993), and Scheinkman (forthcoming), and references therein.} and of little applicability in the current situation. There hasn’t been a crisis as deep as the current one for three-quarters of a century, so how can market participants form rational expectations about how modern economies respond to such a situation, unless they make the leap of faith that responses to a large crisis are similar to responses to smaller perturbations? How can a retired person, who has relied on interest payments from government bonds, form rational expectations about future interest rates when they have *never* been so low? There is no simple empirical evidence on the basis of which he can meaningfully extrapolate what will happen.\footnote{The failure of the rating agencies is related to this quandary: it was argued that newly invented products had fundamentally changed markets. If that were true, there would be no basis for relying on past data to predict future performance. Yet, irrationally, they did. They should have seen that the new products were *worse* than the old.}

The standard rational expectations models not only *assume* that they can do so, but that all market participants have the *same* (rational) expectations. Yet there is little reason to believe that they will formulate the same model—or that the models they formulate,
even on average, are correct; and it is differences in beliefs that drive much of what happens in the economy.15

The problem for the government is even more difficult. For it must base its policies on beliefs about how economic agents are behaving, who in turn must base their behavior on beliefs about how the government is acting. They all have remarkably solved instantaneously for the fixed point—in a context which has never previously occurred! And again, the rational expectations hypothesis is of limited relevance in assessing the consequences of policies that have never (or almost never) been tried before, at least in comparable circumstances.16 Describing and analyzing the full rational expectations equilibrium is complicated enough; to presume that in a rare event such as the current one that all economic and political participants have quickly gravitated to this equilibrium is beyond credence.

Moreover, even if expectations on average were rational, in the presence of financial constraints, market behavior might be markedly different from what would occur if there were a single individual with rational expectations. Consider quantitative easing. It may increase anxieties about future inflation, and if that happens, then long-term interest rates may rise. While an increase in inflationary expectations in excess of the increase in interest rates lowers real interest rates, it may not increase investment: Firms, sensitive to the demands on current cash flows that the higher (nominal) interest rates entail, may curtail investment.

Modern macroeconomics prides itself in combining theory with rigorous empirical work. But all policy analyses are predicated on the belief that data from earlier periods are relevant to the current experience. But whether that is the case is an article of faith—one which may make sense when “today” looks much like earlier periods. But the world today looks markedly different. The hard question is, what aspects of behavior carry over? Are empirical results describing firm behavior when excess capacity is low

15. The combined implications of rational expectations with common knowledge and rational behavior are even more peculiar: under standard assumptions, there would be no trade on the stock market. See Milgrom and Stokey (1982) and Stiglitz (1982b).

16. When there is a unique equilibrium, one could fantasize that somehow they all figured out the equilibrium instantaneously. When there are multiple equilibria (as there typically are), it is hard to envision how they know which equilibrium to coordinate on. Even when there is a single rational expectations equilibrium, there is little reason to believe that the market, on its own, would converge to the rational expectations equilibrium, at least in the time frame that is relevant for short-run macroeconomic analysis (see e.g. Bray 1978, 1981). More recent research has identified conditions under which such convergence holds using standard learning models (see e.g. Evans and Honkapohja 2001). Marce and Sargent (1988) note that when there is an adaptive game between a government and private sector, where each uses a least-square model, the Nash feedback equilibrium to which the economy converges is inefficient. But as those authors note, Bray and Kreps (1987) show elsewhere that “least squares learning schemes are irrational . . . [because] for example, they embody a Bayesian prior that is inconsistent with the law of motion” (Marcet and Sargent 1988, p. 171). There is a similar critique of the hypothesis of intertemporal rationality on the part of individuals. Individuals learn, from repeated experiments, about their preferences. Rationality, as used be economists, refers to the consistency of their choices, that is, where they arise from the maximization of a well-defined set of preferences (satisfying certain restrictions) subject to a budget constraint. But individuals do not have the opportunity to make intertemporal choices in an analogous manner. When they come to the end of their lives, they may regret having saved too much or too little, but there is little they can do with such learning. There is no way that we can test, at an individual level, the consistency of such lifetime choices.
still relevant when excess capacity is at record levels and expected to be for some time? Do empirical results concerning consumer behavior when indebtedness is low still apply when indebtedness is high? We might try to make inferences by looking at the behavior, in more normal times, of firms with high excess capacity or households with high indebtedness. But almost by definition, such firms and households are then “outliers”; we should be cautious in making inferences about the behavior of ordinary firms and ordinary households in these unusual circumstances from observing behavior of outliers in normal times.

*Market Clearing.* In forming expectations, for most workers, more important than future wages and prices is the risk of unemployment; but obviously, such expectations can play no role in a real business cycle model in which the labor market is assumed to clear. Today, for most households, a key question in the solution to their dynamic maximization problem is whether they will be able to refinance their mortgage and if so, at what terms—financial variables that are more relevant than the interest rate at which the government can borrow. For SMEs, the availability of finance is as or more important than the interest rate. In the Standard Model, these questions simply don’t arise, because markets are assumed to clear. Efficiency wage theory has, for instance, provided rigorous microfoundations explaining why labor markets may not clear (Shapiro and Stiglitz 1984; Rey and Stiglitz 1996).

One of the hardest analytic questions is trying to understand which results can be attributed to which assumption. For instance, the result in many variants of rational expectations models, that government policies were ineffective, was attributed to the assumption of rational expectations (with private agents offsetting government actions). But this was wrong. If goods and labor markets don’t clear, say because of wage and price rigidities, and individuals have rational expectations, then not only is government fiscal policy effective, but multipliers are greater than without rational expectations, as consumers respond to rationally expected increased incomes in future periods by increased consumption today (Neary and Stiglitz 1983).

*Institutions Matter.* The Standard Model assumed that institutions don’t matter; but institutional details are often of first-order importance. For instance, to understand mortgage default rates (and the difference between patterns in the United States and some other countries) one has to note that first mortgages in the United States are typically nonrecourse, while those in other countries are not. To understand the difficulties of restructuring (which would seem, in most instances, to be Pareto superior to current practices which often lead to costly foreclosure proceedings and incentives to trash the homes in the process) one has to understand the conflicts of interest that arise between the first and second mortgage holder and the service provider. A model whose structure ignores these issues will give too much credence to the ability of the markets to work everything out for the best, and provide little guidance to what government might or should do.

But even more fundamentally, the Standard Models left out both banks and the shadow banking system, central to the determination of the flow of credit, which in turn is central to the determination of aggregate demand.
2.3. Dynamic Stochastic General Equilibrium Models

Some of the problems that I have outlined have been recognized. Macroeconomists have gone well beyond the representative agent in the DSGE models—especially in the variant known as the New Keynesian DSGE models—that have become so popular. While some of these models have attempted to address some of these concerns, the Ptolemaic approach of attempting to refine a fundamentally flawed model is not, I think, the most promising approach for helping us to understand these issues. Introducing monopolistic competition and nominal rigidities means the market equilibrium is not, in general, Pareto efficient, monetary policy can have real effects, and, more broadly, government intervention can be welfare enhancing (see for example Galí 2008). Still, advocates of the various variants of DSGE have themselves emphasized the overall similarities.\footnote{This is a point of agreement among both the advocates of the New Keynesian DSGE models (Woodford 2009) and its RBC critics (Chari et al. 2009). Early versions of NK DSGE models did not even model unemployment. As Blanchard and Galí (2010) pointed out: “Standard versions of NK paradigm do not generate movements in unemployment, only voluntary movements in hours of work or employment... Paradoxically, this was viewed as one of the main weaknesses of the RBC mode, but was then exported to the NK model.”} As Chari et al. (2009) note, the various versions of the DSGE models even agree on the central policy: “optimal monetary policy... keep[s] inflation low and stable in order to avoid sectoral misallocations” (p. 246). But the social costs of these misallocations are $n^{th}$ order compared to those from disruptions in the financial sector. Indeed, they may be even small compared to those arising from changes in relative prices that result from differences in the short-run price determination processes across sectors (in some markets, prices are set by firms while in others, prices are set by, in effect, auction, Stiglitz 1996a)—effects which were ignored by assumption in virtually all of the DSGE models. Thus, most of the key criticisms leveled against the Representative Agent model are, for the most part, still valid in the DSGE models of whatever variant: Financial sectors are not well modeled, including the banking and shadow banking sectors and the links between monetary policy and credit. Levels of aggregation (key to policy analyses) are similar. Key assumptions, such as market clearing (no credit rationing), rationality, and rational expectations are retained.

Also, as I have noted earlier, many results are implicitly due to particular hard-to-defend (other than as matters of convenience) parameterizations, and many at the various variants of DSGE have continued to employ the same parameterizations. Earlier, I explained how Cobb–Douglas production functions rule out changes in factor distribution, which can be important in determining aggregate demand. But there are other objections: Even if there is unitary elasticity of substitution \textit{ex ante} (before the capital good are constructed), there is not \textit{ex post}.

Seemingly, the one thing that DSGE models seem to have in common is agents that maximize intertemporal utility—often with separable utility functions (though this too is presumably simply an assumption of convenience). The focus of dynamics is on intertemporal substitution effects, mediated through interest rates. There are good
grounds for questioning the significance of these effects for investment or consumption, at least in response to the kinds of variations in interest rates normally observed. Indeed, for long periods of time, real interest rates were approximately constant—making it hard to believe that they were an important channel by which policy affected behavior. It takes considerable massaging of the data to ensure that investment is sensitive to real interest rates and not nominal interest rates. Short-run fluctuations especially are as or more dominated by credit availability, changes in firm or bank equity, and government expenditure shocks.

But even if behavior is significantly affected by interest rates, there are large disparities between lending and borrowing rates (bank deposit rates are close to zero, consumer lending rates through credit cards close to 30%), and between T-bill rates and these rates. What matters is not the interest rate(s) at which the government can borrow, but those at which firms can borrow—if they can get access to funds; the spread between the two is an endogenous variable, that has to be explained. While monetary policy shocks seem to have real effects, and may be reflected in changes in nominal (and/or real) interest rates, that by itself does not necessarily mean that the interest rate effects dominate consumer or firm responses. Rather, the effects may be mediated through banks. Banks can, for instance, raise interest rates and reduce credit availability in response to tightening of credit by the Fed. The observed correlations between interest rates and investment are reduced-form relationships, not structural relationships.

A commitment to the Standard Model with its aggregate production function forces one to conclude that, with real interest rates negative at the current time, in the midst of the economic downturn of the Great Recession, the marginal productivity of capital has suddenly become negative (hard to reconcile with any of the standard aggregate production functions); or else to explain the discrepancy between the negative real interest rate and the seeming positive returns to capital through the imposition of some arbitrary “wedge” in the equilibrium condition. Neither approach is plausible or persuasive. (Note the marked difference between the Great Depression and the Great Recession. In the former, there was worry that a liquidity trap would prevent nominal interest rates from falling to zero, and even if they were very low, rapidly falling prices meant high real interest rates; in the Great Recession, nominal T-bill rates have been brought down close to zero, and prices have been rising.)

It is a positive development that certain imperfections are being introduced into the New Keynesian DSGE models, but how imperfections are introduced matters: not even the advocates of labor market frictions based on “search” believe it can explain current levels of cyclical unemployment. While agency-based theories of credit imperfections are a marked improvement over models with perfect capital markets, they suggest that markets would have shown more restraint in bank leverage and risk-taking than was observed.

The extension of DSGE models to make them more realistic, more consistent with macroeconomic data, has come with a price; as critics of New Keynesian DSGE models, like Chari et al. (2009) point out, the improved macro-performance is accompanied by an increased arbitrariness at least in certain specifications, including
alleged “structural shocks” (say wedges between the marginal rate of substitution and marginal rate of transformation), which may make them not immune from the Lucas critique—the size of the wedges might be affected by the policies themselves. Consider, for instance, the puzzle alluded to earlier, concerning labor supply. To make the models consistent with observed behavior, one can either postulate shocks to consumers’ preference for leisure or to workers’ bargaining power. In one view, markets are efficient, and it is left to psychoanalysis to explain why workers who have decided to enjoy more leisure seem so unhappy. In New Keynesian versions, it is tempting to attribute the outcomes to wage rigidity (increased union power), with obvious implications for the desirability of union busting. But it is more plausible that the wedge between the marginal rate of substitution and marginal rate of transformation changes for other reasons; it is endogenous, and needs to be explained. For instance, changes in the wedge over the cycle may be endogenously generated by (i) decentralized processes of wage and price adjustments (Solow and Stiglitz 1968); or (ii) by changes in the effective (shadow) real interest rate—taking into account financial constraints—in intertemporal models with monopolistic and monopsonistic competition with endogenous mark-ups (not the fixed mark-ups assumed in the Dixit–Stiglitz model).18 Cyclical movements in (shadow) real interest rates, in turn, are related inter alia, to changes in firm and bank equity positions. Thus, the 1991 and 2008 US downturns differ from other post-war downturns, in the large losses in the capital of many banks.

A central thesis of this paper is that the DSGE models have made the wrong trade-offs, focusing on some complexities which are of less importance than those that they ignore (and in some cases employing assumptions that are implausible). For instance, we noted earlier that the complexities of lifetime utility maximization typically forces modelers to employ parameterizations, the implications of which can be rejected. There are marked differences between models with infinitely lived individuals and overlapping generations models—with the latter arguably providing a better description of most households (Benassy 2007). Dynamics are important, but the dynamic effects that were included (arising, say, from intertemporal maximization of an infinitely lived individual) are less important than the dynamics that were excluded.

3. Towards an Economics of Deep Downturns

Part of the problem of modern macroeconomics is that it focused on explaining better the small and relatively unimportant fluctuations that occur “normally”, ignoring the large fluctuations that have episodically afflicted countries all over the world. The fact that a model may do slightly better than straightforward extrapolation in predicting growth rates at \( t + 1 \), say from the vantage point of \( t \), is of little moment: the welfare loss from a typical error in prediction in normal times is small in comparison to that

---

18. There is a large literature, dating back to the Phelps-Winter work on customer markets (1970) and encompassing labor turnover models where firms bear some turnover costs (Phelps 1970; Stiglitz 1972b). For more recent work, see for example Greenwald and Stiglitz (1995, 2003b). Some New Keynesian DSGE models generate endogenous cyclical changes to markups through specifying particular preferences (Ravn et al. 2006).
associated with the failure to anticipate a crisis, the effects of which can persist for years, during which the economy operates well below its potential.

It was as if we had developed a medical science that could treat individuals’ colds, but had nothing to say about serious illnesses. A doctor that said that was good enough, because most of the time individuals were either healthy or suffering from the sniffles, would not be taken seriously; but that was the position taken by much of mainstream economics. Indeed, in medicine, one learns much about the human body in normal times by studying pathology—what happens when things don’t work normally. So too, economists should be learning from the “pathology” of recessions and crises. These are the instances where market inefficiencies cannot be ignored; but these inefficiencies are the tip of the iceberg; beneath are pervasive but sometimes hard-to-detect market failures.

The point may be made in another way: Consider the difference between modern physics and modern macroeconomics. Black holes have played a central role in the development of modern physics. Of course, they “normally” don’t occur. If methodologies analogous to those that prevailed in economics had dominated in physics, black holes would have been dismissed as an irrelevant exception.

What is to be Explained? There are many macroeconomic variables that may be of interest for one reason or another. In the long run, growth matters. In the 1970s, it was understandable that inflation should be the object of concern. Today, as in the Great Depression, it should be deep downturns that should be the focus of attention. Of particular concern is unemployment. Persistent unemployment is, of course, a sign of an important market failure. It represents a massive waste of resources. There is ample evidence too that unemployment gives rise to a loss of well-being that is far in excess of the loss in income, with enormous social consequences (Fitoussi et al. 2010). Any model worth its salt has to be able to explain and predict movements in unemployment. In doing so, a model that assumes that labor markets clear will be of little help; nor will models that simply assume that unemployment arises from arbitrarily specified wage rigidities. Such an assumption pre-ordains the solution: get rid of the wage rigidity. Moreover, such an explanation is suspect: in the Great Depression, wages fell a great deal—they could hardly be called rigid. And in the Great Recession, the United States has been plagued by high unemployment (with one out six workers who would like to get a full-time job not being able to get one), even though it has claimed to have had one of the most flexible labor markets, and has the weakest unions, among the advanced industrial countries.

Credit. This paper focuses on deep downturns; and it is only by understanding credit—how the supply of credit is determined, why at times there can be excessive credit, while at other times the supply of credit can collapse—can we understand this and many of the other major fluctuations that have plagued capitalist economies over the past two hundred years. Even before this crisis, Greenwald and I had argued that understanding changes in the supply and demand for credit was at the heart of understanding economic fluctuations (2003a), and understanding how monetary policy (both convention instruments as well as regulatory instruments) affects the supply of credit should be at the heart of monetary theory. In normal times, money and credit
may be highly correlated, and so data on money supply may do as a surrogate for credit. But in times of crisis, such as the current one or the East Asia crisis, the link is broken. Moreover, secular changes in our financial system can change the linkages between money and credit, and thereby affect how monetary policy works.

This crisis is typically traced to the disappearance of credit after the bursting of the real estate bubble, and especially after Lehman Brothers’ collapse. As the crisis broke, with the credit supply rapidly contracting, Standard Models focusing on money, not credit, where banks and security markets were not well analyzed, provided little guidance on how government could restore the flow of credit. That—not interest rates—was the central issue. The resulting failure to resuscitate lending should thus not come as a surprise.

*Money and Credit.* Unless we understand the relationship between monetary policy (broadly defined, to include regulatory instruments) and credit, we won’t understand the role of monetary policy in responding to credit crises. Summarizing the financial sector into a money demand equation simply won’t do, nor will focusing just on interest rates. Even while some of the Central Bankers admit that their ability to resuscitate the economy is limited, they continue to believe that monetary policy can have some effects on real interest rates, which in turn will have some effects on real activity. Typically, they use *money demand equations*, implicitly based on a transactions demand for money. But today, money is needed for relatively few transactions—credit is all that is required. In the absence of financial market constraints (arising from imperfect information) financial policy (for example maturity structure of debt, quantitative easing) would matter little, if at all (see Stiglitz 1981, Stiglitz (1983, 1988; Greenwald and Stiglitz 2003a). One might, of course, justify the modeling of money demand on the grounds that it is a *good* reduced-form approximation—it works well (except when it doesn’t). Yet, such an ad hoc justification is totally out of the spirit of DSGE modeling, which prides itself on deriving all of the relevant behavioral relationships from more basic primitives (like utility functions and production functions).

Indeed, within the standard model, it would be hard to make sense of much of what has occurred in recent years. We have already noted the seeming irrationality of banks’ demand for leverage. *Collateral-based lending* has played an important role in generating the bubble, and in the bust that followed. But the very practice of requiring collateral only arises because of financial market constraints (and differences in judgments of the likelihood of occurrence of different events, leading to the importance of control). In a neoclassical model, there would be no reason for an individual to borrow, posting an asset as collateral, rather than simply selling the asset (reducing the amount he has to borrow.)

As the crisis has evolved, much has been made of the distinction between *insolvency* and *illiquidity*. But if everyone shared the same (rational expectation)

19. The credit availability doctrine played an important role in discussions of monetary policy within the Bank of England, and the credit channel episodically received attention in discussions in the United States (see e.g. Blinder and Stiglitz 1983). For an early survey of the credit channel of monetary policy see Bernanke and Gertler (1995). For a more relevant analysis of financial intermediate and macroeconomics, see Woodford (2010).
beliefs, an individual who is solvent, who, with probability one, could more than meet his debt obligations could get access to funds.  

In the analysis of deep downturns that follows, we focus on three questions: Why have they occurred? Why do disturbances get amplified? And why are recoveries so slow?

3.1. Bubbles: Explaining the Origins of Fluctuations

The literature during the past couple of decades that has gone under the rubric of business cycle began from the presumption that the origin of fluctuations was exogenous. There were technology shocks. A wave of Alzheimer’s disease passed through the economy in 1929, leading to a regression in technology, to which the economy adjusted.

It is hard to explain in a plausible manner this crisis—or most other major downturns—in terms of exogenous shocks to an economy which in the absence of such shocks would have grown smoothly. This crisis, like most major preceding ones, is man-made: the economic system itself created a bubble, the inevitable bursting of which led to the recession. In terms of general economic theory, there are a variety of conditions under which markets create their own noise, that is, when the equilibrium, sometimes the only equilibrium, entails random behavior (mixed strategies) on the part of market participants.

Macroeconomic structures in which such behavior naturally arises have not received corresponding attention, with one important exception: the proclivity of markets to create bubbles and credit cycles. As we noted earlier, economic historians have noted their repeated occurrence, suggesting that in most instances they were

---

20. Sometimes, a distinction is made between the value that could be achieved if the assets were held to maturity, and the much lower value achieved if the projects are liquidated prematurely (because of liquidity demands). But if everyone were convinced of the long-term value, almost surely there would be someone with resources that would reap the capital gain of the difference between liquidation and long-term value.

21. Before the crisis, advocates of standard New Keynesian DSGE models confidently advised monetary authorities not to worry about asset prices. See Bernanke and Gertler (2001). While the shock was largely endogenous, exogenous shocks may play some role; for instance, high food and energy prices (due to a variety of causes, but including weather and war “shocks”) may have contributed to the timing of the breaking of the bubble. The real business cycles are, themselves, misnomers, for there is no pattern of booms and busts, just a series of idiosyncratic shocks to which the economy responds efficiently. It is perhaps understandable why this new business cycle literature arose in opposition to the older literature, the multiplier accelerator models which gave rise to fluctuations of fixed periodicities. With rational expectations, both private agents and public authorities would undertake countervailing actions. Knowing that there would be excess capacity next period, firms would contract spending this period; and governments should undertake expansionary policies in a timely way to offset the expected contraction.

22. This is typically the case when there are non-convexities; and non-convexities are pervasive, whenever there are problems of information imperfections, R&D, learning, externalities, or bankruptcy costs. See, for instance, the discussion in Stiglitz (2002, 2010d) and references therein. The notion of mixed strategy equilibrium in the presence of non-convexities has been widely discussed. See for example Stiglitz 1975, 1985b, 1987b; Salop and Stiglitz 1977, 1982; Dasgupta and Maskin 1986a, 1986b; Mortenson 2010.
partly the result of irrational exuberance, often following the occurrence of a major innovation. Following such innovations, one cannot simply look to the past to predict the future—rational expectations models are inherently of limited relevance. In the current crisis, there was the irrational belief that innovations in financial markets had allowed risk to be much better managed.

But, regardless of the source of the underlying perturbation (whether exogenous or endogenous), capital market imperfections impede the ability to smooth out fluctuations—and may even amplify them. Individuals who believed that there was a housing bubble had only a limited ability to go short—not enough to “correct” the prices; and they had to have the wherewithal to maintain their position for an extended period of time.

Market irrationalities and financial market constraints interact. While one can construct models with rational expectations with bubbles (Abren and Brunnermeier 2003) (for example through rational herding, e.g. Banerjee 1992), probably more important are irrational herding and collateral-based lending. Individuals, seeing house prices rising, wanted to join the party before it was over (see also Allen, Morris, and Postlewaite 1993). Even if many rationally believed that it would end, they irrationally believed that they could win in the short run, and would outsmart the market, and not be caught in the downdraft. Collateral-based lending (combined with the difficulties of selling short) meant that as prices increased, those who were “long” on housing could borrow more, and take an even bigger position, pushing up prices even further, vindicating their “wisdom”. For discussions of credit-based bubbles and cycles, see for instance Kiyotaki and Moore (1997), Miller and Stiglitz (1999, forthcoming), Minsky (1992).

In fact, even without exogenous shocks, but with financial constraints and lagged responses, it is easy to construct models with fluctuations. Non-linear complex models give rise to interesting patterns of dynamics. Even without exogenous stochastic disturbances there may be oscillations with no regular periodicity, economic fluctuations, chaotic patterns, where the economy neither converges nor diverges, but perpetually oscillates. In fact, if investment is limited by profits (there are no capital markets) with plausible wage dynamics the economy is subject to oscillations (Akerlof and Stiglitz 1969). When wages are low, profits and investment are high, which leads to a larger demand for labor; the resulting rising wages then lead to reduced profits and investment, leading in turn to a lower demand for labor. Were such an economy, buffeted by shocks, it would not necessarily converge rapidly (or ever) or directly to the new equilibrium.

23. Some of these arise with difference equations that give rise to chaotic behavior. For applications to macroeconomics, see for example Christiano and Harrison (1999). For an application of agent-based approaches using the Greenwald–Stiglitz financial accelerator model, see Gallegati and Stiglitz (1992). But one doesn’t have to go to such complex models to generate patterns of oscillations even with rational expectations. Stiglitz (2008a) shows that there an infinite set of paths consistent with rational expectations in a life cycle model, most of which do not converge to a steady state.
3.2. Fast Declines and Amplification

The second major puzzle that has to be explained is why do economic declines sometimes happen so quickly, and why do what might seem to be small shocks get amplified? In the absence of war, state variables (capital stocks) change slowly. Why then can the state of the economy change so quickly? While the economic system sometimes amplifies shocks, standard theory argues that economic systems do just the opposite, through several mechanisms. Price adjustments mean that, say, a shock to the aggregate supply curve results in a smaller change in output than would occur in their absence. Firms create buffers, like inventories, to act as shock absorbers. Speculators do research, anticipating events, putting aside stocks as the probability increases of an adverse shock in which their value might rise. Understanding amplification—how small disturbances can give rise to large effects—should be one of the key objectives of macroeconomic research.

Consider the most recent crisis. Even the major misinvestments in the United States by the financial markets entailed a loss of, say, somewhere between a half trillion and two trillion dollars, a small fraction of the global capital stock. Indeed, the small size—and the belief in the markets’ ability to spread risk throughout the system—probably accounts for Bernanke’s confidence that the risks were contained. He was, of course, badly wrong, but this belief (shared by many other policymakers, and consistent with standard macro-models) helps explain the slowness with which they reacted to the impending crisis.

In the discussion that follows, I describe some aspects of amplification, especially those arising out of capital constraints, that have been uncovered by recent research and experience. While price rigidities have long been blamed for the economy failing to quickly return to full equilibrium after a shock, price disturbances interacting with financial constraints play an important role in amplification. Economies with more flexible prices may actually be more volatile.

Expectations. While the capital stock changes slowly, there can be large and sudden changes in expectations. Before the bubble broke, large numbers had seemingly believed that prices of housing would go up indefinitely (or at least for the foreseeable future); suddenly, the question became only how far down would they go. But to observe that expectations (unlike other state variables, like the capital stock) can undergo rapid or discontinuous changes just pushes the question back further: Why should expectations change so dramatically, without any big news? And especially with rational individuals forming Bayesian expectations? Consider, for instance, the puzzle of October, 1987: How could a quarter of the PDV of the capital stock disappear overnight?

Policy Changes. Another source of large and sudden changes is discrete government policy changes. Lehman Brothers’ bankruptcy can be thought of as an example—suddenly an implicit government guarantee was removed. Similarly, in the last global crisis, there were dramatic increases in interest rates, with large impacts. But like sudden changes in expectations, these discrete policy changes usually (though not always) are a result of sudden changes in state of economy. Though intended to dampen the effects of an exogenous shock, they sometimes have the opposite effect of amplification.
Structural Non-linearities. Modern mathematical modeling (for example, chaos theory) has shown that there can be large changes in the state of economy from small changes in state variables (the butterfly effect). For instance, when the economy is in some part of the state space, it converges to one equilibrium; in another part of the state space, to another. If the economy is near the boundary between one region and the other, a small perturbation can give rise to large effects. Kirman (2010) has noted that economic systems (viewed as complex adaptive processes) may exhibit major phase transitions. Non-linearities have played an important role in traditional business cycles, where a self-supporting expansion in a standard multiplier–accelerator model was suddenly brought to an end as the economy reached labor force constraints (see Goodwin 1951; Kaldor 1951). As these constraints began to bind, growth slowed, so investment slowed; but that meant that aggregate demand itself slowed, and the economy went into a downturn.

Financial Constraints. Financial constraints can give rise to both non-linearities and amplification. Individuals face credit constraints (including borrowing limits which arise endogenously with imperfect information or restraints on lending imposed by regulators); and such constraints can lead to the end of bubble. Regulators in the United States in effect “bent” the standard regulatory constraints, allowing higher loan-to-value and loan-to-income ratios. This allowed the bubble to continue longer than it otherwise would have (with the consequence that the downturn was larger than it otherwise would have been). But there are limits to such regulatory laxness, and eventually financial constraints bind. When that happens, housing price increases are basically limited by the rate of increase of incomes. With most Americans’ incomes stagnating, in the context of the recent crisis, which meant that prices had to stagnate. With prices stagnating, the cost of owning a home, this had been in effect negative (taking into account the expected capital gain) suddenly became very positive (all events that were not “rationally expected”). The demand for housing plummeted. Prices fell. The bubble had broken. Financial constraints meant that the bubble could not have persisted, even with regulatory forbearance.24

Financial Accelerator. The standard multiplier-accelerator model was predicated on a fixed capital–output ratio. With Solow’s 1956 paper, focus shifted to a neoclassical production function, and the multiplier–accelerator model grew out of fashion. Financial constraints give rise to the financial accelerator (derived from capital market imperfections related to information asymmetries), which operates in many ways like the old accelerator.25 Imperfect information explains why it is costly for firms to raise additional equity after a shock which adversely affects firm equity (Greenwald, Stiglitz, 2013; Bernanke and Gertler 1990; Bernanke, Gertler, and Gilchrist 1999). The microfoundations of the financial constraints differ in different models, with somewhat different empirical implications. I believe that those based on adverse selection (signaling), for example Maljuf and Myers (1984) or Greenwald, Stiglitz, and Weiss (1984), are more plausible than those derived from costly state verification (Townsend 1979). Equity constraints combined with bankruptcy costs lead to risk-averse firm behavior, which may differ markedly from the risk-neutral behavior typically assumed.

24. This discussion does not fully explain the sudden abrupt change: with individual heterogeneity, there can (or should) be some smoothing. Models of amplification discussed later explain the forces that counteract such smoothing.

25. See Greenwald and Stiglitz 1993; Bernanke and Gertler 1990; Bernanke, Gertler, and Gilchrist 1999. The microfoundations of the financial constraints differ in different models, with somewhat different empirical implications. I believe that those based on adverse selection (signaling), for example Maljuf and Myers (1984) or Greenwald, Stiglitz, and Weiss (1984), are more plausible than those derived from costly state verification (Townsend 1979). Equity constraints combined with bankruptcy costs lead to risk-averse firm behavior, which may differ markedly from the risk-neutral behavior typically assumed.
and Weiss 1984; Maljuf and Myers 1984) and why firms’ ability and willingness to borrow is limited by their equity and the value of their collateralizable assets. If firms can borrow a multiple of their equity, then a “shock” to equity can give rise to a change in aggregate demand (through investment) and supply (because of limited working capital) that is a multiple of the original perturbation.

**Pro-cyclical Inventory Movements.** Financial constraints lead to amplification through a number of other channels. For instance, inventories have traditionally been thought of as buffers, one of the mechanisms by which the economy absorbs shocks. But in practice, inventories often move pro-cyclically, contributing to economic volatility. When firms face adverse shocks to net worth, given equity and credit constraints, they seek to liquefy their assets, so that they are in a better position to absorb further adverse shocks. One way that they do this is to reduce inventories. In some cases, they may be forced to do so to get cash, when access to credit is restricted. The resulting negative investment weakens the economy further.

**Trend Reinforcement through Interest Rate Changes.** Battiston et al. (2010) have drawn attention to a variety of other trend reinforcement effects: a firm that has a negative equity shock also has to pay higher interest rates, and this means that the expected return on its equity capital going forward is lower.

**Price Changes.** Shocks lead directly and indirectly (through expectations) to price changes. Small shocks can lead to large price changes, which can have large effects on net worth or the value of collateralizable assets, and then through the channels described earlier, those changes are further amplified.\(^{26}\)

Moreover, redistributions of wealth, generated by price changes, can have first-order effects, increasing the magnitude of the effects already noted. (Of course, in a representative agent model, there are no macro-effects from such redistributions.) In fact, with large price changes, and especially with large gambles based on those prices, there can be fast redistributions (large balance sheet effects) with large real consequences, for example if there are large differences between firms, with some facing financial constraints and others not.

One of the insights of the economics of information is that even a small change in prices can have first-order effects on welfare (and behavior). A change in prices can affect the extent to which information constraints (self-selection or incentive-compatibility) bind. This is, of course, not true in the standard model, where market equilibrium is Pareto optimal, and small changes have small welfare effects, by the envelope theorem (see Greenwald and Stiglitz 1986; Akerlof and Yellen 1985; Stiglitz 2009).

**Lending.** Banks can be viewed as firms that specialize in lending (assessing creditworthiness, monitoring, and enforcement). Decreases in bank net worth can lead to contraction of their lending, with effects that are again a multiple of the

---

\(^{26}\) See Miller and Stiglitz 1999, forthcoming; Kiyotaki and Moore 1997, 2002; Korinek 2010a. Similar arguments arise in the context of international exchange rates. These were extensively discussed in the context of the East Asia crisis (see for example Furman and Stiglitz 1998; Bhattacharya and Stiglitz 2000; Stiglitz 1999b, 1999c), with more recent theoretical contributions from Korinek (2010b, 2010c) and Jeanne and Korinek (2010).
original adverse impact on net worth: the decrease in net worth decreases the resources available to lend and their willingness to borrow to get additional resources; moreover, if banks face capital adequacy constraints, the amount that they can lend is reduced by a multiple, unless they raise additional capital, which may be especially costly at such times (partially because the need to raise capital raises questions about the banks’ balance sheet).

**Bankruptcy Cascades.** The bankruptcy of one firm increases the probability of bankruptcy of its suppliers and creditors; this can give rise to a bankruptcy cascade, amplifying the extent of bankruptcy and the systemic costs of the shock that originally gave rise to the bankruptcy (Allen and Gale 2001; Greenwald and Stiglitz 2003a, Chapter 7).

**Contagion.** Bankruptcy cascades are an example of *contagion*, where a problem in one country (firm) spreads, like a disease, to others, with effects (if they are not contained) a multiple of the original disturbance. Contagion is thus one important source of amplification, and is discussed more fully in Section 5.1.

**New Uncertainties.** Amplification can also arise through new *uncertainties* posed by a shock, especially as it evolves through the economic system: Large changes in prices lead to large increases in uncertainties about the net worth of different market participants and hence about their ability to fulfill contracts. More flexible prices thus can serve to amplify the effect of a shock. Changes in risk perceptions (not just means) matter, given the cost of bankruptcy and the risk aversion of firms.

There is another, related reason for amplification: A crisis such as the current one showed that prevailing beliefs might not be correct. Those beliefs had led to a complacency that risk could be, and had been, effectively diversified, so that the economy would be able to handle shocks of any size. The crisis quickly eroded beliefs both in the underlying theories and in the officials responsible for economic management. Beliefs about the possible depth and duration of the crisis accordingly could, and did, change dramatically.

**Control Changes.** There is one more mechanism through which a small perturbation can lead to large effects, and that is through a sudden change in control. Who exercises control matters (unlike standard neoclassical model, where managers simply maximize the value of the firm). The result is that there can be discrete changes in behavior with changes in control, and with bankruptcy and redistributions, there

---

27. This again was evident in the crisis (and in the East Asia crisis before it). Not even well-informed banks knew for sure the balance sheet positions of other banks with which they interacted, realizing that the fall in housing prices and the associated price of mortgage-backed securities meant that there had been large changes in net worth and the risk of default. In both cases, lack of transparency contributed to these problems. American banks had, to deceive both investors and regulators, engaged in off-balance sheet transactions. But they were so successful that they may have even deceived themselves. The difference between Lehman Brothers’ supposed balance sheet before and after filing for bankruptcy bears partial testimony to the magnitude of the uncertainty. The bankruptcy costs—tallying now more than a billion dollars—shows why these should not be ignored. It should be noted that with rational expectations, more transparency can, however, contribute to greater volatility, as the market responds more to changes in circumstances. See Furman and Stiglitz 1998.
can be quick changes in control. In this crisis, there were many changes in control, but whether these were significant enough to lead to macroeconomic changes is not clear.

There were strong feedbacks among these elements: For instance, changes in beliefs led to increased uncertainties directly and indirectly, as they fed into lower prices for real estate, increasing balance sheet uncertainties. These new uncertainties in turn contributed to the “bite” of financial constraints (see also Korinek 2010b, 2011).

3.3. Persistence: Why is Recovery so Slow?

There are large losses associated with misallocation of capital before a bubble breaks. But most of the losses occur after a bubble breaks, in the persistent gap between actual and potential output. Standard theory predicts a relatively quick recovery, as the economy adjusts to the new reality. There is a new equilibrium associated with new state variables (treating expectations as a state variable). The breaking of the bubble does not itself destroy any physical or human capital, so in principle, with efficient markets, these resources should be fully used. Standard theory (and real business cycle theory) says that, given any set of state variables (including expectations of the future, and debts) there exists a set of wages and prices such that all markets clear—including the labor market. Given the elimination of the distortion caused by the “bubble” prices, real output should actually be increased. Sometimes there is a quick recovery (as in the hoped for V-shaped recovery). But sometimes (as in the Great Depression and in this recession) the recovery is very slow. The effects of adverse shocks persist. This section focuses on the third puzzle: Why is recovery so slow.

Slow Price Adjustments. There are two strands of explanations. One, the more traditional, focuses on wage and price rigidities. Many of the standard explanations of these rigidities are not fully persuasive: Staggered wage setting models don’t fully explain why it is that those wages which are fully adjustable don’t fully absorb the shock.28

There are more plausible bases of slow wage and price adjustments, based on firm risk aversion, which itself can be derived, for example from agency theory or financial constraints and the fact that firms know more about their current position (wages, prices) and what might happen were they to alter wages and prices by a small amount than they know about the consequences of large changes.29

28. Elaborations on these models can easily do this, but these elaborations change the model and its policy implications in fundamental ways. For instance, efficiency wage effects may arise from inordinate disparities between wages paid to those hired at time \( t \) and those hired at \( t + 1 \). Moreover, one has to have a compelling reason for why such staggered wage setting persists, when presumably coordinated wage setting would represent such a large improvement to economic welfare.

29. Greenwald and Stiglitz (1989, 1990). The name menu cost theory appropriately trivializes the notion that it is the cost of adjusting prices that is the source of the problem, since the costs of price adjustments are of an order of magnitude smaller than the costs of adjustment of quantities; with a shift in, say, a demand curve, either prices or quantities have to adjust, and given the relative costs, the adjustment should be in prices. Models where sluggish responses are based on “rational inattention” (Sims 2003; Mankiw and Reis 2002, 2010) are more persuasive; but especially large businesses are in fact constantly monitoring both macroeconomic and sectoral conditions.
Slow Recovery of Balance Sheets. In the previous section, I explained how shocks to firm and bank equity can have large macroeconomic consequences; but rebuilding balance sheets takes time. As we have noted, firms’ ability to borrow is limited by their equity; because of capital market imperfections raising equity is very expensive, so that most firms rely on retained earnings. Restoring the lost equity is thus a slow process.

Market Instability: Economic Intuitions. Even if wages and prices were more flexible (or conceivably, perfectly flexible) it would not necessarily imply a quick restoration of the economy to full employment. Especially with information imperfections and incomplete markets, market adjustments to a perturbation from equilibrium may be (locally) destabilizing.\footnote{I do not have space to discuss the underlying mathematics. There has even been a shift in the past 40 years in what is meant by a stable system. The representative agent model is viewed as stable, since the individual, with rational expectations extending infinitely far into the future converges to the long run equilibrium along a saddle-point. But there is a sense in which this is a very fragile equilibrium. If the individual misestimates, he can move along a path satisfying all the short run equilibrium conditions for a very long time, before he realizes that he is not on that single saddle-point trajectory. He then will have to make a large correction. Indeed, Hahn (1966), looking at essentially the same set of equations describing the economy’s dynamics, concluded that the system was unstable (see also Shell and Stiglitz 1967). In standard dynamics, a system exhibits long-run stability if, for all initial conditions (within a range), the system converges smoothly to the steady state. The “trick” here is the assumption that the market miraculously sets the prices of all assets instantaneously so that the economy is on the saddle-point trajectory. Richer models, with heterogeneity and constraints, can give rise to systems which do not display smooth convergence (see e.g. Christiano and Harrison 1999).} Standard economic theory has little to say about out-of-equilibrium adjustment. Walrasian tantamount theory is of little relevance to the real world in which adjustments occur in real time, and there are real consequences (for example capital losses/gains) that result in making transactions at the wrong time.

Traditional Keynesian economics emphasized one aspect of the destabilizing dynamics: cuts in wages in response to unemployment lowered aggregate demand, thus increasing unemployment. Even the supply-side effect (that lower wages lead to increased demand for workers at any given level of employment) may be obviated by adjustments elsewhere in the system. What matters is real wages, not nominal wages, and lower wages shift the supply curve, increasing downward pressure on prices. If prices fall in tandem with wages, real wages will remain little changed. These real wage rigidities are not caused by unions, but follow from natural assumptions concerning the decentralized adjustments to disequilibrium in goods and labor markets (Solow and Stiglitz 1968). Inventory reductions, to “liquefy” balance sheets (described earlier) put further downward pressure on prices.

Fischerian debt-deflation dynamics gives rise to an even stronger set of destabilizing effects (Fisher 1933; Greenwald and Stiglitz 1993, 2003a). Because debt contracts are not typically indexed, what matters is not actual deflation, but simply inflation that is lower than expected: any such outcome adversely affects the firm’s balance sheet, which, given capital market imperfections, may actually lower aggregate demand—increasing the gap between supply and demand. Firms’ demand for investment and their ability to raise finance for investment are impaired. There
are further feedbacks through the financial system. Higher rates of default weaken bank capital, leading to less credit availability and higher lending rates (examples of the phenomenon of trend reinforcement noted earlier, Battiston et al. 2010). There is even the possibility of a bankruptcy cascade. Lower housing prices may lead to more foreclosures, putting further downward pressure on prices. In short, lower prices may lead to further contractionary pressure.

Asset price adjustments too can give rise to an adverse dynamic through other channels: With a reduction in asset prices, in what might be viewed as a salutary correction to a bubble, the value of collateral is reduced, but with less collateral, especially small- and medium-sized firms cannot borrow, either for working capital or investment, so there are further contractionary effects.

Adverse supply effects (for example from asset price adjustments) reduce the demand for labor and put downward pressure on wages. As workers’ incomes fall, so too does their demand for consumption goods, an effect which is not likely to be fully offset by capitalists’ higher consumption as a result of higher profits. Again, aggregate demand is reduced.

Other factors discussed earlier as part of the amplification process too come into play in slowing the process of recovery. For instance, large price changes give rise to increased uncertainty, for example about balance sheets. Increased uncertainty impedes firms’ willingness to make investments and banks willingness to lend. Moreover, because some prices (for example, equities, commodities) are determined in auction markets, and others (much of manufacturing) on the basis of posted prices, adjustments to shocks may, in the short run, lead to relative prices that are badly out of line (Stiglitz 1999a), with large changes in balance sheets and high levels of uncertainty—often in a destabilizing manner. In the absence of distributional effects and financial constraints, such price changes might be of second-order importance. Sellers lose, buyers gain, but in a representative agent model, since the seller and buyer are the same person, nothing happens. But more generally, distribution matters: if prices of agricultural goods fall rapidly, farmers reduce their spending by more than urban workers and rentiers increase their spending. Aggregate demand thus falls. More generally, with both supply and demand concave functions of firm equity, there are real, and potentially large, consequences to such redistributions.

Expectations, too, can have a short-run destabilizing effect. If, as now, interest rates are close to zero, were prices expected to fall, real interest rates would rise, and (at least in the standard theory) the increase in the real interest rate would decrease aggregate demand, reinforcing the shortfall between supply and demand. What matters, of course, is the expected real interest rate, which depends on the expected change in prices. And here again, there can be (locally) destabilizing effects of price adjustments: a fall in the price today can lead to expectations of further decreases. Consumers, rather than consuming more, may postpone purchases, waiting until prices are still lower.

So far, I have focused my discussion on destabilizing movements in prices and wages. But market-based exchange rate adjustments can also be very destabilizing. In the East Asia crisis, worries about the future of the economies led to a decrease in the exchange rate; but these countries had foreign exchange-denominated liabilities, so a
fall in the value of their exchange rate worsened their economic position, facing many firms (and banks) with the risk of default; and these adverse effects more than offset the benefits from increased competitiveness of exports.\footnote{Interestingly, the IMF, normally a believer in market processes, justified its intervention in the market on the grounds that markets’ adjustments in exchange rates were, in fact, destabilizing, and could give rise to contagion. Their interventions were designed to stabilize exchange rates, to slow down adjustment, to prevent “overheating.”}

Some of the instabilities that I have discussed are a natural part of decentralized market adjustment; but some (and others I shall note shortly) are the result of policies. In the East Asia crisis, the IMF and the US Treasury demanded that interest rates be raised to high levels to slow the fall in exchange rates. But the high interest rates pushed large fractions of the firms in the affected countries into financial difficulties (in the worst cases, into bankruptcy); the resulting economic disruption exacerbated the economic decline and worsened the exchange rate, with the follow-on effects described earlier.

One of the objectives of our models should be to provide insight into the design of stabilizing policies. For instance, rigid capital adequacy standards for banks can be destabilizing; countercyclical capital adequacy standards (or what is sometimes call macro-prudential regulation) can be stabilizing (see for example Griffith-Jones et al. 2010).

The Fight over Who Bears Losses after a Bubble Breaks—and over Control of Assets. After a bubble breaks, it is typical that the value of claims on assets (what debtors owe creditors) exceeds the value of assets. Someone has to bear the losses; there is a fight over who bears the losses. But this fight—and resulting ambiguity in long-term ownership—contributes to the slow recovery and the magnitude of losses. When there are no clear owners of an asset, assets are not well maintained. Indeed, there is a risk of asset stripping. One of the costs of home foreclosure is the lack of care (or worse) that occurs in the process. Recognizing this and the high transactions cost associated with bankruptcy, one might have thought that there would be ex post renegotiation: the bank (or owner of the mortgage) will only be able to recoup the current value of the house, so both the lender and the borrower would seem to be better off with a write-down. But one of the reasons that the market is in paralysis is that there are typically two (or more) mortgages. The second mortgage holder, who in the event of foreclosure will receive nothing, has nothing to lose by refusing to renegotiate, unless the holder of the first mortgage agrees to significant concessions. With those responsible for renegotiation, the service providers, being owned by the holders of the second mortgage, there is a clear conflict of interest; and anticipating this, some securitizations limited the scope for renegotiation.

A standard result in the theory of bargaining with asymmetric information is that the renegotiation may entail large inefficiencies, such as strikes in labor disputes; both workers and employers are worse off with a strike. Both the borrower and the lender are worse off with a foreclosure. Still, this is the market equilibrium (Farrell 1987).
Economic policy can, of course, affect the resolution of the problem of excess claims. In the case of a systemic crisis, given the large number of separate debt contracts, restructuring through ordinary bankruptcy processes is difficult, which is why a Super-Chapter 11, an expedited way of handling such workouts, is needed (Stiglitz 2000; Miller and Stiglitz 1999). In some cases, it may be necessary to force restructuring. In the current crisis, there is a need for debt-equity swaps for homeowners (banks would get, say, 80% of the capital gain on the sale of the house, homeowners would pay rent on the basis of the appraised value of the home).³²

To avoid the large distributive consequences associated with open restructuring, many countries have chosen two alternatives. One is to exploit the fact that the claims of the creditors are denominated in nominal terms. Inflation reduces the real value of those claims, enabling the debtors to repay. This is a debtor-friendly approach, which not surprisingly is typically strongly opposed by creditors. The other alternative (the one the United States is choosing) is to muddle through, with non-transparent accounting allowing banks to postpone recognizing losses, and therefore postpone recapitalization. In the meanwhile, banks are being recapitalized through a hidden subsidy, the large spread between the lending rate and the rate at which they have access to funds. But unfortunately, this inefficient approach to restructuring contributes to the slow recovery.

4. How Economic Progress Has Made Our Economy More Vulnerable

In one interpretation of the current crisis, it represents just a large realization of a negative exogenous (technology) shock. It was just a once-in-a-100-year flood. Most observers (see Financial Crisis Inquiry Commission 2011) believe, however, that it was not inevitable, and that the bubble was endogenous, and could have been avoided—or at least its magnitude and consequences lessened. One of the objectives of policy is to reduce the frequency and size of adverse shocks. I believe that changes in our economy—many associated with government policy—have in fact made our economy more vulnerable. In this section, I propose four hypotheses which, with further elaborations, provide, I believe, considerable insights into what has happened, and why the downturn in the West has been so severe, with such a slow recovery. The four hypotheses focus, respectively, on risk, information, credit, and structural change. Each lies outside the boundaries of analysis of the Standard Models.

The picture of the crisis that they provide is different from that of the real business cycle, where the economy efficiently absorbs and adapts to exogenous shocks. This crisis, I believe, is more akin to the credit cycles that have marked capitalism for the last 200 years, but there are some distinctive aspects of this credit cycle. They are likely to result in this downturn being more prolonged than a “normal” credit cycle,

³² Thus, there would be little concern that speculators would take advantage of such a scheme. With restructuring, banks would, of course, be forced to recognize losses; as it is, current management would prefer greater losses in the future rather than recognizing smaller losses today. Capital adequacy requirements combined with lax rules on writing down mortgages provide an incentive not to restructure.
unless appropriate palliative measures are taken. Moreover certain policies—based on misunderstandings of the functioning of the economy—have contributed to creating the crisis; reforms in these policies are desirable.

4.1. Risk

Our first hypothesis is that there have been large (and often adverse) changes in the economy’s risk properties, in spite of supposed improvements in markets.

Central to macroeconomics should be an analysis of the economy’s risk properties, which includes (i) its exposure to risk, (ii) how it amplifies or dampens shocks, and (iii) how individuals and firms are affected by risk. A corollary of the Greenwald–Stiglitz theorem is that privately profitable risk transactions and innovations may be socially undesirable. There is thus a potential rationale for government intervention to regulate and manage economic risk and risk sharing/transferring arrangements.

There are several reasons why, in recent years, in some key respects, the risk properties of the system may have changed for the worse: Ideas, interests, and innovations led to the belief that risk could be handled better, so that more risk could be assumed. But the improvements were less than had been widely thought, with the result that the system actually became more fragile, and especially more sensitive to large and correlated shocks. Indeed, this was perhaps the biggest flaw: there were changes in economic structure that enhanced its performance in handling small risks, but worsened its performance in handling large risks (Haldane 2009; Haldane and May 2010; Stiglitz 2010c, 2010d). Regulations were stripped away and regulators who didn’t believe in regulations were appointed, while financial innovations helped markets circumvent the regulations that remained. Other financial market “innovations”—floating rate mortgages—shifted risk to those least able to bear it, homeowners; while still others (securitization) contributed to information asymmetries (moral hazard), leading to lower-quality mortgages, with more systemic risk.33 The movement from fixed to floating exchange rate systems too has arguably forced businesses to bear more risk associated with exchange rate volatility. In these and other areas, financial markets have not always introduced products that might have helped individuals better bear the risks that they face. In some cases they have actually resisted the introduction of such products (inflation- or GDP-indexed bonds), behavior partly explicable in terms of the difference in their interests (maximizing transaction costs) and social interests. See Stiglitz (2010a) for other examples and alternative explanations of such behavior.

Beliefs and interests combined to lead to policies which exacerbated risk. The repeal of Glass–Steagall led to an enormous increase in banking concentration—too-big-to-fail banks had an incentive to engage in excessive risk taking. The ballooning

33. Some countries have actually banned such mortgages. Some commentators think that Greenspan’s advocacy of such mortgages was not just an instance of poor analysis but a deliberate attempt to transfer large amounts of wealth to the financial sector, given his knowledge that interest rates would be increasing (Taibbi 2010). Other “advances” in financial markets may have contributed to instability (Caccioli et al. 2009), evidenced most recently by the flash crash of 6 May, 2010, associated with flash trading.
derivatives markets were left unregulated. Especially in developing countries, capital and financial market liberalization exposed countries to more risks. Capital market liberalization may increase consumption volatility and lower expected utility.34

In many countries, automatic stabilizers have been weakened, and automatic destabilizers strengthened. The change from defined-benefit to defined-contribution pension systems—motivated in part by mistaken beliefs about the best way to manage risk, without any cognizance of the systemic consequences—left individuals to bear the brunt of volatility in asset prices, with greater sensitivity in aggregate expenditures to these asset price changes than when firms had to bear the risk.35 Widely used managerial compensation schemes encouraged excessive risk taking and short-sighted behavior, with macroeconomic consequences.

Repeated bailouts of the banking system also had adverse consequences (see, for example, Soros 2008). Because those in the advanced industrial countries had, by and large, been spared, some made the wrong inference—that markets worked well on their own, contributing to the view that deregulation was desirable. But they hadn’t really worked well on their own. Moreover, the repeated bailouts gave rise to moral hazard, especially among the too-big-to-fail banks: they had learned that regardless of how reckless they were in their lending, they would be rescued by the IMF and G-7, with the taxpayers picking up the tab.

But perhaps the central problem with the CW was that it paid insufficient attention to the architecture of risk, to the general equilibrium properties of the system. The theory was that diversification would lead to lower risk, and a more stable economy. It didn’t happen, raising the question, where did “theory” go wrong? Part of the answer lies in themes that we have repeatedly stressed: corporate governance and irrationality. Bank managers had perverse incentives, which led them to retain much of the risk, in non-transparent ways. They could thereby record high profits from transactions, reaping large bonuses. Moreover, they (and others in the financial market, such as the rating agencies) systemically and irrationally underestimated the risks.

But there is another part of the answer, reflected in the schizophrenic approach of the conventional approaches to policies toward financial integration, which emphasized the benefits of integration before a crisis, but afterwards, focused on the risks of contagion—which were increased by integration. Because the CW assumed that crises never occurred, in advocating financial integration, there was never a discussion of

34. These results have been shown both theoretically and empirically. In a lifecycle model, for instance, without capital market liberalization, a productivity shock at time $t$ leads to higher wages not just at time $t$, but also, as a result of increased investment at home, higher wages in subsequent periods. With liberalization, the benefits of the productivity shock are not shared across generations (see Stiglitz 2008b). For a more general discussion of the adverse effects of capital market liberalization see Ocampo and Stiglitz (2008) and Stiglitz et al. (2006). In ongoing research with Hamid Rashid, we have shown that financial market liberalization similarly may expose countries to more volatility, as a shock in the home country gets transmitted through the banking system to the host country.

35. There can also be procyclical labor supply responses, as in the current crisis: older individuals, with diminished retirement accounts, have postponed retirement, or re-entered the labor force. The list of risk increasing policy changes is long; for example, rigid enforcement of capital adequacy standards can act as an automatic destabilizer.
the adverse effects that might arise after a crisis occurred. What was so remarkable about this approach was that it was pursued in the face of repeated crises—more than a hundred during the past 30 years. Not only did the models fail to explain the increasing frequency of crises and simultaneously consider the benefits and risks of integration, they made mathematical assumptions in which spreading risk necessarily increased expected utility, ruling out risks of contagion.

It should have been obvious, however, that with non-convexities (for example associated with bankruptcy, R&D) financial integration can lead to lower economic performance. Optimal electric grids recognize that there are advantages of “sharing” capacity over a larger area (the total generating capacity required to achieve a given probability of a brownout or blackout is reduced), but that a fully integrated grid also increases the risk of system wide failure. To reduce such systemic risk, circuit breakers are required. Yet, the CW was adamant in opposition to the analogous concept, capital controls.36

Recent research attempting to reflect both the advantages and disadvantages of integration has upset the conventional wisdom. More generally, full integration is not desirable, and under some circumstances (for example, in the absence of capital controls) no integration may be better than full integration (Stiglitz 2010c, 2010d). Analogous issues arise concerning the architecture of risk sharing within a country, say among financial institutions. Risk-sharing contracts among the banks that were supposed to make the system more stable had just the opposite effect. Interconnectivity can help absorb small shocks but exacerbate large shocks, can be beneficial in good times but detrimental in bad times.

Economic systems can differ not only in the extent of risk sharing, but also in the pattern. Hub systems may be more vulnerable to systemic risk associated with certain types of shocks. Many financial systems have such concentrated nodes. In this perspective, the real problem in contagion is not those countries suffering from crisis (dealing with that is akin to symptomatic relief) but the hubs in the advanced industrial country. More generally, poorly designed structures (including those without circuit breakers) can increase risk of bankruptcy cascades.37

Earlier, I commented on how the Standard Model failed to take into account the incentives for excessive risk taking (arising both out of individual and organizational incentives, including those associated with too-big-to-fail banks). But systemic risk can arise as well from correlated behavior of many financial institutions, each of which is not too big to fail. And there are organizational and individual incentives to engage in such correlated behavior (made so evident by Charles Prince’s remark to the Financial Times in July 2007, about continuing to dance as long as the music plays—as long as

36. There are many other instances where CW adherents took a schizophrenic approach. A standard prescription for handling bank restructuring was to separate out good loans from bad (unmixing), which makes sense only because of non-convexities.

37. Among the contributions to this growing literature, besides those cited earlier, are Greenwald and Stiglitz 2003a; Allen and Gale 2001; Battiston et al. 2007, 2010; Delli Gatti et al. 2006, 2009; De Masi et al. (forthcoming); Gai and Kapadia 2010a, 2010b; Gallegati et al. 2008; Haldane 2009; Haldane and May 2010 and the references cited in these papers.
there is still liquidity; Nakamoto and Wighton 2007). Anyone not going along with the
general set of beliefs would have been punished (Nalebuff and Stiglitz 1983a).

The risk properties of the economic system can be affected by policy frameworks—and
unfortunately, too often, policy makers fail to take this into account. I have already
referred to several examples. Others include: The US bankruptcy law of 2006 which
strengthened creditor rights, provided lenders with fewer incentives to engage in due
diligence in ascertaining credit-worthiness. More competitive banking system lowers
franchise value, and accordingly, may lead to excessive risk taking (Hellman, Murdock,
and Stiglitz 2000).

4.2. Information

Good information is, of course, central to a well-performing economy. Good
information is also critical for managing risks—taking, for instance, appropriate actions
in a timely way to prevent a crisis. Earlier, I explained how information imperfections
and asymmetries are central to understanding economic fluctuations; they explain credit
and equity rationing, which in turn is key to understanding the financial accelerator
and the persistence of the effects of shocks; to understanding why banks play a central
role in our economy and why a loss of bank capital (and bank bankruptcy) can have
large and persistent effects.

Our second hypothesis for understanding the current crisis is that changes in the
financial sector (and the economy more generally) have resulted in a deterioration in
the quality of information, with resulting adverse effects on economic performance.
It is obvious that the financial sector failed to perform well its critical functions of
allocating capital and managing risk. The CW provides little insight into these failures,
which, in the structure of their model, simply couldn’t occur. The discussion of this
section helps explain these failures, and what might be done about them.

There are three aspects of this deterioration in the quality of market information,
which we take up in each of the next three sections.

Moving from “Banks” to “Markets” Predictably Led to Deterioration in Quality
of Information. Market advocates praised how securitization enabled the more efficient
dispersal of risk. Interestingly, the main argument in favor of securitization was never
fully persuasive, for there are many ways of sharing risk besides securitization; for
example, the shares of a local bank can be widely owned. Advocates of securitization
never explained why that particular approach to risk diversification was superior to
alternatives. As I shall explain, it was not.

The most obvious disadvantage of securitization was that, with those originating
mortgages not keeping them, there was a potential moral hazard problem. In principle,
the investment banks who put together the mortgages into securities were supposed
to ensure their quality; further assurance was to be provided by rating agencies. None
of this worked, nor could it, given the numbers of individual mortgages (with the
numbers increasing exponentially in securities consisting of pieces of securities).
Advocates of securitization ignored the technical difficulties of the task, relying instead
on the belief that one could use statistical information. But such statistical information ignored the differences in circumstances between mortgages written in the past and the new mortgages: the mortgages were different, and the incentives for those writing the mortgages were different. The problems are, in fact, to a large extent inherent with securitization. Systems that disperse risk inherently weaken accountability and incentives not just for gathering information, but for ensuring the quality of the financial products being produced. In well-functioning public securities markets, information about the securities is, in effect, a public good—all potential buyers benefit from the information. And it is inherent that, with private provision, there will be an undersupply. Indeed, this observation is the basis of the Grossman–Stiglitz (1980) critique of the efficient markets hypothesis: if markets efficiently transferred information from the informed to the uninformed, there would be no incentive to gather information.38

Banking is an institutional arrangement for internalizing the benefits of information acquisition about borrowers. The shadow banking system based on securitization is not, accordingly, a substitute for the banking system. And the difficulties for the rating agencies in coming up with a convincing business model too are inherent: The current one, with those being rated paying, is rightly criticized for the obvious conflict of interest. But having buyers pay is also problematic, since in reasonably efficient markets, there will be free-riders not willing to pay the rating agencies.

Good financial markets not only are supposed to assess who is creditworthy (a task at which they failed dismally), but they are also supposed to enforce contracts. Good contract enforcement entails renegotiation. A foreclosure is expensive, and when the price of the house has decreased, it is generally Pareto optimal to renegotiate, especially in the context of a non-recourse mortgage. As we have already noted, securitization also greatly increased problems associated with the renegotiation of contracts—especially so, given the conflicts of interest that were built into the system.

Derivatives Market. Developments in derivatives markets provide a second set of examples on how market “advances” led to poorer information, and potentially poorer economic performance. A large fraction of this market consisted of non-transparent, over-the-counter trades, entailing huge exposures, in the billions of dollars. The previous discussion has emphasized the risks posed by this kind of interconnectivity, yet market participants were not in a position to judge even the extent of exposure; and without such information, there is really no ability for markets to exercise any discipline.

Such non-transparency should not come as a surprise: because markets that are fully transparent are more competitive, and less profitable, there are strong market incentives for reducing and impeding transparency. Indeed, some of the arrangements undermined principles of market decentralization. For example, with large credit default swaps not cleared through an adequately capitalized clearing house, knowing

38. Anand, Kirman, and Marsili (2010) construct a model in which whether it pays any individual to gather information and process information, say about the quality of mortgages, depends on whether others do. They show that there may exist an inefficient equilibrium in which no one gathers information. While such behavior may be individually rational, it is “catastrophic at the aggregate level.”
the risk of default of any one firm required knowing the risk position of every firm with which it was financially interlinked—in a vast, difficult, simultaneous equation system.

**Explaining Lack of Transparency.** The lack of transparency within the financial sector was pervasive, and again, for reasons that are easy to understand—and help explain why the CW models are so far off the mark. Incentive pay, with bonuses depending in part on stock market performance or certain other metrics, have been rightly criticized for leading to excessive risk taking and short-sighted behavior. But they also encouraged non-transparency, moving losses or risks off balance sheet, exacerbating already present incentives from within the regulatory system. “Financial innovations” enhanced their ability to do this.

The basic point of this section is that the supposed improvements in markets led to lower information content in markets, and not surprisingly to poorer overall economic performance.

### 4.3. Credit Markets

The Standard Models typically summarized the financial sector in a money demand equation, which determined market interest rates, which in turn affected consumption and investment behavior. There were no equity and no credit markets—and no explanation for sudden changes in the supply or demand for either. They were, presumably, institutional details of little relevance. In fact, of course, this crisis (like many before) is about sudden changes in the supply of credit; and such changes are not necessarily closely linked (in this, as in many other crises) to changes in the money supply.

Elsewhere, Greenwald and I (2003a) have explained the deficiencies in the Standard Model and argued that a principle channel through which monetary policy affects the economy is availability of credit and the terms at which it is available (spread between T-bill rate and lending rates), which is an *endogenous* variable, affected by conventional monetary as well as regulatory policies. Even if one focused only on interest rates (denying the relevance of credit availability), one needs a model to explain that spread—for example, why it may have increased since the crisis began. But the standard approaches not only failed to provide such a model based on plausible microeconomic foundations of the banking sector, but also failed to note the profound changes occurring in the credit systems in some of the advanced industrial countries (such as the United States).

---

39. Each of the underlying hypotheses, for instance, of the standard transaction demand for money are suspect: Money is not needed for transactions: credit is all that is required; the difference between the interest rate paid on CMA accounts and T-bills is determined not by monetary policy but simply by technology and competition; most transactions are, in fact, not related to income-generating activities but rather to transfers of assets. For more recent literature attempting to explain credit spreads, see Curdia and Woodford (2010). Other papers modeling the credit market include Greenwald and Stiglitz (1993), Kiyotaki and Moore (1997), and Gertler and Kiyotaki (2009).
Our third hypothesis is: Changes in the financial sector affected adversely the efficiency of the credit system, and made it more vulnerable to a shock such as that which occurred. These changes will likely slow the restoration of credit.

There were several key changes, all of which were touted as improving the financial system, but most of which adversely affected its efficiency and stability, and none of which were adequately reflected in the Standard Models. One of these we have already discussed: the development of the shadow banking system and securitization, or as it is sometimes put: banks went out of the storage business into the moving business. Many of the seeming advances were really designed to circumvent the remaining regulations intended to ensure the stability of the financial system. CMA accounts, allowing the rapid transfer of money into and out of bank accounts, circumvented reserve requirements on deposits for checking accounts.

Better risk management supposedly justified higher leverage (reflected in Basel II), but somehow, in these discussions, the Modigliani–Miller theorem got lost. What really seemed to be going on was a combination of risk-misperception and risk-shifting to government, especially by the too-big-to-fail institutions. (As we previously noted, the problem that they posed had worsened markedly since the repeal of the Glass–Steagall Act, which contributed to in increased concentration and risk-taking in banking. A working system with concentrated banking is not impossible, but it requires tight regulation and close supervision.)

Banks became more reliant on wholesale deposits—which could leave the bank quickly if confidence in the bank eroded—as opposed to “old-fashioned” deposits that were more sluggish. Those in the financial sector became confused between ensuring that real resources were used more efficiently, and ensuring that “money” earned as high a return as possible. Private and social returns were often markedly different, the former often a consequence of rent-seeking (see for example Stiglitz and Weiss 1990). Indeed, competition for depositors reduced the franchise value of banks, and this too encouraged excessive risk taking.

Other institutional changes may also have had adverse effects. The shift away from partnerships for the big investment banks may have exacerbated short-termism, partially undermining the long-term relationships which have been the core of sound credit.

The changes in credit markets helps explain why the recovery from this crisis may be particularly slow: Banks have lost large amounts of capital, and the process of rebuilding capital can be slow. The lack of transparency in the financial sector (upon which we commented earlier) may make raising additional equity more difficult and costly than it otherwise would be. Bank capital will be restored through profits, the spread between the lending rate and the borrowing rate. Government policies of keeping the rates banks pay for funds low and allowing anti-competitive practices to persist may enable bank equity to be restored faster than it otherwise would have been. Still, the large spread between deposit and lending rates, while good for restoring bank balance sheets, is bad for resuscitating the economy.
Moreover, with many banks having gone out of the old-fashioned lending business, and with the inherent problems in securitization (at least in some areas) finally being recognized, a full restoration of the credit supply will be slow.

### 4.4. Structural Transformation

The last major crisis, the Great Depression, was a period of structural transformation—a move from agriculture to industry. The Great Recession is another period of structural transformation, from manufacturing to the service sector, induced by productivity increases in that sector and changes in comparative advantage brought on by globalization. Such major structural transformations are periods of high uncertainty and occur only episodically, so that rational-expectations models provide little insights in these situations. This brings us to our fourth hypothesis for why this crisis may last longer than most downturns: *structural transformations may be associated with extended periods of underutilization of resources.*

With an elasticity of demand less than unity, a sector like agriculture with increasing productivity has declining income. But, with declining incomes and high capital costs (including individual-specific non-collateralizable investments) associated with transition out of the sector, it may be difficult for those in these sectors to move out. But declining incomes of those trapped in the high-productivity/declining income sector have adverse effects on other sectors, not compensated by offsetting increased spending from those employed elsewhere. Cyclical and structural problems are intertwined. Downturns (and the associated loss of capital) make financing structural transformation more difficult. Booms too can give rise to structural problems, with labor, for instance, “trapped” in the construction sector that the bubble helped bloat.

The United States did not fully emerge from the Great Depression until the massive Keynesian stimulus accompanying World War II provided the impetus finally to move the labor that was trapped in the rural agricultural sector into the urban manufacturing sector. Before the end of the war, there was much worry that with the removal of the stimulus, the economy would return to its under-consumption path. Yet the transition was remarkably smooth: the war had succeeded in pushing through the structural transformation that the economy on its own found so difficult. And the forced savings during the war meant that there was pent-up demand at the end of the war.

Say’s law, that supply creates its own demand, is not true outside of the fictional world of neoclassical economics. Today, the question is, what will make up for the reduced spending resulting from the declining incomes of workers, for example in manufacturing and construction; or the reduced spending that results from the adverse shock to household wealth from the breaking of the bubble, that has left them with their mortgage debt, vastly overleveraged? And to what sectors will those American men—formerly in manufacturing and construction jobs—go? Some countries, like Germany, have maintained a high-tech manufacturing sector, but that requires a different set of educational and industrial policies than those traditionally employed in the United
States. The service sector has been sold as the sector of the future, but the high paying jobs in that sector were in finance—and that sector, it is now recognized, was overbloated, before the crisis receiving more than 40% of all corporate profits. The services people want include especially education and health, but these have been traditionally largely publicly funded, and with current budget stringencies, prospects of such expansion are weak.40

There are, of course, many areas where there are high social returns to investment (large capital shortages in the least developed countries, retrofitting the world for global warming). If funds could be recycled to those uses, there would be no shortage of global aggregate demand. Prior to the crisis, global financial markets did not work well: it is hard to believe that the best use of the world’s scarce capital was constructing homes for Americans beyond their ability to pay. I am not sanguine that global financial markets will succeed in channeling funds to these areas of high social return—with the resulting risk of persistent lack of adequate global aggregate demand.

In this section, I have focused on four ways in which the US economy (and that of many other advanced industrial countries) have been changing that have contributed to the making of the crisis, and will likely contribute to a slow recovery: This crisis combines elements of increased risk, reduced quality of information, a more dysfunctional credit market, and a structural transformation, with two more ingredients:

(a) Growing inequality domestically, which would normally lead to a lower savings rate (though not so in a representative agent model where distributional issues simply don’t arise). The adverse effects were obfuscated by growing indebtedness, itself supported by a bubble based on low interest rates, lax regulation, and irrational expectations.

(b) The build-up of reserves in emerging markets. One of the factors contributing to under-consumption (high savings) is growing global reserves—now in the trillions. The growth increased markedly towards the end of the last century. One factor was the East Asia crisis and the way it was managed by the IMF and US Treasury. Emerging-market countries did not want to lose their economic sovereignty, or to be subjected to policies which forced them into recessions and depressions. To avoid this, they rapidly increased their reserves—that is, there were high levels of global precautionary savings. Moreover, many countries had learned of the benefits of export-led growth, and low exchange rates (with associated trade surpluses and reserve accumulations) promoted export-led growth.

Regrettably, the way this crisis has been managed has shown the virtue of large reserves. And the high level of unemployment in the advanced industrial countries has continued to impose downward pressure on wages. In short, two of the underlying factors contributing to the crisis have become worse.

40. Besides, the United States is now worried about excessive spending in health care, though it is one of the few sectors in which there has been increased employment. Employment in the US health care sector was up 800,000 in December 2010 compared to December 2007 (Bureau of Labor Statistics Current Population Survey).
The interpretation of the crisis provided in this section has some direct policy implications. These include a focus on creating a global reserve system (to obviate the need for individual countries to build up reserves); strengthening that part of the financial system which provides credit to small- and medium-sized enterprises, creating a new mortgage system (perhaps along that which has been so successful in Denmark); and industrial policies aimed at facilitating the restructuring of the economy. All of this will take time and, meanwhile, there is need for government expenditures.

5. Policy

I began this paper by noting the failure of the Standard Models to predict the crisis. But an even more telling criticism of those models is that they provided policy prescriptions which contributed to the crisis, and provided inadequate guidance to what should be done in response. The standard view was that markets could manage risk on their own; low and stable inflation was necessary and almost sufficient to obtain a low output gap, and perhaps even ensure high growth. Because of the belief in the efficiency of markets, there was hesitancy by central banks to use the full arsenal of tools in their toolkit—including regulations that could have dampened or perhaps even prevented the bubble.41 In spite of protests that one couldn’t be sure that there was a bubble until after it broke, prices were so far out of line with historical norms that policymakers should have been very worried—and after all, all policy making is done with uncertainty. The models led policymakers to believe that, even if there were a bubble, with global diversification, the impact on macroeconomic activity would be small. It would be cheaper to clean up any mess afterwards, than to interfere with the wonders of the market.

Admittedly, all of these policy stances look a little ridiculous from our current vantage point. In this lecture, I have tried to explain some of the central reasons that the models went so badly astray: for example, lack of attention to credit, and to the vast array of other problems posed by imperfect information (including agency issues, corporate governance, etc.). These are not second-order refinements: they explain why we have banks and what affects banks, and the limits of securitization. That some central banks employed macro-models in which banking, presumably their raison d’être, was virtually absent is remarkable.

One of the spurious attractions of the DSGE models is that they provide quantitative guides to policymakers who have to make quantitative decisions: for example, about changes in the interest rates. But while they provide numbers, the question is, what confidence should we have in the numbers that they provide? Decision making is sequential: central banks are constantly in the process of reviewing and revising previous decisions, and it is the observation of the short-run responses to previous

41. Like many other tenets, there was some intellectual incoherence: after all, for the government to set the interest rate is a massive market intervention. There is no theory that says that government should intervene in only one place—quite the contrary. (Ramsey’s great contribution was to refute such a belief in the realm of taxation.)
decisions more than numbers provided by models of consumers maximizing utility over an infinite horizon that do, and should, inform this process. (To be sure, real complications arise in complex non-linear systems, with long lags.)

Each major downturn is sui generis, and the art of policymaking entails judgments (never based on rational expectations models, simply because of the distinctive circumstances) about how the distinctive aspects shape current behavior. In this crisis, for instance, behavior is affected by the unprecedented (in modern economic history) legacy of debt, weak banks, underwater homeowners, a securitization market that had come to be responsible for most mortgages in paralysis, excess capacity in real estate, and a manufacturing sector that had lost competitiveness. The distinctive policy setting—near-zero interest rates, some governments, even in advanced industrial countries, facing high debt-to-GDP ratios, high returns to public investments in the United States and some other countries because of a dearth of public investments in the preceding years, political divisiveness—affected firm and household expectations, again in important ways that could not be meaningfully formally modeled with rational expectations, simply because the mélange of circumstances had never occurred. (Note that the list of important factors affecting economic behavior includes political decisions. Politics and economics cannot be separated. I have had to give short shrift to these considerations in this paper.) Putting aside political economy considerations, there are interventions, based on the same limited information available to private parties, which are welfare enhancing. But neither theoretically nor empirically do we have to rely on rational expectations. We can analyze dynamics with adaptive expectations, and, fortunately, there are a wealth of concurrent surveys and instruments that provide information about individuals’ actual expectations, that have predictive power.

In such contexts, policy makers inevitably, and rightly, fall back on the old \(C + I + G + X – M\) analysis, informed, hopefully, by more sophisticated insights, including those that identify relevant supply-side effects, accompanied by quantitative assessments that call attention to salient aspects of the current situation, including expectations data from surveys. Because any short-run macroeconomic analysis focuses on the determinants to changes (in the short run) on aggregate demand and supply, any assessment of a particular modeling framework should ask, to what extent does it enhance our ability to predict and interpret these movements and to design welfare improving (employment increasing, inflation reducing) interventions? It is almost surely not the case that sudden changes in preferences for leisure or adverse technology shocks caused this or any other major downturn in output or employment. Large (too large) changes in asset prices have arguably been more important in leading to volatility than too small changes in nominal wages (see Easterly et al. 2001a, 2001b). There may, in fact, be important changes in the “wedges” between marginal rates of substitution and marginal rates of transformation over the business cycle, but these are not exogenous shocks but endogenously determined, in economies with imperfect competition in labor and product markets.

Key to an analysis of aggregate demand and supply is a sensitivity to the appropriate level of disaggregation, for example among individuals, firms, and assets. Because
individuals and firms differ, redistributions matter. Large redistributions (across individuals, firms) play a role in explaining volatility; and government policies aimed at redistributing income (towards individuals with higher marginal propensities to consume) or access to finance (towards firms which are financially constrained) can lead to increases in aggregate demand. Not all tax cuts or expenditures increases are created equal: some have larger multipliers than others. For instance, increased unemployment benefits have a high multiplier, because those individuals are often finance constrained. Tax cuts or increased transfers for the poor elderly are also likely to have high multipliers: they will spend it all; contrary to the infinite lifetime models, bequests for most individuals are not important. Tax cuts for high-income individuals may have low multipliers. These are the individuals for whom, if Ricardian equivalence ever holds, it holds. Indeed, for high-income older individuals, an increase in the inheritance tax may induce more consumption.

Thus, high-return government investments financed by taxes on households with low marginal propensities to consume or taxes on firms that are not financially constrained may increase aggregate demand, output, and employment in the short run. Even when such investments are financed by deficits, they may reduce the national debt in the long run, because of the increased tax revenue that they generate. DSGE models remind us, of course, to look at potential long-run impacts: in these cases, it is the lowering of the debt (and future tax payments) that reinforces the current impacts in stimulating the economy. That so many of the models came to the opposite conclusion simply reflects that conclusions follow from assumptions: if one assumes that all government spending is for consumption, and constructs models with low short-run multipliers, then there will be adverse long-run effects on the national debt. Indeed, if government capital is complementary with private capital, then government spending can even lead to increased investment for firms that are not financially constrained.

There are other dynamic effects that may be of importance in the medium term, to which policymakers should pay attention, but which were given short shrift in the Standard Models. Firm and bank net worth matter (there are difficulties in raising new

42. There is a large literature within the DSGE framework, surveyed in Heathcote et al. (2009), modeling individual heterogeneity, but largely arising from idiosyncratic income shocks. This literature, while claiming to replicate patterns of income and wealth distribution, ignores the large microeconomic literature in the field (for example, associated with names like Atkinson, Shorrocks, Flemming, Aitchison and Brown, Champernowne), which over a half century has attempted to explain the Pareto tail, the (near) lognormal form, the role of bequests and imperfect annuity markets, patterns of observed savings behavior over individuals’ lives, and changes over time and differences across countries, and the consistency of these patterns with patterns of wealth inequality. More relevant for macro-analysis, there is no discussion of differing marginal propensities to consume, which can serve as the basis of stimulative redistributive policies. Other recent literature attempting to incorporate non-Ricardian effects includes Coenen and Staub (2005) and Galí, López-Salido, and Vallés (2007).

43. In the area of fiscal policy, Ricardian equivalence theories illustrate how direct effects of government deficit spending might be undermined, as individuals set aside money to pay future tax liabilities. But these concerns seem of limited relevance in the years before the crisis. As Bush’s tax cuts created massive new deficits, savings rates fell to record low levels. Do the proponents of Ricardian equivalence really believe that in the absence of the tax cut, household savings would have been, say, substantially negative? In more relevant models, incorporating life cycle effects and financial constraints, Ricardian equivalence does not hold.
equity because of information imperfections). An increase in aggregate demand today increases firms’ net worth, and hence aggregate supply and demand in future periods. An increase in banks’ net worth increases their ability and willingness to lend. That’s why it was such a mistake to allow them in the midst of the crisis to pay out so much in dividends and bonuses. (This is a standard result of banking models; but it was not incorporated in the macro-models). High extended unemployment rates give rise to hysteresis: we saw it in Europe in the 1980s, we are likely to see it in the United States. If so, it suggests that there are particularly high social returns to preventing the unemployment rate from rising to the levels to which it has risen. Labor supply can also be affected by retirements, and large losses in retirement incomes are inducing many to postpone retirement. With fewer flows out of the labor force, the problem of finding jobs for the new entrants becomes worse. Finally, as we noted in Section 4, the overhang of leverage and of excess capacity in housing affects the economy well after the bubble has broken.44

While these observations fall out naturally from models which reflect the presence of financial constraints and firm and individual heterogeneity, they are more than theoretical truisms: There is ample empirical evidence of the relevance of the distinctions; unemployed and poor aged have close to unitary marginal propensities to consume; small firms are more likely to be financially constrained than large firms. So too, since the effects of monetary policy are significantly mediated through the banking system, a central objective of policy analysis should be an understanding of the determinants of banks’ ability and willingness to lend and the terms at which they make loans available. But just as not all households are the same, neither are the banks. Some are more connected to the “real” sector—more engaged, for instance, in lending to small- and medium-size enterprises, where securitization is not an option. Ensuring their quick recovery may be central to rapid economic recovery. These are the issues on which models aiming to give us insight into how monetary and fiscal policy can stabilize the economy, particularly after a deep downturn, should focus.

The appropriate level of disaggregation is also essential in analyzing how the particular situation of the moment differs from previous situations, which form the basis of the estimated empirical relationships upon which we might rely in more “normal” times. For instance, multipliers today may be markedly higher than they have on average been in the past. Highly leveraged households—of which there are many now—may not have access to funds, or have access to funds at very, very high rates (little connected with the T-bill rate). So too, estimates of the elasticity of investment with respect to the real interest rate based on past experience are of limited relevance in the midst of a crisis such as the current one: The questions now are, What is the interest elasticity when there is large excess capacity—and where there is expected to be excess capacity for an extended period of time? How does lowering government interest rates (short or long) affect the availability of capital, and the terms at which it is available, especially to SMEs, when the banking system itself is in trauma?

44. Another example is that cutbacks in spending on education, technology, and infrastructure in response to the increased national debt will lead to lower productivity in the future.
One of the reasons that the situation today is markedly different (and multipliers are likely to be markedly larger) is the policy setting: for many countries, this is the first time that the zero interest rate bound (which was given short shrift in the literature) is binding—and is binding for an extended period of time. When the economy is at full employment, one expects monetary authorities to offset increased government spending, so one expects zero multipliers. But the economy is far from full employment. One source of leakage is savings; but savings leads to future consumption. If the economy is at full employment at that future date, then the increased output at that future date will be offset by monetary policy. But if the economy is not (expected) to be at full employment in that future date, then the cumulative multiplier will be larger—and there will be a feedback (with rational expectations) to the present period.

Moreover, the policy of monetary authorities (critical in determining multipliers) should not be viewed as exogenous, beyond the government’s control (indeed, a major objective of macroeconomic modeling is to prescribe what monetary policy should do).\textsuperscript{45} We should not say that fiscal policy is ineffective simply because it will be offset by monetary policy; we need to ask, would fiscal policy accompanied by appropriate monetary policies be effective in stimulating the economy.

There are other implications of our analysis for policy stances: The zero rate bound means, in particular, that monetary policy cannot offset spending cuts; so the usual response to excessively aggressive fiscal consolidation—if unemployment increases, the monetary authorities can simply lower interest rates—is no longer applicable.\textsuperscript{46}

In short, even if in the past an increase in government spending on average was largely offset by reduced private spending as a result of “wealth” effects (expected future taxes) or policy responses (Fed increases in interest rates), it does not mean that that will be the case now. The kind of spending being contemplated today may be different (investment rather than consumption). Even if in both cases there were investments, the marginal return to public investment or the level of government indebtedness might have been different. Financial constraints may affect firms and households in different ways. And the policy responses (for example, by the Fed) can differ.

Lucas rightly emphasized that behavior might have been different under different policies, which is why ascertaining the right structure is so important. Many of the policy discussions under debate today, for instance, focus on inducing behavioral changes on the part of banks and other financial institutions, seen to be central to creating a more stable macro-economy; but remarkably, little attention has been given within macroeconomics to understanding the behavior of these institutions, including

\textsuperscript{45} With independent central banks captured by financial markets, one should perhaps better model their behavior as reflecting the interests of the financial markets; but a full articulation of this model goes beyond the scope of this paper.

\textsuperscript{46} Remarkably, some economists have even warned about government spending crowding out investment, as more government borrowing drives up interest rates. But with interest rates near the zero lower bound, it is hard to see how this could have happened. If anything, increases in aggregate demand may have reduced deflation/increased inflation, lowering the real interest rate, crowding in investment (if one believes that real interest rates drive investment).
the role of agency problems, accounting standards, and their behavior towards risk. Expectations are one ingredient in determining responses to policy changes. But so too are the structural features, including the differences between firms, households, and financial institutions upon which we have focused.

6. Concluding Remarks

I have provided here a multi-faceted critique of the reigning paradigm in macroeconomics. First, I have raised several methodological critiques. (a) The Standard Models focused on the wrong questions. They focused on explaining the small “normal” variations in the economy—which don’t matter much—and ignored the large variations which matter a great deal. They asked how the economy responded to exogenous shocks, while some of the most important disturbances—the bubbles that periodically occur, and then break—are clearly endogenous. They should have focused on how and why market economies amplify shocks, why the effects persist, and why the economy does not quickly recover.

(b) The approach to model verification was suspect. The focused on how well the model did in the periods in which things worked well; not how poorly the model did when we really cared about the model predictions. Indeed, some currently fashionable strands do little more than curve fitting in an under-identified system, an approach that is a step back from the advances that had occurred in previous decades in statistical inference. (Korinek (2010d) puts the criticism well: “First, the set of moments chosen to evaluate the model is largely arbitrary... Second, for a given set of moments, there is no well-defined statistic to measure the goodness of fit of a DSGE model or to establish what constitutes an improvement in such a framework.”) Moreover, they did not look at all the predictions of the theory, only a selected subset, ignoring the fact that some of the testable predictions could be rejected.

Secondly, the Standard Models make special assumptions—about behavior and about economic and mathematical structures—which limit their generality and relevance. I have noted several that have played a key role in the results. Here, I want to note one more: The mathematical assumptions which necessarily implied

47. He concludes by asking: “Should we have greater confidence in DSGE models that match more moments and that achieve a closer match to the data than other models? Are these likely to provide a more useful guide to reality? There is no scientific basis to answer this question affirmatively. In some instances, the criterion of matching moments may even be a dangerous guide for how useful a model is for the real world. The focus on matching moments creates incentives for researchers (i) to introduce elements that bear little resemblance to reality for the sake of achieving a better fit (ii) to introduce opaque elements that provide the researcher with free (or almost free) parameters and (iii) to introduce elements that improve the fit for the reported moments but deteriorate the fit along other dimensions.”

48. Guzmán (2009, 2010a, 2010b) provides indirect tests of the rational expectations in the context of a macroeconomic model, which strongly reject that hypothesis. She shows (a) forecasts of key economic variables of “experts” can be improved upon by using concurrently available data; and (b) forecasts of different groups (men and women) who presumably have access to the same data (and therefore, with rational expectations, should have the same forecasts) differ. It is also the case that (say, the bluechip) forecasts themselves have predictive power: they can improve upon forecasts that just rely upon past data.
that systemic risk was reduced through diversification is an example of special (and sometimes not obvious) assumptions which bias the result of the analyses. The schizophrenic approach that assumes convexity ex ante—so that risk can be fully diversified throughout the system—and when it evidently has not been, worrying about contagion, is unsatisfactory. We need coherent models that incorporate key nonconvexities.

I have emphasized trade-offs in modeling. The standard approach was too ambitious in some respects, not ambitious enough in others; in the end, it made the wrong trade-offs. Complexities arising from intertemporal maximization over an infinite horizon are far less important than those associated with a more accurate depiction of financial markets and other aspects of household and firm behavior. Future modeling providing greater realism in modeling banking/shadow banking, key distributional issues (life cycle), key financial market constraints may necessitate simplifications in other, less important directions. Many of the limitations on the Standard Model that I have stressed have been recognized by researchers within the field, and as I have noted, there have been significant advances in addressing them. But as the discussion on policy highlighted, I am not convinced of the ability of these models, in their current state, to address key policy issues.49 The attempt to construct dynamic stochastic general equilibrium models is admirable, and perhaps eventually, the profession may achieve the goal of constructing such a model that provides insights into policy issues that really matter—such as how to prevent bubbles and the major economic downturns to which they give rise and how to ensure quick recoveries when they occur. But meanwhile, what is likely to yield the highest returns are better modeling of the critical sectors and behavioral components. In this paper, I have focused particularly on the financial sector. The art and science of macro policymaking will involve blending the insights from these partial models into a consistent macroeconomic framework.

Fortunately, over the past 30 years, we have made enormous progress on this alternative agenda. The task is to put the various pieces together in a tractable manner, to formulate a New Macroeconomics. The New Macroeconomics will need to incorporate an analysis of risk, information, and institutions set in a context of inequality, globalization, and structural transformation, with greater sensitivity to assumptions (including mathematical assumptions) that effectively assume what was

49. This is a view that is shared by Greg Mankiw (2005), who, like me, served as Chair of the Council of Economic Advisers, which has the responsibility, under the Employment Act of 1946, for guiding the Administration in its macroeconomic policies. He wrote that these models have “had little impact on practical macroeconomists who are charged with the messy task of conducting actual monetary and fiscal policy.” My concern, though, is somewhat different: that in mindset, they have had too much influence. Some of the limitations of DSGE models arise from the fact that, while recognizing the limitations of current models, modelers have (understandably) looked for tractable fixes, and those that may help explain normal fluctuations, but not the deep downturns that are the focus of my concern. Thus, as I noted earlier, labor market frictions associated with search (for example Thomas 2008; Merz 1995) do not explain cyclical unemployment rates of 10%. Nor do staggered wage and price adjustments and rational inattention models adequately explain the failure of wages to adjust in a deep downturn. For this, theories based on efficiency wages, inside-outsiders (Lindbeck and Snower), and option value (Greenwald and Stiglitz 1995) are required.
to be proved (for example, with respect to benefits of risk diversification, effects of redistributions). Agency problems and macroeconomic externalities will be central. It will have to be predicated on an understanding that in the presence of imperfect information and incomplete risk markets, market economies are not necessarily either efficient or stable. These problems are not just minor foibles, to be glossed over in the praise of the virtues of capitalism. New policy frameworks need to be developed based on this new macroeconomic modeling. For a start, there will be a focus not just on price stability but also on financial stability. Even such a small change will be a good beginning: the stakes are great, the costs of those who paid excessive attention to economists’ flawed models have been enormous.

References


