



Dædalus

Journal of the American Academy of Arts & Sciences

Fall 2010

on the
financial
crisis &
economic
policy

Benjamin M. Friedman
& Robert M. Solow

Benjamin M. Friedman

Robert M. Solow

Luigi Zingales

Jeremy C. Stein

Lucian A. Bebchuk

Nolan McCarty,
Keith T. Poole,
Thomas Romer
& Howard Rosenthal

C.A.E. Goodhart

Robert E. Hall

Edward L. Glaeser

Barry Eichengreen

Peter Temin

Alexander Hamilton

Introduction 5

Is our financial system serving us well? 9

The bigger they are . . . 22

Learning to live with not-so-efficient
markets 31

Securitization, shadow banking
& financial fragility 41

How to fix bankers' pay 52

Political fortunes: on finance
& its regulation 61

Lessons from the financial crisis
for monetary policy 74

Fiscal stimulus 83

Housing policy in the wake of the crash 95

International financial regulation
after the crisis 107

The Great Recession
& the Great Depression 115

From his writings 125

poetry

Alice Notley

The Codex Eats Me 127





Inside front cover: Traffic passes by a financial institution's headquarters in New York on Tuesday, April 14, 2009. Photograph © Andrew Harrer/Bloomberg via Getty Images.

A foreclosure sign is posted in front of a townhouse in Herndon, Virginia, November, 22, 2007. In the years following the housing market collapse of 2006 and 2007, banks and mortgage firms have filed thousands of foreclosure notices against American homes. Photograph © Paul J. Richards/AFP/Getty Images.

Benjamin M. Friedman and Robert M. Solow, Guest Editors

Phyllis S. Bendell, Managing Editor and Director of Publications

Micah J. Buis, Associate Editor

Erica Dorpalen, Editorial Assistant

Board of advisers

Steven Marcus, Editor of the Academy

Rosanna Warren, Poetry Adviser

Committee on Publications

Jerome Kagan, *Chair*, Jesse H. Choper, Denis Donoghue, Gerald Early,

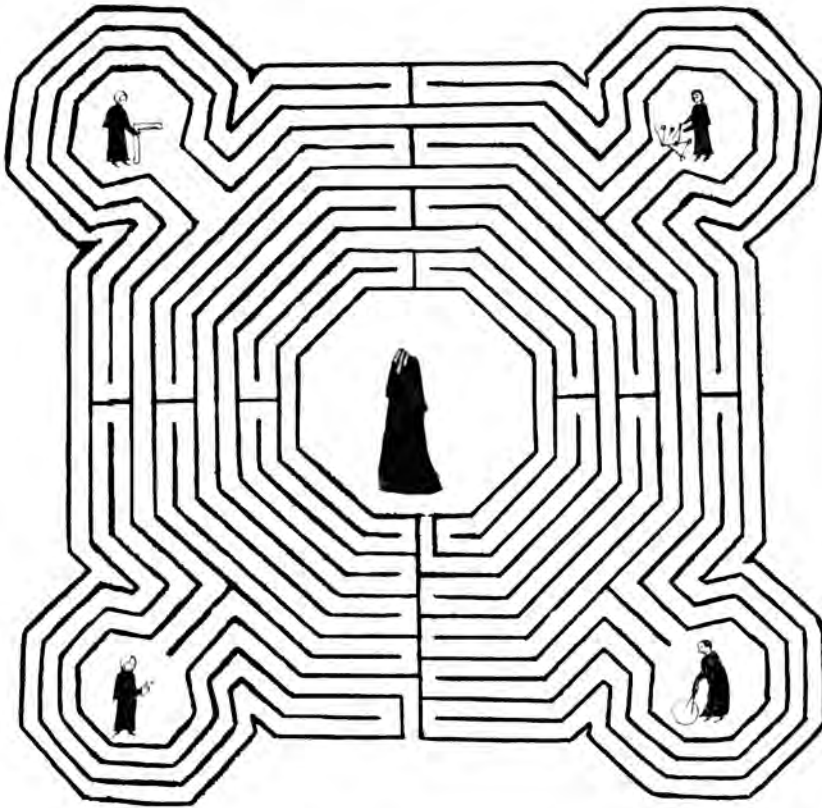
Linda Greenhouse, Jerrold Meinwald; *ex officio*: Leslie Berlowitz

Dædalus is designed by Alvin Eisenman.

The Academy dedicates this issue in memory of Carl Kaysen and Paul Samuelson, noted economists and active and esteemed members of the American Academy of Arts and Sciences.

Dædalus

Journal of the American Academy of Arts & Sciences



The pavement labyrinth once in the nave of Reims Cathedral (1240), in a drawing, with figures of the architects, by Jacques Cellier (c. 1550 – 1620)

Dædalus was founded in 1955 and established as a quarterly in 1958. The journal's namesake was renowned in ancient Greece as an inventor, scientist, and unriddler of riddles. Its emblem, a maze seen from above, symbolizes the aspiration of its founders to "lift each of us above his cell in the labyrinth of learning in order that he may see the entire structure as if from above, where each separate part loses its comfortable separateness."

The American Academy of Arts & Sciences, like its journal, brings together distinguished individuals from every field of human endeavor. It was chartered in 1780 as a forum "to cultivate every art and science which may tend to advance the interest, honour, dignity, and happiness of a free, independent, and virtuous people." Now in its third century, the Academy, with its nearly five thousand elected members, continues to provide intellectual leadership to meet the critical challenges facing our world.

Dædalus Fall 2010
Issued as Volume 139, Number 4

© 2010 by the American Academy
of Arts & Sciences

*Securitization, shadow banking
& financial fragility*

© 2010 by Jeremy C. Stein

How to fix bankers' pay

© 2010 by Lucian A. Bebchuk

The Codex Eats Me, from "Culture of One"

(Penguin; forthcoming, March 2011)

© 2010 by Alice Notley

Editorial offices: *Dædalus*, Norton's Woods,
136 Irving Street, Cambridge MA 02138.
Phone: 617 491 2600. Fax: 617 576 5088.
Email: daedalus@amacad.org.

Library of Congress Catalog No. 12-30299

ISBN 978-0-262-75111-7

Dædalus publishes by invitation only and assumes no responsibility for unsolicited manuscripts. The views expressed are those of the author of each article, and not necessarily of the American Academy of Arts & Sciences.

Dædalus (ISSN 0011-5266; E-ISSN 1548-6192) is published quarterly (winter, spring, summer, fall) by The MIT Press, Cambridge MA 02142, for the American Academy of Arts & Sciences. An electronic full-text version of *Dædalus* is available from The MIT Press. Subscription and address changes should be addressed to MIT Press Journals, 55 Hayward Street, Cambridge MA 02142. Phone: 617 253 2889; U.S./Canada 800 207 8354. Fax: 617 577 1545.

Printed in the United States of America by Cadmus Professional Communications, Science Press Division, 300 West Chestnut Street, Ephrata PA 17522.

Newsstand distribution by Ingram Periodicals Inc., 18 Ingram Blvd., La Vergne TN 37086, and Source Interlink Distribution, 27500 Riverview Center Blvd., Bonita Springs FL 34134.

Postmaster: Send address changes to *Dædalus*, 55 Hayward Street, Cambridge MA 02142. Periodicals postage paid at Boston MA and at additional mailing offices.

Subscription rates: Electronic only for non-member individuals – \$43; institutions – \$113. Canadians add 5% GST. Print and electronic for nonmember individuals – \$48; institutions – \$126. Canadians add 5% GST. Outside the United States and Canada add \$23 for postage and handling. Prices subject to change without notice.

Institutional subscriptions are on a volume-year basis. All other subscriptions begin with the next available issue.

Single issues: current issue – \$13; back issues for individuals – \$16; back issues for institutions – \$32. Outside the United States and Canada add \$6 per issue for postage and handling. Prices subject to change without notice.

Claims for missing issues will be honored free of charge if made within three months of the publication date of the issue. Claims may be submitted to journals-claims@mit.edu. Members of the American Academy please direct all questions and claims to daedalus@amacad.org.

Advertising and mailing-list inquiries may be addressed to Marketing Department, MIT Press Journals, 55 Hayward Street, Cambridge MA 02142. Phone: 617 253 2866. Fax: 617 253 1709. Email: journals-info@mit.edu.

Permission to photocopy articles for internal or personal use is granted by the copyright owner for users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the per-copy fee of \$12 per article is paid directly to the CCC, 222 Rosewood Drive, Danvers MA 01923. The fee code for users of the Transactional Reporting Service is 0011-5266/10. Address all other inquiries to the Subsidiary Rights Manager, MIT Press Journals, 55 Hayward Street, Cambridge MA 02142. Phone: 617 253 2864. Fax: 617 253 1709. Email: journals-rights@mit.edu.

The typeface is Cycles, designed by Sumner Stone at the Stone Type Foundry of Guinda CA. Each size of Cycles has been separately designed in the tradition of metal types.

Benjamin M. Friedman & Robert M. Solow

Introduction

The financial meltdown of 2007 to 2009 was surely a great spectacle. Mighty names toppled like that statue of Saddam Hussein. Lehman Brothers, with a history spanning a century-and-a-half, just disappeared. Bear Stearns and even Merrill Lynch – the same Merrill Lynch that had taught generations of small investors to be “bullish on America” – were sold off at discounts suitable for used furniture. AIG was rescued in the nick of time, but only with \$182 billion of U.S. government assistance. Trillions of dollars of investors’ wealth simply evaporated. One could think, “Oh, well, easy come, easy go.” But still, *trillions* of dollars? It was a spectacle all right, but why did it really matter to the rest of us, who count ourselves merely as citizens of the republic?

After all, the economic well-being of a society and its members depends on “real” outcomes: on the production and distribution of current output, some of it allocated to current consumption of goods and services and some to provision for future consumption through capital investment, research activity, and education (all net of depreciation, resource depletion, and environmental damage). That is what the economy delivers to us

folks. Apart from the sheer theater of it all, the main reason for caring about a financial crisis is the well-founded belief that serious disturbance of the financial system can impair the functioning of the real economy, perhaps drastically. A well-behaved financial system makes the real economy more efficient at producing well-being for its inhabitants (though it may parcel out income and wealth to specific groups among those inhabitants in ways that we and they may find objectionable, and maybe even repugnant). But a breakdown of the financial system can inflict damage on the real economy, damage that may last for years after the breakdown has been repaired. And so it has.

These connections between the real economy and the financial system can be far from simple. In this issue of *Dædalus*, essays by Edward L. Glaeser of Harvard University and Jeremy C. Stein, also of Harvard, show how an episode of overbuilding of houses, the sort of thing that might normally lead to a run-of-the-mill slowdown or minor recession, can be amplified and complexified by the (mis)behavior of financial institutions and the spread of securitization. We therefore end up with a disaster for the real economy – and for millions, or tens of millions, of its inhabitants. Instead of a recession that can be dealt with by rou-

tine monetary and fiscal policy, we face the loss of more than eight million jobs and years of lost output, not to mention the indirect social costs of prolonged recession. It is all the more galling that much of the damage is borne by innocent bystanders, while many of the bad (or stupid or greedy) guys do fairly well.

The meaningful story is about the interaction of financial activity and the real economy. The mechanics of this connection are far from transparent. Many of the articles in this issue focus, to varying degrees, on aspects of this difficult and important matter, and on the ways in which a society can hope to benefit from a highly developed financial system while protecting itself, more or less, against the damage that its proclivity to malfunction can inflict. Essays by Luigi Zingales of the University of Chicago Booth School of Business and Peter Temin of the Massachusetts Institute of Technology aim to elucidate the intellectual apparatus that economists and students of finance have built to help them understand this complex piece of machinery.

One important lesson, driven home with great force by the breakdown itself, is the potential of an elaborate financial system like ours for *instability*, a subject which Benjamin M. Friedman of Harvard University, Robert M. Solow of the Massachusetts Institute of Technology, and Jeremy C. Stein each take up in their essays. The crisis has dramatically demonstrated that our kind of system, instead of responding to errors and disequilibria by self-correcting, gradually or quickly, can magnify initial errors many times, and then spread them by contagion throughout the financial system and even to areas of the real economy that seem remote from the initial source: witness the path from housing bubble to general crisis. The patent contradic-

tion of the assumptions that have stood behind a generation of Reagan/Thatcher “let the markets rule” policies – and the intellectual framework behind those policies, associated with Milton Friedman in an earlier era and Alan Greenspan more recently – is enormous.

A second, less obvious lesson is that an elaborate financial system, with its mysteries, its glamour, its possibilities for instant wealth, can quietly distort the direction of the real economy; it can induce the real economy to spend human and material resources on activities that can lead to immense private profit for some of those engaged while nonetheless making little or no contribution to general well-being. Some of the following articles, in particular, essays by Benjamin M. Friedman, Lucian A. Bebchuk of Harvard Law School, and Barry Eichengreen of the University of California, Berkeley, focus on these dangers and the search for public policies that can fend them off with minimal handicap to efficient financial activity.

As we saw during the yearlong debate over what finally became the Dodd-Frank Wall Street Reform and Consumer Protection Act (passed by Congress in July 2010), attempts to legislate effective limitation and regulation of financial activities – in the search for a viable combination of efficiency and protection – run into the formidable lobbying power and political clout of the financial industry itself. In their essay, Nolan McCarty of Princeton University, Keith T. Poole of the University of Georgia, Thomas Romer of Princeton, and Howard Rosenthal of New York University provide insight into the challenge of regulating finance in the face of political forces. Nor has the lobbying abated now that the numerous agencies charged by Dodd-Frank with making new rules, and otherwise implementing the changes that the law

mandates, have begun their work. The widening inequality of income and wealth in the United States has consisted largely of an enormous increase in the incomes of those at the very top of the heap, the best-off 1 percent or (even more so) one-tenth of 1 percent, with near stagnation lower down. Those tippy-top incomes are often pocketed by the leading figures in the financial services business. They are not likely to give them up without an all-out fight; and those same deep pockets can generate a lot of political leverage, in both parties. The optimistic view is that we are lucky to have got as much as we did by way of improved regulation out of the stonewalling, horse-trading, and dependence on “contributions” that pervade today’s Congress.

With or without new regulation, disruption in the financial system will inevitably disrupt the real economy from time to time. So will other disturbances, like the occasional sharp increase in world oil prices or simply a turn toward pervasive pessimism among business executives deciding on their firms’ capital spending programs. Two of the following articles, by C.A.E. Goodhart of the London School of Economics and Robert E. Hall of Stanford University, reconsider the standard defenses against shocks to the real economy: monetary policy and fiscal policy. Now that the worst of the financial crisis is behind us, and the debate over financial reform has been consummated in the Dodd-Frank legislation, these policies have become the focus of much of the current economic debate.

Monetary policy, which today means mostly interest rate policy – the domain of the U.S. Federal Reserve System and other central banks elsewhere – is naturally more directly connected to financial events; C.A.E. Goodhart thus argues that future monetary policy-makers need

to pay more explicit attention to the risk-iness, as well as the volume, of assets held by the financial sector. In the meanwhile, as of this writing, despite record-low long-term interest rates and no sign of inflation anywhere, many observers of U.S. monetary policy, especially in the financial markets, seem terrified that any move toward monetary expansion would soon unleash irresistible inflationary forces. The same fear paralyzed monetary policy-makers in Japan for more than a decade following that country’s financial crisis in the early 1990s, with the predictable result of ongoing economic stagnation. To conclude that, with short-term interest rates at zero, monetary policy can do nothing further is arguable if not necessarily correct; to argue that monetary policy *should* do nothing further seems, under today’s circumstances, hard to fathom.

The analysis of fiscal policy is, in many respects, the same no matter where the shocks to the real economy come from. But Robert E. Hall makes the important point that the hallowed tendency of most central banks to “lean against the wind” means that any expansionary impulse, whether from fiscal policy or elsewhere, is likely to be partially resisted by the Federal Reserve through an increase in short-term interest rates. (One is entitled to ask why this should be so: aren’t the Federal Reserve and the fiscal authorities in Congress part of the same government, trying to do what is right for the same economy? Why should the right hand undo what the left is doing?) In the aftermath of a financial crisis like the one we have just experienced, however, when the Federal Reserve has pushed short-term interest rates all the way to zero and presumably wishes they could go even further, this monetary policy reaction to any fiscal stimulus seems unlikely. Fiscal policy therefore becomes a substantially more effective tool in a deep recession.

This insight makes the immediate state of affairs around the industrialized world more visibly dysfunctional. The United States and the major European economies have barely started to climb out of a serious recession, and they have not even begun to make good the two years of lost growth. Unemployment and excess capacity are still high and persistent. Nobody believes that the next two years look promising. One would think this represents just the sort of circumstance when fiscal policy is at its most potent, and most needed. Yet most of the talk in both the United States and Germany, the Western world's two bellwether economies, is, incomprehensibly, about fiscal contraction: about deficit reduction, to be achieved via lower public spending and, possibly, higher taxes.

True, the United States has a long-term budget imbalance that needs to be addressed (maybe even with legislation enacted now, to take effect later, when the economy has more nearly returned to full employment); but pursuing fiscal retrenchment *now* seems at best perverse, and in sufficient force, a suicidal recipe for renewed and protracted economic downturn. In Europe, some countries – Greece and Portugal, for example – probably will have to undergo deep and lasting recessions, not just because their tax and spending policies have got so far out of line with one another in recent years, but because their prices and costs have drifted far compared with those of other European economies and, as members of the common-currency (Euro) group, they cannot simply alter their exchange rates to correct the problem. But the likelihood of severe weakness in those European economies that have to have it is all the more reason for those that don't, like Germany, to remain strong; not for them to pursue recession-inducing policies, too, merely for their own sake.

In the wake of the financial crisis, it is understandable that consumers and businesses, conscious of vanished wealth and uncertain about the future, would be reluctant to spend. As economists have recognized since observing what happened in the 1930s, such behavior may be individually rational, even if in the aggregate it makes things worse for everyone. Governments, by contrast, are supposed to think about the health of the economy as a whole, and to offset, or more than offset, such temporary deficiencies in demand for what the economy normally produces. Today they instead seem to be poised to reinforce them. Four-legged lemmings no doubt look on with astonishment.

Benjamin M. Friedman

Is our financial system serving us well?

In 1772, at the height of Scotland's worst banking crisis in two generations, David Hume wrote to his close friend Adam Smith. After recounting the bank closures, industrial bankruptcies, spreading unemployment, and even growing "Suspicion" of the soundness of the Bank of England, Hume asked Smith, "Do these Events any-wise affect your Theory?"

They certainly did. Smith's analysis of the role of banking in *The Wealth of Nations*, published just four years later, clearly reflected the lessons he took away from the 1772 crisis. In contrast to the doctrinaire antiregulatory ideology with which he is usually associated by today's economists, Smith favored such measures as usury laws – specifically, no lending at interest rates above 5 percent – and restrictions on the obligations that banks could issue.¹

Large-scale and unusual events, especially when they bring unwanted consequences, provide an opportunity to ask basic questions. Even if no one is at fault for causing an event (an earthquake, for example), it is only natural to ask what might be done differently to mitigate the consequences should a similar catas-

trophe recur. When disaster is the result of human action, the question at issue is not merely about containment; it is also a question of prevention. It is no surprise, therefore, that the recent financial crisis has prompted a flood of proposals to reform the regulation of financial markets and financial institutions in the United States and elsewhere. The Dodd-Frank Wall Street Reform and Consumer Protection Act, which Congress passed in July, made some of them into U.S. law.

What is missing from this conversation, however, is any significant probing of the more fundamental aspects of how well our financial system is serving us, and at what cost. To date, the discussion has focused on the specific symptoms of this particular episode of financial malfunction: the losses incurred by investors, the need for taxpayer-financed bailouts, the disruptions triggered by Lehman Brothers' failure, and the like. Most people are aware, too, that the associated loss of jobs, incomes, and profits was a result of what went wrong in the financial sphere. But no one seems to be relating these costly manifestations to the role that financial markets should play in our economy in the first place or asking how well the markets are performing that role.

The crucial function of the financial markets in an economy like that of the

United States is to allocate scarce investment resources. Typically, about one-fifth of what the U.S. economy produces (in recent years the fraction has been somewhat smaller) is devoted to investment of all kinds: factories and machinery that allow our firms to produce physical goods, office buildings and computers to house and equip a vast service sector, homes for a population of more than three hundred million, and inventories on the shelves of supermarkets and clothing stores. No one simply decrees that one-fifth of our nation's total output is the ideal share to invest for all these purposes. That outcome is the result of countless decisions made every day by individual businesses and homeowners as they interact with the banks, insurance companies, stock buyers, mortgage lenders, and other providers of the funds they need to carry out their desired investment programs. Similarly, no central agency dictates that such-and-such a percentage of the overall investment should go into the computer industry, some percentage into opening new restaurants, and another percentage into putting up new apartment buildings. That allocation is also the result of countless individual-level market interactions.

What guides these interactions and, therefore, what determines both the total amount of investment our economy undertakes and the allocation of this total among different potential applications is the combination of signals and incentives created by the prices set in the financial markets. Is the price of computer company stocks high? Then firms in that industry have both the incentive and the ability to issue more new shares and use the funds for further expansion. Is the interest rate on mortgages low? Then home buyers have the incentive and the ability to take out more loans and buy more houses. Conversely, when the price

of GM and Chrysler stock fell to near zero, and the interest rates on their bonds rose to record highs, in the months before those companies declared bankruptcy, these firms had neither the incentive nor the ability to expand; without the government's rescue they would have contracted – maybe all the way to nonexistence.

To be sure, modern financial institutions perform other important functions as well. They provide checking accounts and other elements of the economy's payments mechanism, vehicles for families to keep their rainy-day funds and to save for retirement, and group and individual life insurance, for example. But while those services are essential to have, they are not essential to the make-up of a free-enterprise economy. In fact, there are numerous models for providing them as a public utility.

By contrast, the function of allocating investment capital – determining the economy's aggregate amount invested, as well as the allocation of that investment across different industries, among different firms, between business installations and homebuilding – *is* essential to the make-up of a free-enterprise economy. This is what our financial system is supposed to be doing. We are therefore entitled to ask how well or poorly it is performing this role, and at what cost.

By now the main outlines of the recent financial crisis and the economic downturn that it caused are well known. Beginning in the late 1990s, and then more so once the relatively mild 2001 recession had ended, house prices rose rapidly. Increasingly lax mortgage underwriting standards – high loan-to-value ratios, back-loaded repayment schemes, little if any documentation – were both a cause and a consequence of the rise in prices. Less onerous lending conditions spurred demand for houses while the rising value

of the underlying collateral lessened concerns for borrowers' creditworthiness. Securitization of a large fraction of the newly issued loans further diminished the originators' interest in borrowers' integrity. In turn, investors that purchased the created securities either deluded themselves (for example, by counting on rising housing prices to nullify the implications of borrowers' lack of creditworthiness) or were misled by rating agencies that carried out shoddy analysis while also facing serious conflicts of interest. Importantly, many of the investors that bought these ill-supported securities were non-U.S. entities.

Three additional developments contributed to the vulnerability of the U.S. financial system. First, within the banking system, the distinction between banking and trading had for the most part disappeared. This development was not simply a consequence of the formal repeal of the Glass-Steagall Act in 1999; the Depression-era separation between commercial banking and investment banking had largely eroded long before. Most of the big commercial banks, seeking to raise their own capital in speculative securities markets, increasingly relied on trading profits to enhance their returns, in effect turning themselves into hedge funds (otherwise they would have had little reason to retain, on their own balance sheets, shares of the mortgage-backed securities that they earned a fee from packaging and selling). Meanwhile, most of the big investment banks, which already had significant trading operations, increasingly funded themselves by rolling over short-term obligations.

Second, the pressure to boost the returns they provided to their shareowners led many of the largest institutions to increase their leverage – the amount of assets they held compared to the base

of invested capital that supported those assets – to record levels. Leverage of twelve- or fifteen-to-one was not uncommon among the big U.S. commercial banks, and many investment banks had ratios of twenty-five- or even thirty-to-one. As a result, these firms had little cushion with which to absorb whatever losses they incurred on their trading operations.

Third, the ongoing development of the market for financial derivatives – instruments based on the value of other financial instruments, which in many cases themselves depended on the value of still other financial instruments – moved beyond the role of enabling financial institutions and other investors to hedge their existing risk exposures; instead, the market provided vehicles for investors to take on new, unrelated risks. As a result, many of the risks to which investors of all kinds became exposed bore little or no connection to fluctuations in any component of the economy's actual wealth, such as housing prices or the value of companies issuing shares. More and more, the risks borne were merely one side or the other of zero-sum bets.

In light of these vulnerabilities, it is not surprising that some catalyst would set off a widespread crisis. The turnaround in housing prices – declines at nearly 20 percent per annum on average across the country, and far more in some states and in many local residential markets – provided that catalyst. (Because what matters for any individual mortgage is the specific house collateralizing that one loan, greater dispersion of house-price changes around a given average rate of decline exacerbates the frequency of default.) Delinquencies and defaults increased rapidly, especially in the market for “subprime” mortgages. The value of securities backed by packages of these mortgages declined. Leveraged derivative

claims against these securities declined even more. The investors that held these instruments took losses. Highly leveraged financial institutions saw their capital erode, in many cases to the point of probable failure in the absence of government assistance. Banks stopped lending; the market in which many companies regularly issued commercial paper shut down. Unable to borrow, businesses and families cut their spending.

Public discussion of the economic fallout from these unfortunate events has focused on four components: the losses that homeowners suffered on their houses, the losses booked by investors in mortgages and mortgage-backed securities, the recession that the financial crisis triggered, and the money put up by government on behalf of the taxpayers. All are important. From year-end 2006 to year-end 2009, U.S. households saw the value of their real estate fall from \$23 trillion to \$17 trillion (and the value of their holdings of corporate stock, including mutual funds, fall from \$14 trillion to \$12 trillion) – in all, a 13 percent reduction in their total net worth. Banks and other institutional investors suffered more than \$4 trillion in losses on their holdings of mortgages and mortgage-backed securities alone (roughly half of that total outside the United States), not to mention another \$6 trillion in losses from declining stock prices and further losses that they have not yet tabulated on their non-mortgage loans. The recession held the expansion of total economic output in the United States to just 0.4 percent in 2008, followed by a 2.4 percent *decline* in 2009; compared to a more normal economic trajectory with even modest growth of 2.5 percent per annum, the loss of production – and therefore of incomes and profits – was \$1.3 trillion for just those two years. The U.S. Treasury

and the Federal Reserve System together put nearly \$2 trillion into either rescuing individual institutions (\$182 billion for just one firm, insurance company AIG) or buying private-sector obligations in order to support troubled markets. With some luck, the government will get most of that money back when the loans are repaid and it sells off the assets that it bought; any remaining losses will come out of taxpayers' pockets.

Important though these losses are, however, what this discussion has failed to address is the misallocation of the economy's resources that led to the crisis in the first place. The lax mortgage underwriting standards and unsustainable expectations for ever-increasing housing prices encouraged interest rates on many forms of mortgage lending to be unrealistically low. Once housing prices started to decline, and borrowers ran late with their payments and even began to default on their loans, the interest rates on those loans rose. This upward adjustment of mortgage interest rates, and the decline in the prices of mortgage-backed securities that produced the losses suffered by banks and other investors, were, in effect, the same phenomenon. (For an instrument such as a bond that promises a fixed annual payment, whether it is backed by mortgages or not, the price and the effective interest rate vary inversely. If a security promises to pay \$60 per year, the rate of return earned by the investor who buys it for \$1,000 is 6 percent; if the price falls to \$800, the effective interest rate for a buyer at that price rises to 7.5 percent.)

But the problem is not that the interest rates on these loans rose so much as that they were too low to begin with. The prices set in financial markets provide signals and incentives for investment behavior whether or not they are right, and even whether or not they are sensible. In

this case, because mortgage rates were so low, and mortgage loans so accessible even to noncreditworthy borrowers, Americans bought and built far more houses than they otherwise would have. On average, during the 1980s and 1990s, the U.S. construction industry built 1.4 million new homes per year. Although the rate of new family formation remained roughly unchanged, by 2003 the industry was building 1.8 million new houses; in 2004, 2 million; in 2005, 2.1 million; in the first half of 2006, still 2 million at an annual rate.

Building all these extra houses used resources in two parallel senses. First, the physical activity of constructing them absorbed labor and materials, both of which are expensive. (Despite the industry's depressed state, last year weekly earnings in construction averaged \$852 versus \$726 in manufacturing and only \$389 for retail jobs.) Second, like any investment, homebuilding absorbs saving. If the prevailing low interest rates on mortgages had not steered so much labor and material and so much of the economy's scarce saving into homebuilding, the United States could have deployed those resources for some other purpose – new manufacturing plants, perhaps, or new gasoline refineries, airports, or school buildings. Furthermore, the allocation of too many resources to homebuilding had consequences beyond the forgone uses to which they could have been put at the time. Once the turnaround came, the presence of all those extra houses on the market pushed housing prices down lower than they otherwise would have fallen. It also depressed new home construction – to only five hundred thousand in 2009 – with a further consequent loss of jobs and lower sales of associated products such as furniture and appliances.

This misallocation of resources, because of what in retrospect were clearly incorrect prices set in financial markets, is hardly unique to the recent housing bubble. In the stock market boom of the late 1990s – which lifted the Standard and Poor's 500 index from 542 on average in 1995 to a high of 1,553 in August 2000 and the Nasdaq index from 925 to 5,132 – one of the hottest areas was telecommunications. This surge in stock prices did not apply just to highly speculative new companies but to large, well-established ones, too. AT&T stock, for example, rose from 21 at the beginning of 1995 to 60 in 2000. Not surprisingly, once the market turned, those same stocks suffered some of the largest declines. By late 2002, AT&T had returned to 21. As of mid-2010, it remained at 25; the anomaly was the run up to 60, not the subsequent fall.

The ensuing public discussion primarily focused on the losses that investors in these stocks had incurred. But in that episode as well, no less than in the recent housing boom, losses on financial assets were a reflection of wasted resources and misallocated saving. In retrospect, \$60 per share was not the right price for AT&T; other telecom stock prices were too high as well. But the fact that the stock prices were too high meant that the cost of capital to the firms issuing the shares was too low. Those firms therefore had the incentive and the ability to issue more shares than they should have. And they used the proceeds to expand more than they should have: in the 1990s, telecom firms laid hundreds of millions of miles of fiber-optic cable that have never been lit and probably never will be.

As with any device that performs a useful function – a car, for example, or a dishwasher or clothes dryer – the performance of the mechanism that allocates

our economy's investment capital is best assessed against some notion of what it costs. Everyone understands that a Chevy Aveo provides neither the pickup nor the comfort, not to mention the styling, of a Ferrari. But few people are willing to pay what it costs to buy a Ferrari and then keep it running. Cost considered, the cheaper option suits many people very well.

In the case of a mechanism for allocating investment capital – where the whole point is that the capital, once allocated, be productive in the usual economic sense – there is a built-in benchmark against which to compare the associated cost: the return on the invested capital itself. If a new fertilizer or irrigation system offers a farmer the prospect of a higher crop yield, it is only common sense for the farmer to compare the economic return from the enhanced harvest to the cost of achieving it. If what it costs to buy, transport, and spread the fertilizer exceeds the sale price of the additional crops it helps to produce, then it's a poor investment for the farmer – despite the enhanced technical efficiency. A financial system that allocates investment capital is no different. Even if it allocates the economy's capital more efficiently than some alternative, if it costs more to run than what the superior allocation produces, it's only so much overpriced manure.

In recent decades, the U.S. economy's mechanism for allocating its capital has been getting a lot more expensive, even compared to the total returns earned on the capital being allocated. From the 1950s through the 1980s, profits earned by financial firms (not counting insurance companies and firms in the real estate business) represented 10 percent of all profits earned in the U.S. economy – hardly an excessive charge as long as this capital allocation system delivered modest benefits. In the 1990s, however,

financial firms' share of total profits rose to 22 percent. During the years 2001 to 2005 – that is, until just before the surge in borrowing, securitization, and derivatives finance began to transform itself into a world-class crisis – these firms' share of all U.S. profits reached 34 percent.

Although this astonishing drain on the profits earned by U.S. business received a fair amount of attention at the time, it is far from the total cost of running the economy's financial system. That cost also includes the salaries financial firms pay to their workforce and the rents they pay for their office space (including the rental equivalent for firms that own their own buildings), as well as more mundane elements – the associated utility bills, travel tickets, and advertising budgets, for instance.

The large salaries and bonuses paid to U.S. financial executives have recently attracted ample public attention, especially once so many major firms found themselves floundering during the crisis. But the phenomenon extends beyond the outsized incomes of a few individuals at the top, which arouse public anger but count for little in the aggregate. Like its share of economy-wide profits earned, the finance industry's share of all U.S. wages and salaries paid has also been rising in recent decades. Fifty years ago it was 3 percent; more recently it has been 7 percent. The standard argument is that high salaries (including the eight-figure bonuses for those at the top) are necessary to attract the talent that enables these firms to do their job. If this argument is true, it means that the economy's capital allocation mechanism is inherently all the more expensive to operate. The same principle applies to the financial sector's other expenses. It may be true that without lavishly furnished offices in choice locations, or lots of prime-time television advertis-

ing, the capital allocation mechanism would not be able to serve its function. If so, the necessary cost of running it is that much greater.

Here as well, the basic principle of a market economy holds: expenses paid are the counterpart of resources used. The fact that financial firms pay more than other companies do (not to mention the government and the nonprofit sector) means that, on average, they attract the most talented and energetic workers. Workers in the financial sector have long been more likely to have a college education than those elsewhere in the U.S. economy. In pace with the widening wage differentials, the extent to which financial-sector workers are more likely to be college-educated has more than tripled over the last thirty years. The changing pattern is easily visible in our colleges and universities. At Harvard (where I teach), more than one-fourth of the graduates in recent classes have gone to work at investment banks, hedge funds, private equity firms, and the like.

At the individual level, no one can blame any of these graduates. They are responding to the incentives presented to them. Furthermore, these incentives are not merely a matter of how the for-profit economy works. These young people are observant enough to understand that despite their university presidents' pious pronouncements about the value of humanitarian and other less-well-paid careers, even their universities will shower them with attention, and in time be more eager to admit their own children, if they earn enough to become sizable donors. That so many top students, acting individually, respond to such incentives can only be expected.

The question, rather, is whether *in the aggregate* the direction of such a large fraction of our most skilled, best educated,

and most highly motivated young citizens to the financial sector constitutes the best use of what is surely one of our nation's most valuable resources. These young people could of course be doing something else. If they are not really needed in the financial firms that employ so many of them – if what they do there actually adds little or no economic value – then something is seriously wrong with yet another market that allocates our economy's resources (in particular, the labor market). But if the financial sector *is* the best place to use their talents and energies, that need is yet another part of what makes our economy's capital allocation mechanism so expensive to run.

What makes these matters so perplexing, and potentially of such huge practical importance, is that the respective cost and efficiency trends in capital allocation in the United States seem to be moving in opposite directions. Over a long enough period of time – since World War II, say, or even longer – the U.S. economy's growth and the rising living standards this growth has brought certainly suggest that a dynamic enough system exists. The important role played by start-up companies in that growth is further evidence that whatever mechanism has been allocating the economy's capital must have been doing a pretty good job. Apple, Microsoft, and Google are only a few of the new entrants that over the years have not only generated substantial profits for their owners, and incomes for their employees, but also changed how the majority of Americans go about their daily activities.

More recent trends have been less reassuring, however. At the same time that the financial sector has been growing more expensive to operate – absorbing a larger fraction of the economy's total profits, claiming a larger share of the most talented workers, and so on –

Is our financial system serving us well?

the economy's performance has been disappointing. Even apart from the recent housing bubble and all the economic costs associated with its demise (costs that must be subtracted from the economic gains the financial sector delivers), there is less to cheer about in the gains accruing to capital investment.

In the first quarter-century following the restoration of a post-World War II peacetime economy (1948 to 1973), the total output of U.S. nonfarm businesses increased, on average, 2.8 percent per annum faster than the growth of labor input. During the next twenty years (1973 to 1993), productivity growth in the nonfarm business sector slowed to just 1.4 percent. Developments in the mid-1990s stirred hope that the spread of electronic technology had brought a new era of rapid productivity growth, but, more recently, the gains have been only mediocre. For 1993 to 2007 overall – that is, even omitting the effects of the recession brought on by the financial crisis – productivity growth averaged 2.3 percent. Although the correspondence is far from exact, over time, movements in productivity roughly correspond to movements in wages and, therefore, in living standards. Allowing for price inflation, the average wage in the economy's nonfarm business sector increased by 2.7 percent per annum during 1948 to 1973 but only 0.8 percent during 1973 to 1993 and 1.4 percent during 1993 to 2007.

No one believes that the performance of the economy's capital allocation mechanism is responsible for all of this variation. Indeed, a major part of the problem is that no one knows what part of this variation to attribute to the performance of the financial sector. Shifting trends in technology, in the education and training of U.S. workers, in the regulation of business, and in tax and other government policies affecting the overall volume of

investment: all presumably play an important role. So, too, do influences from abroad, especially the often wide fluctuation in the price of oil and other commodities (recall the years of chaos following the OPEC cartel's 1973 and 1979 price hikes) and the intensity of competition from foreign producers (such as Japan in the 1970s and China since the early 1990s).

But our inability to know what part of the deterioration over time in the growth of productivity and living standards to attribute to the performance of the U.S. economy's capital allocation mechanism does not imply that the capital allocation mechanism is doing its job just as well as it always has. It does not mean that all of the deterioration in final outcomes can be attributed to other causes. Not knowing means not knowing.

To determine the adequacy of our current financial system, the first task is to gain a well-grounded quantitative understanding of how successfully the financial sector is allocating our economy's investment capital and how much that allocation of resources costs. The more straightforward task – although no one (to my knowledge) has done it – is to measure the all-in costs of running the capital allocation mechanism. Again, those costs include not just the profits earned and the wages and salaries paid but also the building rents and all other costs of doing business in the U.S. financial sector. (They exclude, however, costs associated with providing other services such as running the payments mechanism.) Just how large does this all-in cost bulk in relation to the total profitability of capital of all kinds invested in the U.S. economy? What fraction of the overall return to our invested capital are we paying for the mechanism that allocates it? Knowing the answer would be a useful

first step. The calculation would require time and painstaking effort. But, conceptually at least, the task is clear-cut. Someone – some part of the Commerce Department’s statistical apparatus perhaps, or a team of university-based researchers – should be commissioned to provide an answer.

A somewhat more challenging task is to include in the overall cost estimate the risk of the occasional meltdowns to which our modern financial system exposes us. Imagine that we wanted to buy insurance against the kind of losses – not just losses of paper wealth but forgone output and therefore incomes and profit – that the recent financial crisis and its aftermath have entailed. How much would we be willing to pay for that insurance? Here, too, the quantitative assessment would require serious work, but the idea behind it is conceptually straightforward.

The most challenging assignment is to evaluate the performance side of the relationship: how well is the financial sector allocating our economy’s investment capital? The challenges to this task are not just computational but conceptual. The chief problem is the absence of an obvious counterfactual: how well is the financial sector allocating our economy’s capital *compared to what*? If we did not rely on our banks, other private lenders, and stock and bond markets to allocate our capital, with all the costs that they entail, how else would we perform this function?

In this respect, today’s protracted debate over how to reform our creaky and increasingly accident-prone financial system is eerily reminiscent of Cold War-era laments about the politics of many European countries. In some countries, even supporters of the then-dominant center-right parties readily acknowledged these groups’ sclerotic character,

empty programs, and chronically corrupt office holders. But, so the argument went, what was the alternative? Only the Communists, and that was no alternative at all.

In the wake of the recent financial crisis, everyone is now keenly aware of many of the shortcomings of our current financial system. But most people also recognize the important role that the financial system plays and the credit it can rightly claim for fostering the dynamic, technologically ever-evolving economy that we have so highly valued during much of our nation’s past. What, then, is the alternative? Even within the highly limited confines of the recent debate over financial regulatory reform, the usual argument is that genuinely binding regulation would stifle the markets’ ability to do their job, while softer attempts at new regulation could drive our vital institutions to relocate in more lightly policed jurisdictions offshore.

The more fundamental task of positing an alternative against which to assess how well our financial markets are allocating our economy’s scarce capital is a far greater challenge. Surely no one wants to contemplate central planning. The technological stagnation and grand-scale dissipation of resources in economies that have tried that route are all too familiar. The resulting conceptual challenge is to measure the efficiency of our system for allocating investment capital in the absence of a clearly specified alternative. Suppose, for example, that the U.S. financial sector had been smaller, and in some specific ways. What if some specific market, or some financial instrument – say, mortgage-backed collateralized debt obligations (CDOs) – did not exist? What would have been different – other than the wages and salaries and rents saved? Would the total amount of capital we invested in our economy have been different? Would

the allocation of our capital among the multitude of competing uses, and therefore the performance of U.S. industry, have differed? The questions may seem relatively straightforward; the route to shedding quantitative light on them is highly complex.

Yet another way to approach this challenging assignment is historical comparison (pertaining to recent history, that is): how has the result of the U.S. economy's allocation of capital over, say, the most recent twenty years stacked up against that of the prior twenty-year period? We know that our capital allocation process has become much more expensive over the more recent period. But has it become more efficient, too? Might it even have become less efficient? Without the necessary research, the only honest answer is that no one knows.

There are, however, at least a few concrete steps we can take in the meanwhile, in the absence of new empirical measurement and new conceptual thinking that may provide sufficient basis for some more fundamental reassessment. First, we can move the liability for the damage banks and other institutions incur when their bets go wrong away from the taxpayer and onto the financial sector's account. Several recent proposals (alas, none included in the recently enacted financial reform legislation) have this aim. For example, one proposal would require financial firms to issue at least some fraction of the liabilities with which they fund themselves in a form that would automatically be converted into equity shares in the event of a specified erosion in the firm's capital position. Presumably, these liabilities would be more expensive for firms to issue than ordinary bonds or commercial paper. The cost difference would be a

form of insurance premium; in effect, the issuing firm would be buying its "rescue insurance" from the market rather than receiving it free from taxpayers.

Critics of this proposal object that some firms might not be able to make ends meet if they had to finance themselves to some extent in this way. But this argument is just another way of saying that a firm would be unable to survive without the subsidy provided by taxpayers in the form of the rescue insurance that it now receives for free. In this case, perhaps the firm should not be in business. Alternatively, if the service the firm provides is sufficiently essential that the taxpayers *should* subsidize it, the fact that they are doing so bears potential implications for such matters as how the firm is allowed to price its product, how the firm and the taxpayers share the returns when its bets go right, and how the operators of a publicly subsidized activity are paid.

A second familiar proposal that does not hinge on any profound new thinking or even any new measurements of costs and benefits is to discourage some forms of trading activity that absorb large amounts of resources, and expose the economy to serious risks, but clearly serve little economic function. Perversely, at the same time that the costs of running the financial sector have mounted in recent years – costs that include the absorption of ever more of the economy's best young talent – the highest rewards, and therefore the greatest attraction for this scarce human resource, have been in some of the areas of financial activity that contribute the least. The leading example in recent years is high-speed trading that exploits computer-based technologies to earn profits from small and fleeting departures of securities prices from normal patterns.

Given the salaries and bonuses that firms in this line of business pay, it is no surprise that many of our country's best young mathematicians and physicists – graduates whose education has been paid for mostly by either government funds or university endowments – flock to them. Yet the activity of resolving micro-departures of securities prices within a nanosecond time frame adds little to the financial system's ability to perform any of its economic functions. At the same time, should one of these firms fall into major difficulty, the financial markets as a whole are then exposed to substantial risk. (In 1998, when Long-Term Capital Management was unable to meet its obligations, the Federal Reserve ended up organizing a consortium of banks to take over the firm; unlike in the more recent crisis, that rescue involved no direct use of taxpayer money.)

One strategy for reducing this form of economically useless activity, proposed years ago by Yale University economist James Tobin, would be to impose a tiny tax either on all transactions that are reversed within some stated period of time or, more simply, on all financial transactions. For an investor buying a stock in the hope that its price will move significantly higher over some period of time, the tax would be a trivial subtraction from the return to be earned. For a firm operating a computerized trading system that moves in and out of thousands of companies' shares many times per day, the tax would be a significant impediment. Further, the tax revenue could be used to reduce the tax rate on long-term capital gains (under current law, "long-term" means more than one year), thereby discouraging economically unproductive trading while at the same time decreasing the burden of taxation on the kind of investment activity that presumably does help allocate the economy's capital.

As is the case for many such regulations, a transactions tax of this form would be difficult to implement unless other countries followed suit; no one wants to see U.S. firms merely move offshore. But many other countries now face these same problems, and therefore have similar incentives to act. International coordination presents a challenge, but it need not be a decisive impediment.

Third, public policy could distinguish between losses incurred that are the financial counterpart of genuine losses of wealth to the economy and losses that are merely the losing side of zero-sum bets (in which for every dollar lost by one institution, someone else has won a dollar). This distinction could be reflected in financial firms' accounting and capital requirements, in how these firms are taxed, and even in deciding which ones to rescue in the event of a crisis.

The U.S. economy's total physical wealth, \$57 trillion (\$186,000 per person) at year-end 2008, consisted of the houses, apartment buildings, factories, office buildings, machines, inventories, computers, software, automobiles, and other durable goods owned by all families and businesses combined, as well as the offices, school buildings, hospitals, military installations, and the military and civilian equipment owned by government at all levels. Buildings, both private and government (excluding only the land value of government buildings), made up \$44 trillion of this total. Equipment and software totaled \$7 trillion, consumer durables of all kinds \$5 trillion, and business inventories another \$2 trillion. Intangible assets, such as the education that makes the U.S. workforce so productive, scientific patents, and other aspects of business know-how, are probably just as important but are nearly impossible to measure in dollar terms. The value of a corporation's intangible assets

is presumably built into its share price, but separating this component from the rest of what a firm is worth – including the factories and machines that it owns – is problematic not just practically but conceptually as well.

In addition, of course, U.S. households, businesses, and governments all own lots of financial instruments – as of year-end 2008, \$140 trillion of them. The \$16 trillion of corporate stock outstanding (down from \$26 trillion just two years earlier) represented ownership of the underlying firms' assets and earnings and thus constituted a claim against one key component of the economy's overall wealth. But most other financial instruments traded in U.S. markets are not part of our overall wealth. The great majority are debt instruments, representing one party's asset but another's liability. They are wealth in the eyes of whoever owns them, but for the economy as a whole, the owner's wealth is offset by the debtor's obligation. Even bank deposits are not wealth for the economy as a whole: they are assets to the depositors, to be sure, but liabilities to the bank.

Following the events of the recent crisis, the most obvious example of a loss that constitutes a genuine loss of wealth to the economy is a fall in the price of an individual's house. When the value of a house falls, that is a loss of wealth to the economy as a whole. If the homeowner continues to service the mortgage (or owned the house free and clear), he or she bears the entire loss; the person's net worth is diminished by the full amount of the price decline. If the homeowner defaults on the loan, then someone else – either the bank that originally lent the money or some investor to which the bank has by now sold the loan – also bears part of the loss. If the government steps in and reimburses the bank or investor, then taxpayers bear part of it, too. Much

of the debate surrounding the recent crisis is about how these losses should be divided among homeowners, banks, loan-purchasing investors, and the taxpayers. But however the losses are divided, someone must bear them, and the U.S. economy is poorer because of it.

By contrast, many of the largest losses that U.S. financial institutions sustained in the recent crisis had nothing to do with losses of wealth to the economy. They were merely the losing side of zero-sum bets, in which one side turned out to be right and the other wrong, and the amount the winner won was identical to what the loser lost. The most transparent example of this phenomenon, and the most important in the crisis, is credit default swaps.

A credit default swap is a contract that pays off if a designated company (neither of the two parties entering into the contract) defaults on its debt. When a bank that has lent money to a company uses a credit default swap to protect itself from the loss that it would incur if the company were to default on its loan, the loss that the other party to the contract incurs if the company defaults is a reflection of a loss of wealth to the economy – the creditors' share of the decline in the company's overall value. But in the case of most credit default swaps traded in the United States today, the volume outstanding far exceeds the amount of debt being insured. As of 2009, the value of credit swaps outstanding in U.S. markets was \$36 trillion – three times the entire amount of bonds issued by all U.S. corporations combined and a far larger multiple of the indebtedness of the specific companies against which the swap contracts were written. The vast majority of these swaps, therefore, had nothing to do with how participants in the financial markets spread the risk of genuine losses of wealth. Instead, their purpose was

simply to create gains for the firms that bet correctly on how the contracts' prices would move, exactly matched by losses for whoever bet incorrectly on the other side.

The distinction between losses that reflect actual declines in the economy's wealth and what are instead merely the losing side of zero-sum bets has a number of potentially important policy implications. There is no reason that either financial accounting or the tax code need treat the two in the same way. (Individual taxpayers, for example, do not get the same deduction for their Las Vegas gambling losses that they do for losses on their stock portfolios.) The distinction should also matter for regulatory purposes. In particular, firms that might be eligible for taxpayer rescue if they get into trouble should be restricted in how much exposure to this kind of loss they are allowed to take on. Finally, once a financial institution does become impaired, whether it has done so as a result of sharing in genuine losses to the economy or merely from having chosen the

losing side of a zero-sum bet should matter in deciding whether its situation merits a taxpayer bailout.

These measures, and others like them, would discourage the financial activities that seem least likely to contribute to the economic function we look to the financial sector to perform. They would also make the financial sector less costly to operate, in that some parts of what our financial firms currently do – parts that add little or no economic value – would be curtailed. Some of these measures would also help reduce the risk of events, like the current crisis, that impose costs in the form of either the incomes and profits lost during a resulting downturn or the burden imposed on taxpayers.

The more fundamental need, however, is to assess how well our financial system is performing its most basic function – allocating our scarce investment capital – and at what cost. Without that information, neither a defense of the status quo nor the consideration of potential alternatives has much basis.

*Is our
financial
system
serving us
well?*

ENDNOTES

¹ For an analysis of Smith's views, see Hugh Rockoff, "Upon Daedalian Wings of Paper Money: Adam Smith and the Crisis of 1772," NBER working paper no. 15594 (National Bureau of Economic Research, December 2009). I first became aware of Hume's letter to Smith by reading Rockoff's paper.

Robert M. Solow

The bigger they are . . .

Any discussion of financial policy and regulation should begin with an urgent reminder that the financial system is a means, not an end. Otherwise, it is all too easy to become wholly engrossed in the hopes and fears, successes and failures, of financial enterprises and the people who love them, as if that were what really matters.

One socially useful function of the financial system is to intermediate between savers and investors. Many diverse individuals, enterprises, and other institutions save – spend less on their current needs than they take in – and it is economically important that their savings be made available to those firms, governments, investors, and other units in the financial system that can make the most profitable (or otherwise valuable) use of such savings. Because most savers lack the information and understanding they would need and because they cannot easily diversify, financial institutions perform this function for them. When something hinders the performance of the financial system, the “real” economy of production and employment suffers. The economy invests too little or too

much, or it invests in the wrong industries. If it is true today that many viable businesses are unable to obtain credit on reasonable terms, the system most likely is not functioning well.

The other socially useful function of the financial system is more complicated and recondite. In the course of real economic life, an enormous variety of risks arises. Bank A may have made a large loan to company B, with the survival of both of them depending on the uncertain success of B’s new line of products. A retired couple with no heirs has to allocate their accumulated savings over their uncertain lifetimes; if they spend too much, they may run out of funds and suffer, and if they spend too little, they may die with useless wealth, having skimmed their golden years.

Some individuals and institutions don’t mind bearing economic risk because their attitudes, their wealth, the nature of their incomes, or their ability to diversify makes it relatively easy. There are also those whose circumstances make substantial risk-bearing painful or intolerable. The financial system can arrange to transfer many risks from the second group to the first, with appropriate compensation all around. Consequently, the real economy works better. Company C may have the ideas and the skills to undertake some-

© 2010 by the American Academy of Arts & Sciences

thing potentially very valuable but cannot bear the inevitable risks; something useful may happen only if the risks can be off-loaded.

But a complication arises: a financial system that is elaborate enough to do this job of reallocating the risks of real economic life is also capable of *creating* risks that have no connection to real economic life, rather like gambling at a casino or betting on football games. For instance, recall the “credit default swap” (CDS) that played such a central role in the AIG debacle. Suppose that lender D has made a large loan to the company C mentioned earlier. The loan seems worth making, but the risk of default is more than D can bear. The CDS is a way of spreading that risk around. D pays E, F, G, and so on a fixed annual fee, and E, F, G each agree to pay something to D if and only if C defaults. The risks associated with C’s business have been transferred to a willing home with E, F, G. This form of insurance allows the real economy to take advantage of opportunities that might otherwise go to waste.

Once the concept of the CDS is available, there is nothing to prevent H and J from writing the same contract: H pays J a fixed premium and J pays H if C defaults on its loan from D. Now H and J are simply making a bet on the outcome of the C-D transaction, though neither of them has any connection with C’s business venture. This is called a “naked CDS,” and there have been many of them. The functioning of the real economy is in no way improved by this transaction, which has merely created a risk that was not there before, and would go away if this transaction were canceled. Moreover, such a transaction would likely not be valid in a normal insurance context. I could not buy insurance against the possibility of a fire destroying someone else’s house; I have no “insurable interest” in

the house, and the contract could not be enforced. The dictionary definition of “insurable interest” is an interest (as based on blood tie or likelihood of financial injury) that is judged to give an insurance applicant a legal right to enforce the insurance contract against the objection that it is a wagering contract and therefore contrary to public policy.

With this background in mind, I turn to policy issues and the “too big to fail” (TBTF) question. Economic policy is often more complicated than it looks. Any significant policy action creates winners and losers, even if the distributional effects are not part of the intended purpose of the policy. For analytical purposes, economists usually avoid these distributional side effects by imagining that they can be canceled by a well-chosen set of lump-sum taxes and transfers. Lump-sum taxes and transfers are those that cannot be avoided or enhanced by any deliberate act of the taxpayer or beneficiary; there are no incentive effects on behavior. But this is a purely imaginary fix. Lump-sum taxes and transfers are implausible in practice.

The vehement backlash provoked by the taxpayer-financed bailout of large financial institutions in the course of the recent meltdown and the ensuing recession illustrates this problem. Even if the bailout was necessary to fend off a much more damaging economic collapse, innocent bystanders resent seeing taxpayers’ money in the pockets of the very bankers, stockholders, and creditors whose greed, shortsightedness, and overconfidence brought on and deepened the recession. Such political-economy considerations are an ever-present constraint on practical economic policy.

All of this is relevant to a discussion of the issue familiarly summarized by the catch phrase “too big to fail.” In the run-up to the recent bailouts, the responsible

federal agencies were faced with the potential failure (inability to meet contractual obligations) of some very large financial institutions that were interconnected with other large and small financial institutions through lender-borrower and analogous relations. The threat of their falling into default can so threaten the solvency of their creditors (and *their* creditors, and so on) that much of the financial machinery might grind to a halt, and with it much of the economy. These institutions, and some nonfinancial corporations, were regarded as so central to the economic life of the country that they could not be allowed to fail.

How can the likelihood of such situations be eliminated or minimized in the future? A necessary first step is to consider closely what makes a financial institution TBTF.¹ Size, certainly, is part of the picture. The insolvency of a few small banks or nonbank financial institutions does not threaten a breakdown of the system that provides credit for viable businesses and redistributes the risks of real economic activity. Given the existence of federal deposit insurance, the potential losers in the failure of a smaller bank are mainly the stockholders and the nondeposit creditors; prudence should impel these parties to take into account the possibility of such contingent business losses when they buy stock and make loans. In practice, the regulator of a “problem bank” often arranges for it to be taken over by a stronger neighbor, thus minimizing disruption.

Should nature be allowed to take its course in the case of very large banks and nonbanks? If they are too big to be taken over, should they just be allowed to go broke? A practical obstacle has stood in the way, at least in the past. Large banks often operate with large leverage; in other words, they borrow a

lot in order to acquire assets in amounts that far exceed their own capital. They hope to profit from the difference between the cost of borrowing and the higher return on the assets they acquire. Since higher return usually goes along with greater risk, there is potential for trouble. Suppose that a bank with thirty-to-one leverage – unexceptional by recent standards – has \$1 billion in capital and has borrowed \$29 billion to acquire \$30 billion in at least slightly risky assets. It takes only a \$1 billion loss to wipe out the owners of the bank; a loss of \$2 billion renders the bank insolvent. On the upside, a gain of \$1 billion doubles the owners’ money, which explains why leverage is so attractive.

As part of their function in mediating between savers and investors, banks are typically engaged in “maturity transformation.” They borrow at short term because savers generally want quick access to their money. But they make longer-term loans because they are financing real business investment. The persistent question, then, is about liquidity, or the ability to convert even sound assets into cash when necessary. In parlous times, liquidity problems can become solvency problems when the soundness of assets is uncertain. Even without much leverage, troubles may arise; greater leverage signals a clear possibility of cascading disaster.

The difficulty with very large banks is not only that they are big, but that they are interconnected with other financial institutions. When large institutions are highly leveraged, the interconnectedness looms as a danger to the whole system. The lenders to a large bank regard those loans as assets. If bad news about the borrowing bank’s assets threatens its solvency, then its lenders see their own balance sheets deteriorating; the value of those putative “assets” becomes

uncertain. But the lending banks have also borrowed, and so a third layer of banks is caught up in uncertainty and pessimism about asset values.

This combination of sheer size, interconnectedness, and leverage can endanger the financial system's ability to perform its socially useful functions. The temptation for large, highly leveraged financial institutions to engage in "wagering contracts contrary to public policy," using borrowed money, adds to the potential for systemic instability without contributing anything to the efficiency of the real economy. These circumstances impel – or force – governments to bail out the banks, essentially to guarantee the value of the assets of financial institutions that are considered TBTF.

It is now widely understood that this kind of situation is fraught with "moral hazard." Highly leveraged purchases of risky assets create opportunities for spectacular profits on the relative small amount of own capital invested. They also create opportunities for disasters so large and extensive as to threaten the functioning of the system. If this threat forces governments to bail out the occasional disasters to protect the creditors, then the opportunities for large profits belong to the risk-takers and the worst of the occasional losses belong to the taxpayers. Banks are encouraged, or rather driven by competition, to take those system-threatening risks, and other banks are encouraged to lend to them for that purpose. They have little or nothing to lose, and a lot to gain. *And the successes probably add little or nothing to the efficiency of the real economy, while the disasters transfer wealth from taxpayers to financiers.* This reality is "heads I win, tails you lose" writ very large.

There is another cost of the TBTF phenomenon that is even less visible,

although it occurs regularly, even in the absence of crisis. A bank or other financial institution that is perceived as TBTF can borrow in the market at a lower interest rate than other, otherwise similar, banks. Compensation for default risk is built into any interest rate; default-free U.S. Treasury bonds, for instance, carry a lower interest rate than corporate bonds of the same maturity. Thus, TBTF banks are subsidized every day by the taxpayer. The subsidy is not borne by taxpayers in the form of a continuing cash outlay; it takes the shape of an implicit promise to bail out a TBTF bank when it might otherwise have to default.

This experience, now so clear in the collective memory, is not only costly to taxpayers, but also hair-raising to workers and small businesses whose livelihoods hang by a thread when the economy threatens to dissolve, and irritating to those who do not like to see high-level vice and stupidity rewarded. Attempts to improve the regulation of the financial system are in development in the United States and Europe. Included in these blueprints for reform are various proposals for dealing with the TBTF problem.

It may be useful to start with the (hopelessly) idealized laissez-faire solution. After all, this is what former Chairman of the Federal Reserve Board Alan Greenspan famously believed in, only to be shocked by the grim reality. Suppose nothing were TBTF; suppose the government could credibly state that it would bail out no failing bank, no matter how big or how interconnected with others. In principle, this stance warns potential creditors that lending to a large (or small) bank with a risky balance sheet is itself an act with considerable downside risk. If the borrowing bank defaults, the creditors will take the loss. Lenders to banks, especially big lenders to big banks, are

sophisticated, knowledgeable people. They know how to read a balance sheet and understand complex securities. They will not endanger their own capital by lending large amounts to a large bank that will use the borrowed capital to buy excessively risky assets. In this view, banks that are too big to fail will not fail or, if they do, their failure will not endanger the system.

There are a few problems with this picture. Perhaps the most important is that governments cannot credibly abjure bailouts. In a capitalist system, even a reasonable balance sheet will carry some risks. When a large bank is on the verge of defaulting, thereby threatening the viability of the financial system, a responsible government cannot step aside – *c'est la vie!* – and let the real economy tumble into depression. Preventing the collapse of the financial system does not imply weakness; the government is doing what has to be done. The notion that the long run is best served by letting two or three catastrophes happen cannot be taken seriously.

Second, there is evidence that those potential creditors are not always as sophisticated and knowledgeable – or as effective – as presumed. They may be prone to act on foolishness, incompetence, laziness, greed, overconfidence, and the herd instinct. Granted that our observations come from a world of moral hazard induced by the TBTF doctrine, one would not be quite comfortable betting the health of the real economy on the unfailing intelligence and self-discipline of real-world financiers. The *laissez-faire* solution, therefore, is probably a nonstarter – and for good reason.

The most direct solution to the TBTF problem would be to disallow the existence of banks that are TBTF in the first place. A regulatory body could require a bank either to divest or sell off some

risky part of its business, or shrink its balance sheet to an acceptable size in some other way, once the bank exceeded a predefined limit. The goal would be to achieve a landscape in which any bank that seems about to fail could be allowed to fail because, by definition, that failure would not threaten the satisfactory functioning of the system.

This proposal is premised on the belief that the expansion of a bank beyond the acceptable size limit brings at best negligible gains in efficiency for the real economy: even if the achievement of TBTF size adds to the private profitability of a financial institution, this private gain does not correspond to any net contribution to society. This argument seems plausible. Indeed, recent history suggests that the main consequence of megasize may be unmanageability. I, for one, have not seen any convincing arguments for real economies of scale at extreme size.

The unmanageability of very large banks reflects something deeper than mere bureaucracy: there is a fundamental incentive problem. Individual traders in a large institution can enrich themselves fantastically by taking on risks whose downsides endanger not themselves but the firm. Not many individual bankruptcies have made the headlines. Better-aligned incentives would help, but such restructuring is not easy in a large, variegated organization run by clever individuals.

Nevertheless, there are genuine problems with this approach to TBTF. If the largest acceptable size is still fairly large, as I imagine it would be, then even if no single bank is TBTF, the threatened failure of two or three large banks would still require the bailout response that the scheme is designed to prevent. Alternatively, the cut-them-down-to-size proposal may be interpreted as the partitioning of a large bank into many small

banks. In that case, (a) some genuine economies of scale in saving-investment intermediation and risk allocation may be lost; and (b) the failure of many small banks – as happened in the 1930s – can be a problem as well. Deconstructing a large financial institution into a number of small ones does not create a cluster of statistically uncorrelated banks, such that mass failure is unlikely. The danger is not analogous to tossing a separate success-or-failure coin for each bank; rather, it stems from the fact that all or most of them are hit simultaneously by a common shock – such as the burst of a housing bubble – and tend to fail together. In this context, imposing a size limit on banks that would otherwise be TBTF can be a helpful and not very costly assist, but is unlikely by itself to solve the problem.

The difficulty is that size is functioning as a symptom of something else, and it is that underlying factor that really creates the problem. Imagine a bank or financial institution that simply invests its owners' wealth or capital in a collection of business ventures of varying risk. The owners' profit is the return on those investments minus any administrative costs. The size of such an institution is of little consequence. It would not "fail" unless all or most of its investments failed. Even if that unlikely event were to happen, the only consequence would be that the owners (shareholders) would have lost their stakes. That might be hard on their heirs, but not on the financial system. The bank is interconnected in the sense that it has lent to many enterprises, but not in the relevant sense that its debts appear as assets on the balance sheets of other banks. In the limiting case that it has no debts, its leverage ratio would be one to one.

In fact, it is leverage – borrowing in order to buy risky assets – that is the

fundamental problem. Extreme leverage underlies extreme bigness. The megabanks would not be nearly as oversized, or as interconnected in the relevant sense, without leverage ratios of twenty-five to one, thirty to one, or greater. Therefore, the best way to control the TBTF problem may be to control leverage – which is no easy task.

In principle, limitations on leverage should be conditioned on the riskiness of the assets to be acquired. Any assessment of riskiness will inevitably contain a large element of judgment, presumably to be exercised by a changing cast of regulators: some strong, some weak, some strict, some lax. Practice is likely to be even more unreliable. Any collection of specific criteria and regulations, especially if embalmed in statute or code, will be vulnerable to the attentions of clever lawyers and creative accountants. Regulators are usually unable to keep up with the athleticism of the highly motivated. Thus, controlling leverage necessarily involves three steps: it must be cut back sharply, regulated closely, and safeguarded against evasions and loopholes.

What form could those fail-safe preparations take? This is perhaps a good place to mention the Volcker Rule, which has been on the radar since it was proposed by Paul Volcker, Chairman of the Economic Recovery Advisory Board under President Obama. The general idea of Volcker's proposal is that "true" banks, that is, institutions that accept deposits and make loans, should be prohibited from trading in securities for their own account (though they might be permitted to do so as agents for their customers). This proposal has implications for "true" commercial banks.

First, if implemented, the Volcker Rule would effectively control the leverage assumed by depositary institutions. Some leverage is necessary: a bank that subsists

by taking deposits and making loans earns its profit only through the difference between the interest rate it charges for loans and the (lower) rate that it pays its depositors. Unless its earning assets (its loans) exceed its capital, it probably cannot earn a high enough return on its capital to keep itself in business. But such banks are thoroughly regulated already; an overextended bank would be reined in by its regulator. In any case, what tempts a bank to leverage itself excessively (that is, to borrow in the capital market) is the prospect of large trading profits, which would be forbidden under the Volcker Rule. There is not much profit in borrowing at risk-adjusted capital-market rates in order to lend at what would be roughly risk-adjusted capital-market rates.

Second, there is particular reason to limit the leverage of commercial banks. The danger of high leverage is that a small adversity can bankrupt a highly levered institution. If that institution is a commercial bank, there is automatic disruption of an important channel through which ordinary businesses – and consumers – routinely obtain credit to carry out standard activities. Thus, adverse effects on the real economy are immediate. When Gary Stern, then-president of the Federal Reserve Bank of Minneapolis, published his book *Too Big to Fail* in 2004, he was thinking entirely in terms of commercial banks. Today, the TBTF problem is much more a matter of non-bank financial institutions – investment banks, insurance companies, and so on – which would not come under the Volcker Rule. The Volcker Rule would be a useful part of a comprehensive attempt to protect the real economy from financial instability, but it cannot be the whole story, nor did Paul Volcker intend it to be.

Suppose we accept the inevitable: regulators are fallible or worse, and statuto-

ry requirements can be gamed. So from time to time banks that are TBTF will be headed for failure. The real economy has to be protected from damaging disruption. Is there an alternative to bailing out the banks' uninsured creditors at the expense of the taxpayers, with all the moral hazard problems that a bailout entails?

There is broad agreement that nontrivial costs must be routinely inflicted on creditors to take away their free ride and induce them to exercise some discipline on the risk-taking of large banks to which they lend. Ordinary bankruptcy serves as a deterrent – in the laissez-faire process mentioned earlier – but it is plausibly argued that ordinary bankruptcy is a process so lengthy and its outcome so uncertain that it makes the real economy vulnerable to disruption. Several schemes have been suggested that would deal with TBTF by prepackaging and automating a form of bankruptcy-equivalent that would progress quickly and predictably.

One class of such schemes comes under the picturesque heading of a “living will.” This provision would require that every large financial institution – and maybe some nonfinancial firms – file a detailed, binding statement of how its assets will be allocated in case of impending default: after allowing for insured creditors, like ordinary depositors, the firm would designate which party has first claim on the remaining assets, which party comes next, until the common equity shareholders bring up the absolute end of the line and presumably get nothing at all.

If sufficiently large, leveraged, and interconnected banks were to fail, even with living wills, the TBTF problem would not quite go away. Many creditors would find their own balance sheets damaged, and therefore so would *their* creditors, and so on. Healthy financial activity could be

inhibited. The idea behind the living will is that with the consequences defined in advance and without ambiguity, potential creditors would shy away from highly leveraged, risky borrowers and either refuse to lend or demand such high interest rates that the borrowers themselves would find the game unprofitable. The proposal is fundamentally an attempt to make market control of leverage more effective.

Another version of this general idea is sometimes called “bailing in.” Instead of the government holding the bag, each class of creditor, preferred-stock owner, and so on would be contractually obligated in a certain order and under certain conditions to convert its claim to common equity. When the original common stockholders are wiped out, the next designated class would walk the plank. Eventually, the lowest-ranked surviving class of creditors would become the equity owners of the business. Again, as with the living will, the procedure is presumed to have adequate clarity and visibility to discourage the capital market’s willingness to accept highly leveraged risk-taking.

Yet another version would require banks, in addition to holding a certain proportion of equity capital against their liabilities, to issue a certain proportion of contingent bonds – contingent in the sense that they automatically convert to equity shares when the wolf appears at the door. This proposal has the advantage that such contingent bonds would certainly bear a higher rate of interest than bonds without the contingency. Borrowing banks with higher leverage would incur a higher cost of finance.

The last proposal of this general class differs from the others because it involves the federal government directly. In this scheme, a regulatory body would have the “resolution authority” to step in early

in the case that a “systemically important” (that is, TBTF) bank were moving toward default. The authority would essentially take over the bank, replace some or all of the management, wipe out some or all of the equity, and impose necessary losses on creditors.

The goal would be a quick dispatch that would keep the bank operating with little or no interruption and allow it to emerge as a viable institution. To this end, the authority might need to have the resources to inject new capital into the bank, acquiring an ownership interest in return – preferred, convertible, or even common shares – that could later be sold in the market when the resolved (that is, newly solvent) bank’s prospects have been restored. The authority would need money, perhaps in hefty amounts. The bill that originally passed the House proposed to fund the resolution authority by levying a fee on the (risk-adjusted) assets of large financial institutions. In that way, the financial system would bear the costs of its own risk-taking. This stipulation would make borrowing more expensive for all firms. Why not? Those extra borrowing costs are a measure of what taxpayers are bearing now.

The House proposal has attracted much opposition; at this writing it looks as if it will disappear. The Obama administration did not favor it. Critics have argued that the very existence of such a fund, even a very small one, appears to validate the idea of bailing out the TBTF banks. My own hypothesis is that this fund is not important enough to justify the uproar (which may be mainly decoy, anyway). The point of the fund is that once a really big domino – a TBTF domino – is about to fall, enough money will be found to prevent collapse of the real economy. If that outcome is to be avoided, the real defense must occur at an earlier stage. A functioning resolution authority seems

like an excellent idea. At a minimum, it could backstop the more market-oriented schemes described earlier.

My own (weak) preference is for a combination of the contingent-bond device *and* a federally operated resolution authority. It would be useful to have the interest cost of risky behavior quoted daily (rather than having to be inferred) as a clear signal to lenders and borrowers. In addition, wherever a fail-safe is possible, it should exist, provided that a resolution authority enforces the measure in a somewhat orderly way.

The above discussion centers on only one aspect of stabilizing the financial system, not on the full range of considerations. Still, the TBTF phenomenon was a critical part of the recent economic downturn and, as a result, deserves careful attention. What are the main lessons to be gleaned from the TBTF problem?

Excessive leverage appears to be the key destabilizer, and limiting it is the main remedy. Limiting leverage will tend to shrink the financial system, but if, as I suspect, there is a sizable amount of financial activity that adds little or nothing (or perhaps less than that) to the efficien-

cy of the real economy, then we should cheerfully let it shrink. (If a reduced financial sector leads more clever graduating seniors to materials science and fewer to investment banking, all the better.)

There are several kinds of regulatory reform that could place limits on leverage, preserve the essential functions of finance, and diminish the burden on taxpayers. Some of them are more market-oriented, others more state-oriented. A well-designed system could make use of several of them, as long as priority is clear. One reason for welcoming the presence of several layers of protection is that *laissez-faire* won't do, paper regulations are vulnerable to the creation and exploitation of loopholes, and the political process will sometimes lead to neurasthenic regulators. We are probably better off with defense in depth, even with the risk of some bureaucratic interference.

It is worth adding that international cooperation and alignment are necessary in a globalized world. The temptation to set up pseudo-shop in places where regulations are feeblest would be irresistible – in which case the alternative to tough international agreement could be the Cayman Islands.

ENDNOTES

¹ In this essay, I indiscriminately refer to financial institutions as “banks,” ignoring the distinction between commercial banking – taking deposits and making loans – and investment banking, as well as differences between other kinds of financial firms that play a role in the economy's flow of credit, such as insurance companies. When the distinctions are important, I refer to them explicitly.

Luigi Zingales

Learning to live with not-so-efficient markets

For most non-economists, the biggest intellectual casualty of the 2007–2008 financial crisis is the efficient market theory (EMT). Newspapers and talk shows have analyzed the theory's apparent demise; new books with titles like *How Markets Fail*, *The Myth of the Rational Market*, and *A Failure of Capitalism* abound. Yet, in academia, the EMT has been challenged since the 1987 stock market crash, and the theoretical and empirical shortcomings of the theory have been well established. The marginal contribution of the 2007–2008 financial crisis was to weaken the already-losing side of hardcore believers in the EMT.

Thus far, the recent crisis has not provided any crucial new evidence on the deviations of markets from fundamentals, only evidence of the costs of these deviations, especially when coupled with very high leverage. The rise and collapse of the real estate market, with all the consequences these events produce, are much easier to explain as the result of variation in fundamentals than is a 22.6 percent stock market drop in a day with no major news. Hence, those who have remained unchanged in their beliefs since the 1987 stock market crash have been little af-

fectured in their faith by the 2007–2008 financial crisis or by the evidence collected in the intervening twenty years. Yet the increasing proportion of academics who question the EMT have found in the recent crisis a painful example of the costs of ignoring potential market inefficiencies. A major rethinking of the validity of the theory's implications is all but inevitable.

From courtrooms to boardrooms, from policy cabinets to classrooms, the EMT provided the intellectual foundation for an entire generation. The current debate has thrown this foundation into question. Rethinking the theory, however, does not mean abandoning it: the approach still holds many useful insights. Because it has provided a coherent framework with great practical value, rethinking the EMT will be painful; nonetheless, it is necessary. This journey will inevitably be full of mistakes, but the biggest mistake of all would be not to undertake it.

While its intellectual origin can be traced back to the early twentieth century, the EMT hypothesis gained importance in academia in the mid-1960s. In its early formulation, the EMT was simply the idea that if prices adjust rapidly to new information, they should be unpredictable. This hypothesis, validated

by economist Eugene Fama's dissertation at the University of Chicago in 1965, has three very important implications. First, if prices are unpredictable, then all technical analysis, which uses past price behavior to predict future trends, is useless. Second, if prices are unpredictable, then investors should not pay for investment advice and active money management, but should simply buy mutual funds that passively track the stock market index. Last but not least, if future prices are unpredictable, current prices are the best available estimate of the true value of any asset. If this were not the case, so the theory goes, speculators would intervene to drive prices equal to fundamentals.

Early on, the EMT enjoyed wide empirical support. Subjected to rigorous testing, most technical rules for investing in the stock market, it was found, did not outperform a random strategy. On average, actively managed mutual funds did not outperform a stock market index, and there was no evidence of any persistence among the winners. The only evidence of persistence in mutual funds' performance was in the negative domain: this year's laggards were likely to be next year's, too, because they were charging above-average fees. If one combines a random performance with above-average fees, statistically the consequence is an abnormally low after-fee performance!

The only implication of the EMT that was difficult to test was the last one: that stock prices are the best available estimate of fundamentals. Fama, the main supporter of the EMT, said in 1970:

The primary role of the capital [stock] market is allocation of ownership of the economy's capital stock. In general terms, the ideal is a market in which prices provide accurate signals for resource alloca-

tion: that is, a market in which firms can make production-investment decisions, and investors can choose among the securities that represent ownership of firms' activities under the assumption that securities prices at any time "fully reflect" all available information.¹

This implication proved difficult to test because one needs a model of fundamental values in order to test whether prices are equal to fundamentals. Yet if one rejects this equality, it is unclear whether that entails rejecting the asset-pricing model that assesses the fundamental value or rejecting the EMT. This joint hypothesis problem has made it difficult to reject the most important implication of the EMT: that current prices are the best estimate of future fundamental values.

Since Friedrich von Hayek's work, economists have recognized the importance of prices in aggregating information and in directing the allocation of resources. Financial markets play a crucial role in this respect: stock prices provide a major indicator of where resources should be allocated. In the late 1990s, seasoned executives abandoned their corporate positions and M.B.A. students their classrooms to join Internet start-ups because of the high valuations these ventures commanded in the marketplace. Thus, market efficiency is not just about how difficult it is to make money in the stock market, but how the stock market provides useful information about how resources should be allocated.

In 1978, empirical support for the EMT was such that Michael Jensen, one of the most prominent researchers in finance, wrote that "there is no other proposition in economics which has more solid empirical evidence supporting it."² In science, however, hypotheses cannot be

proved, only rejected. As soon as the EMT became widely accepted, researchers began gathering evidence against it. By 1983, the *Journal of Financial Economics*, the primary journal in finance, dedicated an entire issue to financial “anomalies,” a politically correct term for violations of the EMT.

The first challenge to the EMT arose from the fact that prices are not unpredictable. With longer and more frequent time series, researchers established that there is some predictability in prices in the very short term and even in the very long term. (These latter findings, however, were less statistically robust, due to the lack of long-term time-series data.) As is typical of scientific progress, these initial findings led to modification of the EMT rather than its rejection. In an environment where investors are risk averse, today’s prices do not simply reflect expected future prices; rather, they reflect the expected value for the investors of these future prices. Because investors are risk-averse, this expected value reflects not only the expectation of future prices but also their variability, which impacts investors’ utility. Hence, if spikes in risk aversion (intense moments of panic) are followed by reversals, today’s change in prices can predict tomorrow’s change in prices without violating the EMT. Risk aversion is hard to observe and measure, and this justification has protected the EMT from being clearly rejected.

The second blow to the EMT came from the few cases in which fundamental value can be measured without reference to an underlying asset-pricing model. The primary case is closed-end mutual funds. Unlike the more popular open-ended mutual funds, closed-end mutual funds cannot be redeemed daily for their net asset value; they can only be bought and sold in the stock market. Their price, however, tends to differ from the value

of the underlying assets, which is reported daily, because their assets are traded stock. Even more puzzling is the fact that the difference between price and net asset value oscillates over time, with occasional premiums and more frequent discounts, often reaching 30 percent. This evidence, however, was not considered a clear rejection of the EMT, since the difference between closed-end funds prices and net assets value could reflect some unique ability of the closed-end fund manager (when this difference is positive) or some value appropriated by the manager (when negative).

The EMT was also attacked from a theoretical point of view. If prices indeed reflect all available information – as the EMT claims – there are no incentives for market participants to collect any information. How can the market reflect all available information if no one has any interest in collecting it? In this respect, the EMT was modified to resist the challenge. Prices, it was argued, should reflect information up to the point where the marginal benefits of acting on the information (the expected profits to be made) do not exceed the marginal costs of collecting it.

Finally, the link between the various implications was broken. Unpredictability of stock prices does not equate with prices equal to fundamentals; the difference can be a slow-moving fad that cannot be easily detected with the short time series available.

As economists Andrei Shleifer and Larry Summers, two of the main critics of the EMT, wrote in 1990, “The stock in the efficient markets hypothesis – at least as it has traditionally been formulated – crashed along with the rest of the market on October 19, 1987 and its recovery has been less dramatic than that of the rest of the market.”³ EMT

stock crashed because of the lack of any plausible theory of the stock market crash. What fundamental news could have caused aggregate stock prices to drop 22.6 percent in one day? Even after a very aggressive search, the worst the supporters of the EMT could find were a potential announcement of a larger-than-expected trade deficit, the revelation that a key committee of the U.S. Congress would support the elimination of tax benefits for leveraged buyouts, and press speculation that the Federal Reserve would raise its discount rate. How could this news have changed fundamental values by 22.6 percent in a day? This single fact in isolation accounted for more than all the empirical evidence against the EMT collected up to that point.

Not surprisingly, shortly after the 1987 stock market crash the intellectual foundations of the EMT were challenged. Besides empirical evidence, the strongest argument in favor of the EMT was the logical argument originally advanced by Milton Friedman. When prices deviate from their fundamental value, rational investors can profit from these deviations by buying the undervalued asset and selling the overvalued one. The very act of buying undervalued assets and selling overvalued ones pushes prices closer to their fundamental values. No matter how many irrational traders there are in the marketplace, as long as there is a sufficient number of rational investors with access to large pools of money, they will drive prices close to fundamentals. Where but Wall Street can you expect to find smart, rational people with access to plenty of money?

The theoretical counterattack to Friedman's argument took place along two directives. First, even rich and rational agents are risk averse. If "[m]arkets can

remain irrational far longer than you or I can remain solvent," as John Maynard Keynes remarked, then rational but risk-averse individuals may choose not to undertake some arbitrage, even when these arbitrages, carried over a sufficiently long period of time, are risk free. Consider the case of closed-end mutual funds selling at a discount with respect to their net asset value. If carried over until the fund is liquidated, buying the closed-end fund and selling short its assets represent riskless arbitrage, that is, a way to make money without any risk. Yet in any one period, the arbitrageur can incur losses if the deviation widens before it shrinks.

If the risk that the difference between a closed mutual fund price and its net asset value widens is not correlated with the risk that other deviations widen, then this risk can be diversified away by creating large portfolios of these trades. If these events are correlated, though, this diversification will not work, and this risk can prevent rational traders from bringing prices back to their fundamental levels.

The other direction of attack concentrated on potential arbitrageurs' access to funds. Few investors are as rich as George Soros; most need to raise other people's money to trade. In allocating their money to various potential traders, though, the investors look at their performance. This induces a very short horizon even among farsighted rational traders. If a trader risks seeing the money he invests withdrawn every time his performance deteriorates for a few months, his time horizon shrinks to a few months. Furthermore, if his performance is judged vis-à-vis an index, the trader will become very cautious in deviating his investment away from his benchmark index. In other words, if an equity fund's performance is benchmarked against the S&P 500 index, the

manager will be very cautious about taking any arbitrage opportunity that induces him to tilt his portfolio away from S&P 500 stocks. By staying mostly invested in S&P 500 stocks, he faces very little risk. If those stocks go down, so does his benchmark; therefore, it is difficult for him to be fired. If he feels that the S&P 500 stocks are overvalued with respect to similar stocks outside the index, he will be reluctant to try to arbitrage that difference away. Even if he is right, but the deviation increases for a few months before it decreases, he will run the risk of being fired. Benchmarking, while optimal from an investor's point of view, might reduce the willingness of traders to lean against broad movements of stock prices away from fundamentals.

Long before the recent crisis, a majority of academic economists – at least those trained after the 1987 stock market crash – were ready to admit that stock prices could deviate from fundamental value for extended periods of time. The spectacular rise and fall in real estate prices that characterized this recent crisis did very little to change the opinion of academic economists on market efficiency. For one, even the strongest believers in the efficiency of the equity market are ready to concede that there is not a similar presumption for the housing market. Unlike the equity market, the housing market is not populated by smart traders trying to arbitrage away possible differences from fundamentals; most of the buyers and sellers in this market are unsophisticated in their market knowledge. And unlike in the equity market, in the real estate market sophisticated traders face very high cost of arbitrage. If one trader thinks that Microsoft's stock is overvalued with respect to Apple, he will buy Apple and sell short Microsoft. This operation can be completed

in the space of a few seconds with minimal transaction costs. But if a real estate developer thinks that house prices in the suburbs are overvalued, the only opportunity he has to make money is to build more houses in that area. This operation will take years, not seconds, and will involve sizable transaction costs. Thus, Friedman's argument does not apply to the housing market.

Second, most economists were already willing to admit that prices could drift away from fundamentals, and the faith of the remaining economists could not be shaken by the spectacular rise and fall in real estate prices, even if the housing market were similar to the equity market. The reason is that staunch supporters of the EMT are very creative in devising rational explanations for the most extreme events. Even the Internet boom and bust, which most non-economists see as the quintessential bubble, has been justified as the rational expectation of a few of those stocks turning into the new Microsoft. If a 22.6 percent drop in a day with no news does not shake their beliefs, what could?

The dramatic drop in mortgage-backed securities prices during the crisis cannot be considered very strong evidence of market inefficiency either. With the benefit of hindsight, it looks incredible that securities backed by mortgages offered to the riskiest part of the U.S. population could be considered safe, granted the best rating, and priced accordingly. Yet short of a generalized drop in house prices throughout the United States (an event that had not occurred since the Great Depression), those securities appeared to be priced correctly. For markets to be efficient, they do not need to be perfect. Thus, one mistake over the occurrence of a very rare event is hardly evidence against market efficiency.

Paradoxically, the aspect of the crisis that has shocked economists the most has been the diffuse violations of fundamental arbitrage conditions during the peak of the crisis. Take the covered interest rate parity, for example. If I borrow in yen, exchange the proceeds in dollars, invest them in dollars, and then buy yen with delivery when my debt is due to cover my exchange rate risk, I should not be able to make any money. If I can consistently do so, I have found a money machine. For an economist, the violation of this (and other) riskless arbitrages is tantamount to observing numerous \$100 bills on a crowded sidewalk and people passing by without collecting them.

During Fall 2008, many such fundamental relations were violated. The emerging consensus is that these violations were due to a lack of capital in the hands of smart arbitrageurs. Having faced major losses and even larger capital withdrawals by scared investors, smart traders who knew how to exploit these differences were running out of available money. And when there are many opportunities to make money but very little time and money to exploit, smart traders focus on the most profitable one. In other words, if I am running to put out a fire in my house, I might rationally choose to leave \$100 bills on the sidewalk.

While the conditions in Fall 2008 were extreme, they illustrate the importance of access to capital and the limits of the arbitrage process. If limited access to capital can prevent smart traders from exploiting simple, risk-free opportunities, it is not hard to imagine that it might prevent, even in more normal times, the exploitation of complex and risky opportunities. Thus, limited access to capital is a major flaw in Friedman's argument.

What the recent crisis did change is the sense of how costly the violations of

the EMT hypothesis can be, especially when combined with very high leverage. Very few doubted that violations existed. But as long as the documented violations were limited to the fact that stock with positive returns today tend to have a bit of positive momentum in the next six months, or that closed-end funds can trade at a 30 percent discount with respect to their net asset value, this is hardly earth-shattering. For most practical purposes, markets can be considered as if they were efficient.

These minor violations were the easiest ones to document empirically. Unfortunately, the deviations that really matter – like whether the level of stock prices in general is close to fundamental values – are the most difficult to prove because of the joint hypothesis problem: we can always claim that we do not have the adequate model of fundamental values. In other words, this implication of the EMT is not a good theory in the Popperian sense (after philosopher Karl Popper) because it is almost impossible to reject. This is the sense in which the crisis started to change the perspective of people inside and outside academia about the costs associated with ignoring these possible deviations.

In 1996, then-Federal Reserve Chairman Alan Greenspan questioned whether some “irrational exuberance” had taken the stock market to too high a level. The Shiller price-earnings ratio, based on the average of earnings of the previous ten years, was at 25 when Greenspan delivered that speech. It rose 77 percent, to 44.2, at the end of 1999, before it started to drop. During those three years, many commentators ridiculed Greenspan's remarks.

The lesson Greenspan learned was how politically costly it was to lean against the wind. He dutifully applied this lesson when real estate prices rose.

In fact, he elevated the lesson to a principle (the Greenspan Doctrine) that it was not a responsibility of a central banker to try to lean against the formation of potential bubbles. In academia, the staunchest supporter of this approach was a Princeton University macroeconomist little known in the political world at that time: Ben Bernanke! In a 1999 article with fellow economist Mark Gertler, Bernanke analyzed the impact of monetary policy when prices move away from fundamentals. That this contingency was the object of their analysis illustrates how the EMT was losing ground. Their conclusion, however, was that the Fed should not intervene, not only because it is difficult to identify the bubbles but also because “our reading of history is that asset price crashes have done sustained damage to the economy only in cases when monetary policy remained unresponsive or actively reinforced deflationary pressures.” Thus, the case against intervention was not based on the idea that the market always gets it right, but on the premise that the costs of these deviations are relatively minor, with respect to the cost of wrong interventions.

This is what the 2007–2008 financial crisis has changed: the comfortable belief that even if the EMT is not exactly true, it is a sufficiently close approximation to reality that we can safely use it for most practical approximations. This realization is not limited to academia; it is mostly shared by an entire generation of financial economists, accountants, and lawyers who have been trained in the fundamental belief that market prices reflect the best estimates of fundamentals. This belief is the reason why accountants like to use market prices as much as possible in compiling companies’ financial statements. It is the rea-

son why lawyers use stock price movements to assess damages. It is the reason why corporate boards judge and compensate their executives on the basis of changes in stock prices. It is the reason why in most business schools M.B.A.s are taught that any active investment management is a waste of time and resources and that chief financial officers should not try to outguess the market in their financing decisions.

Besides the intellectual appeal of Friedman’s argument and the supporting empirical evidence, the EMT was also very attractive for its practical implications. By providing an objective metric for performance, the shared belief in efficient markets made everyone’s life easier. It was not one expert’s opinion against another’s: the market was the absolute and impartial judge.

Once we lose this objective criterion, we are back in the world of self-interested opinions. All corporate executives will argue that the market undervalues their stock and that any drop in their stock price is due to market irrationality (while any increase is entirely justified by their actions). All politicians will claim that the rise in the yield of their government’s bonds is not the result of their profligacy, but of evil speculation or market irrationality (while ready to take credit for any reduction in the yield). Finally, all charlatans will feel entitled to claim that they have an investment strategy that can beat the market systematically. This is the uncharted territory where the crisis leaves us: a world where confidence in the rationality of the market is shaken but where there is no clear, viable alternative.

Eventually, a grander theory will emerge, one that will enable us to understand when we should expect mar-

ket prices to deviate from fundamentals. At the moment, we can grasp only some elements of this theory. Unfortunately, this quasi-EMT will be less elegant and less clear-cut, making its day-to-day application more difficult. As a synthesis between the original EMT and some of the arguments against it, this emerging theory will displease both sides. It will be considered apostasy by true believers in the EMT, while radical opponents will view it as too conservative. Nevertheless, it is not only the right compromise, it is the only reasonable way forward.

The necessary starting point is always Friedman's argument that deviations of prices from fundamentals create money-making opportunities. Yet this argument should be tempered by the understanding that these opportunities are not as easy to capture as \$100 bills on the sidewalk. Even in the simplest cases, lack of capital can prevent arbitrageurs from doing their job, particularly when it comes to leaning against the entire market.

In normal times we can expect relative prices across similar securities to reflect differences in fundamental values. If the price of Google rises while the price of Yahoo drops, we can be fairly confident that the fundamentals of the two companies have moved accordingly. Nevertheless, in the short run, relative prices can deviate from their relative fundamentals, especially when institutional or technical factors create a significant imbalance between demand and supply. Consider, for instance, the case of acquiring companies. It is well known that on average the stock price of companies engaged in large acquisitions drops. Historically, this drop has been interpreted as a signal that the market values most of these acquisitions negatively. In fact, this drop could be due, at least in part, to the selling pressure of merger arbitrageurs, who buy the target and hedge by selling short

the bidder. Over several days, one expects this deviation to be corrected, but at the announcement of the bid, the imbalance can be very strong and can temporarily alter relative prices. These temporary misalignments have no serious impact on the efficiency with which market prices direct the allocation of resources; still, they must be considered when we make inferences from events studies.

When it comes to the absolute level of prices, however, or the relative price between very different asset classes, we cannot have the same level of confidence in market efficiency. Arbitrages between asset classes or directional arbitrages against the overall level of the markets are extremely risky for any trader. Only a very self-confident trader, one whose job is not at risk because he bets mostly with his own money, can afford to engage in these arbitrages. Warren Buffett is one of the few investors who fit this description. Not surprisingly, he has an amazing record of performance. The paucity of this type of arbitrageur should give us pause in taking the overall level of the market at face value.

That markets may deviate from fundamentals does not necessarily imply that they always do or that it is easy to spot deviations when they occur. The scarcity of investors with long-term success in timing the market (like Warren Buffett) suggests the extreme difficulty of this task. The very reasons that push market prices away from fundamentals – irrational exuberance, lack of capital for arbitrage, or a combination of the two – are likely to affect the market timer as much as the rest of the market. (One has to be very rich, live in Omaha, Nebraska, and not use the Internet to be a good contrarian.) A quasi-EMT, thus, would have practical advice very similar to the EMT when it comes to financial invest-

ment strategy. Yet its implications for real resource allocation would be very different. Without the confidence that the market always provides the best guess, the risk that resources will be misallocated is very real.

Once we recognize these occasional deviations we should not disregard market information, but adjust it or complement it with fundamental analysis. It is safe to start by assuming that market prices reflect fundamentals, but we must be open to changing this assumption when there is overwhelming evidence to the contrary. While a price-earnings ratio of twenty-five might not be irrationally high, it is hard to explain a price-earnings ratio of forty-five – just as it is difficult to justify that the land underneath the imperial palace in Tokyo could cost as much as the entire state of California, as was the case at the peak of the real estate bubble in Japan. While our current knowledge does not provide us with the certainty that these situations are bubbles, it does suggest that completely disregarding these indicators is very risky and leads, on average, to bad decision-making. Faced with these aberrations, central bankers would be foolish not to lean against the wind, especially after seeing the costs a bubble's burst can have on the real economy.

At the same time, in designing executive compensation one should recognize the limitations of stock-based performance. In fact, a significant component of change in a company's stock price is due to variations in the market risk premium, not in the actual fundamentals of the company that can be influenced by its executives. The change in the price of a company relative to the change of an index of its competitors is a much better measure of the actual performance, because we have more confidence in the relative efficiency of the

market rather than in the absolute efficiency. Similarly, before making any inference from an event study, it is necessary to adjust the price changes for temporary order imbalances that might have temporarily moved the price.

As for financial reporting, the faith in marking to market (that is, assessing the accounting value of assets at their market price) should be less absolute. That most of the arguments against marking to market are self-serving ones aimed only at covering bad performance does not imply, however, that marking to market is always the right choice. In this respect, providing two valuations in financial statements, one marked to market and another based on fundamental analysis, with clear description of the underlying hypotheses, can be a useful compromise.

Even when market prices deviate from fundamental values, they still provide useful information about the demand of various assets, information that should not be ignored. When closed-end funds of a particular sector, for example, start selling at a large premium with respect to their net asset value, we should infer that the demand of small investors (the main buyers of closed-end funds) for that sector is booming and is likely to drive the stock prices of the firms in that sector above their fundamental value.

Market prices are a very important source of information. Our only mistake was to consider them the only source of information, in any moment in time, regardless of the conditions under which those prices were obtained. The way forward is not to dismiss market prices, but to consider them in context.

When the definitive history of the EMT is written, the 2007 – 2008 financial crisis will not emerge as a major turning point, only a catalyst of an intellectual

journey that started at full speed after the 1987 stock market crash. Rather than demonstrating market inefficiency, the 2007–2008 financial crisis stands out as an example of the potential costs that deviations of assets' prices from fundamentals can have on the real economy, especially when these deviations are

amplified by high leverage. This realization is forcing experts both inside and outside academia to rethink the robustness of the implications of the EMT and openly accept the possibility of some inefficiency. While just beginning, this process promises to be very challenging but, eventually, very useful.

ENDNOTES

- ¹ Eugene F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance* (May 1970): 383.
- ² Michael Jensen, "Some Anomalous Evidence Regarding Market Efficiency," *Journal of Financial Economics* 6 (2–3) (1978): 95–101.
- ³ Andrei Shleifer and Lawrence H. Summers, "The Noise Trader Approach to Finance," *Journal of Economic Perspectives* 4 (2) (1990): 19.

Jeremy C. Stein

*Securitization, shadow banking
& financial fragility*

Many of the most dramatic and memorable moments of the recent financial crisis involved the failures or near-failures of some of the nation's biggest financial institutions: Bear Stearns, Lehman Brothers, and AIG, to name a few. Much of the subsequent policy response has been shaped by a desire either to avert such failures in the future – for example, by imposing higher capital requirements on systemically important financial firms – or to lessen the adverse consequences of failures if they do occur – for example, by improving the methods available to resolve large institutions in distress.

Yet from the perspective of credit creation and impact on the rest of the economy, one of the most damaging aspects of the crisis was not just the problems of these big firms, but also the *collapse of an entire market*, namely the market for asset-backed securities (ABS). For example, the market for so-called traditional or consumer ABS – those based on credit-card, auto, and student loans – averaged between \$50 and \$70 billion of new issuance per quarter in the years leading up to the crisis. (The total issuance for calendar year 2007 was \$238 billion.) However, in the last

quarter of 2008, following the bankruptcy of Lehman, total issues in this category fell to slightly more than \$2 billion. Given that banks were suffering their own problems and were not easily able to step into the breach, the disappearance of this market represented a major contraction in the supply of credit to consumers, and may well have played a central role in the steep drop in aggregate consumption that occurred at this time. The traditional ABS market only began to rebound in mid-2009, with the implementation of the Federal Reserve's Term Asset-Backed Lending Facility, or TALF, which made tens of billions of dollars of Federal Reserve loans available on attractive terms to investors seeking to buy newly issued consumer ABS.

In what follows, I explore the role that the ABS market plays in the broader process of credit creation, focusing on four sets of issues. First, I describe how the market works: how pools of loans (for example, mortgages or credit-card and auto loans) are packaged and structured into ABS and how investors such as hedge funds, pension funds, and broker-dealer firms finance the acquisition of these ABS. Second, I outline the economic forces that drive securitization; these include both an efficiency-enhancing element of risk-sharing and a less

desirable element of banks trying to circumvent regulatory capital requirements. Third, I discuss the factors that appear to have contributed to the market's fragility during the crisis period. Finally, I suggest some policy reforms that might help mitigate this fragility going forward.

Suppose you buy a new car and take out a \$10,000 loan to finance the purchase. This loan could come from a bank or from the financing arm of an auto company. After the loan is made, the lender has two options: it can hold on to the loan it has originated, or it can securitize it and thereby sell it off. In the first step in the securitization process, known as *pooling*, the lender gathers your loan with tens of thousands of other loans like it – from other cars sold around the country at about the same time – and assembles these loans together in a trust. The cash payments coming from all the designated loans from that point forward go into the trust.

The second step in the process is *tranching*. Tranching involves designating different classes of claimants on the cash flows to the trust, some of whom are given higher priority than others. Said differently, tranching spells out who loses money and under what conditions if some of the loans that make up the pool go bad. Consider this simplified example: a pool of ten loans of \$10,000 each all come due at the same time. Each borrower can either pay off his loan in its entirety, or default and pay nothing. One possible structure would divide the pool into ten layers, or tranches, each of which is owed \$10,000. The most junior or lowest priority tranche – call it T₁ – would recoup its money only if all ten loans were repaid; if even a single loan went bad, T₁ would see its investment wiped out. Thus, T₁ stands at the bottom

of the pecking order and is a very risky security. The second most junior tranche, T₂, regains its money if nine or more of the loans are repaid; equivalently, it gets wiped out if there are two or more defaults. At the top of the hierarchy stands T₁₀, the most senior tranche, which is very well protected and loses only if all ten of the loans in the trust go bad.

As the example suggests, the senior-most tranches of securitizations are likely to have high credit quality. That is, they experience losses only under rare circumstances, when a large fraction of the loans in the underlying pool is hit with defaults. One reflection of this tendency is the AAA rating typically assigned to these senior-most tranches by the rating agencies. This rating in turn makes them attractive investments for institutional investors that are looking for safe places to put their money but are either unable or unwilling to expend the resources required to do loan-level due diligence. Such institutions may not have the expertise to evaluate individual applicants for auto loans, but given the reduction in credit risk associated with the process of pooling and tranching, may be comfortable buying the senior tranches of auto-loan securitizations even after relatively little investigative effort.

While the subprime crisis has called into question the whole idea of turning risky loans into apparently near-riskless AAA-rated ABS – a process many have dubbed “alchemy” – the flaw is not so much with the basic concept of securitization, which has been around for a long time, but rather with the reckless and excessively complex way in which it was applied to subprime mortgage loans. The problem in the subprime sphere was not that the most senior T₁₀-like tranches of subprime pools were rated AAA; it was that many of the less well-protect-

ed tranches (say the T3s) were as well, though only after a series of machinations that were less than transparent to most market participants.¹

Figure 1 provides some perspective on this point. It plots quarterly issuance of ABS over the period 2000 to 2009, broken into two categories. The first category, “traditional” ABS, comprises securitizations based on consumer credit: auto, credit-card, and student loans. The striking characteristic of this series is how stable it is in the several years prior to the crisis; one gets the sense of a market that has functioned steadily and unremarkably, with only modest trend growth, for a long period of time. The second category, “nontraditional” ABS, includes securitizations based on subprime mortgage loans as well as second- and third-generation *resecuritizations*, in which the collateral going into the trust is not a pool of actual loans, but rather a collection of tranches created from earlier rounds of securitizations of subprime loans and other assets.² This latter, more highly engineered market shows the signs of a bubble, with the volume of issuance growing explosively in the period 2003 to 2007, before completely collapsing during the crisis period.

Beyond pooling and tranching, a final important element in the securitization process is *maturity transformation*. Often, the investors that purchase ABS tranches rely heavily on borrowed money, with much of this borrowing being very short-term in nature. In the period leading up to the crisis, entities known as “structured investment vehicles” (SIVs) or “conduits” were prominent examples of this behavior. These entities, which in many cases were affiliated with sponsoring commercial banks, held ABS tranches and financed those holdings by issuing commercial paper, which typically had a maturity

of only days or weeks and therefore had to be rolled over frequently. Another example came from the hedge funds and the broker-dealer firms (like Bear Stearns and Lehman Brothers) that financed their holdings of ABS with repurchase agreements (commonly known as “repo”), which are a form of overnight collateralized borrowing.

Collectively, the various investors that acquire ABS and finance them with short-term borrowing are often referred to as the *shadow banking system*. Just as traditional commercial banks invest in long-term loans and finance these loans by issuing short-term deposits, the shadow banking system invests in securities based on the same sorts of long-term loans (for example, mortgages or auto loans) and finances this investment by issuing short-term claims such as commercial paper or repo. On the one hand, the shadow banking system performs an economic function that looks much like that performed by the traditional banking system: it borrows on a short-term basis to fund longer-term loans; this is what is meant by maturity transformation.³ On the other hand, it does not face the same set of regulations because the institutions involved are generally not banks per se. And it does not benefit from the same safety net. For example, unlike bank deposits, the short-term financing used by shadow banks is not federally insured. Nor do shadow-banking players typically have the right to borrow from the Federal Reserve’s discount window.

Not all investors that hold ABS finance these holdings with short-term borrowing. Anecdotal evidence suggests that pension funds and insurance companies tend to hold these securities on an unlevered basis, that is, without borrowing against them. Remarkably, however, little is known about the relative magni-

Figure 1
Quarterly Issuance of Asset-Backed Securities, 2000 – 2009

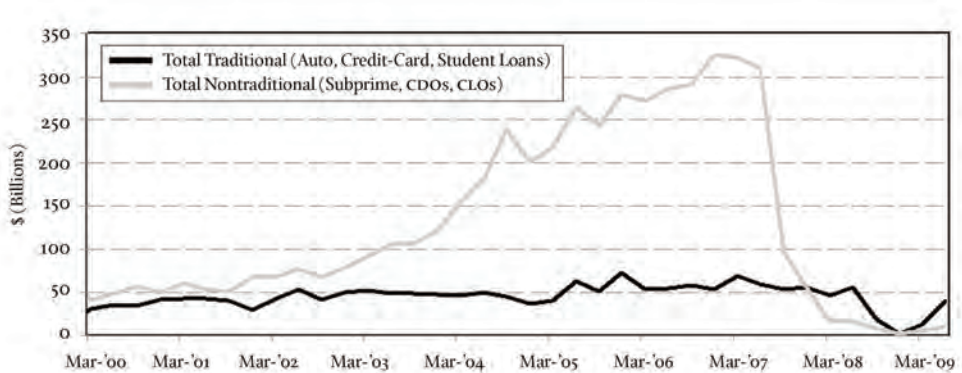


Figure created using data from Thompson SDC.

tudes. In particular, as far as I am aware, there are no good empirical estimates to address the following question: what fraction of the AAA-rated ABS in a given loan category (credit cards, for example) is owned by investors that finance their holdings of these securities primarily with short-term debt? This is an important question because, as argued below, the heavy use of short-term debt financing by ABS investors contributes to the fragility of the market.

With securitization, loans originated by banks are packaged and sold off to a variety of other end investors. What economic forces encourage banks to off-load their loans in this way, as opposed to keeping them on their books? There are two broad stories that one can tell.

The first, more benign story depends on the principle of risk-sharing. When banks sell their loans into the securitization market, they distribute the risks associated with these loans across a wider range of end investors, including pension funds, endowments, insurance companies, and hedge funds, rather than tak-

ing on the risks entirely themselves. This improved risk-sharing represents a real economic efficiency and lowers the ultimate cost of making the loans. Moreover, as noted above, the pooling and tranching process, if done properly, makes the senior tranches of ABS relatively easy to evaluate, even for nonspecialized investors that do not have much ability to judge the credit quality of the individual loans that underlie these securities. Therefore, it is particularly conducive to expanding the buyer base.

According to this story, the securitization of consumer loans and mortgages closely parallels the natural transition that many growing companies make when they reach a certain size and reputation and begin to shift their borrowing away from exclusive reliance on banks and toward the corporate bond market. In either case, if the securities in question (either ABS or corporate bonds) can be easily evaluated by a broader set of investors, it makes sense to tap into this broader market, as opposed to relying exclusively on the banking sector for financing.

While this wholesome-sounding story undoubtedly captures some of what drives the securitization market, it is also incomplete. It has become apparent in recent years that another important driver of securitization activity is regulatory arbitrage, a purposeful attempt by banks to avoid the rules that dictate how much capital they are required to hold. Particular attention in this regard has focused on the bank-sponsored SIVs and conduits mentioned above, vehicles that held various types of ABS and financed these holdings largely with short-term commercial paper. What is striking about these shadow-banking vehicles is that many of them operated with strong guarantees from their sponsoring banks. Indeed, when the SIVs and conduits got into trouble, the banks honored their guarantees, stepping up and absorbing the losses.

This outcome runs counter to the spirit of the risk-sharing story, since rather than widely distributing the risks associated with the ABS they created, banks retained them, albeit in an opaque, off-balance-sheet fashion. The most obvious alternative explanation is that banks exploited a regulatory loophole: if they held the loans directly on their balance sheets, they faced a regulatory capital requirement on these loans; but if they securitized the loans and parked them in an off-balance-sheet vehicle (albeit one with essentially full recourse to the banks in the event of trouble), the regulatory capital requirement was much reduced.⁴

While this particular loophole will no doubt be closed going forward, the more general concern remains. Securitization and the shadow banking system enable bank-like maturity-transformation activities – specifically, the pairing of long-term assets with a short-term funding structure – to take place out of the reach of traditional banking regulation. To the

extent that bank regulation is burdensome, it creates a powerful pressure for banking assets to be securitized and to migrate out of the formal banking system. Absent some form of harmonization that puts shadow banks and traditional banks on more of an equal regulatory footing, this pressure is likely to intensify as capital requirements on banks are raised in the wake of the crisis.

Figure 1 illustrates the complete meltdown of the ABS market during the financial crisis, with issuance in both the traditional consumer (auto, credit-card, and student loans) and nontraditional (subprime) categories falling essentially to zero. The nontraditional market started to come apart first, in August 2007, as the extent of losses on subprime loans became more apparent. The traditional consumer market held up better for a time, but then also disappeared in the wake of the failure of Lehman Brothers in September 2008.

While less spectacular from a quantitative perspective, the decline of the traditional consumer ABS market is in many ways the more challenging phenomenon to explain. The demise of the subprime-related ABS market represented the deflation of a classic bubble; many of the loans involved were so poorly underwritten that they should never have been made in the first place. Not surprisingly, when market participants finally began to understand this point, the issuance of new subprime loans dried up.

It is more surprising that ABS issuance related to auto, credit-card, and student loans was hit so hard. There is much less reason to believe that these were bad loans to begin with: again, the superficial evidence in Figure 1 does not suggest a bubble in this part of the market. Indeed, *overall lending* in

these categories did not completely vanish in the same way that it did in the subprime area; rather, some fraction of this lending reverted back to being done in a nonsecuritized fashion by the banks, which suggests that they still viewed the loans as worth making. However, given the limited capacity of the banks, whose capital by this point was badly impaired, the inability to securitize and thereby off-load some of their loans no doubt contributed to a sharp contraction in the overall supply of credit available to consumers.

If the underlying auto, credit-card, and student loans were still worth making, what caused the market for ABS based on these loans to contract so sharply? A prominent emerging hypothesis is that, effectively, there was the analogue of a widespread bank run on the shadow banking system.⁵ Recall that many ABS investors finance their positions with short-term borrowing, either in the form of commercial paper or overnight repurchase agreements. In this sense, they are very much like banks, which finance long-term loans with short-term deposits. But unlike bank deposits, the short-term financing that supports the ABS market is not insured by the government. This difference makes the shadow banking system vulnerable to something that looks like a classic bank run from the days before deposit insurance: as short-term lenders lose confidence and refuse to roll over their loans, investors in ABS are forced to liquidate some of their holdings to come up with cash. The liquidations in turn depress the price of these ABS via a “fire sale” effect. Moreover, short-term lenders view ABS as less attractive collateral as their prices fall and become more volatile. The lenders then pull back even further, leading to another round of liquidations and price

declines. Once under way, this vicious cycle is very difficult to arrest.

One concrete manifestation of the dramatic withdrawal of short-term lending to the ABS market comes from the behavior of what are called “haircuts” in repurchase agreements. When an investor borrows from the repo market to finance its holdings of ABS, it is required to post a margin, or down payment; this is the haircut. Haircuts on ABS were extremely low prior to the crisis, on the order of 2 percent. What this means is that if, say, a hedge fund wanted to acquire \$1 billion of auto-linked ABS, it only needed to put up \$20 million of its own capital as a down payment. The other \$980 million could be borrowed on an overnight basis in the repo market; in many cases, the ultimate lenders of this short-term money were money-market mutual funds looking to find slightly higher yielding short-term investments than, for example, Treasury bills.

In the midst of the crisis, haircuts skyrocketed. Even haircuts on traditional consumer ABS – those not linked in any direct way to the housing and subprime problems – rose to more than 50 percent.⁶ From the perspective of the hedge fund holding \$1 billion of auto-linked ABS, suddenly it could borrow only \$500 million, and instead of having to post a \$20 million down payment, it now had to post \$500 million. If it did not have the cash on hand, it was forced to liquidate its holdings. These forced liquidations, and the powerful impact they had on both the level and volatility of ABS prices, in turn justified the increased skittishness of the lenders in the repo market, because their protection was entirely predicated on the collateral value of the assets they were lending against.

The bank-run analogy offers what feels like a compelling account of the fragility

of the securitization market. However, it would be premature to call it a fully empirically validated explanation for why the market dried up so dramatically during the crisis. For one, as emphasized above, it is not known what fraction of ABS was held by investors that financed themselves in a vulnerable bank-like way – that is, largely with short-term debt. If the fraction turns out to have been, say, 40 percent instead of 80 percent, this finding would temper the force of the theory.

There is an alternative, more behavioral hypothesis for the fragility of the securitization market that does not rely on a predominance of short-term debt financing.⁷ This alternative hypothesis begins with the observation that a large proportion of ABS tranches – in both the traditional and subprime sectors – was rated AAA. The AAA rating may have encouraged investors such as pension funds or insurance companies to think of these securities as essentially riskless and therefore to treat them as equivalent to Treasury bonds when constructing their portfolios. When the problems in the subprime area became apparent, this premise was utterly destroyed, and investors that were determined to allocate a fraction of their portfolios to safe assets realized that they had to dump their holdings of AAA-rated ABS and buy actual Treasuries instead. Thus, instead of a short-term-debt-driven bank run, we have what might be called a widespread buyer's strike. In this account, the rating agencies' failures with respect to the subprime market undermined their credibility more generally, so that any AAA-rated tranche of an ABS, be it linked to subprime or credit cards, was no longer considered a virtually riskless asset.

Of course, the two theories are not mutually exclusive, and may interact in

interesting ways. For example, what begins as a simple strike on the part of unlevered buyers may evolve into a run-like phenomenon since the buyer's strike puts downward pressure on ABS prices, making the position of short-term lenders more precarious and thereby encouraging these lenders to withdraw from the market.

To frame the policy issues with respect to securitization and the shadow banking system, it is useful to begin by emphasizing three key points. First, we are almost certainly heading in the direction of imposing significantly higher capital requirements on large banks. Second, while this is undoubtedly a valuable and much-needed reform, and one that holds the promise of making the banking sector itself more robust in future episodes of financial volatility, it will also have the effect of encouraging more credit-creation activity to migrate away from the banks and toward the shadow-banking sector, in an effort to evade the burdens associated with more stringent regulation. Third, we have seen that the shadow-banking sector can be a powerful source of fragility in its own right, one that can lead to damaging disruptions in the flow of credit to households and businesses. Thus it would be a mistake to pursue a set of policies that focuses heavily on insulating our large banks but that pays insufficient attention to potential vulnerabilities in the rest of the financial system. Rather, the goal should be a balanced approach that addresses all elements of the system in an integrated fashion.

What concrete steps might be taken in this regard? I present three specific ideas for regulating the securitization and shadow-banking markets. To be transparent about my own prejudices, I label the ideas *the good*, *the bad*, and *the maybe*.

The good: regulation of haircuts in the ABS market. To mitigate the incentives for regulatory evasion, and to help reduce the fragility of credit creation no matter where it takes place, a systematic effort must be made to impose similar capital standards on a given type of credit exposure, irrespective of whether it is a bank, a broker-dealer firm, a hedge fund, or any other entity that ends up holding the exposure. This is not an easy task, but one tool that would help is broad-based regulation of haircuts (that is, minimum margin requirements) on ABS.⁸

Consider the case of a consumer loan. If this loan is made by a bank, it will be subject to a capital requirement; that is, the bank will have to put up some amount of equity against the loan, rather than borrow all the money. Now suppose instead that the loan is securitized by the bank and becomes part of a consumer ABS whose tranches are distributed to various types of investors. The regulation I have in mind here would stipulate that whoever holds a tranche of the ABS would be required to post a minimum down payment against that tranche – with the value of the haircut depending on the seniority of the tranche, the quality of the underlying collateral, and so forth.

For example, before the current crisis, market-determined haircuts on AAA-rated consumer ABS tranches were very low, in the neighborhood of 2 percent. With no further regulation, they are likely to return to these levels as markets re-normalize. However, the new regulation might instead impose a minimum haircut requirement on AAA-rated consumer ABS of at least 10 percent, independent of market conditions. That is, any investor in such a security would be required to post and subsequently maintain a 10 percent margin at all times.

Such a requirement is nothing conceptually new and should not be difficult to enforce; indeed, it is closely analogous to the initial and maintenance margin requirements that are currently applicable to investors in common stocks.

If well structured, these haircut requirements have two important benefits. First, they go a long way toward achieving harmonization across organizational forms, in that there would no longer be an obvious regulation-avoidance motive for moving the consumer loan off the balance sheet of the bank and into the shadow-banking sector. This benefit is especially important as we move toward significant increases in the capital requirements imposed on banks. The goals of these higher bank-capital requirements are likely to be partially frustrated if they drive significant amounts of activity outside the banking system.

Second, for that portion of credit-creation activity that does end up residing in the shadow-banking sector, haircut regulation can help dampen the bank-run-like crisis dynamics described above. The problem is that if haircuts begin at 2 percent before the crisis, and then jump to more than 50 percent during the crisis, this increase creates a powerful forced-selling pressure on the owners of ABS. If the haircuts are instead set at a more prudent value before the crisis – again, say 10 percent – so that investors are required to put up more of their own cash at the outset, this forced-selling mechanism, and the vicious spiral it unleashes, might be substantially attenuated.

The bad: extension of the federal safety net to shadow banks. Some observers have taken the analogy between the traditional commercial banking sector and the shadow banking system one step further, arguing that in order to prevent run-like

panics in the latter, it should be covered by the same federal safety net as the former. This coverage would entail giving shadow-bank entities access to the Federal Reserve's discount window, as well as possibly insuring some of their short-term debts. Thus, when a specialized investment vehicle is set up to buy a portfolio of ABS financed largely with short-term commercial paper borrowing, the commercial paper issued by the investment vehicle might be explicitly federally insured, much as some bank deposits are today. Instead of trying to lean against the private market's propensity to finance ABS with large amounts of short-term debt – as the haircut regulation described above would do – this alternative approach amounts to embracing the use of such short-term financing and attempting to use government insurance to make the world safer in its presence.

What makes this policy unattractive is the moral hazard that it invites, as private actors seek to exploit government-provided insurance by using it to finance riskier-than-expected activities. This concern is particularly acute when the insurance is attached to the kinds of highly engineered financial products that were held by some shadow-banking investors prior to the current crisis – products for which risks are often not easily understood or accurately measured ahead of time. For example, one can imagine a government insurer trying to devise a formula for risk-based pricing of the insurance it provides to a specialized investment vehicle, in an effort to deter excessive risk-taking. But should we expect any such formula to do better than those of the rating agencies, which so spectacularly misjudged the risks embedded in complex ABS based on subprime mortgages? Indeed, one can argue that the mind-bending complexity of some of these structures emerged pre-

cisely as a means of gaming the rating-agency formulas. Thus, although a government insurance agency would not face the same overt conflicts of interest as the rating agencies, it seems reasonable to worry about how it would fare when pitted against Wall Street's best financial engineers.

With this bit of pessimism in mind, I would argue that in order to entertain the idea of expanding the safety net, one would have to believe that the short-term debt claims created by the shadow-banking sector are of substantial social value – so much so that sustaining them with moral-hazard-prone insurance, rather than trying to constrain them with haircut regulation, is a first-order imperative. I don't think that we have nearly enough empirical evidence to meet this burden of proof. Hence I would be strongly inclined to steer clear of any expansion of the safety net.

The maybe: limiting the creation of “pseudo-riskless” securities. As discussed above, an alternative theory for the fragility of the ABS market during the crisis is that, even absent short-term debt financing of ABS positions, the proliferation of so many “pseudo-riskless” securities is inherently dangerous. By pseudo-riskless, I mean AAA-rated securities that appear so safe in good times that investors are lulled into a sense of complacency whereby they treat these securities as equivalent to truly riskless Treasuries. Only in a crisis do investors discover that this was a false equivalence, which leads them to panic and dump their holdings of the AAA-rated securities.⁹

If one takes this point of view, it is tempting to think about ways to constrain the production of those ABS tranches that can be represented to investors as being near-riskless. One way might be to require the credit-rating

agencies to use a coarser set of ratings when evaluating ABS than when they evaluate corporate bond issues. For example, instead of a finely tuned scale that goes from AAA to AA+, AA, AA-, A+, and all the way down to CCC, the ratings for ABS might be restricted to one of three broad buckets: A, B, or C.¹⁰ While this idea admittedly has a bit of the feel of deploying the language police, it might prevent any ABS tranche from being thought of as near-riskless, since even the highest rating category would now encompass securities with a wide range of credit qualities.

An alternative approach would be to leave the current ratings categories in place but to impose on the creators of any ABS an upper limit on the amount of highly rated securities that they could manufacture from any given underlying pool of loans. For example, one rule might be that only a maximum of 50 percent of the dollar value of tranches coming from any pool of consumer loans could ever seek a rating of AA or higher; all other subordinate tranches would have to be targeted at lower ratings categories.

I put this last set of ideas in the “maybe” category because I view them as interesting and worthy of further thought, but I am not at this point confident that their virtues outweigh their potential for unintended consequences. On the one

hand, they highlight the logical implications of taking a more behavioral perspective on the ABS market’s fragility – of positing a world in which investors are overly prone to seek out pseudo-riskless investments and in which financial innovators actively try to exploit this tendency. On the other hand, the specific proposals I have sketched raise some fairly obvious flags as well. For example, restricting the vocabulary available to the rating agencies may have meaningful effects in the short run, but over time it is easy to imagine industry conventions evolving so as to work around any such restrictions. If so, it would be a mistake to place much long-term faith in this approach.

The overarching goal of financial reform must not be only to fortify a set of large institutions, but rather to reduce the fragility of our *entire system* of credit creation. This system involves a complicated interplay between banks and non-banks and between traditional forms of lending and securitization. Thus far, more effort has been devoted to the banking side of the equation. This imbalance is perhaps not surprising given the accumulated expertise of many of the regulators involved in the reform process. But the difficult issues associated with securitization and the shadow banking system demand equal attention.

ENDNOTES

¹ More precisely, the riskier lower-rated tranches of subprime securitizations were themselves used as the raw material (in place of the original mortgage loans) to create second- and third-generation *resecuritizations*. Many of the biggest problems in the crisis arose from the fact that large fractions of these resecuritized vehicles were also rated AAA, in spite of the dubious collateral supporting them. This is where the most extreme alchemy can be said to have taken place. For a discussion, see Joshua Coval, Jakub Jurek, and Erik Stafford, “The Economics of Structured Finance,” *Journal of Economic Perspectives* 23 (2009): 3–25.

- ² The data in the figure come from Thompson SDC. While the “nontraditional” category includes securitizations based on subprime mortgage loans, it does not include securitizations based on prime mortgage loans, such as mortgage-backed securities guaranteed by the government-sponsored enterprises Fannie Mae and Freddie Mac. I am grateful to Sam Hanson, who put together the data and shared it with me.
- ³ See Gary Gorton and Andrew Metrick, “Securitized Banking and the Run on Repo,” NBER working paper 15223 (National Bureau of Economic Research, 2009).
- ⁴ For a detailed study of this phenomenon, see Viral Acharya, Phillip Schnabl, and Gustavo Suarez, “Securitization without Risk Transfer,” NBER working paper 15730 (National Bureau of Economic Research, 2010).
- ⁵ The bank-run hypothesis is articulated in Gorton and Metrick, “Securitized Banking and the Run on Repo,” as well as in Daniel Covitz, Nellie Liang, and Gustavo Suarez, “The Anatomy of a Financial Crisis: The Evolution of Panic-Driven Runs in the Asset-Backed Commercial Paper Market,” working paper (Board of Governors of the Federal Reserve, 2009).
- ⁶ For more detail on the evolution of repo-market haircuts during the crisis period, see Gorton and Metrick, “Securitized Banking and the Run on Repo.”
- ⁷ A version of this hypothesis is presented in Nicola Gennaioli, Andrei Shleifer, and Robert Vishny, “Financial Innovation and Financial Fragility,” working paper (Harvard University, 2010).
- ⁸ Such haircut regulation is discussed in John Geanakoplos, “The Leverage Cycle,” discussion paper 1715R (Cowles Foundation, 2010); and in Jeremy Stein, “Monetary Policy as Financial-Stability Regulation,” working paper (Harvard University, 2010).
- ⁹ Again, see Gennaioli, Shleifer, and Vishny, “Financial Innovation and Financial Fragility,” for elaboration on this argument.
- ¹⁰ This scale is used by both Standard and Poor’s and Fitch. The other major rating agency, Moody’s, has a somewhat different alphanumeric convention, albeit with similarly fine-grained categories: Aaa, Aa1, Aa2, Aa3, A1, A2, etc.

Lucian A. Bebchuk

How to fix bankers' pay

The financial crisis has highlighted the importance of discouraging excessive risk-taking by financial firms. The extent to which firms take on risks can be expected to depend partly on their executives' incentives. In the aftermath of the financial crisis, G20 leaders announced their commitment "to implement strong international compensation standards aimed at ending practices that lead to excessive risk-taking."¹ But how should bankers' pay be fixed? The devil, as is often the case, is in the details.

Below I sketch some key principles for reforming bankers' pay, drawing on my academic work² as well as on written testimonies before the Financial Services Committee of the House of Representatives.³ I describe two distinct sources of risk-taking incentives: first, executives' excessive focus on short-term results; and second, their excessive focus on results for shareholders, which corresponds to a lack of incentives for executives to consider outcomes for other contributors of capital. I discuss how pay arrangements can be reformed to address each of these problems and conclude by examining the role that government should play in bringing about the needed reforms.

Standard pay arrangements have incentivized and rewarded short-term results. Jesse Fried and I warned about this problem and its consequences in our book *Pay without Performance: The Unfulfilled Promise of Executive Compensation*, published six years ago.⁴ Under the standard design of pay arrangements, executives have been able to cash out large amounts of compensation based on short-term results, which has provided executives with incentives to seek short-term gains even when doing so creates excessive risk of a later implosion.

Modern pay packages are largely performance-based, consisting of both equity-based compensation (such as options and restricted stock) and bonus compensation. Such provisions are intended to provide incentives for improved performance. In the past, however, these critical components of standard pay arrangements have also produced perverse incentives.

Bear Stearns and Lehman Brothers, the firms whose respective meltdowns ushered in the current financial crisis, are illustrative of the problem. Various observers, including *New York Times* and *Wall Street Journal* columnists,⁵ largely assumed that the wealth of these firms' executives was wiped out, together with that of their firms. This "standard narra-

© 2010 by Lucian A. Bebchuk

tive” led commentators to dismiss the potential role of flawed pay arrangements in the firms’ risk-taking. Assuming that the executives’ wealth was wiped out alongside the shareholders’, these commentators inferred that the firms’ risk-taking could not have been an upshot of deliberate choices produced by perverse incentives; rather, it must have been the result of the executives’ failure to perceive risks – or their hubris. Alternative pay structures, they argued, would not have made a difference.

In our study “The Wages of Failure: Executive Compensation at Bear Stearns and Lehman Brothers 2000 – 2008,”⁶ Alma Cohen, Holger Spamann, and I examine this standard narrative and find it to be incorrect. We piece together the cash flows derived by the firms’ top five executives using data from Securities and Exchange Commission filings. Our analysis indicates that, notwithstanding the 2008 collapse of the firms, the bottom lines of those executives for the 2008 – 2009 period were positive and substantial.

Most important, the firms’ top executives regularly unloaded shares and options and thus were able to cash out much of their equity before the stock price of their firms plummeted. Indeed, the top five executives unloaded more shares during the years prior to their firms’ meltdowns than they held when disaster came in 2008. Altogether, during 2000 to 2008, the top executive teams at Bear Stearns and Lehman cashed out through equity sales about \$1.1 billion and \$850 million, respectively.⁷

These payoffs to top executives were further increased by large bonus compensation. During 2000 to 2007, the top executives’ aggregate bonus compensation reached \$300 million at Bear Stearns and \$150 million at Lehman. Of course, the earnings that provided the basis for

these bonuses evaporated in 2008. But the firms’ pay arrangements did not contain any “claw-back” provisions that would have enabled the firms to recoup bonuses that had already been paid.

Combining the figures from equity sales and bonuses, we find that, during 2000 to 2008, the top five executives at Bear Stearns and Lehman pocketed about \$1.4 billion and \$1 billion, respectively, or roughly \$250 million per executive. The CEOs – James Cayne of Bear Stearns and Richard Fuld of Lehman Brothers – pocketed about \$380 million and \$520 million, respectively. These cash proceeds are substantially higher than the value of the holdings that the executives held at the beginning of the period. Thus, while earnings for their firms’ long-term shareholders were largely decimated, the executives’ performance-based compensation kept them in positive territory.

The divergence between how the top executives and their companies’ shareholders fared raises a serious concern that the aggressive risk-taking at Bear Stearns and Lehman (and other financial firms with similar pay arrangements) could have been the product of flawed incentives. The concern is not that the top executives expected their aggressive risk-taking to lead to certain failure for their firms, but that the executives’ pay arrangements – in particular, their ability to claim large amounts of compensation based on short-term results – induced them to accept excessive levels of risk.

To address the problem of short-term focus, financial firms should reform compensation structures to ensure tighter alignment between executive payoffs and long-term results. Senior executives should not be able to collect and retain large amounts of bonus compensation when the performance on which the

bonuses are based is subsequently sharply reversed. Similarly, equity incentives should be subject to substantial limitations aimed at preventing executives from placing excessive weight on their firm's short-term stock price. Had such compensation structures been in place at Bear Stearns and Lehman, their top executives would not have been able to derive such large amounts of performance-based compensation for managing these firms in the years leading up to their collapse.

Equity-based compensation is the primary component of modern pay packages. Under standard arrangements, equity-based awards vest gradually over a period of time. Once they vest, however, executives typically have unrestricted freedom to cash them out and, indeed, often unload such equity incentives quickly after vesting. This right to freely cash out vested equity incentives has contributed substantially to the creation of short-term distortions.

In a recent article, Jesse Fried and I, building on the approach we put forward in *Pay without Performance*, proposed a detailed blueprint for preventing equity-based compensation, the primary component of modern pay packages, from producing an excessive focus on short-term results.⁸ To begin, it is desirable to separate the time that options and restricted shares can be cashed out from the time in which they vest. As soon as an executive has completed an additional year at the firm, the options or shares promised as compensation for that year's work should vest; it should belong to the executive even if he or she immediately leaves the firm. The executive, however, should not be free to cash out these vested equity incentives; rather, he or she should be permitted to do so only after a substantial passage of time.

Some shareholder proposals and compensation experts have called for allowing executives to cash out shares and options only upon retirement from the firm. Such a "hold-till-retirement" requirement, however, would provide executives with a counterproductive incentive to *leave* the firm in order to cash out their portfolio of options and shares and diversify their risks. Perversely, the incentive to leave will be strongest for executives who have served successfully for a long time and whose accumulated options and shares are especially valuable. Similar distortions arise under any arrangement tying the freedom to cash out to an event that is at least partly under an executive's control. To avoid this problem, firms should adopt limitations on unwinding that are not a function of such events.

In particular, unwinding should be subject to a combination of grant-based and aggregate restrictions. Grant-based limitations would require executives to hold equity incentives awarded as part of a given grant for a fixed number of years after vesting. For example, an executive receiving an equity award could be prevented from unwinding any awarded equity incentives for two years after vesting, with each subsequent year freeing another 20 percent of the awarded incentives to be unloaded.

These grant-based limitations, however, might not be sufficient to secure adequate long-term focus. With only grant-based restrictions in place, long-time executives might amass large amounts of equity incentives that they could immediately unload, which could induce them to pay excessive attention to short-term prices. Therefore, grant-based limitations should be supplemented with aggregate limitations restricting the fraction of an executive's otherwise unloadable equity incentives that could

be sold in any given year. To illustrate, executives could be precluded from unloading, in any given year, more than 10 percent of their total portfolio of otherwise unloadable incentives. By construction, such limitations would ensure that executives would not place substantial weight on short-term stock prices.

Finally, firms must adopt robust limitations on executives' use of hedging and derivative transactions, a practice that can weaken the connection between executive payoffs and long-term results. An executive who buys a "put" option to sell his or her shares at the current price, as executives are generally free to do under standard pay arrangements, is "insured" against declines in the stock price below current levels. Empirical evidence indicates that executives engage in a significant amount of hedging, and that such hedging is at least partly motivated by their inside information.⁹ Furthermore, the adoption of significant limitations on unwinding would likely increase executives' incentives to engage in hedging and derivative transactions, as doing so could neutralize the effects of those limitations. Indeed, if executives continued to enjoy the freedom to engage in such transactions, the practical impact of limitations on unwinding would be much reduced.

Whether or not they are motivated by the use of inside information, executives should not be at liberty to use hedging and derivative transactions to undo the effects of equity-pay arrangements put in place by their firm's board at a cost to shareholders. Rather, executives should be precluded from engaging in any hedging or derivative transactions that would reduce or limit the extent to which declines in the company's stock price would lower executive payoffs. In Fall 2009, following the anti-hedging approach that Jesse Fried and I advocated, the Special

Master for TARP Executive Compensation Kenneth Feinberg required companies subject to his jurisdiction to adopt such an anti-hedging requirement.¹⁰ This approach should also be followed by public firms in general. Whatever equity-plan design is chosen by a given company's board, executives should not be allowed to unilaterally use hedging and derivative transactions that undo the incentive consequences of this design.

In addition to equity compensation, bonus plans also must be redesigned. Under standard pay arrangements, executives have been able to cash bonus compensation based on short-term results and retain it even when those results were reversed. To address the short-term distortion arising from such arrangements, bonuses should not be cashed right away; instead, the funds should be placed in a company account for several years and adjusted downward if the company subsequently learns that the bonus is no longer justified. The need for such a downward adjustment is not limited to firms in which financial results are restated. Even if results for a given year were booked consistent with accounting conventions, executives should not be rewarded for profits that are quickly reversed. Rewarding executives for short-term results distorts their incentives and should be avoided by well-designed compensation arrangements.

Thus far, I have focused on the insulation of executives from long-term losses to shareholders – the problem that has received the most attention following the recent crisis. However, as Holger Spamann and I analyze in detail in recent work,¹¹ there is another type of distortion that should be recognized: payoffs to financial executives have been shielded from the consequences that losses could impose on parties other than share-

holders. This source of distortion is distinct from the “short-termism” problem discussed above and would remain even if executives’ payoffs were fully aligned with those of long-term shareholders.

Equity-based awards, coupled with the capital structure of banks, tie executives’ compensation to a highly levered bet on the value of banks’ assets. While bank executives expect to share in any gains that might flow to common shareholders, they do not expect to bear (in the event losses exceed the common shareholders’ capital) any part of losses borne by preferred shareholders, bondholders, depositors, or the government as a guarantor of deposits. This state of affairs leads executives to pay insufficient attention to the possibility of large losses sustained beyond the shareholders’ equity; it thus incentivizes excessive risk-taking.

Insulation of executives from losses to parties other than shareholders can be expected to produce at least two types of risk-taking distortions. First, it encourages executives to make investments and take on obligations that can contribute to “tail” scenarios, in which the bank suffers losses exceeding the shareholders’ capital. Second, it creates reluctance to raise capital and fosters excessive willingness to run the bank with a capital level that provides inadequate cushion for bondholders and depositors. The more thinly capitalized banks are, the more severe these distortions become.

How could pay arrangements be re-designed to address this distortion? To the extent that executive pay is tied to the value of specified securities, such pay could be tied to a broader basket of securities, not merely common shares. Thus, rather than tying executive pay to a specified percentage of the value of the common shares of the bank holding company, compensation could be tied to a spec-

ified percentage of the aggregate value of the common shares, the preferred shares, and all the outstanding bonds issued by either the bank holding company or the bank. Because such a compensation structure would expose executives to a broader fraction of the negative consequences of risks taken, it would encourage greater prudence in evaluating risky choices.

The structure described above would not by itself cause bank executives to internalize fully the adverse consequences that risk-taking might have for the interests of the government as guarantor of deposits. Achieving such fuller internalization would require broadening further the set of positions that aggregate value executive payoffs are tied to. One could consider, for example, schemes in which executive payoffs were tied not to a given percentage of the aggregate value of the bank’s common shares, preferred shares, and bonds at a specified point in time, but rather to this aggregate value minus any payments made by the government to the bank’s depositors, as well as other payments made by the government in support of the bank, during the specified time period.

Alternatively, executive payoffs could be tied to the aggregate value of the bank’s common shares, preferred shares, and bonds at the specified time minus the expected value of future government payments as proxied by the product of (i) the implied probability of default inferred from the price of credit default swaps at the specified time, and (ii) the value of the bank’s deposits at that time. Even if such schemes are not used, however, tying executive pay to the aggregate value of common shares, preferred shares, and bonds would by itself produce a significant improvement in incentives compared with existing arrangements.

Similarly, in firms where executives receive bonus compensation tied to speci-

fied accounting measures, bonuses could be linked instead to broader metrics. For example, the bonus compensation of some bank executives has been dependent on accounting measures that are of interest primarily to common shareholders, such as return on equity or earning per common share. Such plans could be redesigned to be based on more expansive measures, such as earnings before any payments made to bondholders.

Ensuring that executives internalize perfectly the expected losses that choices would impose on contributors of capital other than shareholders is far from straightforward. But doing so imperfectly would likely be better than not doing so at all. Requiring financial executives to expand their focus beyond consequences for shareholders would significantly improve their risk-taking incentives.

Having discussed changes in pay arrangements that would curtail incentives for banks and other firms to take excessive risks, I turn to the question of what role, if any, the government should play in implementing reform. Some will argue that, however desirable such reforms may be, introducing changes should be left to the marketplace. According to this standpoint, private players can be fully expected to draw lessons from past problems and adopt whatever reforms are desirable. In my view, however, the government has two important roles to play.

The first function the government should perform – both in the financial firms that are the focus of this essay and in other public firms – is to improve governance arrangements. In particular, the government should ensure that shareholders have sufficiently strong rights to discourage choices adverse to their interests, including the adoption of pay arrangements that insulate executives from long-term losses of shareholder value.

The new Dodd-Frank bill, the Wall Street Reform and Consumer Protection Act, introduced “say on pay,” which enables shareholders to express their views on the pay arrangements of public firms in advisory votes. To the extent that directors would seek to avoid a negative say on pay vote, the introduction of such votes could help prevent some egregious cases. But advisory votes by themselves cannot ensure that directors are sufficiently attentive to and focused on shareholder interests. To provide directors with such incentives, a broader reform is necessary – and indeed long overdue.

Shareholders in the United States have long had weaker rights than they have in the United Kingdom and other common law countries. As I discuss in detail in a series of recent articles,¹² shareholder rights should be strengthened in two important ways. To begin, the power of shareholders to replace directors should be turned from myth into reality. Existing rules and arrangements provide incumbent directors with a substantial advantage over outside challengers. For example, whereas challengers from outside the firm must bear their campaign expenses themselves, incumbents have their expenses fully covered by the company. Moreover, the boards of many companies are “staggered,” which requires challengers seeking to gain control of the board to win two elections one year apart. Such existing impediments to shareholders’ ability to replace directors should be dismantled.

In addition, shareholders should have the power to propose and vote to adopt “rules of the game” decisions – that is, decisions to amend the corporation’s governance arrangements. Current arrangements give directors the sole power to initiate changes in the company’s charter or state of incorporation, and severely limit the scope of bylaw pro-

visions that shareholders may initiate and adopt. This state of affairs gives directors an unwarranted role in determining the rules that regulate their own conduct. Shareholders should be given the power to shape governance arrangements.

A reform that strengthens shareholder rights would limit deviations from shareholder interests. Among other things, it would discourage pay arrangements that reward executives for short-term gains and insulate them from losses to long-term shareholder value. This development would help eliminate risk-taking incentives that are excessive even from the shareholders' perspective. Still, there would remain the problem of risk-taking incentives that are favored by shareholders but that, once the interests of other stakeholders are taken into account, are undesirable from a society's standpoint. This point brings us to the second role that the government should play.

Outside the financial sector, the government should limit itself to strengthening shareholder rights; it should not intervene in the substantive terms of pay arrangements. In the case of banks, however, the government's role should extend beyond governance reforms: financial regulators should monitor and regulate firms' compensation structures. Such pay regulation is justified by the same moral hazard reasons that underlie the long-standing system of prudential regulation of banks.

When a bank takes risks, shareholders can expect to capture the full upside, but part of the downside may be borne by the government as guarantor of deposits. Because bank failure imposes costs on the government and the economy that shareholders do not internalize, shareholders' interests may be served by greater risk-taking than is in the interest of the government and the economy. This moral

hazard problem provides a basis for the extensive body of regulations that restrict the choices of financial firms with respect to investments, lending, and capital reserves.

Aligning the interests of executives with those of shareholders, which some governance reforms seek to do, could eliminate risk-taking that is excessive even from the shareholders' perspective. But it cannot be expected to get rid of incentives for risk-taking that are excessive from a social standpoint but not from the shareholders' perspective.

Shareholders' interest in greater risk-taking implies that they stand to benefit when bank executives take excessive risks. Given the complexities of modern finance and the limited information and resources of regulators, the traditional regulation of banks' actions and activities necessarily is imperfect. Regulators are often one step behind banks' executives. Thus, executives with incentives to focus on shareholder interests can use their informational advantages and whatever discretion traditional regulations leave them to take excessive risks.

Because shareholders' interests favor incentives for risk-taking that are socially excessive, substantive regulation of the terms of pay arrangements – that is, limiting the use of structures that reward risky behavior – can advance the goals of banking regulation. Regulators should focus on the *structure* of compensation – not the amount – with the aim of discouraging excessive risk-taking. By doing so, regulators would induce bank executives to work for, not against, the goals of banking regulation.

The regulation of bankers' pay could well supplement and reinforce the traditional direct regulation of banks' activities. Indeed, if pay arrangements are designed to discourage excessive risk-taking, direct regulation need not

be as stringent as would otherwise be necessary. Conversely, as long as banks' executive pay arrangements are unconstrained, regulators should be stricter in their monitoring and direct regulation of banks' activities.

At a minimum, when assessing the risks posed by any given bank, regulators should take into account the incentives generated by the bank's pay arrangements. When the design of compensation encourages risk-taking, regulators should monitor the bank more closely and should consider raising its capital requirements.

Before concluding, it is necessary to respond to the main objections that could be raised against a government role in this area. First, pay regulation in banks could be opposed on grounds that *it's the shareholders' money*; the government does not have a legitimate interest in telling bank shareholders how to spend their money. Choices of compensation structures, it might be argued, inherently belong to the province of private business decisions where regulators should not trespass. This objection is not persuasive, however, because the government does have a legitimate interest in the compensation structures of private financial firms. Given the government's interest in the safety and soundness of the banking system, intervention here would be no less legitimate than the government's established involvement in limiting banks' investment and lending decisions.

Second, opponents of regulating executive pay in banks could also argue that *regulators know less* – that is, regulators will be at an informational disadvantage when setting pay arrangements. But the knowledge required of regulators to effectively limit compensation structures that incentivize risk-taking would be no

more demanding than that which is requisite to regulators' direct intervention in investment, lending, and capital decisions. Furthermore, setting pay arrangements should not be left to the unconstrained choices of informed players inside banks; these players do not have incentives to take into account the interests of bondholders, depositors, and the government.

Third, opponents may also argue that *bankers will flee* – that is, pay regulation will drive talent away, and financial firms will lose valuable employees. As I stressed, however, regulation of pay in financial firms should focus on pay structures, not on limiting compensation levels. (Prudential regulation may, of course, impose such limits to the extent that compensation level might result in cash outflows that would leave the bank with insufficient capital.) Indeed, the bill passed by the House of Representatives, and the Federal Reserve Board's proposed guidance, explicitly rules out intervention in pay levels. Thus, to the extent that the use of pay structures that eliminate perverse incentives would be less attractive to some executives, banks would be able to compensate those executives with higher levels of expected pay. Even when such an increase proved necessary, however, providing more efficient incentives would be worthwhile.

To reduce the likelihood of future financial crises, we should not only constrain what banks may do but also pay close attention to the incentives provided to the executives who will make choices within these constraints. There are simple but useful ways to induce executives to focus on long-term rather than short-term results, and to take into account the consequences of their decisions for all those contributing to the bank's capital, rather than only for shareholders. Mon-

itoring and encouraging such compensation structures should be an important instrument in the toolkit of financial reg-

ulators. Fixing bankers' pay would reduce the likelihood and costs of future financial crises.

ENDNOTES

- ¹ See Leaders' Statement: The Pittsburgh Summit, September 24 – 25, <http://www.pittsburghsummit.gov/mediacenter/129639.htm>.
- ² Lucian A. Bebchuk and Jesse M. Fried, "Paying for Long-Term Performance," *University of Pennsylvania Law Review* 58 (2010): 1915 – 1960; Lucian A. Bebchuk and Holger Spamann, "Regulating Bankers' Pay," *Georgetown Law Journal* 98 (2) (2010): 247 – 287; Lucian A. Bebchuk, "Fixing Bankers' Pay," *The Economists' Voice* (November 2009).
- ³ Lucian A. Bebchuk, written testimony, hearing on compensation in the financial industry, January 22, 2010, Committee on Financial Services, U.S. House of Representatives; Lucian A. Bebchuk, written testimony, hearing on compensation structures and systemic risk, June 11, 2009, Committee on Financial Services, U.S. House of Representatives.
- ⁴ Lucian A. Bebchuk and Jesse M. Fried, *Pay without Performance: The Unfulfilled Promise of Executive Compensation* (Cambridge, Mass.: Harvard University Press, 2004).
- ⁵ Jeffrey Friedman, "Bank Pay and the Financial Crisis," *The Wall Street Journal* online, September 28, 2009; Floyd Norris, "It May be Outrageous, But Wall Street Pay Didn't Cause this Crisis," *The New York Times*, July 31, 2009.
- ⁶ Lucian A. Bebchuk, Alma Cohen, and Holger Spamann, "The Wages of Failure: Executive Compensation at Bear Stearns and Lehman 2000 – 2008," *Yale Journal on Regulation* 27 (2010): 257 – 282.
- ⁷ All dollar figures are in 2009 dollars.
- ⁸ Bebchuk and Fried, "Paying for Long-Term Performance."
- ⁹ Carr Bettis, John Bizjak, and Swaminathan Kalpathy, "Insiders' Use of Hedging Instruments: An Empirical Examination," unpublished manuscript (March 2010), <http://ssrn.com/abstract=1364810>.
- ¹⁰ See testimony of Kenneth R. Feinberg, the Special Master for TARP Executive Compensation, before the House Financial Services Committee, February 25, 2010, <http://www.ustreas.gov/press/releases/tg565.htm>. Feinberg reports that one of the principles incorporated in evaluating pay at subject firms was that "employees should be prohibited from engaging in any hedging, derivative or other transactions that undermine the long-term performance incentives created by a company's compensation structures."
- ¹¹ Bebchuk and Spamann, "Regulating Bankers' Pay."
- ¹² Lucian A. Bebchuk, "The Case for Increasing Shareholder Power," *Harvard Law Review* 118 (2005): 833 – 914; Lucian A. Bebchuk, "Letting Shareholders Set the Rules," *Harvard Law Review* 119 (2006): 1784 – 1813; Lucian A. Bebchuk, "The Myth of the Shareholder Franchise," *Virginia Law Review* 93 (2007): 676 – 732.

Nolan McCarty, Keith T. Poole,
Thomas Romer & Howard Rosenthal

Political fortunes: on finance & its regulation

The financial crisis of 2008, as well as many earlier crises, had an important political dimension.¹ Government not only failed to intervene to restrain a bubble but also directly abetted the expansion of the bubble. After the bubble popped, political considerations limited and delayed appropriate policy changes.

Both during and after bubbles, political outcomes reflect the same forces that operate in normal times. Such forces include decision-making based on ideologies: free market conservatism, egalitarianism, and populism. Political decisions (and non-decisions) also reflect checks and balances across branches of the federal government and across layers of government in a federal system, as well as institutional checks, such as bicameralism and filibusters. Interests favoring current arrangements benefit from the status quo bias inherent in our political institutions. How these forces operate is greatly influenced by how ideology has contributed to polarization between the two major political parties.

These forces are part of the normal flow of politics in our democracy. Politicians respond to the politically active – campaign contributors and lobbyists in

particular. Policy is swayed by the self-interest of financial firms and, more broadly, creditors and debtors. So-called independent regulators, including the Federal Reserve, respond not only to their own ideology and expertise but also to elected officials with political power over them.

The crisis of 2008 followed the latest in a long history of real estate bubbles in the United States. Bubbles often spill over into other sectors, so that a pop in an asset bubble frequently engenders a banking crisis.

Before 2008, the most recent example of a real estate bubble's devastating effect on financial institutions was the savings and loan (S&L) crisis of the 1980s. This earlier crisis differed in important respects from the 2008 crisis. First, the political system was substantially less polarized than it is today. Second, the economic shock was much smaller. These two factors contributed to a less contentious eventual resolution of the crisis. Moreover, the crash in real estate prices at that time was much more geographically concentrated than in the 2000s, and the thrifts (as the firms in the S&L industry were called) were more Main Street than Wall Street. Nonetheless, the interaction between the finan-

© 2010 by the American Academy of Arts & Sciences

cial and the political sectors strongly foreshadowed the events of 2008.²

Prior to the late 1970s, the S&Ls had a specialized niche: for the most part, they took in short-term deposits and offered long-term fixed-rate mortgages. Trouble might arise if the interest they had to pay to attract depositors rose above the rates on older mortgages, but market conditions and government regulation of deposit interest rates helped S&Ls avoid the potential losses that might result from a mismatch between the short and long terms. A serious mismatch did arise after the financial innovation of money market mutual funds in the 1970s; the funds drew deposits away from S&Ls and eroded their profitability.

The interest rate mismatch became critical as a consequence of the abrupt increase in interest rates initiated by the Volcker Fed and the ensuing severe recession of 1981 to 1982. Real estate values collapsed, especially in the Oil Patch and the Farm Belt. By 1982, more than two-thirds of the thrifts had become unprofitable. In the aggregate, the S&L industry had negative net worth by the regulatory capital standards of the time. The industry's regulator, the Federal Home Loan Bank Board (FHLBB), recognized the situation; however, the deposit insurer, the Federal Savings and Loan Insurance Corporation (FSLIC), had insufficient funds to shut down the insolvent S&Ls.

The political response to the severe economic shocks facing the thrifts was to relax regulatory standards and expand the scope of assets that thrifts could hold. The key elements of the legislation affecting S&Ls in the early 1980s were deregulating interest rates; allowing for adjustable-rate mortgages (ARMs); and permitting S&Ls to expand their loan products from home mortgages to commercial real estate,

junk bonds, and other risky investments.³ Moreover, the FSLIC was not granted additional funding authority, which ensured a policy of regulatory forbearance against failing thrifts.

The profitability of some S&Ls temporarily improved. But the extra risk-taking encouraged by regulatory forbearance soon took its toll. By 1987, the magnitude of the industry's insolvency problem had increased dramatically, yet political action enshrined continued regulatory forbearance by extending the use of lenient accounting rules and weakened capital standards. At the same time, legislation reaffirmed the "full faith and credit" backing of FSLIC-insured deposits without providing the agency additional financing authority to act aggressively against the owners of essentially bankrupt S&Ls.⁴ Equity holders and management in such insolvent firms bore virtually no downward risk. But, as long as they were allowed to operate, they would benefit from any success of risky gambles that restored profitability.⁵ These "zombie thrifts" continued to stay open, making increasingly risky bets as they gambled for resurrection.

Congress and the White House opted for forbearance for two related reasons. First, ending the S&L debacle required an unplanned and unbudgeted expenditure. The funds could come either from assessments on the industry or from general government taxes and borrowing. Sick thrifts had a strong interest in continued forbearance; the remaining healthy thrifts resisted the additional fees that they may have been assessed under a recapitalized FSLIC. The immediate beneficiaries of lax regulation were politically active. They pressed elected officials to support their case with the regulators and in framing legislation. There was no significant constituency in favor of confronting the full magni-

tude of the S&L problem when it was still relatively small, or of curtailing – rather than extending – forbearance. Importantly for connecting the dots to the current crisis, these looser regulations remained in effect even after the S&L meltdown.

Second, incumbents of both parties were concerned that dealing with the crisis through a large recapitalization of FSLIC would look like a bailout and would be unpopular with the voters in the November 1988 elections. Both parties supported banking deregulation and regulatory forbearance. Even though experts both inside and outside the Beltway widely recognized the gravity of the S&L industry's situation, the candidates mostly avoided discussing the crisis during the presidential election campaign.

By early 1989, the need to shut down failed S&Ls and to recapitalize the deposit insurance fund could no longer be put off. Relatively rapid congressional action produced the Financial Institutions Reform Recovery and Enforcement Act of 1989 (FIRREA), which President Bush signed that August. FIRREA led to a bailout of depositors and disposal of the assets of failed thrifts by a new entity, the Resolution Trust Corporation (RTC). The FHLBB was folded up and a new regulator, the Office of Thrift Supervision (OTS), put in place. The OTS represented a soft reform: by the 2000s, it was a regulatory venue of choice for mortgage institutions seeking to escape the stricter regulatory arm of the Fed. A concession to the industry in 1989, OTS in more recent years accommodated the likes of Washington Mutual, IndyMac, and AIG.⁶

In brief, unintended consequences of financial innovation and deregulation produced a crisis; interest-group politics and populist abhorrence of bailouts pro-

longed it. The media coverage of the S&L crisis in full bloom focused on colorful stories of malfeasance and alleged political corruption (most notably Charles Keating of Lincoln Savings and Loan and his five senators, including future presidential nominee John McCain). But the true highlight tape should recognize that most of the S&L gambles were completely legal actions in the regulatory environment created by elected officials responding to interest-group pressures and to electoral concerns with the mass public.

The resolution of the S&L debacle did not include tightening the looser regulation of interest rates and other dimensions of the home mortgage industry. Not surprisingly, then, financial markets witnessed an explosive growth of financial innovation in a generally permissive regulatory environment. Before the 2008 meltdown, this financial innovation was generally seen as a positive force that delivered real benefits. The pace of innovation accelerated during a period of global economic growth and, in particular, the American economic ascendancy of the 1990s. Financial sector profitability rose and financial sector profits became a much larger share of total corporate profits. Financial sector wages increased much faster than wages in other sectors.⁷ Highly skilled individuals flocked to financial engineering.

While the financial sector often attributed this innovation and growth to the virtues of unbridled capitalism and free markets, the shadow of politics was never far behind. First, much of this innovation was designed to create instruments that would optimize profits around regulatory constraints. Notably, banks invented off-balance-sheet vehicles that created leverage without violating the Basel I capital requirements.⁸

Second, the importance the industry placed on maintaining a favorable political environment is underscored by a massive increase in political involvement. Between 1992 and 2008, political campaign contributions from the financial sector nearly tripled, even after adjusting for inflation. (Only the legal profession had a faster growth rate of contributions.) The current magnitude of giving is also remarkable. Four subsectors of finance (securities and investments, real estate, insurance, and miscellaneous finance) are now in the top ten of all industry contributions, and two of them (securities and investments, real estate) have dominated the growth in contributions.⁹

Although Republicans and Democrats both have been blessed with contributions from the financial sector, the implications for the Democratic Party have been especially significant. Over the past forty years, voting behavior and partisan identification in the United States have become highly structured by income. The Republican Party has fared somewhat better among voters with middle income and higher, while the Democratic Party has received the majority of the votes from those with lower incomes.¹⁰ But even as the Democrats were depending more heavily on the votes of lower-income citizens, the party came to rely more on the financial resources of wealthier supporters and interest groups. Consequently, the Democratic Party now has two distinct wings: the “money wing” and the “votes wing.” The financial services sector has become an increasingly important part of the money wing.

Like many corporate contributors, the financial sector tends to shift its campaign contributions based on which party controls Congress. From 1987 to 1994, a majority of the money went to the Demo-

cratic Party, which had majorities in both the House and the Senate. Following the 1994 elections, there was a large shift in contributions to the newly empowered Republicans. Following the Democratic takeover of Congress in 2006, the money switched back to about where it was in the early 1990s.

But even during the period when it had relatively little power in Washington, the Democratic Party did well with the securities and real estate industries. From the party’s perspective, finance makes an almost ideal money wing (almost as good as Hollywood). Unlike many other industries with potential claims on political attention, financial sector firms do not pollute the environment (at least not directly) and do not have especially contentious labor relations. During normal times and during economic booms, these facts make for much less conflict with other Democratic constituencies like environmentalists and labor unions. Of course, an ideal alliance in good times is not necessarily one that can survive a bust.

In a Washington epitomized by Alan Greenspan’s nearly nineteen-year tenure as Fed chairman, regulatory constraints were viewed in a dim light. There was either bipartisan support for a hands-off policy or lack of a winning coalition that would revamp regulation to address a new and highly complex environment. Indeed, a considerable amount of financial deregulation occurred through agency decisions and legislation during the Clinton administration. In 1998, Brooksley Born, Clinton’s appointee as chair of the Commodity Futures Trading Commission (CFTC), proposed regulating off-exchange trades in swaps and other derivatives. Her initiative was shot down by Greenspan, Treasury Secretary Robert Rubin, Deputy Secretary Larry Summers,

and Securities and Exchange Commission (SEC) Chairman Arthur Levitt. The Commodity Futures Modernization Act of 2000 (CFMA) – passed by an overwhelming 377 to 4 roll-call vote in the House and a voice vote in the Senate – codified the marginal role of the CFTC and facilitated active expansion of over-the-counter derivatives markets.

Previous executive and deregulatory measures of the financial sector culminated in the Financial Services Modernization Act of 1999, which also passed with large bipartisan majorities.¹¹ It repealed the Glass-Steagall Act of 1933, which had separated commercial banks, investment banks, and insurance companies; it also explicitly prohibited the SEC from regulating securities-based swap agreements. During the decades when these boundaries were contested, commercial banks, investment banks, and insurance companies were frequently on different sides of disputes over financial regulation. The eventual erosion of legal and regulatory boundaries between segments of the financial sector also removed some of the political conflicts over economic turf within the sector. As deregulation took these issues off the table, previously separated parts of the sector now had more closely aligned interests.¹²

The turn of the century was not without warning that deregulated financial markets posed dangers. The collapse of the hedge fund LTCM in 1998 revealed the potential for global contagion and systemic risk. A second warning shot was fired in 2001 with the failure of Enron, WorldCom, and other firms with accounting scandals. It became apparent that the specifics of deregulation were designed to benefit financial and corporate interests – such as the “Enron loophole” in the CFMA.¹³ There was also evidence that deregulation’s fundamental

premise – that accounting firms, Wall Street analysts, and rating agencies would provide unbiased and transparent information to investors – was false. Nonetheless, the spirit of deregulation continued relatively unabated. Much like FIRREA, the Sarbanes-Oxley Act (2002) did not address forms of accounting arbitrage, such as off-balance-sheet vehicles that avoided capital requirements, employed by financial institutions as they marched toward the 2008 crisis. During the George W. Bush administration, the SEC essentially eliminated its capacity to inspect investment banks and weakened its enforcement.¹⁴ Ignoring the warning signals of the turn of the century was facilitated by continued economic growth. The crisis of 2008 arrived in an environment where most financial transactions were unregulated and regulated activities were largely unmonitored.

By 2000, the foundation had been laid for a housing market bubble that would reflect the alignment of two disparate ideologies: free market conservatism and redistributive egalitarianism. The foundation consisted of a crazy quilt of legislative, executive, and judicial decisions in the 1980s and 1990s that are overwhelming in their details. A brief recap would note that (1) a variety of financial institutions were permitted entry into various loan and insurance markets, (2) usury laws were stripped away, and (3) the ability of states to regulate financial products, like credit cards, offered by out-of-state firms was largely eliminated while other features, such as loan-to-value limits, remained with the states. States could also compete in the chartering of banks. The result was the expansion of predatory lending practices, loans that would be quickly underwater if housing prices declined, and

another round of “zombie” banking that encompassed institutions large – Washington Mutual, for example – and small, including forty Georgia banks eventually forced into receivership by the FDIC.¹⁵

A Republican twist to free market capitalism was a belief in the political advantages of the “ownership society.” In brief, homeowners were thought to be more likely to vote Republican than renters.¹⁶ The post-1980 support for low- and moderate-income housing by Republicans represented a historical shift. In 1977, conservatives, supported by lenders, had mustered considerable backing for a Senate amendment that unsuccessfully sought to delete Title VIII, the Community Redevelopment Act, from the housing bill that passed that year. The later shift away from partisanship on housing policies came after the deregulation of the 1980s and 1990s created a profitable market in loans to low-income families.¹⁷ By the administration of George H.W. Bush, there was bipartisan support for legislation such as the National Affordable Housing Act (1990), which included a variety of initiatives directly aimed at expanding homeownership among low-income households, and the Federal Housing Enterprises Financial Safety and Soundness Act (1992), which set minimum percentage-of-business targets for Fannie Mae and Freddie Mac purchases of mortgages issued to low-income households.

On the Democratic side, redistributive egalitarianism sought to increase homeownership among the poor and minorities. Fannie Mae stated in its 2003 Annual Report, “[A]s long as there is a gap in minority and non-minority homeownership rates, Fannie Mae and Countrywide will continue to make sure all Americans have the chance to realize the dream of homeownership.” While it was not politically feasible to make homeownership

an entitlement like food stamps and Medicare, political pressure and economic incentives encouraged lenders in the mortgage market to accomplish egalitarian goals. Promoting egalitarianism through a market led by government-guaranteed enterprises Fannie Mae and Freddie Mac appealed not only to the Democratic base but also to highly compensated Democrats. The CEO of Fannie Mae from 1991 to 1998 was James Johnson, an executive assistant to Vice President Walter Mondale, member of the John Kerry vice presidential selection team, and cashiered member of the Obama vice presidential selection team.¹⁸ From 1998 to 2004, Fannie Mae was headed by former director of the Office of Management and Budget during the Clinton administration, Franklin Raines.¹⁹ Raines presided over Fannie Mae during a major accounting scandal; his executive compensation has been extensively criticized.²⁰

Housing policy was one area largely devoid of partisan conflict in an era characterized by highly polarized politics.²¹ The American Homeownership and Economic Opportunity Act of 2000 – directed at easing the financing of mortgages, including reverse mortgages, and increasing financial assistance for homeownership by the poor, elderly, and disabled – was passed by voice vote in the House and unanimous consent in the Senate. The American Dream Downpayment Act of 2003 was passed by unanimous consent in the Senate and without objection in the House.²² President Bush enthusiastically signed the act into law, as a measure that would build the “ownership society” by providing “\$200 million per year in down payment assistance to at least 40,000 low-income families.”²³

Both “American” acts relaxed standards for lenders. In contrast to these measures directed at expanding homeownership,

from 2000 to 2006 only one of sixteen legislative bills aimed at curbing predatory lending and enforcing other aspects of consumer protection passed the House; none passed the Senate; and, of course, none became law. Lobbying against these bills came disproportionately from riskier lenders that originated mortgages with high loan-to-income ratios, used securitized instruments, and had fast-growing mortgage loan portfolios.²⁴ As in the S&L fiasco, lenders in the weakest economic condition were most likely to lobby against regulatory constraints.²⁵

The “What, me worry?” approach to regulation had been made possible by the agreement of free market capitalists and redistributive egalitarians on policy in the mortgage market. This coalition was blown apart by the unraveling of the housing bubble that preceded the financial crisis. Housing prices started to fall in 2006, subprime mortgage bonds plummeted in early 2007, and residential foreclosures accelerated. Dealing with the foreclosure problem meant moving from the “win-win” scenario of the pumped-up housing market to the hard choices involving helping borrowers through an on-budget bailout or forcing lenders to take a haircut via a moratorium or an imposed reduction in either principal or interest.

Making these choices took place in a far more polarized political environment than was present at the time FIRREA resolved the S&L crisis. Partisan “blame game” politics prevailed over any grand crisis coalition in the national interest. The first move in this game was the American Housing Rescue and Foreclosure Prevention Act (AHRFPA), signed into law by George W. Bush in July 2008, after he had threatened to veto an earlier bill passed by the House in May. The July

vote on final passage in the House showed a sharp ideological split, with unanimous support from Democrats, support from more moderate Republicans, and opposition from conservative Republicans. Salient local issues can at times trump ideology, lobbying, and contributions from out-of-constituency interests. In the AHRFPA case, Republican House members also proved sensitive to the extent of foreclosures in their districts, especially foreclosures in Republican voting areas of the districts.²⁶ AHRFPA, however, left so much discretion to lenders that it did little to avoid foreclosures.

The bubble definitively popped with the failure of Lehman Brothers in September 2008. Political circumstances strongly influence how the executive and regulatory branches handle a pop. The S&L crisis did not endanger the entire economy; for electoral considerations, both the Reagan administration and the Democrat-controlled Congress chose to sweep the crisis under the rug in 1987–1988. It is instructive to compare the 2008 crash with the stock market crash of October 1929 that occurred just seven months into the four-year term of Herbert Hoover. His administration, largely personified by Treasury Secretary Andrew Mellon, responded passively. In a European parliamentary system, Hoover would not have been expected to survive for four years. The American institution of four-year terms may well have contributed to the severity of that crisis. (The Democrats did capture the House in the 1930 midterm elections.)

In contrast to the setting of the Depression, the crisis of 2008 occurred less than two months before an election and less than four months before the inauguration of the new president. Not only was the incumbent president a lame duck, his presidency was tarnished by Iraq

and other events. Bush had very limited sway over Republicans in a Democrat-controlled Congress. In the push for the financial reform bill that passed in July 2010, President Obama was out front, with Treasury Secretary Geithner in a supporting role and Fed Chairman Bernanke largely off-stage. But when the housing bubble popped, Treasury Secretary Henry Paulson and Bernanke were front and center. In the absence of presidential leadership, Paulson and Bernanke were especially concerned that their actions be legitimated by Congress – or at least not overly attacked. The financial markets' response to the initial failure of the Troubled Asset Relief Program bill (TARP, "the bank bailout") showed that those markets could be roiled by either the capitalist right, fixated on moral hazard, or by the progressive left, bothered by the distributional implications of a bailout.

Opposition from both extremes of the liberal-conservative spectrum, reelection concerns, and a proposal from a Republican administration facing a Democratic Congress combined to blur ideological voting on the TARP bill in Fall 2008. When the auto bailout bill was voted on in December, there was strong ideological polarization, a pattern that reached perfection when the stimulus bill was considered early in the Obama administration. In the House, there was perfect party separation, except for six, nonpivotal Democrats who were "allowed" to defect. In the Senate, there was perfect ideological separation; the administration made just enough concessions to buy the votes of the three least conservative Republicans: Arlen Specter, Olympia Snowe, and Susan Collins. This pattern, likely to reappear on all aspects of financial reform, was again manifest when Edward Kennedy's replacement, moderate Republican

Scott Brown, demanded and obtained modifications to the conference report on the 2010 reform bill. The modifications benefited such financial institutions as State Street Corporation and Fidelity Investments, headquartered in Brown's home state of Massachusetts.²⁷

Not only is polarization problematic in initial reform efforts, the gridlock it produces impedes the routine legislative maintenance required for a robust regulatory environment.²⁸ Any reform legislation that may be forthcoming will undoubtedly need such future fine-tuning.

Ideally, regulators would have the resources and expertise to monitor closely developments in the regulated sector and the incentive to promulgate and implement policies that are in the public's interests. The conditions for such regulatory performance are hard to meet in any domain, but the problem of regulatory capacity is particularly acute in the case of financial regulation.

The most obvious difficulty stems from the complexity of modern finance. Armies of rocket scientists are employed to develop and implement increasingly complicated financial products and trading strategies. Many of the products are not well understood by Wall Street executives, much less outside regulators. This problem might be mitigated somewhat if the regulatory agencies could easily draw from the same talent pool as Wall Street. But the salary differentials make this difficult. The highest paid financial regulators (the president of the New York Fed and the chairman of the Fed) make a fraction of a middling trader's annual bonus. Even where regulatory agencies can hire individuals with the background to understand the intricacies of modern finance, such individuals are usually on their way from or their way back to Wall Street. Such a revolving

door undermines the autonomy of regulatory agencies from the industry they are supposed to regulate. Even if such regulators are not motivated by a future Wall Street payout, they may still be inclined to share Wall Street's worldview.

The most direct implication of low regulatory capacity is that it will be hard to sustain a regulatory regime that depends too heavily on the delegation of discretionary power to regulators. This concern speaks directly to the debate about whether a council of super-regulators can monitor the financial sector for emerging systemic risks and react effectively with new capital requirements, leverage limits, or conversion of contingent bonds. Such a system requires that regulators have very high levels of information and expertise as well as the incentive to act in ways that may be adverse to the financial services industry. The recognition of low capacity argues against sophisticated discretionary regulatory management of the industry and in favor of blunter approaches such as banning the most systematically dangerous products and practices or capping the size of financial institutions. Blunter, less complex, and less lengthy legislation would not only reduce opportunities for financial innovators to find loopholes but also focus the attention of regulators.

Beyond the technical problems that plague low-capacity agencies, there is important political feedback. Low capacity makes it harder to hold agencies accountable to congressional and presidential oversight because it is harder to distinguish between bad policies and poor implementation. This may cause elected leaders to be reluctant to endow agencies they cannot control with substantial discretionary power.

Recently enacted legislation strives to mitigate these problems by strengthening the informational capabilities of reg-

ulators (such as the creation of the Office of Financial Research in Treasury) and by enhancing accountability mechanisms (especially those relating to the Federal Reserve, including presidential appointment of the head of the New York Fed and restriction of the Fed's emergency lending authority with increased oversight by the Government Accountability Office). But there are strong reasons to believe that these reforms by themselves will not significantly improve financial regulation. Better information and policy analysis can go only so far if agencies lack the resources to act effectively on that information and analysis. The lack of expertise and information also reduces the value of increasing regulatory capacity. Why increase the ability of an agency to implement uninformed policies? Expertise and capacity are complements; this creates a bureaucratic "reform trap."²⁹

A reform trap also exists with respect to improving oversight and accountability. Investing in greater oversight of agency decisions is most valuable when the links between agency policy and outcomes are the most transparent, because it is then easier to detect policies Congress does not approve of. Since low capacity distorts the relationship between policies and outcomes, more oversight is not very helpful. Conversely, when oversight mechanisms are poor, raising capacity is not very valuable since the political overseers do not benefit from the increased transparency of the policy-outcome link. Reforms, therefore, might well consider restrictions on firm behavior that would simplify rather than expand regulatory authority. For example, a relatively limited (and visible) menu of exchange-traded derivatives might be preferable to a giant (but semi-secret) smorgasbord of over-the-counter special orders. Similarly, bank supervision

might benefit from the Volcker Rule, which, among other things, would prohibit banks from running private-equity and hedge funds.

Although the current economic situation is much more complex than that of the S&L crisis, there are some similarities. There is conflict between large and small banks and between interests that were cashiered (Lehman Brothers, Washington Mutual) and those that were rescued (Citicorp, Bank of America). Both crises were influenced by regulatory venue shopping. Existing regulators have fought to maintain their turf, without regard to competency. Both crises were influenced by “zombie seeking” of high returns without due regard to risk. The reluctance to deal with the S&L crisis when the 1988 elections were on the horizon appears to be paralleled by a failure to address reform of Fannie Mae and Freddie Mac before the 2010 elections. More generally, the inadequacy of the S&L reforms gives pause as to the effectiveness of the recently passed Dodd-Frank Wall Street Reform and Consumer Protection Act.

There are several difficulties in the political process and in a “regulatory-financial complex” that present severe hurdles to effective regulation of the financial sector. First, we should keep in mind an observation made by the political scientist E. E. Schattschneider: in normal times, policy reflects the balance of political clout among moneyed interests; but if an issue becomes broadly salient, as after an economic crisis, then all bets are off. Politicians care about being reelected, so if the “folks” get upset the politicians will cater to their wishes (or at least what the politicians perceive these to be). Once an issue sinks back below the surface, the moneyed interests reassert themselves.³⁰ With the new reg-

ulatory reform legislation, many of the key decisions will be made by regulators months and years afterward – when financial regulation is not nearly so salient – and therefore can be expected to be much more deferential to the industry than the spirit of the legislation.

Second, because political polarization appears to decline and rise with income inequality,³¹ and income inequality in turn partly reflects the strong increase in financial sector profits and income,³² future efforts at dealing with U.S. financial crises are likely to continue to occur in a highly polarized environment. Polarization inhibits timely and effective responses, both by causing legislative gridlock and by increasing the willingness of those out of power to block all change until they get back in, rather than compromise on centrist policies. Intensive lobbying from affected interests further reinforces ideological opposition.

Third, financial markets are now extremely complex, with myriad products. Regulatory incentives, staffs, and budgets are not aligned toward successful monitoring and enforcement.³³ Consequently, even if financial products are limited and regulated, private parties will continue to innovate within the constraints of existing regulation. Technological change, as in data processing and mathematical models, will bring forth innovations. Regulators, in contrast, are unlikely to be proactive with respect to innovations.

Fourth, there is a taboo against rigorous enforcement. In his remarks at the Pittsburgh G20 meeting in 2009, President Obama spoke of the financial crisis as the result of the “reckless few.” After Enron and WorldCom, President Bush claimed “a few bad actors can tarnish our entire economic system.” And following the Panic of 1907, President

Roosevelt comforted the public by declaring, “Dishonest dealing and speculative enterprise are merely the occasional incidents of our real prosperity.”³⁴ Once the immediate crisis has passed, there is little continued political payoff in either stating or acting on the premise that most of those in the financial sector will, at best, bend the rules as much as they can and push politicians to get rid of the rules when the rules are inconvenient.

Finally, not only have financial markets become more interconnected, so have the worlds of finance, politics, philanthropy, and academia. For example, the outside directors of AIG between 2005 and 2008 included Obama and Clinton diplomat Richard Holbrooke, Clinton Defense Secretary William

Cohen, Reagan White House advisor and Harvard economist Martin Feldstein, Ford Housing and Urban Development Secretary and Bush trade representative Carla Hills, Ford “energy czar” Frank Zarb, American Museum of Natural History President Ellen Futter, and public television executive George Miles. Feldstein, Futter, Holbrooke, Miles, and Zarb served on the board for all or part of the years 2005 to 2008. Over that period, their individual compensation as director ranged from \$792,000 to \$1,136,000.³⁵ Again, once the crisis is past, criticism of the financial sector from government, academia, and nonprofits is likely to be muted. We appear to be stuck with the regulatory-financial complex.

ENDNOTES

¹ Poole thanks his colleagues at the University of Georgia for offering many helpful comments. Romer is grateful to the United States Studies Centre at the University of Sydney, where he was on sabbatical leave during the North American winter months of 2010, for its gracious hospitality. It was an excellent environment for stimulating conversations about the global financial crisis. Rosenthal thanks his colleagues at the California Institute of Technology in the spring quarter of 2010, especially Peter Bossaerts and Jean-Laurent Rosenthal, for discussion and comments, as well as participants in conferences at Columbia University and Università Cattolica del Sacro Cuore.

² Our summary of the politics of the S&L crisis is based on Thomas Romer and Barry R. Weingast, “Political Foundations of the Thrift Debacle,” in *Politics and Economics in the 1980s*, ed. Alberto Alesina and Geoffrey Carliner (Chicago: University of Chicago Press, 1991).

³ The two principal pieces of legislation were the Depository Institutions Deregulation and Monetary Control Act of 1980 and the Garn-St. Germain Depository Institutions Act of 1982.

⁴ Competitive Equality Banking Act of 1987.

⁵ See David H. Pyle, “The U.S. Savings and Loan Crisis,” in *Handbooks in Operations Research and Management Science*, vol. 9, *Finance*, ed. Robert A. Jarrow et al. (Amsterdam: Elsevier, 1995).

⁶ The RTC is now generally thought to have done a good job, though it had some rocky years. In disposing of the assets of failed thrifts, the RTC faced conflicting political mandates (including social policy objectives such as minority contracting and affordable housing, as well as getting the highest value for the assets). See Lee Davison, “The Resolution Trust Corporation and Congress, 1989–1993. Part II: 1991–1993,” *FDIC Banking Review* (2006), <http://www.fdic.gov/bank/analytical/banking/2007apr/article1/index.html>.

- 7 Thomas Philippon and Ariell Reshef, "Wages and Human Capital in the U.S. Financial Industry: 1909 – 2006," NBER working paper 14644 (National Bureau of Economic Research, 2009).
- 8 See Markus K. Brunnermeier, "Deciphering the Liquidity and Credit Crunch 2007 – 08," *Journal of Economic Perspectives* (Winter 2009).
- 9 Lawyers and law firms, considered a separate industry, contributed \$126 million in 2008. Many of these contributors may also have had interests in financial sector regulation.
- 10 See Nolan McCarty, Keith T. Poole, and Howard Rosenthal, *Polarized America: The Dance of Ideology and Unequal Riches* (Cambridge, Mass.: MIT Press, 2006).
- 11 The law is more commonly known as the Gramm-Leach-Bliley Act, after its three Republican cosponsors.
- 12 See Luigi Zingales, "Capitalism After the Crisis," *National Affairs* (Fall 2009): 30 – 31.
- 13 Certain energy derivatives contracts were exempt from regulation by the CFTC. One of these exemptions was for trades conducted over electronic trading platforms such as the one developed by Enron; hence, the "Enron loophole." See Mark Jickling, "The Enron Loophole," report RS22912 (Congressional Research Service, 2008).
- 14 Ross Levine, "An Autopsy of the U.S. Financial System," NBER working paper (National Bureau of Economic Research, April 2010).
- 15 <http://www.fdic.gov/bank/individual/failed/banklist.html> (accessed July 21, 2010).
- 16 This belief extended to support for self-directed defined contribution pension plans and privatization of Social Security.
- 17 See Edward M. Gramlich, *Subprime Mortgages: America's Latest Boom and Bust* (Washington, D.C.: The Urban Institute Press, 2007), 5.
- 18 Johnson resigned his position on the selection committee after it was revealed that he received a loan on very favorable terms from Countrywide Financial Corporation, Fannie Mae's largest mortgage provider and a key player in the subprime lending crisis. See John M. Broder and Leslie Wayne, "Obama Aide Quits Under Fire for His Business Ties," *The New York Times*, June 12, 2008.
- 19 Clinton also named former Arizona Senator Dennis DeConcini to the Freddie Mac board, where he served from 1995 to 1999. DeConcini was one of the Keating Five and was formally criticized in 1991 by the Senate Ethics Committee for improperly interceding with the FHLBB on Charles Keating's behalf.
- 20 See Lucian A. Bebchuk and Jesse M. Fried, "Executive Compensation at Fannie Mae: A Case Study of Perverse Incentives, Nonperformance Pay and Camouflage," *The Journal of Corporation Law* (2005).
- 21 See McCarty, Poole, and Rosenthal, *Polarized America*.
- 22 Some conservative groups, such as the Heritage Foundation, opposed the bill as "fiscally irresponsible." See <http://www.heritage.org/Research/Reports/2003/12/American-Dream-Downpayment-Act-Fiscally-Irresponsible-and-Redundant-to-Existing-Homeownership-Programs> (accessed April 22, 2010).
- 23 See remarks of President George W. Bush, December 16, 2003, <http://www.presidency.ucsb.edu/ws/index.php?pid=64935> (accessed April 22, 2010).
- 24 See Deniz Igan, Prachi Mishra, and Thierry Tresselt, "A Fistful of Dollars: Lobbying and the Financial Crisis," working paper 09/287 (International Monetary Fund, 2009).
- 25 The development of the subprime mortgage market can also be traced back to the deregulation of the early 1980s. Subprime mortgages were based on charging higher interest rates and offering ARMs to risky, often low-income and minority, borrowers. If borrowers had

- been informed of the risks, then allowing for the “freedom to choose” among a wider variety of mortgage products would make sense. Instead, naive borrowers were faced with predatory, fraudulent, and unsupervised mortgage originators. See Gramlich, *Subprime Mortgages*.
- ²⁶ Conservative Republicans, ideologically opposed to government intervention, were less likely to be swayed by economic conditions in their districts than were their more moderate colleagues. See Atif Mian, Amir Sufi, and Francesco Trebbi, “The Political Economy of the U.S. Mortgage Default Crisis,” *American Economic Review* (forthcoming).
- ²⁷ Damian Paletta, “Finance Bill Close to Passage in Senate,” *The Wall Street Journal*, July 13, 2010.
- ²⁸ For example, although Glass-Steagall is now often hailed as instrumental in maintaining financial stability through the 1980s, the regulatory regime it created needed to be patched and extended by other important pieces of legislation, such as the Public Utility Holding Company Act of 1935, the Investment Act of 1940, and the 1956 Banking Act.
- ²⁹ John Huber and Nolan McCarty, “Bureaucratic Capacity, Delegation, and Political Reform,” *American Political Science Review* (2004).
- ³⁰ Elmer Eric Schattschneider, *The Semisovereign People: A Realist’s View of Democracy in America* (New York: Holt, Rinehart and Winston, 1960).
- ³¹ McCarty, Poole, and Rosenthal, *Polarized America*.
- ³² Philippon and Resheff, “Wages and Human Capital in the U.S. Financial Industry.”
- ³³ In the context of investigating the one thousand-point plunge of the Dow Industrial Average on May 6, 2010, SEC Chairman Mary Schapiro stated that “the technology for collecting data and surveilling our markets is . . . as much as two decades behind the technology currently used by those we regulate.” See “Warp-speed Trades Outpace SEC,” Politico.com, June 1, 2010.
- ³⁴ President Barack Obama, press conference, September 25, 2009, <http://www.whitehouse.gov/the-press-office/remarks-president-g20-closing-press-conference>; President George W. Bush, weekly radio address, June 29, 2002, <http://www.whitehouse.gov/news/releases/2002/06/20020629.html>; President Theodore Roosevelt, October 24, 1907, in *The Letters of Theodore Roosevelt*, vols. 1–8, ed. Elting E. Morison (Cambridge, Mass.: Harvard University Press, 1952), 839, cited in Robert F. Bruner and Sean D. Carr, *The Panic of 1907: Lessons Learned from the Market’s Perfect Storm* (Hoboken, N.J.: Wiley, 2007).
- ³⁵ From AIG’s DEF14A filings, available at <http://www.sec.gov>. For 2006, 2007, and 2008 we used AIG’s pricing of the value of stock grants and options. For 2005, we priced stock grants at the closing price on December 30, 2005. We did not include a value for 2005 stock options. Much of the value of stock awards was erased in the financial crisis.

C.A.E. Goodhart

*Lessons from the financial crisis
for monetary policy*

The years preceding the financial crisis that began on August 9, 2007, were in many respects the most economically successful in human history. These years have been called the Great Moderation, as output growth, inflation, unemployment, and interest rates stabilized at comfortable levels and remained steady. Most developed countries experienced remarkably little volatility. Japan's lost decade was an exception to the several other emerging Asian countries – especially China, but also South Korea, India, and Indonesia – that grew even faster, taking millions out of poverty.

Expert commentators awarded much of the credit for this improved performance to a new regime of monetary policy. In academia, the extent to which the better outcome was the result of good policy, good luck (for example, the effect China's entry into world markets had on limiting inflation), or yet other factors is still debated. Under this new regime, a nation's central bank raised interest rates whenever inflation was forecast to rise above a predetermined low target level, and lowered them if inflation was expected to fall below that level. Indeed, in most countries – excepting the United States –

this policy was termed “inflation targeting.” Although policy in the United States was managed in very much the same way as elsewhere, the objectives for the Federal Reserve Board's policy had been established earlier, in the Humphrey-Hawkins Act of 1978, and included a requirement that the Federal Reserve System maintain maximum sustainable output as well as stable prices.

In practice, this difference in stated objectives made little difference to policy outcomes. Changes in the rate of inflation are believed to be determined by whether output – otherwise called the pressure of demand – is above the “natural” level the economy can generate when employment/unemployment is at a normal level, or equilibrium. The margin between actual and equilibrium output is called the output gap, though this is not accurately measured, and there are continuous disputes about the size, sometimes even the sign, of this gap.

Thus, whether a country (like most) was an explicit inflation targeter or not (like the United States), the standard operational behavior of central banks was essentially the same. Central banks raised interest rates whenever the output gap (actual minus equilibrium) was positive and inflation was expected to be above target. Rates were generally

© 2010 by the American Academy of Arts
& Sciences

calibrated by a Taylor reaction function, named for John Taylor, the Stanford University economist. In the Taylor reaction function, the current interest rate decision responds to the output gap – as the gap will drive inflation further up (or down) if output is assessed to be above (or below) its equilibrium level – and to the deviation of inflation from its target level.

This policy relied on the general assumption that as long as the monetary authorities (that is, central banks) kept the macroeconomy stable, then broad and efficient financial markets could be expected to monitor financial stability. Although most economies, including the United States, had remained remarkably stable over the fifteen years from 1992 to 2007, experience had shown that even in stable macroeconomic conditions, financial markets can be subject to waves of optimism and pessimism, greed and fear, bubble and bust. Indeed, the financial climate oscillated even during the Great Moderation: the information technology (IT) and NASDAQ bubble-and-bust cycles of 2000 and 2001 are the most recent and perhaps best remembered.

Prior to the start of the financial crisis in August 2007, the monetary authorities, notably the Fed, were able to defang the “bust” in each case and limit its virulence by a program of quick and aggressive interest rate reductions. These swift reactions gave most people in financial markets confidence not only that the general economy would remain more stable than in the past, but also that any residual financial tremors would soon be righted by a vigilant and powerful Federal Reserve System; Chairman of the Federal Reserve Board Alan Greenspan would always support financial markets in any collapse, so the belief went.

Central bankers, like most of us, believed in this context that financial sta-

bility was under control, and a few, including the Bank of England (BoE), even deemphasized their financial stability divisions. Generally forgotten was the analysis, developed earlier by economist Hyman Minsky, that price and macroeconomic stability do not equate to financial stability; rather, the former may have an inverse effect on the latter. Minsky argued that confidence in financial markets incentivizes financiers to take on additional leverage and more risk, while sometimes failing to notice the existence of such risks. Margins and liquidity ratios decline so much that when something goes wrong (such as the subprime mortgage crisis in 2007), the system lacks sufficient equity and liquidity buffers to absorb the loss and the system becomes overextended. Banks are forced to sell assets into unwilling markets in order to restore their own liquidity; these sales so reduce the market value of other banks’ assets that they, too, come under pressure to sell their own assets. This scenario and similar factors converge to amplify the collapse. Earlier serious financial collapses (such as the Great Depression in 1929 and Japan’s collapse in the 1990s) were also preceded by periods of strong growth, enhanced confidence, and rising asset prices.

Prior to the recent recession, moreover, there had been a long-term trend of credit expansion growing faster than retail bank deposits. Banks and investment houses had financed the credit expansion by reducing holdings of public sector liquid assets; relying more heavily on (short-term) borrowing from wholesale markets; and selling earning assets to various other nonbank financial intermediaries. The latter is known as securitization (the originate-to-distribute model). The start of the crisis was marked by a malfunction in the wholesale money market, whereby banks could access such

funds only at elevated spreads or costs, if at all, and by halting securitization.

Thus the banking system had become steadily more fragile. High reported profits disguised this decline during the preceding asset price boom, a façade enhanced by changes in accounting practices that allowed bank assets to be revalued at current market prices (mark-to-market accounting). Neither bankers nor regulators were really aware of the fragility underlying the financial system (though some evidence suggests that bankers had miscalculated risk and were not, in fact, cynically betting on a taxpayer bailout). A combination of untoward confidence in the strength of the financial system; an unfortunate coincidence of an inflationary bubble in commodity prices, notably in oil; and excessive concern about moral hazard led central banks around the world to delay aggressive cuts in interest rates until after the sky fell, on September 15, 2008, with the Lehman Brothers bankruptcy.

We have learned that the achievement of price stability does not in itself guarantee financial stability: these are two separate objectives for central banks. How can they be simultaneously achieved, and what changes, if any, should this assessment make for the conduct of monetary policy?

There are two, not necessarily mutually exclusive, main adjustments that have been proposed. The first is to relax the focus on targeting the price inflation for current goods and services and to give more weight to asset prices, a strategy termed “leaning into the wind.” The second recommendation is to devise a second set of macro-prudential, counter-cyclical regulatory instruments that can be *directly* deployed to diminish asset price bubbles and busts.

I favor the second approach. As of 2010, the idea of relaxing adherence to an inflation target has little support. Many fear the likelihood of severe deflation, sharp inflation, or even one followed by the other. The extraordinary stimulus to both fiscal and monetary policies, the sharply rising debt ratios (both public and private), and intensifying political difficulties for most incumbents at a time of economic recession and high unemployment all provide a fertile basis for an inflationary upsurge. At the same time, the (broader) money stock and bank lending continue to decline as banks, under pressure from public and regulatory authorities, seek to retrench and reduce their dependence on (wholesale market) debt finance (deleveraging). Both the public and private sectors are also trying to raise savings and pay off debt. The crisis is not over, particularly as evidenced by the debacle in Greece, and further debt or deflation remains possible. Against this background of uncertainty and the fear of both inflation *and* deflation, it is now more than ever important to provide reassurance that monetary policy will remain focused on achieving price stability.

Identifying unsustainable asset price bubbles at the time they occur is extremely difficult – although it always looks easy *after* the fact. Asset prices could not go to such heights if there were not large parts of the market that expected them to go higher yet, and there are always persuasive reasons why they might. During their expansionary phase, asset price bubbles are extremely popular, not only with almost everyone in the market, but also with politicians, who see the rise in asset prices as the due market response to their own successful policies; when you believe that you have abolished boom and bust, as then-Prime Minister Gordon

Brown announced in 2008, you would expect a stronger market.

It would take an extraordinarily self-confident and brave central banker to raise interest rates at a time when inflation seemed under control simply because he felt that some asset price was too high. Increasing interest rates often depresses the real economy, while likely having little effect on the asset market in question. And even if those asset prices did subside, the event would provide ammunition for claims that the initial rise in asset prices was not so dangerous after all. In my assessment, “leaning into the wind” never amounts to much in practice – not enough to make a difference, at least. The central bank that came closest to leaning into the wind was the European Central Bank (ECB), which, because it paid greater attention to time trends in the monetary aggregates (its second monetary pillar, the first pillar of its analysis being the usual forecast of current inflationary developments within the eurozone), appeared to have a generalized concern about financial expansion; close observers of the ECB have, however, doubted whether following the movements of this pillar ever made more than a difference of twenty-five or, at most, fifty basis points (a basis point is equal to one one-hundredth of a percentage point) to the ECB’s interest rate judgment.

This is not to state that there is no case for any alterations in the practice of inflation targeting. One consideration raised by the recent crisis is whether the interest rate entered into central banks’ Taylor reaction function should be a *risk-adjusted* short-term interest rate, rather than the official rate. The primary risk in this case is that a borrower may default, that is, not be able to pay back the loan. In order to offset that risk, lenders will increase the interest rate charged to bor-

rowers. In a recession, especially one triggered by a financial collapse, most borrowers seem considerably riskier. Consequently, the margin – or spread – of average borrowing rates above the riskless rate, say, on U.S. Treasury bills and bonds, increases. In extreme cases, the average rate facing private sector borrowers may rise at the same time that official rates fall.

In the standard macroeconomic models employed up until 2007, however, the risk (of counter-party default) was assumed to be low – in effect, zero – and constant. In practice, under the influence of bubble confidence, increased leverage, and enhanced competition for business, risk margins were gradually pared down over the course of 2002 to midsummer 2007. Thus, *effective* rates facing private sector borrowers were declining over the years 2002 to 2006 while some commentators (again John Taylor) had already criticized the Fed for keeping official rates too low. The fact that such risk margins reached their nadir in Summer 2007, just before the financial system entered into crisis, should also temper faith in the wisdom of markets.

After the start of the crisis, and especially after the Lehman bankruptcy in September 2008, risk margins spiked higher; so did the effective cost of capital to private sector borrowers. If central banks had placed more emphasis on risk-adjusted rates, and less on the basic official rate, they might have raised official interest rates earlier and faster before August 2007, and then lowered them more quickly and by a larger amount in the period from August 2007 to October 2008. During its first few months, the crisis was generally (but mistakenly) seen as a pure liquidity emergency, rather than as a solvency problem as well. A liquidity crisis occurs when a firm does not have suffi-

cient ready cash to meet a payment that is due; a solvency crisis arises when the value of a firm's liabilities is greater than that of its assets. Such crises tend to occur together because firms known to be clearly solvent can almost always borrow money to meet upcoming payments. Moreover, there was never any shortage of overnight liquidity. Instead, bankers were concerned with how they might replace their one- or three-month wholesale market borrowing once it matured and likely might not be rolled over in a crisis, largely because of lurking concerns about default.

For the foreseeable future – two or three years ahead at most – banks will be under pressure to reduce their reliance on wholesale funding, to delever, to become safer and smaller. As a result, they will be stricter and more cautious in extending new credit. This prudence will be evidenced by higher margins between loan rates and the official rate. Whatever the level of official rates, the risk-adjusted effective rate that private sector borrowers face will be higher than in the past. Put differently, the “equilibrium” (real) official interest rate likely will need to be lower in the next few years than it should have been in 2001 to 2007 in order to keep average private-sector borrowing rates at a level consistent with a recovery to economic equilibrium.

The standard macroforecasting and analytical models used up until 2007 defined no role for banks, financial intermediaries, money, or default and risk. Precisely how the economics profession allowed itself to be sidetracked into this extraordinary dead end will be a subject of fascination to students of the history of thought for decades. Be that as it may, the implication is that central banks *should* have a continuous concern for monetary developments. At Milton Friedman's ninetieth birthday celebra-

tion, Ben Bernanke, Greenspan's successor as chairman of the Fed, acknowledged that the Fed was partially to blame for the debacle of 1929 to 1933 and promised that (his) Fed would not allow that to happen again: “I would like to say to Milton and Anna [Schwartz]: Regarding the Great Depression. You're right, we did it. We're very sorry. But thanks to you, we won't do it again.” What is odd about this statement is that Friedman and Schwartz, in their epic *Monetary History of the United States*, primarily blamed the Fed for allowing the (broad) money supply to diminish sharply during 1929 to 1933. Meanwhile, as of Spring 2010, the Fed ceased credit expansion at a time when bank lending to the private sector has been declining, and broad money supply appears almost stagnant. (The monetary aggregate M2 declined from an annual rate of 8 to 9 percent in Spring 2009 to just over 1 percent in 2010.) Certainly, current methods of interpreting the monetary data are problematic. But if inflation *and* deflation are monetary phenomena, central banks must put more effort into data interpretation. Focusing on the transmission of official interest rates into real expenditures, and then inflation, without concern for financial intermediation along the way, is no longer a sufficient means of analysis and should be discarded. This is not to suggest that analyzing monetary developments is easy or can be reliant on a single statistic from one or another of the several monetary and credit aggregate statistics, which only rarely move in lockstep.

But whether or not central banks choose to increase their focus on risk-adjusted interest rates, they should continue to set rates with the primary objective of maintaining price stability and providing an anchor for the private sector's inflation expectations. Given that

inflation expectations are now accorded a key role in determining how the economic system works, the achievement of financial stability must be realized by other means, namely, a combination of financial structure reform and the introduction and use of specific macro-prudential controls, such as counter-cyclical required capital ratios.

In the 1960s, when almost everyone, including Richard Nixon, professed to being believers in Keynesian theory, governments generally tried to steer the economy with fiscal policy. They were skeptical of the effectiveness of monetary policy (apart from its influence on international capital flows), and preferred to keep interest rates low and steady to encourage productive investment. Between then and 2007, Keynesian theory went out of fashion. Monetary policy – via setting the official interest rate – was to be steered so as to keep the ship of state on the straight and narrow, while fiscal deficits were to be kept low and stable, once again, to encourage productive investment in the private sector. An extreme version of this latter dogma was incorporated in the Stability and Growth Pact in Europe, which proposed that eurozone countries should never run a deficit greater than 3 percent of GDP, a limit that has been regularly transgressed, and by most of the countries hardest hit in the current recession.

The 2007 crisis led simultaneously to a collapse in private sector investment, especially in housing, and to a realization among consumers that they had overextended their indebtedness, triggering a rise in the desired savings ratio. The private sector surplus of savings over investment escalated sharply. Of course, all policy-makers would like to see the counterbalance to the net savings of the private sector in their own

economy be a current account surplus (that is, with the rest of the world running a deficit). But across the global economy as a whole, current accounts sum to zero. So the counterpart to the greater private sector surplus *has to be* a larger public sector deficit. The only real question is whether this relationship is achieved by deflation, via reducing tax receipts *and* private sector savings, or by Keynesian fiscal stimulus, thereby maintaining income levels. Besides Hyman Minsky, the other great intellectual winner from this crisis has been John Maynard Keynes.

The problem herein is that the private sector's shift into surplus has been so great in scale that its counterpart, the public sector deficit, has ballooned to levels never before seen in peacetime. Moreover, the aging population and its associated medical needs offer a dire outlook for public sector deficits in the longer term; in the immediate future, the crisis has left the financial system so weakened that the economic recovery is expected to be slow and hesitant. In these circumstances, the so-called sovereign risk of public sectors is everywhere under strain, especially in countries, such as Greece, and subsidiary states, such as New York and California, that are not fully "sovereign," insofar as they do not have their own independent currencies and cannot, at an inflationary pinch, print more of it to meet the interest on their debt.

How soon, then, must states begin to rein in the extent of fiscal deficit? Too soon and one risks a "double dip" deflation; too late, and markets, notably the bond and foreign-exchange market vigilantes, will take fright, as they already have in Greece, and thereby deliver a "double dip" deflation by another route.

With respect to monetary policy, the various arms of policy, fiscal and mone-

tary in particular, must move in coordination. The more slowly and timidly fiscal policy is retrenched, the faster and more robustly interest rates will have to be normalized and credit (or quantitative) easing reversed. Per contra, any reduction in fiscal stimulus must be offset by further, even enhanced, easing in monetary policy.

A problem with fiscal policy is that it is far less flexible and reversible than monetary policy. Fiscal easing, either via extra expenditures or lower taxes, tends to create perceived entitlements. It often takes a long time to shift from announcement to economic effect; there are few “shovel ready” projects on hand. Given the lags and difficulties of getting fiscal retrenchment under way, together with the market risks of not doing so, my own preference (which is influenced by conditions in the United Kingdom) falls on the side of making an early start on fiscal retrenchment; in this case, the adverse effects on incomes and expenditures must be offset by further monetary expansion. But is monetary policy capable of offsetting fiscal retrenchment? Now that interest rates have effectively dropped to zero, would any additional attempt at monetary expansion be “pushing on a string”?

The zero interest rate bound (or zero interest rate policy [ZIRP]) was reached fairly quickly after the financial panic in September to October 2008. Fortunately, this event did not mark the limit of expansionary monetary policy. As financial markets and the economy weakened progressively until March 2009 – in a downturn worse than that of 1929 – central banks turned to unconventional policy measures: long-term loans by the ECB, quantitative easing by the BoE, and credit easing by the Fed. Although other (fiscal) stimulatory policies were introduced concurrently, the timing of these additional monetary policies coincided with

the recovery of financial markets, from their low point in March 2009 to the hesitant recovery in the real economy in the middle quarters of 2009 (Q2 and Q3). Studies have indicated that these quantitative- and credit-easing measures succeeded in bringing about some further easing in longer-term interest rates and revivifying those private sector markets, such as the mortgage-backed securities (MBS) market, at which they were targeted.

The additional unconventional measures were, therefore, partially successful. But they were not as successful as some may have initially hoped, though the prior experience of Japan – in which similar measures were applied in 2001 to 2006 – should have given warning not to expect too much. These measures primarily involve open-market operations, whereby a country’s central bank buys assets of one kind or another and pays for them by writing a check on itself. These checks are paid into commercial banks, which then hold the receipts as deposits with the central bank, thereby increasing the commercial banks’ own cash reserves. Under normal circumstances, when a commercial bank receives extra cash reserves, it uses them to make an extra loan, or to buy some other additional earning asset. But in the crisis conditions, the banks sat on the additional reserve base. Despite a vast increase in the banks’ cash base, the increase over the same period in their holdings of either private or public sector assets was minute; the multiplier had failed to work.

There were, of course, reasons why the banks were happy to accumulate cash reserves rather than putting them to work. During this recession, many corporations preferred to strengthen their balance sheets by repaying bank loans. In general, a recession makes

projects offered to banks by prospective borrowers seem, almost by definition, riskier. In view of the size and expected persistence of the fiscal deficit, the interest rate (and inflation) risk of buying longer-term government debt remains palpable, while the yield available on shorter-term debt has been exceptionally low. The interest rates are expected to rise in the future, the more so if inflation should also experience an uptick. When interest rates rise, bond values fall commensurately. Banks that hold such bonds suffer a capital loss.

This relative failure was in part due to policy errors made by the central bank. In particular, because the aim was to induce the banks to make use of the cash base that the central bank created, it was a mistake to offer an attractive deposit rate on such holdings. When the Fed activated the permission to pay a positive interest rate to commercial banks on their deposits held at the Fed itself in October 2008, the move could not have come at a worse time. Just as the Fed increased required reserve ratios in 1936, fearing the possible inflationary consequences of banks' "excess" reserves (now perceived as a mistake), the Fed in 2008 likewise began to pay a positive interest rate on bank deposits, fearing the same potential inflationary consequences. Indeed, the exact opposite policy – seeking to impose a slight penalty on deposits, above requirements – would have been desirable (though also somewhat unfair, since the banking system in aggregate cannot avoid holding the cash base that the Fed creates). The ability to set and vary the upper limit of rates at which banks can obtain cash from the central bank, and the lower limit at which the banks can place deposits with it – a "corridor system" – is a new, useful, and flexible tool. But the experience of the last two years indicates

that central banks have, in general, not yet learned how to make full and best use of this capacity.

Central bankers are sensitive to the criticism that, in the past, they reacted to each financial bust by encouraging, via aggressive and maintained cuts in interest rates, a new asset price bubble. Eventually, so the argument goes, the super-credit cycle became so top-heavy that interest rate cuts could no longer cope; public sector (taxpayer) support was needed. But the capacity and the willingness of the public sector to provide further support to the financial sector are now exhausted. Against this background, the unconventional measures of monetary expansion are being phased out by the ECB, BoE, and Fed, (though not in Japan, where the government is pressing the Bank of Japan to restart such measures).

What we have currently is a continuation of the fiscal stimulus (until the recovery is fully established) and a cessation of the unconventional monetary measures (though the ZIRP is being maintained at least for the time being). My own preference, as earlier indicated, would be for the reverse, with some early fiscal retrenchment offset by continuing – indeed, reinforced – monetary expansion.

Monetary expansion will, however, be hard to achieve at a time when everyone wants to reform the banks into safer and smaller entities. Financial regulators (and even the general public) are like generals poised to fight the last war over again, especially since that battle ended in a major defeat. As a result, banks will be faced with:

- 1) New taxation on a base (whether on profits, assets, wholesale borrowing or transactions) and at a rate still to be determined;

- 2) Higher required capital ratios;
- 3) More restrictive leverage ratios, especially in Europe, which previously had no such restrictions; and
- 4) Enhanced liquidity requirements.

In this context, and after the traumatic shocks of the crisis, which will also make it more difficult to find creditworthy borrowers, the banking system is likely to shrink. The forthcoming bank tax – whatever its details – will be quite popular and will seem, superficially, almost painless, but will actually be paid mostly by depositors, in the guise of lower deposit rates, and through higher charges to borrowers. This scenario, together with the necessity of enacting fiscal retrenchment and the remaining debt overextension in the personal sector, points to a sluggish recovery at best. Monetary policy will need to remain expansionary for a considerable amount of time into the future.

The basic monetary policy regime of inflation targetry, achieved by adjusting the official interest rate, should remain. However, it should be subject to the following technical changes:

- a) In Europe, housing prices should be included in the relevant index.
- b) Central banks should target a risk-adjusted interest rate rather than the official rate.
- c) More attention should be paid to a range of monetary aggregates, despite the undoubted difficulty of interpreting their signals.

In recent years, the macroeconomy has been in large part steered by monetary policy adjustments, with fiscal policy primarily aimed at longer-term structural objectives. During the recent crisis, Keynesian demand management via

fiscal policy has, however, returned to center stage. How, then, must the coordination of the two arms of policy be achieved? I believe a quick start to fiscal retrenchment offset by continuing, possibly enhanced, monetary easing is the best available approach. This tack, however, is unlikely to be adopted for several reasons:

- 1) The unconventional expansionary monetary policy measures have not been as effective as might have been hoped.
- 2) It is so difficult to claw back fiscal expansion, especially when the recovery remains so hesitant, that it will not happen soon, except under extreme external pressure, as in Greece.
- 3) The imposition of extra regulatory constraints on banks, while fully understandable in view of past excesses, is hardly calculated to encourage current monetary expansion.

The implication of delayed fiscal retrenchment is that economic recovery will be slow-moving at best. The financial crisis is not behind us; there could easily be further alarms. It is likely that monetary policy will have to remain in its expansionary mode for longer than most commentators now appreciate.

Fiscal stimulus

During the calm period in recent U.S. macroeconomic history, from the mid-1980s to mid-2000s, stabilization policy relied mainly on monetary policy to deal with recessions, which were mild and many years apart. Besides relatively small tax rebates, there seemed to be little need for fiscal stimulus. The government did not launch major public works or infrastructure projects to try to restore full employment. Even monetary policy fell short of its full stimulus: the Federal Reserve did not push its target interest rate all the way to zero in either of the recession years 1991 or 2001. It wasn't until the extraordinary conditions facing the incoming Obama administration in January 2009 that monetary policy was at its maximal stimulus in terms of the traditional tool, the Fed funds interest rate. The bottom had fallen out of the economy in the last months of 2008. The new administration made fiscal stimulus a key part of its program for saving the economy.

Fiscal stimulus has two arms. One is the government's direct purchase of goods and services. Though the government buys a huge variety of products and employs millions of workers, the focus

of stimulus is usually public construction, called "public works" in the past and now known as "infrastructure." I include all levels of government in infrastructure stimulus because it is common for the federal government to pay for projects that state and local governments build. I will refer to the first arm as infrastructure stimulus, though I will show that the federal government delivered almost no increase in infrastructure spending and that state and local governments cut spending during the recession. A greater effort on the part of the federal government to prevent the decline in state and local purchases would have served the purpose of the stimulus effort.

The second arm of fiscal stimulus pays increased benefits to the public. Expanded unemployment insurance is a leading form, but many other types of public benefits grow during recessions as well. The second arm also includes tax rebates and other tax cuts that put more cash in the hands of the public. I do not include tax cuts in this article, however, because the modest tax rebate in 2008, though a response to the mild contraction that started at the beginning of 2008, was not an important part of the government's response in 2009 to the Great Recession.

One simple measure of the effectiveness of fiscal stimulus – the multiplier –

receives the most attention from economists and often enters public debate as well. The multiplier records the number of dollars of increase in total national output and income per dollar of stimulus spending. Much of this article reviews current thinking among economists about the size of the multiplier. A weak consensus holds that in normal times, including earlier recessions, the infrastructure multiplier is about one: each dollar of infrastructure stimulus boosts output by a dollar. Put differently, when the government buys more highways and schools, output rises by enough to permit other categories of spending – such as consumption and private investment – to remain unchanged. In times of extreme recession, such as 2009, there is widespread agreement that the infrastructure multiplier is higher – perhaps twice its normal value.

A multiplier greater than one occurs in an economy with strong feedback effects. The feedback effect stressed in elementary macroeconomics is the increase in consumption that results from higher income when production is higher. More complete macro models describe the interplay of a variety of feedback effects, some positive and some negative. The positive effects are more likely to prevail in a severely depressed economy.

Any consensus about the multiplier relating benefits spending to total output is even weaker than that of infrastructure spending. An increase in benefits has a first-round effect on spending from the fraction of the increase that recipients spend rather than save. Evidence on this fraction is truly mixed. The benefits multiplier is reduced in comparison to the infrastructure multiplier by this fraction. For example, if 30 percent of a benefits increase is consumed and the remaining 70 percent is saved, the benefits multiplier is roughly 0.3 in normal times when

the infrastructure multiplier is one and 0.6 in deep recession when the infrastructure multiplier is two.

Before plunging into a discussion of the sizes of the various multipliers, I consider the amount of infrastructure and benefits stimulus the United States actually applied during the Great Recession. Infrastructure spending is part of total government purchases of goods and services. Therefore, the multiplier for infrastructure need not be separated from the multiplier for other goods and services that the government purchases. I begin by looking at total purchases, which make up one of the major categories of government spending; the other categories are benefits and interest on debt.

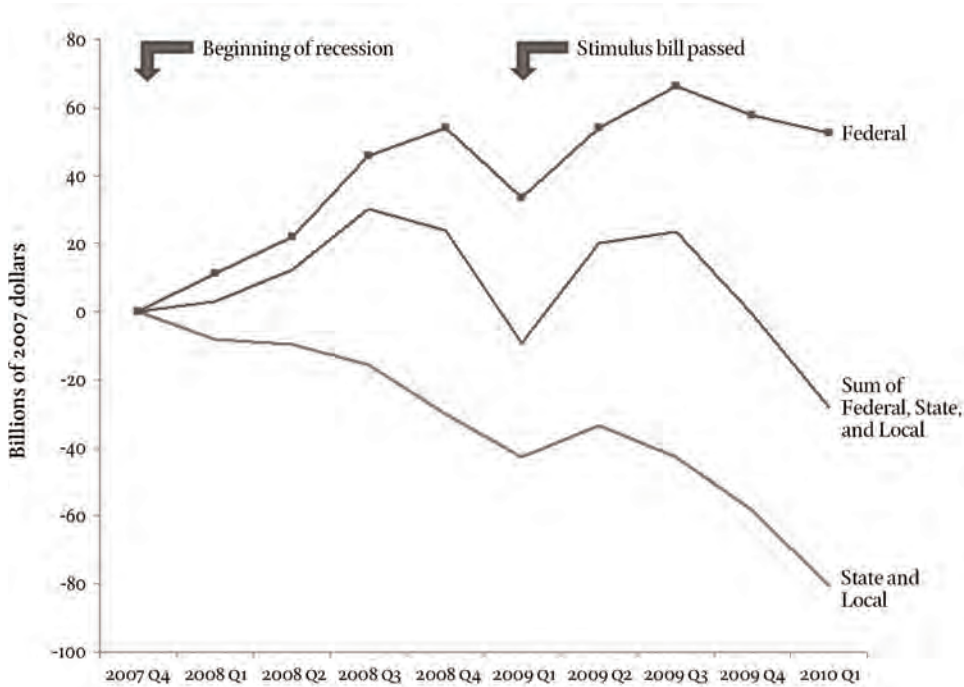
My approach here is to examine government purchases as they actually played out, rather than trying to isolate the purchases that resulted from the stimulus program. For example, I explore the details of a surge in federal military spending in 2008 that provided a well-timed stimulus but certainly was not part of any stimulus program.

Figure 1 shows the amount of extra federal and state-local government purchases in constant 2007 dollars from the onset of the Great Recession at the end of 2007 to the first quarter of 2010.¹ I define “extra” as the amount in excess of an extrapolation of purchases from the end of 2007 at the growth rate recorded from 1999 to 2007. Two facts are immediately apparent: only the federal government contributed to the stimulus from government purchases, and even that stimulus was not very big.

Federal purchases escalated as soon as the recession began at the end of 2007. The initial expansion was almost as rapid as it was immediately after the passage of the stimulus bill a year later at the outset

Figure 1

Extra Purchases by the Federal Government and State-Local Governments, 2007 Q4 to 2010 Q1



Source: U.S. National Income and Product Accounts, as adjusted by author.

of the Obama administration. Both federal and state-local purchases plunged at the beginning of 2009 and then expanded to their highest levels soon after the stimulus bill was enacted in February 2009. At that time, economists expressed concern that the bill’s extra infrastructure spending would ramp up slowly, with a much larger increase in 2010 than in 2009. Instead, the peak in purchases occurred in mid-2009. The perspective shown here is quite different from that taken by the bill’s proponents; the figure illustrates the excess of all categories of government purchases over its trend path, not just the specific projects that the bill funded.

Although the bill was intended to boost state and local government-infrastructure spending, the negative force of large declines in state and local revenue swamped

that effect. Throughout the recession and early recovery, state and local purchases fell below their normal trend path, dragging down the economy compared to their normal effect. State and local governments lack the federal government’s authority and ability to borrow aggressively to raise spending at the same time that revenue collapses. By the fourth quarter of 2009, the decline in state-local purchases relative to trend was enough to offset the federal stimulus more than fully. Taken together, all levels of government were retarding the recovery.

All the changes in government purchases shown in the figure are small in comparison to the overall size of the U.S. economy. The peak of an extra \$66 billion in purchases was less than one-half of one percent of total output of goods and services. The stimulus from government

Table 1
Extra Government Purchases, by Category, in 2009 Q3

Category	Extra Spending, Billions of 2007 Dollars
Federal Nondefense, Noncapital	13.7
Federal Nondefense, Capital	3.6
Federal Defense, Noncapital	34.2
Federal Defense, Capital	14.4
State and Local, Capital	-34.6
State and Local, Noncapital	-8.5

Source: U.S. National Income and Product Accounts, as adjusted by author.

purchases could not have had a visible effect on the overall economy.

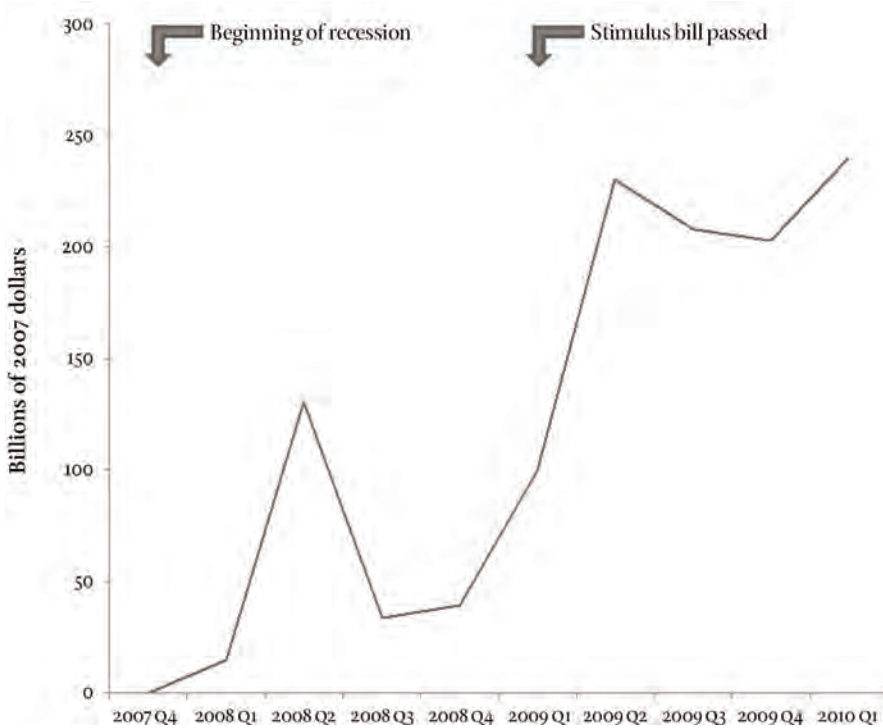
Table 1 breaks down extra government purchases by government level and type of spending. Noncapital spending consists primarily of wages and salaries of public employees but also includes materials and supplies that are used quickly and do not become part of government capital stock. Capital spending denotes spending on buildings, equipment, and software. The table shows that by far the largest spending category was noncapital defense purchases – a rather different outcome than the authors of the stimulus bill had intended. The single biggest negative factor was the decline, relative to trend, in state and local noncapital purchases.

The other type of fiscal stimulus provides income to families in the form of government benefits. Part of the stimulus comes from the automatic operation of the benefit systems, without special features, and part from special increases legislated to help stressed families and

support continued consumption spending in the face of recession-induced declines in earnings. A leading example of an automatic benefit increase is the rise in unemployment compensation that resulted from the doubling of the unemployment rate during the recession. And a leading example of a special increase is the extension of unemployment benefits to long-term unemployed who would have lost their benefits sooner under the standard formula. The data I review include both automatic and legislative increases in benefits.

During the Great Recession, benefit increases added substantially to disposable income. Figure 2 is derived from the same approach as was applied in Figure 1 to determine the extra spending on benefits beyond the normal upward trend. The measure combines the automatic increases, which pushed benefits above trend, and the legislated increases, which added to the automatic ones. The total extra contribution to disposable income exceeded the rate of \$200 billion per year starting in the second quarter of 2009.

Figure 2
Extra Government Benefits to Individuals, 2007 Q4 to 2010 Q1



Source: U.S. National Income and Product Accounts, as adjusted by author.

This spending made an important contribution to plugging the hole in total disposable income excluding benefits, which exceeded \$500 billion per year in the first quarter of 2009 and rose to more than \$700 billion in the first quarter of 2010.

How does fiscal stimulus raise output and employment? The Great Recession revived long-standing debates among economists on this point. I will start with the analysis of an infrastructure stimulus.² One polar view maintains that an increase in public use of resources must inevitably displace an equal amount of private activity, so that an increase in infrastructure spending has zero effect on total output and employment. On the

other end of the spectrum is the view that various amplification mechanisms result in an increase in output several times greater than the government's direct purchase of output. It is convenient to express these opinions in the form of the multiplier: the number of dollars worth of total output that a one-dollar increase in infrastructure spending causes. Thus, assessments of the multiplier range from zero to perhaps four. There is, however, a concentration of estimates in the range from one to two, together with some agreement that whatever the value of the multiplier, it is higher when the economy is extremely slack and monetary policy has spent its expansionary power by driving the short-term interest rate to zero.

Professional discussions of the multiplier include its dynamics. First, most views hold that the expansionary effect of an increase in infrastructure spending depends on how it is distributed over time. Second, the effects on output and employment continue past the end of the stimulus program. The later effects are likely to be negative: if the stimulus cuts into private capital formation, as many believe it will, the economy will have lower productive capacity once the stimulus is removed than it would have had absent the stimulus. I avoid these issues by focusing on the immediate – that is, first-year – effects of a stimulus like the one that occurred in 2009 as a temporary burst of extra government purchases.

My discussion will enter the territory of what many economists call Keynesian properties of the economy. Among professional economists, the term Keynesian has lost most of its political overtones. We use the term to mean properties of the aggregate economy that depart from the economic principles we emphasize in our accounts of individuals and markets. The opposite of Keynesian is not conservative, but rather neoclassical, meaning that aggregate economic behavior is analyzed according to the accepted principles for individuals or individual markets. As neoclassical principles have it, prices and wages are flexible and markets always clear. Unemployment is absent. In the face of events during the Great Depression that defied neoclassical principles, John Maynard Keynes was a pioneer in exploring alternatives, some of which have made their way, with important alterations and improvements, into modern macro analysis. In particular, contemporary theory includes Keynes's observation that prices and wages are not flexible in the short run, so markets are not clear and unemployment can occur.

The financial press equates the term Keynesian to advocacy of fiscal stimulus to offset recessions, while among economists, Keynesian refers to the adoption of some principles of behavior of the aggregate economy that are not neoclassical. There are plenty of Republican Keynesian economists these days. Even libertarians – numerous among economists if not in the population in general – can be Keynesian. The popular and professional meanings interact. Because Keynesian principles lead to higher estimates of multipliers, fiscal stimulus is more effective in a Keynesian than in a neoclassical world.

It is useful to start with the neoclassical analysis of the effect of infrastructure stimulus in order to illustrate its limitations and the need for Keynesian features. The neoclassical model gives prices and wages free rein to cushion against shocks and prevent unemployment. An increase in infrastructure spending raises the interest rate, which causes workers to choose longer hours of work. The future tax increase that will finance the spending also raises hours of work because workers choose to offset some of the loss of purchasing power due to higher taxes by raising their earnings. Investment and consumption fall because of the tax and interest-rate effects. Total output rises, but by only a fraction, about 30 percent, of the increase in infrastructure spending. The multiplier is just 0.3. Displacement of private spending by public spending is a substantial issue, though it is not complete, as some economists have claimed. The critique that appears in *The Wall Street Journal* whenever fiscal stimulus is under consideration is on point in a neoclassical economy: stimulus spending drives up the interest rate and displaces private investment.

What is wrong with the neoclassical model? First, it neglects unemployment.

It has taken a long time, but modern economics finally has a theoretically respectable and empirically reasonable view of unemployment. Adding unemployment to the neoclassical model has the expected effect of raising the multiplier. A primary reason for the low neoclassical multiplier is that, in the neoclassical economy, stimulus results in more output because people feel that they are worse off under the burden of taxation and because a higher interest rate rewards immediate work by increasing the amount of deferred consumption from an hour of current work. Neither of these effects is strong, according to a large body of research on household behavior. Factoring in unemployment, governed by a model that emphasizes employers' incentives for job creation, adds a powerful third force. A stimulus draws workers out of unemployment and puts them to work. The multiplier is substantially larger with this important modification, at around 0.6.

The second reason that the neoclassical model delivers a low multiplier is that it lacks an amplification mechanism based on sticky or rigid prices. Sticky prices are the hallmark of modern thinking that calls itself Keynesian. (In fact, to distinguish itself from older ideas, this vibrant modern school calls itself New Keynesian.) Amplification works in the following way: At all times, the economy is held back because of pervasive market power. The suppliers of productive inputs – labor and capital – receive inefficiently low rewards because businesses extract profits derived from market power prior to passing on their revenue to labor and capital suppliers. At all times, the economy produces less than it could. But when a stimulus expands the economy and drives up the rewards to labor and capital while prices do not respond because they are sticky, the price/cost margin

contracts. In effect, the economy becomes more competitive. Rewards to labor and capital improve and the economy expands more than it would in the neoclassical model, where prices are flexible. Without invoking an unreasonable extent of price stickiness, a model that includes both unemployment and sticky prices can deliver a multiplier around one.

Are prices sufficiently sticky to generate substantial amplification? Research on this topic has been intense in the past decade. The Great Recession has generated a raft of new evidence on how prices respond to slack, as the economy had unprecedented slack in 2009. In a flexible-price economy, prices should fall in times of slack, as merchants cut prices to take business away from their rivals and to take advantage of the lower cost of inputs. But prices hardly fell at all as the economy collapsed. Table 2 presents data for categories of output that declined from the end of 2007 to the end of 2009. The most striking example of a sticky price is in the category of business equipment: although output declined at 9.1 percent per year, the price actually rose slightly.

Although prices seem to be quite sticky, research has not so far been able to document that market power increased during this or earlier contractions. Evidence that the economy becomes less competitive in a recession is only circumstantial.

We teach college freshmen quite a different version of the Keynesian explanation of the multiplier. There is no mention of unemployment, variable market power, or interest rates. Instead, when the government spends more on infrastructure, income rises, consumers spend more, income rises further, and so on. The multiplier expresses the cumulation of this process; it depends on the feedback operating through the propensity

Table 2
Annual Rates of Output and Price Changes, 2007 Q4 to 2009 Q4

Category	Annual Rate of Output Change (%)	Annual Rate of Price Change (%)
Consumer Durables	-4.1	-1.3
Business Structures	-12.2	-1.2
Business Equipment	-9.1	0.4
Homebuilding	-16.9	-3.4
Goods Exports	-1.9	-1.0
Service Exports	-2.5	0.6

Source: U.S. National Income and Product Accounts, as adjusted by author.

of consumers to spend more when their incomes rise. A high propensity to consume is not required for a modern model, based on the two key elements of unemployment and amplification, to generate a reasonable multiplier. Some investigators have included fairly high values of the propensity to consume in their models, but the mechanism described in freshman economics has little to do with the numerical value of the multiplier in most models in use today.

Recent interest in the infrastructure multiplier has caused researchers to re-evaluate the direct evidence on how total output is affected when the government buys more output. It has been more than fifty years since the government made a sudden change in its purchases; since the end of the Korean War, government purchases have grown smoothly along with the overall economy. The Vietnam War and the Reagan military buildup were gradual and small. Direct evidence, then, derives from spending changes during World War II and the Korean War. These rapid buildups and build-downs show fairly clearly that total out-

put rises by between 50 and 70 percent of the amount of increase in government purchases. The evidence, however, is by no means dispositive; many other factors influenced output during those periods. In particular, during World War II the government held back private spending with rationing and direct controls, and during the Korean War heavy taxes may have inhibited expansion. A reasonable conclusion is that the two wartime experiences do not refute a multiplier of one. More recent experience in the United States or other advanced countries certainly does not refute that value.

The year 2009 was unique in post-Depression U.S. experience because the short-term safe interest rate that the Fed uses as its policy instrument – the Fed funds rate – was at its most expansionary value of zero throughout the year. The multiplier was larger than under normal circumstances, when the Fed has the capacity to adjust its policy. The Fed, along with other properly managed central banks, steers the economy by adjusting the interest rate. By appropri-

ate adjustments, the Fed accomplishes its mission of delivering low and stable inflation in the longer run and leaning against recessions in the shorter run. The Fed sets a higher interest rate if inflation is above the target of 2 or 3 percent annual inflation. The Fed sets a lower interest rate if unemployment is above 6 percent. Thus, if the economy enters a boom with low unemployment and high inflation, the Fed sets a high interest rate to cool off the economy. Conversely, in a recession with high unemployment and low inflation, it sets a low rate to stimulate. In stagflation, with both inflation and unemployment above target, the Fed balances one goal against the other. The Fed has to make an intelligent compromise when high inflation calls for restraint and high unemployment calls for expansion.

In normal times, if fiscal policy tries to expand the economy with an infrastructure stimulus, unemployment will fall and inflation will rise. The Fed, following its stabilization principles, will raise its interest rate and contract the economy. The models supporting the conclusion that the normal value of the infrastructure multiplier is one describe the behavior of an economy where the central bank responds to all forces, including fiscal policy, that alter inflation and unemployment by following these stabilization principles.

Sometimes, as in 2009 and 2010, the stabilization principles call for a negative interest rate. If unemployment is high and inflation is low, a great deal of expansion is desirable. But the Fed is incapable of making its interest rate negative, for the reason that the interest rate the Fed controls is the rate at which banks borrow and lend reserves. Banks have the right to convert reserves to hundred-dollar bills in unlimited amounts. The rate that the bills pay is zero, so if the

reserves had a negative rate – if banks had to pay to hold reserves – they would simply convert reserves to bills.

The Taylor rule (so named for economist John Taylor) calls for the Fed to raise the interest rate when a fiscal stimulus goes into effect. The higher rate would offset part of the expansionary effect of the stimulus. But if the interest rate is pinned at zero when the Taylor rule asks for a negative rate, the Fed will keep it at zero unless the fiscal stimulus is so effective as to raise the interest rate – dictated by the rule – to a positive level. In 2009, the Taylor rule yielded a deeply negative rate, but in the absence of practical fiscal policy that would raise the rate above zero, the best the Fed could do in the face of fiscal expansion was to keep the rate at zero. Consequently, the Taylor rule ceased to inhibit fiscal expansion. Thus, in an economy where the central bank is governed by a Taylor rule (or a central bank that raised interest rates for any reason during economic expansion), the bank's reaction to the expansionary effect of fiscal stimulus blunts the effect. The infrastructure multiplier is lower in an economy with a responsive central bank than in an economy with a central bank that keeps the interest rate constant.

A number of investigators have studied the elevation of the infrastructure multiplier when the interest rate is pinned at zero. All have concluded that it is substantially larger than the value of around one that is the weak consensus for the multiplier in normal times. The range of estimates for the zero-interest-rate multiplier is wide, even within the same study, because it is sensitive to the timing of the stimulus and the duration of the period when the interest rate will remain pinned at zero. That said, a value for the zero-interest-rate multiplier of around two is representative of recent research.

The benefit multiplier is quite another matter. If the government can induce the public to consume more by providing higher benefits, the effect of higher consumption purchases is, as far as can be determined, equal to the effect of higher government purchases. As I noted above, the benefit multiplier is the product of the marginal propensity to consume and the infrastructure multiplier. The new central issue in measuring the benefit multiplier, then, is determining the public's propensity to consume out of increased benefits.

A large number of observers, including many economists, believe that most of an increase in benefits goes straight into higher household spending. Given that a substantial part of the increase in benefits in a recession takes the form of higher unemployment insurance payments, and a presumption that the unemployed have their backs to the wall financially, the hypothesis of a high propensity to consume benefits commands the high ground of plausibility. One factor pointing in the opposite direction is that people with their backs to the wall borrow to the hilt. Credit cards make this kind of borrowing very easy. A worker who has avoided falling into the credit-card trap while working may well start borrowing after losing a job. Much research shows that one of the first things that people do with benefits is pay off debts. Reducing debt is a use of benefits that does not result in a boost to consumption spending. Rather, reducing debt is a form of saving; the effect on the economy of a family paying off some debt is equivalent to that of another family raising its savings by the same amount.

It has proven surprisingly difficult to resolve the issue of new spending versus paying off debt with hard research. The United States lacks a body of data that tracks consumption spending accurately

– or even roughly – at the household level. A baseline model of consumption-smoothing implies that families will spread the spending resulting from a one-time receipt of income over many future years. In other words, the household will consume only 5 to 10 percent in the first year and save the rest to boost future consumption. Some ingenious studies have focused on showing that people tend to consume more than 10 percent of small cash injections from tax rebates and other distinctively one-time sources. These studies do not generally find 100 percent-immediate consumption. Another generalization from a large number of studies is that the bigger the amount of a sudden cash receipt in relation to normal income, the smaller the fraction that goes out as immediate consumption. All the research finds that the context of an income receipt is an important determinant of the amount that is consumed in the first year. Providing assistance to families with their backs to the wall financially raises their spending by more than would giving the same amount of assistance to those with ample liquid savings. Thus, raising unemployment benefits generates more added consumption spending than does a general tax reduction.

The application of all this research to the question of the marginal propensity to consume given an increase in benefits during a recession is controversial and uncertain. The bulge in benefits is not confined to a single year. As of the first quarter of 2010, it was still growing, despite the recovery that appears to have begun more than six months earlier, in mid-2009. The longer duration points in the direction of a higher propensity to consume. A defensible guess for the marginal propensity to consume from a benefits stimulus is around 0.4, but opinions range from almost zero to one.

Table 3
Effects on GDP

	Purchases	Benefits	Total
Average Federal Stimulus, 2009 Q2 to 2010 Q1, Billions of 2007 Dollars	58	220	
Multiplier	2.0	0.8	
Effect, Billions of 2007 Dollars	115	176	291
GDP, Billions of 2007 Dollars			14,338
Percent of GDP	0.8	1.2	2.0
Average GDP Shortfall, Percent of GDP			8.2
Counterfactual GDP Shortfall, Percent of GDP			10.2

Source: U.S. National Income and Product Accounts, as adjusted by author.

Combining the two elements, I conclude that the benefits multiplier is around 0.4 in normal recessions, in which the interest rate is never pinned at zero, and around 0.8 in deep recessions like the one that began at the end of 2007, during the period starting at the end of 2008 and continuing to this writing, when the interest rate was pinned hard at zero.

What was the overall contribution of the increases in federal purchases and benefits? In spite of my previous warnings about the uncertainty in the multipliers, I make estimates in Table 3. These calculations refer only to the federal purchases component, as I have no way to measure the federal contribution to state and local finances that prevented even greater shrinkage in their purchases. I reiterate that some of the increase in

federal purchases resulted from a military buildup unrelated to the recession. And the benefits estimates apply not only to the increased benefits resulting from the stimulus legislation, but also to the increases that occurred because the benefit programs helped more people as a result of the recession.

The combination of increased federal purchases and benefits raised output and income by 2 percent. Had the two elements not been in place, the shortfall of GDP from its trend path would have been 10.2 percent rather than the actual 8.2 percent. Stimulus worked in the sense that the recession would have been substantially worse without the stimulus considered here. But the stimulus moved the economy only a bit of the way toward its normal growth path. It left an economy badly injured by the recession.

ENDNOTES

- ¹ A computer file containing the data and calculations is available at <http://www.stanford.edu/~rehall>.
- ² For a discussion of this topic aimed at professional economists, with many references, see my paper, “By How Much Does GDP Rise If the Government Buys More Output?” *Brookings Papers on Economic Activity* 2 (2009): 183 – 231. For a detailed discussion of the effect of fiscal policy in a slack economy, see Lawrence Christiano, Martin Eichenbaum, and Sergio Rebelo, “When Is the Government Spending Multiplier Large?” NBER working paper 15394 (National Bureau of Economic Research, October 2009).

Edward L. Glaeser

Housing policy in the wake of the crash

Between 2000 and 2010, the U.S. housing market experienced a convulsion more extreme than in any previous recorded cycle. From May 2001 to May 2006, the Case/Shiller Standard & Poor's twenty-city housing price index, which controls for changes in housing quality by comparing prices from repeat sales of the same homes, rose 54 percent more than consumer prices rose. In the three years that followed, housing prices, measured by the same index and corrected for inflation, fell more than one-third. Across metropolitan areas, the correlation between boom and bust was almost perfect. For every 10 percent that an area's prices increased between 2001 and 2006, that area's prices fell by 7.6 percent between 2006 and 2009. As prices plummeted, financial institutions that had exposed themselves to sizable housing-market risk became insolvent. The entire U.S. banking system seemed at risk.

The extreme housing-price swings were mirrored by equally oversized fluctuations in the construction industry. America typically produces about 1.5 million housing units annually, which is about 200,000 more than the rate of

household formation. In both 2005 and 2006, builders completed more than 1.9 million new units, at least 580,000 more than the rate of household formation in each year. By a conservative estimate, America erected at least 2 million extra units, relative to historic standards, during the boom. By 2009, there were upwards of 5 million more vacant homes in the United States than there were in 2000. That glut, in turn, explains the decline in new construction and jobs: fewer than 800,000 units were built in 2009; the unemployment rate in the construction sector exceeded 20 percent in early 2010.

While the extremity of the cycle is clear, the causes of the great housing convulsion are more mysterious. Some economists have emphasized the role of easy credit and low interest rates. Certainly, as real and nominal rates fell, buyers benefited from cheaper mortgages, and credit-constrained purchasers were able to cover the costs of more expensive houses. But there is little evidence to support the view that changes in interest rates were large enough to explain the housing market's recent crash. On average, as interest rates decline by one percentage point, housing prices increase by roughly 6.8 percent. Even at very low interest rate levels in markets

where housing supply is limited, a one-hundred-basis-point decrease in real rates does not push prices up more than 8 percent. Between 2000 and 2006, real interest rates fell by 1.3 percent, which suggests that falling rates can explain, at most, a 10.4 percent housing price rise.

Another view implicates aggressive mortgage approval or low loan-to-value ratios. Similarly, compelling evidence to support this explanation is hard to find. My work with economist Joshua Gottlieb and finance expert Joseph Gyourko¹ has estimated that changes in approval rates explain about one-fifth to one-third of the price boom. Loan-to-value levels appear to have an even smaller impact, but the difficulty of controlling for changes in the pool of mortgage applications limits the value of this finding. Moreover, the results of our research are not meant to clear the Federal Reserve System, or anyone else, of error, but rather to emphasize that we cannot identify with any certainty what caused the boom or why it ended.

We do know, however, that individual buyers had wildly unrealistic expectations about future price growth and little understanding of the basic economics of housing markets. Typically, housing prices adjust back to their historic norms, so that if prices rise, relative to historic trends, by an additional \$10,000 over one five-year period, those prices should be expected to fall (again, relative to trend) by \$3,200 over the next five-year period.² Yet during booms, economists Karl Case and Robert Shiller have found, buyers typically expect growth to continue.³ Such irrational exuberance was clearly present in Las Vegas in 2006, when home buyers paid more than twice as much as buyers had only six years earlier.

The price booms in Las Vegas and Phoenix are particularly hard to recon-

cile with informed housing acquisition. These areas have few limits on new construction, and builders supply new housing for less than \$100 per square foot. The power of unfettered building has long kept prices low despite decades of enormous growth. But while these markets did not see price growth during the boom of the late 1980s, buyers in Phoenix and Las Vegas must have assumed that their regions were headed for a permanently higher price plateau. Subsequent events have proven that, to the contrary, unrestricted construction will, in the medium run, keep prices low.

While we do not yet understand the reason for the whirlwind that coursed through the housing markets, we can draw a major policy lesson from the cycle. For decades, the U.S. government has pushed homeownership with policies that subsidize mortgage rates and encourage borrowing. The bust reminds us that housing prices move down as well as up. Promoting homeownership by subsidizing borrowing now seems as likely to create a “default nation” as it does an “ownership nation.” Federal policies that encouraged home buyers to leverage themselves to the hilt in order to bet on housing were clearly complicit in the debacle. As the crash has demonstrated, it is time to rethink these policies.

The home mortgage interest deduction, the sacred cow of federal housing guidelines, was not originally meant to serve as a housing policy. It was introduced with the income tax law in 1916, when essentially all interest payments were deductible. Home mortgages were a minor form of borrowing at that time; the deductibility of interest was largely used to make business expenses tax deductible. Although the Tax Reform Act of 1986 eliminated the deductibility of

non-mortgage interest payments, by then the majority of Americans were homeowners, and the home mortgage interest deduction had become untouchable. Today, the deduction is restricted to mortgages of less than \$1 million, or to the first \$1 million of mortgages above that amount, though in 2005, President Bush's tax reform panel unsuccessfully proposed a substantial reduction of that upper limit.

In the wake of the Great Depression-era housing crisis, the federal government pursued a number of interventions to reinvigorate the moribund credit market. The National Housing Act created the Federal Housing Administration (FHA) in 1934; it encouraged home buying by insuring mortgages against default. The FHA has typically charged borrowers a significant "up-front" mortgage insurance charge (equal to 2.25 percent of the mortgage during mid-2010)⁴ and has financed itself from that fee.⁵ In fact, during the early postwar period, the organization received criticism from the left for being too stingy and refusing insurance to borrowers and neighborhoods that the FHA deemed higher risk. During the same period, the Veterans' Administration (VA) also became a significant insurer of new mortgages at below-market rates; recent evidence uncovered by economist Dan Fetter suggests that the VA had a significant impact on homeownership. The FHA was relatively inactive during the recent boom, as private insurers offered better terms; in 2009, however, one-third of all new mortgages were FHA-insured.

In 1938, the FHA was joined by the Federal National Mortgage Association (Fannie Mae), which was also intended to encourage bank lending. While the FHA insured mortgages, Fannie Mae actually bought mortgages from banks, freeing capital and allowing banks to

make more loans. When Fannie was privatized in the 1960s, the Federal Home Loan Mortgage Corporation (Freddie Mac) was created to generate competition. During the 1970s, Freddie and Fannie moved – with private-sector partners like Salomon Brothers – into the securitization business. They moved from holding mortgages on their own balance sheets to becoming pass-through agencies that bundled and insured mortgages for resale on the secondary market. Fannie Mae and Freddie Mac used a variety of criteria, such as debt-to-income ratios, credit rating, and mortgage size, to determine whether a particular loan conformed to quality standards and could be securitized. Like the FHA, Fannie Mae and Freddie Mac primarily impacted housing markets by allowing better credit terms. A third player, the Government National Mortgage Association (Ginnie Mae), also insures mortgage-backed securities and does so with the full faith and credit of the U.S. government. Ginnie Mae's role is considerably more modest, however, because it insures securities representing mortgages that already have a guarantee from the FHA, VA, or another federal source.

During most of their post-privatization history, these enterprises steadfastly claimed that their profits reflected the benefits that came from the large, liquid market in mortgage-backed securities created by the vast scale of their operations, not implicit government guarantees. Those claims were never terribly plausible, and they certainly are not so today, as the agencies are now kept alive only through tax dollars. Although these entities continue to be subsidized primarily because their failure would further damage the economy, historically, Fannie Mae was justified as a means of solving market imperfections in the secondary mortgage market. Because banks

have an incentive to keep high-quality mortgages on their books and sell only the worst mortgages on any secondary market, the fact that Fannie Mae and Freddie Mac guaranteed mortgages was supposed to prevent a market breakdown. Today, however, a large market in private mortgage insurance permits securities backed by jumbo mortgages, credit card debt, and commercial credit to trade comfortably in the secondary market without an implicit federal guarantee.

The majority of federal housing policies are either less significant or less problematic. Section 8 housing vouchers provide housing-related aid to poorer Americans, who typically rent. Economists generally accept this intervention as a reasonable form of market-oriented in-kind redistribution. The Low-Income Housing Tax Credit subsidizes the production of affordable housing nationwide. The credit, however sensible in places where housing is expensive, such as along America's coasts, is less tenable in the case of construction-crazy Atlanta or depressed Detroit. I am no fan of the tax credit, but I will restrict my discussion to the even bigger interventions – the home mortgage interest deduction, the FHA, and the Government-Sponsored Enterprises (GSEs) – that encourage people to buy homes by borrowing cheaply.

In 2009 and 2010, the Federal Reserve bought \$1.25 trillion worth of mortgage-backed securities, allegedly “to provide support to mortgage and housing markets and to foster improved conditions in financial markets more generally.”⁶ The vast size of this purchase certainly suggests the large public cost of Fannie Mae and Freddie Mac, but the intervention itself had far more to do with avoiding a financial market meltdown than with supporting housing markets. This essay focuses on the housing policies

that helped create the crisis, not on the interventions – like the purchase of mortgage-backed securities – that tried to prevent the Great Recession from turning into a second Great Depression.

Broadly speaking, there are three public objectives that underlie housing interventions and, in theory, justify the existence of the FHA, Fannie Mae, and Freddie Mac. The first is narrowly related to imperfect information in credit markets. For at least four decades, economists have understood that a “lemons” problem – whereby borrower and lender have asymmetric information – may hamper insurance or lending markets. In principle, the lending market could break down if some borrowers face a higher cost (psychic or otherwise) of default than others and if those borrowers (but not their lenders) know the costs of default. If lenders cannot distinguish between high and low risk and, in turn, charge all borrowers the same rate, then it is possible that high-risk borrowers will be more likely to borrow while low-risk borrowers may choose to self-finance or, less plausibly, rent. In this scenario, the high-risk borrowers are willing to pay the high interest because they have a greater chance of defaulting, or exercising their option not to repay the loan. If only risky borrowers go to banks, then interest rates will be extremely high, in which case the safe borrowers who will never default may prefer to avoid the market altogether. A vicious circle results in which risky borrowers cause high interest rates, and high interest rates deter safe borrowers. The FHA is tasked with preventing this scenario by creating a more liquid market that does not charge prohibitively high rates.

A similar breakdown can occur in the secondary market for mortgages if mortgage issuers know more about a mort-

gage's quality than any prospective buyer. The U.S. Securities and Exchange Commission case filed against investment banking and securities firm Goldman Sachs in April 2010 hinges squarely on whether the firm had material private information about mortgage-backed securities that it failed to disclose. More generally, mortgage issuers, like banks, have an incentive to keep the good mortgages on their own balance sheets and sell only the bad mortgages to outside investors. This lemons problem in the secondary market can impede the development of a risk-sharing mortgage market and lead banks to hold housing risk in their own portfolios. If bank defaults carry systemic consequences, either because of federal deposit insurance or because a failure of the banking system threatens the entire economy, then the country has an interest in developing a secondary market to spread mortgage risk more widely. By creating a subsidy so that all mortgages, at least within a given class, are resold, Fannie Mae and Freddie Mac can help that lemons problem disappear.

These asymmetric information problems are real, but are they large enough to justify the vast public subsidies that now exist? Do the public subsidies actually solve the problems? A remarkably large body of information about borrowers and their credit histories is available to banks and could be provided to purchasers of mortgage-backed securities. The fact that many lenders were sloppy in the run-up to the crash and ignored this information does not help the case for public insurance, which only mutes the incentives for due diligence. Moreover, an abundance of lending occurs without the benefit of public insurance. Many mortgages are securitized without the help of Fannie Mae and Freddie Mac. Arguably, the FHA and Fannie Mae made

sense during the Great Depression but are less relevant today. It is particularly hard to justify the existence of for-profit entities that deal in mortgages with the aid of an implicit government guarantee.

A second reason for public intervention in the mortgage markets is the desire to encourage Americans, especially veterans, to consume more housing. In 1940, 45 percent of America's housing units lacked complete plumbing facilities, and 20 percent housed more than one person per room.⁷ Nostalgia for prewar housing is misguided: the attractive older homes that remain are the homes that were nice enough to keep around. The typical prewar home was terrible by modern standards.

Prewar America provided at least three defensible reasons for the government to encourage housing consumption. First, tiny homes that lacked plumbing were often unsanitary and helped spread disease. The externalities from contagious illness were a primary motivation for the progressive reformers who focused on improving housing conditions in America's cities. Second, Americans have often been more comfortable with in-kind transfers, like food stamps, then with outright cash grants. It was surely an easier political sell to support better housing than to support higher welfare payments. Even "Mr. Republican," Robert Taft, was an advocate of federal housing assistance. Third, some have argued that better housing implicitly helps children by providing a healthier and safer environment. Recent research on Section 8 housing vouchers has found that voucher recipients move to neighborhoods with less crime and less exposure to environmental irritants that may trigger asthma.⁸

It is hard to justify mortgage-related interventions with this line of reasoning today. Essentially all Americans have ad-

equate plumbing, and even the poorest quintile of the population enjoys an average of 855 square feet of living space per capita. Indeed, the average amount of housing space consumed among the poorest fifth of Americans is nearly double the population-wide average in France, Germany, and the United Kingdom.⁹ Americans are now, by either historic or world standards, spectacularly well housed.

The third standard justification for mortgage-related interventions is that they encourage homeownership and its (alleged) attendant social benefits. Certainly, homeowners are more likely to vote, go to church, garden, know the name of their U.S. representatives, and say that they “work to solve local problems.”¹⁰ Economists have struggled to identify whether these effects represent a causal effect of homeownership, whereby owning a house creates better citizens, or whether the reverse is true: people who are inclined to good citizenship are more likely to become homeowners.

One approach to this problem has been to look at people over time, tracking whether their citizenship-related activities change with homeownership status. Using data on homeownership and civic involvement in Germany, economist Denise DiPasquale and I have found that a change indeed occurs when people become homeowners, but the effect is much smaller than that estimated by comparing homeowners and renters.¹¹ We have also attempted other – perhaps less convincing – means of estimating the causal link between ownership and citizenship. For example, we find that much of the impact of homeownership appears to come simply from length of time lived in the community. When people stay in one place, they develop ties and work more for communal aims.

The case for homeownership is also supported by the view that it leads to asset accumulation. Certainly, homes are the primary asset for many Americans, and rising home values have made many people wealthier. Various political leaders, including former President George W. Bush, have thought that America would be better off if more people had more capital and as a result have promoted homeownership as a means of achieving an “ownership society.”

Economist David Albouy has provided a more original argument for the home mortgage interest deduction. Income taxes influence many personal decisions, such as how many hours to work or how much money to save; economists typically consider these artificial influences, called distortions, to be unfortunate.¹² Income taxes also distort decisions about where to live by disproportionately taxing people who live in highly productive, high-wage areas like Boston and New York. Albouy argues that the home mortgage interest deduction makes up for some of this distortion by providing greater benefits for people living in high-cost areas, which tend to offer higher wages.¹³

The real benefits of federal housing policy must be stacked against the costs. Perhaps the most obvious disadvantage of the home mortgage interest deduction is its highly regressive character. Economists James Poterba and Todd Sinai have estimated that the average homeownership household earning between \$40,000 and \$75,000 per year receives an annual tax benefit of \$523 from the deduction. An average homeownership household that earns more than \$250,000 per year receives a benefit of \$5,459.¹⁴ These estimates assume – no doubt incorrectly – that behaviors would not change if the de-

duction were eliminated. However, we have no good way of estimating how many people would pay off their mortgages were the deduction eliminated.

Less attention has been paid to the progressivity of subsidies implicit in the FHA, Fannie Mae, and Freddie Mac. The FHA provides insurance only for less expensive homes and moderate-income individuals, though in the past, it has been accused of bias against minorities and poorer neighborhoods. Likewise, Fannie Mae's and Freddie Mac's benefits typically go to a middle section of the U.S. income distribution. These programs certainly are not strongly progressive – the very poor are less likely to be homeowners – but they are also far less regressive than the home mortgage interest deduction.

The regressive nature of the deduction should be part of any public discussion about reform. To an economist, however, regressivity is not intrinsically damning. After all, the income tax as a whole is highly progressive, and the regressivity of a particular deduction must be considered within that larger context. Even more important, the tools of economics do not naturally allow us to consider the virtues of taking from one group of citizens and giving to another. That discussion is the province of philosophers and politicians.

More aptly, the tools of economics are designed to judge the impact of these mortgage subsidies on potential homeowners' behavior. The subsidies do more than just encourage homeownership; they have significant, usually negative, side effects – at least from the perspective of mainstream economics, which generally views unjustified distortions of behavior as negative.

Perhaps the most obvious distortion is that the subsidies encourage people to invest excessively in housing relative to

other forms of capital. Indeed, providing incentives for people to buy bigger homes may not make much sense in a country where homes are already extremely large. Larger homes typically use more energy,¹⁵ and if carbon emissions damage the environment, then pushing people to buy large, energy-intensive dwellings has costs.

By pushing homeownership as a source of asset accumulation, the government ensures that individual portfolios will be highly skewed toward housing wealth. As we have just seen, housing wealth is hardly risk-free. Moreover, individual wealth levels are often enormously sensitive to changes in the housing market. Private buyers and lenders ultimately bear primary responsibility for the defaults and bankruptcies that have followed the housing collapse, but by encouraging individuals to borrow to bet on housing, the government bears some auxiliary responsibility.

Any housing subsidy would encourage individuals to invest too heavily in housing – relative to uninfluenced decision-making – but mortgage subsidies in particular encourage people to leverage themselves to the hilt. Even before the recent boom, in 1998, 3.9 percent of new mortgages had loan-to-value ratios above 95 percent. Enabled by subsidized mortgage interest, individuals typically refinance their homes to get cash out. In turn, home buyers have debt-financed portfolios, the value of which fluctuates wildly with the state of the housing market.

As Todd Sinai and Nick Souleles¹⁶ have argued, owning a home is itself something of a hedge. People are born into the world short of housing. Renters bear housing-price risk associated with changes in rental costs. At its best, homeownership provides protection against these fluctuations in rental costs,

but for people who anticipate moving or trading down, homeownership also creates portfolio risk. Policy that encourages people to borrow as much as possible to invest in just one asset class is of dubious value.

Subsidizing homeownership and housing consumption also impacts urban form. There is a tight connection between structure type and ownership type. More than 85 percent of people who live in single-family detached homes own their residences. More than 85 percent of people who live in structures with five or more units rent. This pattern provides good incentives for maintaining single-family homes and avoiding inter-owner conflict in multi-family dwelling. Single-family detached houses depreciate by an additional 1 percent per year when they are occupied by renters who fail to provide the sweat equity needed for maintenance. As those familiar with co-op boards can attest, the costs of spreading ownership of one large structure across a large number of owners underscore the benefits of having one owner for each roof.

But single-family detached houses are typically in suburbs while apartments are typically in cities. High land values in the urban core push structures upward. Seventy-six percent of Manhattan residents rent; 64 percent of residents in nearby Westchester County own.¹⁷ By subsidizing homeownership, the government implicitly encourages people to leave cities and move into single-family detached houses elsewhere.

There are good reasons to question the many government policies, including subsidies to agriculture and highway construction, which tilt against urban areas. In many cases, suburbs are far more homogeneous than cities, and encouraging suburbanization pushes people to live in segregated polities.

Suburban living normally entails more energy use because of larger homes and longer commutes.¹⁸ Much of the world's economic and cultural innovations occur in cities, and if such urban-centered activity generates wider social benefits, then discouraging city living carries further social costs.

Finally, homeowners are less mobile and less able to respond to economic shocks. Economist Andrew Oswald has found that, across European regions, unemployment is highest where homeownership is highest. This fact does not hold in the United States, but his argument has merit, especially in declining regions. Should public policy really be encouraging poorer Americans to buy property, tying them to cities that seem locked in long-run decline?

A housing market crash creates two opposite pressures on government policy. On one side, it exposes the folly of government policies that push people to bet on housing markets. On the other, it produces a dangerous climate in which to attempt new policies that could depress the housing markets even more. One force pushes toward reform, the other toward the status quo.

As of May 2010, the housing market seemed somewhat less poised on the brink of further disaster. Between May 2009 and January 2010, housing prices remained relatively steady in much of the country. In Las Vegas and Phoenix, prices had returned to their historic norms, which are closely related to construction costs. If recent trends provide any guide, the housing market can be quite stable for many years after prices level off. For example, between March 1991 and May 1997, the Case/Shiller ten-city index neither rose nor fell more than 2.5 percent in nominal terms relative to its March 1991 level. It is therefore rea-

sonable to contemplate appropriate housing reforms as long as those reforms are implemented gradually and with care.

The most promising target is the home mortgage interest deduction. Even if the goal of homeownership is considered sacrosanct, there are far better ways to promote it. The deduction not only has a number of harmful side effects, like encouraging borrowing and extra housing consumption, but it is also poorly targeted. The biggest benefits of the deduction accrue to the wealthy, who are likely to be homeowners regardless of tax policy. Less prosperous Americans, who are as equally likely to own a home as they are to rent, often fail to itemize the deduction even if they are homeowners. In fact, more than three-quarters of homeowners do not itemize; the poorest four-tenths of Americans take the standard deduction.¹⁹

For this group, the mortgage interest deduction does little, if anything. Over time, the deduction greatly increases in value during expected periods of high inflation because nominal interest can be deducted, yet there is hardly an observable increase in homeownership during periods when the deduction becomes more valuable. Across states, the size of the benefit (or changes in the size of the benefit) has no apparent effect on homeownership.

To increase homeownership without encouraging excessive borrowing or bigger homes, the natural tool would be a flat homeowner's tax credit. This tax credit could be independent of the size of the house or mortgage, would reduce tax payments for every American who owns his or her own home, and could be included on top of the standard deduction. This clear, tangible tax benefit would accrue to the owner regardless of the size or value of the home, and would

thus be far less tilted toward the rich. Inevitably, homeowners will tend to be wealthier, and a tax credit likely will not benefit people who do not pay taxes. Nevertheless, even though a flat homeowner's tax credit would still encourage migration from cities and would still give more benefits to the rich, it would be far less regressive and distortionary than the current policy.

As a somewhat more politically feasible policy approach, the cap on the home mortgage interest deduction might be lowered from \$1 million to, say, \$300,000, as suggested by the 2005 Tax Reform Panel. A simple, gradual method of implementing this change would be to reduce the deduction cap by \$100,000 per year for the next seven years. To a small degree, the cap would limit the incentive to invest in larger homes or borrow extreme amounts, but this approach would be decidedly second-best relative to the straight tax credit; for most Americans, it would preserve strong incentives to borrow and buy large homes.

The FHA, Fannie Mae, and Freddie Mac present a thornier problem. It is reasonable to argue that the FHA does solve a genuine market failure. Moreover, it provides for lower income Americans and is thus the least attractive candidate for major reform. The FHA has taken on a great deal of risk, but by slightly raising fees, it can easily accumulate enough capital to continue to cover its costs.

The case for a major overhaul of Freddie and Fannie is considerably stronger. At the time of this writing, many estimates for the taxpayer losses paid by Fannie and Freddie were upwards of \$300 billion.²⁰ Any subsidy to homeownership can be created directly through the tax system; there is little apparent need for a second,

less visible subsidy operating through government-sponsored enterprises.

These public agencies may not need to exist at all, but if they do, they will probably function best as slow-moving, highly regulated public utilities, with a mandate limited to insuring, bundling, and selling mortgage-backed securities. One model suggests reinstating them as a strong, independent public entity, with a governance structure vaguely resembling that of the Federal Reserve Board. The GSEs have often been thought to wield much weight on Capitol Hill, and lawmakers have often supported an expansion of their activities, as long as this expansion was joined with other activities that lawmakers found desirable. A better model would establish a national mortgage insurer with a greater degree of separation from politics and the profit motive.

National mortgage insurers should be required to refrain from all activities other than the insuring, bundling, and selling of mortgages. There is no reason for these entities to hold hundreds of billions of dollars of mortgage-backed securities in their own portfolios, as they did before the crash. If the agency is meant to solve market failures that challenge a purely private secondary mortgage market, then it should stick to that function.

A revamped public mortgage insurer would continue to charge a guarantee fee based on borrowers' loan-to-value and credit score. However, fees can be conservatively high. Moreover, the national mortgage insurer, like the Federal Reserve Board, should recognize that housing prices are not like stocks. They do not typically follow a random walk but, rather, show considerable mean reversion over five-year horizons. To correct this phenomenon, guarantee fees should be based not only on the ratio

of loan-to-purchase price, but also on the ratio of loan-to-expected resale value, which may be considerably lower.

For example, in markets that have risen considerably, expected resale value will be lower than current prices because prices tend to revert to historic norms. In markets where there is little restraint on building and land is abundant (which is the case in much of America), expected resale value should be tied to construction costs – despite the vicissitudes of current prices – with the expectation that prices will eventually return to the costs of delivering housing. While it is a mistake to put too much faith in any forecasting model, the evidence for mean reversion is sufficiently strong that requiring higher guarantee fees in places that have experienced significant recent price increases is a solid policy approach. If this policy had been followed during the boom, then guarantee fees would have become much higher during the booms of Phoenix and Las Vegas. The higher fees would have both protected the GSEs against risk and perhaps also helped check the wild swings in prices.

It is, of course, quite possible that an excessively conservative national mortgage insurer will charge too much, but that is a fairly minor problem as long as there are no barriers to free entry in the market for mortgage insurance. If private insurers offer better and cheaper products, then the public entity will not have much business, which is perfectly acceptable, at least as long as the private insurers cannot offer an implicit government guarantee of their own. The national mortgage insurer should guarantee that the secondary mortgage market does not disappear; that is its primary function.

The great national housing convulsion bears two major lessons: housing mat-

ters and prices fall as well as rise. Ordinary home buyers and the global financial system can suffer enormously when housing markets crash. It is neither possible nor desirable to legislate housing bubbles out of existence; the human capacity for over-optimism is just too strong. But it is both possible and desirable to limit the public role in encouraging those bubbles and to reduce the suffering they cause. We can reduce the public subsidies that encourage people to borrow as much as possible to make risky investments in housing.

The tax code is currently geared to induce mortgage debt. Every buyer who bought during the boom had the blessing of a government that encouraged homebuying through rhetoric and the tax code. Additional subsidies

were provided by the government-sponsored enterprises, which insured mortgages at a cost that now seems too low. These policies artificially encouraged borrowing.

Even if America wants to encourage homeownership, it does not need to encourage borrowing or buying big homes. A simple, moderate homeowner's tax credit would encourage homeownership more fairly and with fewer distortions. A plodding, conservative public mortgage insurer that charges relatively high fees can ensure the existence of a secondary mortgage market with less risk to taxpayers and less excessive borrowing. The housing crash exposed the folly of subsidizing leveraged bets on housing. It would be a shame if we ignore that lesson.

ENDNOTES

- ¹ Edward L. Glaeser, Joshua D. Gottlieb, and Joseph Gyourko, "Can Cheap Credit Explain the Housing Boom?" NBER working paper 16230 (National Bureau of Economic Research, 2010).
- ² Edward L. Glaeser and Joseph Gyourko, "Housing Dynamics," NBER working paper 12787 (National Bureau of Economic Research, 2006).
- ³ Karl E. Case and Robert J. Shiller, "The Behavior of Home Buyers in Boom and Post-Boom Markets," in *Stock Market Crashes and Speculative Manias*, ed. Eugene N. White, Elgar Reference Collection, International Library of Macroeconomic and Financial History, no. 13 (Cheltenham, U.K.; Brookfield, Vt.: Elgar, 1996), 314–330 (distributed by Ashgate).
- ⁴ <http://www.hud.gov/offices/adm/hudclips/letters/mortgagee/files/10-02ml.pdf>.
- ⁵ This fee, called the "up-front" fee, is paid at closing, but lenders often roll the fee into the loan, so that the "up-front" fee is paid off over many years. The FHA also charges an annual mortgage insurance fee.
- ⁶ http://www.newyorkfed.org/markets/mbs_faq.html.
- ⁷ Edward L. Glaeser and Joseph Gyourko, "Rethinking Federal Housing Policy" (Washington, D.C.: AEI Press, 2009).
- ⁸ Lawrence F. Katz, Jeffrey R. Kling, and Jeffrey B. Liebman. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics* 116 (2) (May 2001): 607–654.
- ⁹ Glaeser and Gyourko, "Rethinking Federal Housing Policy."
- ¹⁰ Denise DiPasquale and Edward L. Glaeser, "Incentives and Social Capital: Are Homeowners Better Citizens?" *Journal of Urban Economics* 45 (2) (March 1999): 354–384.

- ¹¹ Ibid.
- ¹² David Albouy, “The Unequal Geographic Burden of Federal Taxation,” *Journal of Political Economy* 117 (4) (August 2009): 635–667.
- ¹³ The correlation across metropolitan areas between median incomes and median property values in the 2000 U.S. Census was 67 percent. Joseph Gyourko and Todd Sinai provide a summary of the geography of the home mortgage interest deduction, demonstrating that the benefits accrue primarily to richer areas; Joseph Gyourko and Todd Sinai, “The Spatial Distribution of Housing-Related Ordinary Income Tax Benefits,” *Real Estate Economics* 31 (4) (Winter 2003): 527–575.
- ¹⁴ James Poterba and Todd Sinai, “Tax Expenditures for Owner-Occupied Housing: Deductions for Property Taxes and Mortgage Interest and the Exclusion of Imputed Rental Income,” *American Economic Review* 98 (2) (May 2008): 84–89.
- ¹⁵ Edward L. Glaeser and Matthew E. Kahn, “The Greenness of Cities: Carbon Dioxide Emissions and Urban Development,” *Journal of Urban Economics* 67 (3) (May 2010): 404–418.
- ¹⁶ Todd Sinai and Nick S. Souleles, “Owner-Occupied Housing as a Hedge Against Rent Risk,” *Quarterly Journal of Economics* 120 (2) (May 2005): 763–789.
- ¹⁷ The U.S. Census, American Factfinder, is the source of both figures, which refer to the 2006–2008 period.
- ¹⁸ Glaeser and Kahn, “The Greenness of Cities.”
- ¹⁹ Edward L. Glaeser and Jesse M. Shapiro, “The Benefits of the Home Mortgage Interest Deduction,” *Tax Policy and the Economy* 17 (2003): 37–82.
- ²⁰ http://www.bloomberg.com/apps/news?pid=email_en&sid=a2Z5GnTAPcuo.

Barry Eichengreen

*International financial regulation
after the crisis*

What was once almost quaintly referred to in the United States as the Subprime Crisis eventually came to be known, in America and abroad, as the Great Global Credit Crisis of 2008 – 2009. The change in terminology is itself indicative of the international spread of the crisis. It reflects how, in the end, no country was immune to the global reach of financial instability.

Now that the worst effects are past, officials have begun drawing lessons and formulating policy responses. The United States has focused on strengthening mortgage underwriting standards, moving transactions in derivative financial instruments onto organized exchanges, and curbing proprietary trading by depository institutions insured by the Federal Deposit Insurance Corporation (FDIC). The United Kingdom has emphasized the perverse incentives created by bonus-based compensation of financial executives and sought to reform executive pay. France and Germany have singled out the risks created by lightly regulated hedge funds and private equity firms. Officials in other countries have prioritized still other issues. On what should take precedence there is little agreement.

Herein lies the problem: to be effective, regulation of financial markets and institutions must be coordinated across countries. Most big banks operate in more than one country, affording them the opportunity to relocate their operations and employees. When the United Kingdom moved to tax bankers' bonuses, the bankers in question threatened to move to Geneva. When the European Union began to contemplate strict regulation of hedge funds, fund managers proposed moving their operations to New York. When the EU then mooted the possibility of prohibiting residents from investing in hedge funds and private equity firms regardless of where they were located, the U.S. Treasury complained that such measures unfairly discriminated against the U.S. financial services industry and violated international treaties like the General Agreement on Trade in Services.

Not only is disagreement among national regulators over priorities and strategies a source of potential conflict, but it threatens to vitiate their efforts to make the world a safer financial place. In a financially integrated world, many regulatory restrictions are impossible to effectively enforce purely at the national level. Whatever the regulatory response – taxing bank size, bank bonuses, or specific

bank activities – limiting evasion requires a significant degree of international cooperation.

The above examples are specific instances of the general phenomenon known as *regulatory arbitrage*. When restrictive regulation raises the cost of doing business (as it is designed to do when business activities have social costs), businessmen have an incentive to relocate to more permissive jurisdictions, frustrating regulators' efforts and raising the costs to society of the business activity in question. In the United States, regulatory arbitrage may mean shifting an activity between affiliated firms that are subject to different regulations and overseen by different regulators – from a bank to an affiliated insurance company or a structured investment vehicle (SIV), for example. Regulatory arbitrage can also mean shifting high-risk activities from a country with more stringent regulation to another where regulation is more permissive. To limit this behavior, international cooperation in establishing a common regulatory standard is the obvious way forward.

While the case for cooperation is straightforward in principle, the mobility of finance creates a temptation for regulators to undercut one another in practice. Competition for business may create a race to the bottom. Competing jurisdictions, seeking to attract foot-loose financial firms, have an incentive to offer more permissive regulation or lax enforcement. More than one country has launched “Big Bang” reforms, liberalizing burdensome regulation and, sometimes, weakening enforcement in an effort to enhance its attractions to internationally mobile financial firms. To address this problem, competing countries might enter into an international agreement that requires consent-

ing parties not to undercut one another's regulatory efforts.¹

A further argument for international cooperation is that individual countries, making decisions in isolation, lack adequate incentive to engage in rigorous supervision and regulation of domestic financial firms and markets. While doing so is costly, the benefits accrue not just to the individual country but also to its neighbors. The virulence of financial “contagion” – the speed and extent to which instability can spread, infecting the financial systems of other countries – such as that which was evident in the aftermaths of the Bear Stearns, American International Group, and Lehman Brothers crises in March and September 2008, illustrates the point. In these episodes, inadequate supervision and regulation in the United States spawned a crisis that engulfed the entire world. As Thomas Mayer, chief economist of Deutsche Bank, put it, “In this day and age, a bank run spreads around the world, not around the block.”²

The implications for regulation are direct. If a government invests in regulation, some of the benefits may accrue to other countries, giving the initiating country inadequate incentive to invest. The problem is not unlike that of residents of a fire zone. It is in the self-interest of each resident to clear the underbrush around his home in order to enhance his own safety. But the individual homeowner may not consider the additional benefits, from a social point of view, of his brush-clearing exercise; he may not see that doing so also enhances the safety of his neighbors. In the urban context, municipal regulation requires everyone to clear additional brush. In the context of global finance, the solution rests on international standards and coordination of financial regulation.

Finally, cooperation could create a viable alternative to the uncontrolled bankruptcy of troubled financial institutions on the one hand and emergency rescues on the other. Emergency rescues are criticized on both equity and efficiency grounds. To the extent that a rescue is financed by taxpayer money, present or future, it is rightly seen as unfair. Moreover, because banks know they will receive assistance in the event that the bets they make go bad, they have an incentive to place bigger and riskier bets. Regular recourse to rescues creates moral hazard, which has social costs.

But the alternative – allowing a big bank to declare bankruptcy – is not tenable if doing so threatens the stability of the financial system. The troubled institution will have borrowed from other financial institutions. It will have other contracts outstanding, many of which will be frozen when bankruptcy is filed and an automatic stay is imposed. A stay may therefore cause liquidity problems – and worse – for the bank’s counterparties, forcing them to call in their own loans in order to raise funds. Asset prices may collapse in a fire sale, and the liquidity crisis may cascade through the financial system. Lack of attention to third-party effects in Chapter 7 and Chapter 11 bankruptcy proceedings coupled with the slow pace of court proceedings render these problems especially pervasive and this approach to resolution particularly problematic in the case of financial firms.³

In the United States, the FDIC can step in, seize, and ring-fence the operations of a bank to which it extends deposit-insurance coverage. It can pay off a bank’s obligations to its counterparties. But the FDIC lacks this authority with regard to the big, complex bank holding companies that pose the most serious threat to systemic stability. Insurance companies

and other large nonbank intermediaries are also beyond its jurisdiction.

The Dodd-Frank financial reform bill adopted in Summer 2010 gives federal regulators new authority to seize and break up large troubled financial firms. What it fails to recognize, unfortunately, is that most large financial companies – the presumed targets of these procedures – operate internationally. A case in point is Lehman Brothers, which had consequential operations in the United Kingdom as well as in the United States. The conflicting claims of creditors in the two jurisdictions, together with differences in resolution regimes, created serious difficulties for courts and regulators seeking to limit the systemic consequences of the institution’s failure in 2008.

Establishing an orderly resolution regime as an alternative to bailouts is desirable. But meaningful progress will require, at a minimum, that provisions adopted at the national level be coordinated internationally. And if cooperation proves inadequate, regulators will have to contemplate creating a supra-national resolution authority.

Not everyone is convinced, however, by the case for international regulatory coordination. Regulatory oversight, to be effective, must be tailored to local needs. National financial structures and systems differ, requiring differences in the structure and application of regulation. In some countries, financial systems are predominantly bank-based. In others, the United States and the United Kingdom being prominent examples, the securitization of financial claims is more extensive. In these countries, the growth of securities markets has led to “disintermediation” – the displacement of bank borrowing and lending to securities markets. Financial systems must be regulated differently according to whether they

are dominated by banks or securities markets. Further differences include whether the state owns a stake in the country's leading financial institutions; whether Internet banking is prevalent; and whether the country is predominantly Islamic, as Islamic banking prohibits the payment of interest, requiring lending to be structured in other ways. All these are reasons why one-size-fits-all regulation, which is what tends to come out of international agreements, is undesirable.

An additional danger is that international coordination may lead to lowest-common-denominator regulation. Agreements on matters such as minimally adequate capital ratios for internationally active banks, to be effective, must be accepted by all countries with consequential banking systems. As a result, agreements tend to be by consensus. For all concerned countries to agree, any provision that one country finds objectionable must be removed from the agreement or, at the least, watered down. The result is a weak and ineffectual agreement. As a case in point, critics of international agreements point to the Basel Accord for Capital Adequacy for Internationally Active Banks. The Basel Accord is designed to ensure that banks have buffers to cope with adverse circumstances, but it has not prevented bank capital from falling to near zero and bank solvency from being threatened in each of our recent financial crises.

Part of the problem may be that international agreements negotiated in far-distant places like Basel are prone to be captured by the regulated. Banks with a preference for relatively permissive regulation have the resources and expertise to influence the Basel Committee on Banking Supervision. This is not so for the man or woman in the street concerned that regulation should be suffi-

ciently stringent to protect his or her interest. This is another reason that international agreements may lead to weak and ineffectual rules.

Finally, some argue that regulatory diversity is desirable for the same reasons that biodiversity is desirable. Efforts at international coordination, whereby governments converge on a single set of standards, may in fact leave the global financial system more exposed. The standards in question may turn out to be poorly designed and inappropriately targeted. If they are designed to contain one set of risks and a different set materializes, the fact that all jurisdictions have adapted their regulatory policies in the same way may only contribute to financial instability. From this standpoint, given the possibility of diverse shocks to global financial markets, the presence of diverse regulation may enhance the stability of the financial ecosystem.

There is something to these arguments – which is precisely why there is a debate about the efficacy of internationally coordinated reform. Yet, even conceding these points, the case for cooperation is overwhelming. For one, the implications of differences in financial structure should not be overstated. Despite variance across national systems, over time there has been a strong tendency toward convergence. Furthermore, countries with bank-based financial systems and those with market-based systems can adopt the same approach to regulation of their banks and securities markets; the consequential difference would be to which set of regulations they devote the bulk of their enforcement effort.

Second, the fear that agreement by consensus leads to lowest-common-denominator regulation should not preclude cooperation. Rather, countries most concerned with risks to financial stability should move ahead with coordinated

reforms and apply sanctions that discourage their financial institutions from doing business with the countries that lag behind. They should similarly prohibit financial firms chartered in less regulated jurisdictions from entering their markets.

Finally, if nonfinancial interests are inadequately represented in international negotiations, then the appropriate response is not to abandon those negotiations but to open them up to additional stakeholders.

Historically, the most prominent international institution concerned with regulatory reform has been the Basel Committee on Banking Supervision, which is made up of representatives of central banks and other banking supervisors. The Basel Committee meets on the premises of the Bank for International Settlements in Basel, Switzerland, a minimum of four times a year. When it was founded in 1974, only the United States, Canada, Japan, and seven European countries were represented. Membership has since expanded to twenty-seven countries, including the emerging markets of Latin America and Asia.

While the Basel Committee has traditionally focused on ensuring that banks have enough capital to weather shocks, over time it has also considered a variety of other stability-related issues, including, most recently, liquidity risk. Its signature achievement is the Basel Accord for Capital Adequacy for Internationally Active Banks, as virtually all large banks today are internationally active.

Unfortunately, the Basel Accord, and especially the updated version, Basel II, published in 2004 and adopted by a growing list of countries subsequently, are now understood to be seriously flawed. Basel II allowed banks to hold less capital against assets that received

investment-grade ratings from commercial credit rating agencies. Because ratings are revised upward in good times and downward in bad times, this practice was strongly procyclical. It encouraged even more lending when lending was booming and more retrenchment when financial institutions were retrenching. It ignored the conflicts of interest to which the rating agencies were subject. It also allowed banks to use their own internal models – later shown to be problematic – to evaluate the risk of losses and the amount of capital that had to be held against those risks. It bought into the bankers' arguments that they could safely reduce their own funds held in reserve (so-called Tier 1 capital) to little more than 2 percent of bank assets.

Some of these problems issue from the Basel Committee's origins and their influence on its remit. The Committee was created in 1974 in response to problems with a major cross-border *bank* (Germany's Herstatt Bank); the focus of the Basel Committee on Banking Supervision, therefore, is on *banking* supervision. Once upon a time, the perimeters of the banking and financial systems were roughly coincident because banks were the dominant providers of financial services. But with the growth of securities markets and nonbank financial firms, this is no longer the case.

As a result, the Basel Committee set capital adequacy standards for commercial banks; meanwhile, in countries such as the United States, it was not just the large commercial banks (and commercial bank holding companies) that posed potential risks. Once upon a time, investment banks like Lehman Brothers invested only their own partners' capital, but more recently, they began leveraging their operations with borrowed funds. Broker dealers like Bear Stearns that

booked and cleared the trades of others increasingly engaged in proprietary trading, using an even higher ratio of borrowed money to own capital than was typical of investment banks. Insurance companies like the American International Group (AIG) were overseen only by state insurance commissioners – to the extent that they were overseen at all. Markets in structured derivative securities were often entirely unregulated. The members of the Basel Committee were aware of these gaps, but awareness does not equate to the capacity to act.

Early recognition of these problems led to the 1999 formation of a second body, the Financial Stability Board (FSB; originally named the Financial Stability Forum). The FSB, which is supported by a small secretariat also housed at the Bank for International Settlements, has a mandate to assess vulnerabilities affecting the entire financial system and to oversee action to address them. It seeks to monitor market activity, highlight regulatory developments, and identify systemic risks. Some two dozen countries are represented. Members include not just central banks and other regulators but also financial standard-setting bodies like the International Association of Insurance Supervisors and the International Organization of Securities Regulators. The FSB has created committees concerned with a range of issues that potentially pose risks to financial stability, including international capital flows, hedge funds, and offshore financial centers. Its deliberations, when agreement is reached, result in a set of recommendations and a published report.

The FSB's mandate to monitor the global financial system as an integrated whole constitutes an important step forward. The Board's limitation is that it is essentially a talk shop: a place to exchange information and pontificate on

what is desirable, after which regulators are free to go home and do more or less as they please. Like the Pope, the FSB has no army. It has even less sanctioning power.

Its members are aware of this problem. Indeed, one need only consult the FSB's antiseptically titled January 2010 report, "Framework for Strengthening Adherence to International Financial Standards." This report succeeds in identifying only three mechanisms for ensuring adherence to its standards: leading by example on the part of member jurisdictions, peer review, and a vaguely specified "toolbox of measures" (specific tools presumably being too sensitive for the lid on the box to be lifted). It seems unlikely that the FSB entertains any illusion that these limp instruments will get national regulators to sit up and listen.

Then there is the Group of 20 (G20), which, recognizing the emergence of a more multipolar world, has assumed the role of steering committee for the world economy (a role played previously by the Group of 7 and Group of 10 advanced countries). Its twenty members include advanced countries and emerging markets alike, as well as the European Union. At recent meetings, it has also included a twenty-first member, Spain, a large country that was inexplicably excluded when the G20 was formed but whose attendance was insisted upon by the European Commission, and a twenty-second, The Netherlands, a member of previous groupings that insisted on inviting itself. The G20 is a mechanism to ensure that not just regulators but leaders (finance ministers and prime ministers, who presumably give marching orders to the regulators) buy into the process of policy reform. Following its biannual meetings of heads of state and government, it issues communiqués that include commitments on financial reform. At its June 2010 summit

in Toronto, for example, leaders committed to phasing in higher capital standards for banks. Similar to the FSB, the G20 forms working groups to investigate financial problems and offer recommendations.

But the G20 has a legitimacy problem: it was formed in response to an earlier crisis in the late 1990s, essentially in ad hoc fashion. (The equally ad hoc participation of Spain and The Netherlands, whether desirable or not, is indicative of this fact.) No one appointed this particular set of countries to make decisions for the world. Nothing ensures that their recommendations will be respectfully received and acted upon by countries represented in their deliberations.

A final institution concerned with financial stability is the International Monetary Fund (IMF). The IMF has a written constitution, the Articles of Agreement, to which its members are bound. It organizes countries into constituencies, each of which is represented by a member of its Executive Board, ensuring that all 180-plus national members have voice; it therefore does not suffer from the same kind of legitimacy deficit as the G20. It publishes a Global Financial Stability Report designed to provide a synthetic assessment of risks to the international system. In conjunction with the FSB, it conducts early-warning exercises designed to anticipate financial problems. Together with the World Bank, its sister organization, it conducts regular financial-sector assessments intended to provide outside reviews of the financial strength and regulatory practices of its members. In the course of this review, it recommends adherence to international best-practice standards. Countries that borrow from the Fund are subject to an enforcement mechanism; the IMF can deny disbursement of the next quarterly installment

of its loan to countries it judges to have made inadequate progress in fixing economic and financial problems.

But the IMF possesses no such power over other countries. Illustrative of this limitation is the fact that countries must agree before a potentially embarrassing assessment of their financial sectors can be conducted. Shortly before the Subprime Crisis, the IMF and World Bank reportedly approached the U.S. government to request an assessment and were refused.

Evidently, there is no dearth of studies and no shortage of committees, boards, and organizations concerned with international aspects of regulatory reform. There is, however, a shortage of consequences for countries whose regulatory policies are not adequate. The question is how to address this problem.

One option would be to create a new body, a World Financial Organization (WFO), membership in which would create concrete obligations whose violation would have significant consequences. In the same way that the World Trade Organization (WTO) establishes principles for trade policy (such as nondiscrimination, reciprocity, transparency, binding and enforceable commitments) without prescribing outcomes, the WFO would establish principles for prudential supervision and regulation (capital and liquidity requirements, limits on portfolio concentrations, adequacy of risk measurement systems and internal controls) without attempting to prescribe the structure of regulation in detail.

But once the WFO defined obligations for its members, the latter would be obliged to meet them. Membership could be made mandatory for all countries seeking freedom of access to foreign markets for their investors and domestically chartered financial institutions. The WFO

could appoint independent panels of experts to determine whether countries were in compliance with their obligations. In cases of noncompliance, other members would be within their rights to restrict the ability of financial institutions chartered in the offending country to do business in their markets. Not only would this measure protect members from the negative consequences of inadequate regulation abroad, but doing so would provide a real incentive to comply.

Critics will undoubtedly object that governments, not least the U.S. government itself, will never allow an international organization to dictate their national financial policies. However, the WTO would not dictate; the specifics of implementation could be left to the individual country. Furthermore, the equivalent already exists for trade. The United States is among the countries that have signed WTO agreements with specific

obligations. The WTO has the power to establish dispute settlement panels and determine whether national law complies with a country's WTO obligations. Violators have the choice of changing that legal provision or facing trade sanctions. If the United States and other countries accept this authority in the case of trade, one might ask, why shouldn't they accept it in the case of finance?

There is no reason why the Basel Committee, the FSB, the IMF, and the others should not continue their useful work studying regulatory problems, encouraging their correction, and promoting the international coordination and harmonization of regulatory initiatives. It has become clear, however, that more comprehensive, binding, and coordinated international regulation will be crucial to financial stability worldwide, now and in the future.

ENDNOTES

¹ To be sure, some would assert that restrictions on financial firms and markets are excessive and that regulatory competition is desirable as a countervailing force. But their position is less tenable in the wake of the crisis.

² Quoted in Mark Landler, "Europe Catches What's Ailing U.S. Financial Sector," *The New York Times*, October 1, 2008.

³ Chapter 7 and Chapter 13 are the provisions of the bankruptcy code under which the operations of an insolvent entity are liquidated and reorganized, respectively.

Peter Temin

The Great Recession & the Great Depression

In the depths of the Great Depression, John Maynard Keynes wrote that “[p]ractical men, who believe themselves to be quite exempt from any intellectual influence, are usually the slaves of some defunct economist.”¹ This acute observation is applicable to our current Great Recession as well. In fact, the newly discredited ideas are not all that different from the old, suggesting that Keynes may have overestimated people’s ability to learn from their mistakes.

I pursue the parallels between these two watersheds in recent economic history along three paths: the causes of the crises and their relation to economic theory; the spread of the crises on a global scale; and, finally, recovery – at least as far as we can see it at this point. As Karl Marx famously said, history repeats itself “the first time as tragedy, the second as farce,”² a criticism that also fits our current condition.

Both of these dramatic and costly economic crises emerged from the interaction of economic imbalances in the world economy and the ruling ideology of financial decision-makers who confronted these imbalances. World War I,

a paroxysm of violence that brought the long economic expansion of the nineteenth century to a sudden end, produced the imbalance that led to the Depression. Britain, the workshop of the prewar world, was exhausted by the struggle. America, the rising economic behemoth, was unprepared to take responsibility for its new role in the international economy. Germany, having unsuccessfully challenged the Allied Powers, refused to acknowledge its defeat.

Patterns in the international movement of capital reveal this imbalance. During the postwar decade, one of the most important reasons for approving resumed capital flows was the ruling economic theory of the gold standard. In the eighteenth century, philosopher, historian, and economist David Hume explained how currencies valued in gold remain stable relative to each other. If, for instance, a shock to one country decreased its exports, the result would be an outflow of gold, which would lower prices in the exporting country. Lower prices would encourage exports and decrease imports, leading to an inflow of gold. Prices would rise again, re-creating the previous equilibrium. This argument is known among economists as the price-specie-flow mechanism.

© 2010 by the American Academy of Arts & Sciences

Hume's contribution is still useful today, although we are now aware of the many assumptions that underlie the mechanism's proper functioning. In particular, the theory presupposes that prices are fully flexible and determined in competitive markets. These assumptions express a view of the economy – often attributed to Adam Smith – that has become characteristic of economic models in the years since Hume's writing. This view is taught in introductory economics classes; it is the starting point for many journal articles; and it is referred to as a perfectly competitive economy. When conditions cross the line between descriptive and normative, however, they are transformed from description – which may or may not be accurate – to prescription – which in turn affects public policy.

These conditions may have been fulfilled in the eighteenth century, but they were not accurate in the 1920s postwar world. In contrast to Hume's assumption of a fixed link between gold flows and prices, central bankers thought themselves responsible for inflation and deflation. Business firms had become larger, and many product markets were no longer fully competitive. As the size of production units, whether mines or factories, became larger, the ability of labor markets to be fully competitive also diminished. Large employers yielded little bargaining power to workers to negotiate wages and working conditions. If a factory, for example, was the only large employer in town, the options for workers were even more limited and the market power of the employer more obvious. Workers formed unions to countervail the market power of employers, and wage bargaining and strikes supplanted the individual wage negotiations implicit in Hume's and Smith's analyses.

Nevertheless, after World War I, policymakers could think of no better way to reorganize the international economy than to restore the gold standard – that is, to fix one price (the exchange rate) while assuming all others were flexible. Freezing exchange rates in this fashion reduced countries' ability to adapt to new conditions; this defect, however, was deemed preferable to the anticipated chaos of alternative arrangements. When England attempted to reduce prices to sustain the value of sterling, a general strike resulted, revealing both the inaccuracy of the gold standard's underlying assumptions and the strength of the economic policies based on those assumptions.

In this context, the United States took over the position of leading international lender and exported massive amounts of capital to Germany in the 1920s. The loans were meant to help Germany maintain the gold value of its currency, and they enabled Weimar Germany to pay reparations owed to the victors in World War I and enjoy a consumption boom. Higher prices after the war also put strain on gold currencies, and while England and Germany struggled as a result, economists in the more prosperous United States proclaimed the advent of a new economy in which stability and prosperity would continue indefinitely. Hind-sight suggests that these conditions were not sustainable; rather than celebrating the promised strength and vitality of a new economy, countries should have been concentrating on how to avoid a rough landing from the high-flying results of the previous shocks.

Economic troubles appeared in Germany and the United States in the late 1920s. The former's consumption growth produced a boom in municipal expenditures that began to fizzle; in the latter, both housing and stock market booms

eventually crashed. As recession spread to other countries, international trade decreased, but prices did not fall rapidly enough to equilibrate markets in the fashion Hume described. Prices were sticky, and rather than deflation, a lack of foreign reserves led to unemployment. When all countries found their exports falling, the processes of deflation and depression chased a moving target.³

A similar international imbalance developed after the end of the Cold War. The new world lender, the United States, traded roles and became the world's largest borrower. China, a "loser" in the Cold War, became the United States' primary lender. Just as the inflow of capital to Weimar Germany had fueled expansion, the inflow of capital from China financed a consumption boom in the United States that developed into a housing boom.

This global imbalance was apparent, and economists feared that a crisis would ensue. Because the United States no longer adhered to the gold standard, the value of the dollar could change freely from day to day. The question was whether there would be a smooth decline in the value of the dollar, in the fashion of the price-specie-flow mechanism, or an abrupt fall. These concerns were misplaced; even though the international imbalance created crisis conditions, short-run booms and busts precipitated economic calamity in the interwar and recent years. One such boom was surging housing expenditures in the 1920s. The housing market was only a minor player in the drama of the Great Depression, but it had a starring role in our current crisis.

The housing boom flourished in recent years, nourished by the availability of Chinese capital and the ruling economic theory of the Washington Consensus. This term, coined in 1989, referred to a set of economic policies that ranged from

stable exchange rates and responsible fiscal policies to deregulation and privatization.⁴ It was an adaptation of the gold standard to current conditions, stipulating stable – instead of fixed – exchange rates to avoid the rigidities of the gold standard that proved to be harmful in the 1930s. Other requirements marked a departure from the era of large government that followed the Great Depression and World War II. The terms of the Consensus favored diminished government influence, so as not to impede the progress of private finance and industry; competition would ensure continued growth and prosperity. Like the gold standard, the Washington Consensus was based on the Enlightenment ideas of David Hume and Adam Smith and promulgated as a way to organize the postwar world. It was the economic component of the new world order that the first President Bush was looking for.

More explicitly than the gold standard mentality, the Washington Consensus spelled out the conditions needed to maintain stable exchange rates. It acknowledged that most economies in the later twentieth century did not resemble the eighteenth-century conditions analyzed by Hume and Smith and argued that policies designed to re-create these earlier conditions would lead to economic growth and prosperity. Using familiar theories of competition and flexible prices, the underlying theory showed how the competitive process of allocating resources in individual markets would generate stable conditions for society as a whole.

Banks and associated businesses in the United States extended the underlying reasoning to the creation of new assets known collectively as structured finance. The Washington Consensus was designed to reduce risks, and innovative securities provided a means of allocating risk to

those investors who wanted to take it on. Just as banks can hold fractional reserves on the assumption that people draw on their deposits randomly and independently, the creators of new securities reasoned that homeowners default on mortgages randomly and independently. Collateralized debt obligations (CDOs) allowed financial institutions to benefit from the fact that only a few homeowners default in any given time, so that most mortgages are safe. Combining mortgages into “tranches,” banks could separate the safe part of mortgages from the risky parts without knowing which mortgages would be defaulted – just as banks do not know which deposits will be withdrawn but can safely assume that only a fraction will be withdrawn at any given time.

Based on the ability to sell mortgages to be securitized, mortgage brokers expanded, encouraged homeownership, and promoted the ownership society championed by the second President Bush. Banks and other financial intermediaries holding securitized assets took on more and more leverage, which was justified by their calculations that the risks of many of these assets were vanishingly small. But when the resulting housing boom burst and many mortgages failed, the assumption that defaults occurred randomly and independently turned out to be false. CDOs were much more risky than they had appeared to be, and the separation of risky and safe assets proved to be invalid.⁵ Investors refused to buy the CDOs, and credit markets seized up. Countries that had adopted the policies of the Washington Consensus found themselves mired in a worldwide financial crisis.

The Great Depression and the Great Recession were both caused by policies derived from nostalgia for the world of the Enlightenment. Drawing on theories

from the eighteenth century, hard-headed policy-makers either assumed or tried to re-create the idealized conditions described by Hume and Smith. These policy-makers ignored both the growth of economies of scale in modern economies and the work of behavioral economists that has shown that people do not behave as *homo economicus*. Their efforts produced the new economy of the 1920s and the Goldilocks economy of recent decades that turned into booms and busts. Was it inevitable that these economic expansions would end badly? According to the late economist Hyman Minsky, people become more complacent with prosperity and more willing to take on risks they often know are highly suspect.⁶ More recently, economists Carmen Reinhart and Kenneth Rogoff analyzed historical evidence and reached a similar conclusion: booms typically precede financial crises, just as pride goes before a fall.⁷

More formally, people in both expansions miscalculated the risks they faced. Their models were based on shocks to individual countries or homeowners and did not allow for collective actions. The gold standard model explained how to deal with a shock to an individual country, implicitly assuming that other countries would be immune to whatever disturbance affected the distressed country. The interaction between the country in crisis and other countries would lead back to stability; a collective shock to many economies was not considered. Similarly, the model behind the Washington Consensus considered individual risks. Structured financial obligations were valued as if the underlying risk of mortgage foreclosure was the result of random and independent shocks to individual homeowners. As with the gold standard, no consideration was given to collective shocks; housing prices were expected to rise continually. It was

assumed that homeowners experienced financial difficulty and defaulted on their mortgages randomly. The randomness of defaults enabled financial designers to reduce the risk to any security by diversification, that is, combining many mortgages the same way a bank combines many bank deposits. When the housing boom ended and housing prices fell, however, many homeowners began to default, and the risk that was supposed to be protected for through diversification was now present in securities previously thought to be risk free. Investors could not discern safer assets from those more at risk, and the prices of all structured finance fell. Prices of some securities fell rapidly because there were no buyers for them. Financial markets froze in September 2008.

The second step of this comparative analysis is to evaluate the spread of each crisis. In the early 1930s, countries pursued a moving target as their economies contracted to deal with what appeared to be budget and current-account deficits. Consequently, a series of currency crises in Summer and Fall 1931 turned a bad recession into the Great Depression. The German mark collapsed when the chancellor put domestic politics ahead of sensible finance.⁸ The Bank of England abandoned the gold standard after a subsequent speculative attack. And the U.S. Federal Reserve raised its discount rate dramatically in October 1931 to preserve the value of the dollar; the Fed kicked the American economy when it was down and drove it further into depression.

Many countries continued to maintain deflationary policies in the early 1930s as they tried to hold on to the gold standard or, in the case of Germany, follow its prescriptions even after abandoning the gold standard. Some countries followed England off gold and created room

for expansive policies, which were neither large nor expansive enough to stimulate recovery in countries that remained in thrall to gold. It has become common to attribute the continued economic decline to banking crises, but banks failed only in countries that adhered to the gold standard.⁹ As long as countries set policies to maintain the value of their currency, their banks were at risk; bank failures were a damaging outcome of the depression, not its cause. Governments and central bankers – not commercial banks – led the way into depression in country after country.

This process is illustrated clearly in the United States' experience. Banks continued to fail as the government clung to the gold standard. For instance, in early 1933, the Reconstruction Finance Corporation refused to help a prominent Michigan bank holding company for reasons that are not clear (anticipating the failure of Lehman Brothers in 2008). States declared bank holidays, and the New York Federal Reserve Bank lost gold as investors speculated against the dollar. Franklin Delano Roosevelt took office in early March 1933; he immediately instituted what he called a federal bank holiday to protect the banking system from complete collapse.¹⁰

At the center of the financial panic that initiated the Great Recession were the banks and other private financial institutions that had accumulated large portfolios of mortgage-backed assets, in which mortgages were combined and then separated – at least in theory – into securities of differing risks. When the housing boom ended in 2006 and 2007, homeowners began to default on mortgages at an increasing rate. These defaults were not the random defaults assumed in the construction of mortgage-based securities, and investors could no longer distinguish between the various assets.

Efficiency-promoting securities were transformed into toxic assets as it became progressively harder to sell them. High leverage, initially a way to multiply earnings, became a company hazard as the price of assets fell.

Apparent in American and European financial markets by Summer 2007 were the deleterious effects of the inability to sell these toxic assets. Pressure continued during the fall, and the Fed lowered its discount rate by more than a percentage point between September 2007 and January 2008. (The National Bureau of Economic Research later concluded that a recession had started in December 2007.) Fed Chairman Ben Bernanke, Treasury Secretary Henry Paulson, and President of the New York Fed Timothy Geithner rescued the New York investment house Bear Stearns at the point of collapse with Fed funds and purchase by another investment house in March 2008. They took over the two quasi-governmental mortgage brokers, Fannie Mae and Freddie Mac, in August. Even at this late date, the Fed and other public figures argued that the pressure was largely limited to the housing sector and that the measures taken up to that point were sufficient to maintain financial health.

Bernanke and Paulson asserted after the fact that they tried in Summer 2008 to get Congress to take action to forestall a crisis. This effort proved futile for several reasons. For one, the financial leaders were making reassuring statements to the public at the same time they were appealing to Congress, a mixed message that did not lend persuasiveness to the arguments they presented. In addition, Congress was not convinced that the financial system was on the verge of a meltdown and was reluctant to act outside of an emergency. This reluctance may illustrate a general problem: it is hard to prepare for a hypothetical crisis

absent evidence that a crisis is indeed about to emerge. Certainly, if palliative action forestalls the putative crisis, people will ask what all the pressure was about. Congress is a large and unwieldy body; all these complexities precluded preventive action in Summer 2008.

In September, another investment house, Lehman Brothers, found itself unable to borrow. It tried selling assets to pay its obligations but could not sell its toxic assets and fell short of its needs. Creditors wanted to be paid and investors wanted to sell Lehman Brothers stock. Though an investment rather than a commercial bank, Lehman was in a process that resembled nothing so much as an old-fashioned banking panic. Bernanke, Paulson, and Geithner did not wish to repeat their rescue of Bear Stearns and therefore allowed the firm to fail on September 15, 2008. After the fact, none of these articulate policy-makers was able to tell a coherent story about why they had not avoided bankruptcy for the firm. The event reprised the government confusion that led Michigan banks to fail in early 1933, precipitating the bank holiday.

The financial triumvirate had tried to find a buyer for Lehman Brothers, as they had done for Bear Stearns, but was unable to do so. They apparently reverted to the gold standard mentality as expressed in the free-market ideology of the Washington Consensus: Lehman Brothers had taken large risks and now had to pay the penalty for losing too many bets. But hard on the heels of Lehman Brothers' failure came American International Group (AIG). Although it was not an investment bank, this multinational insurance company also had taken too many bets on what were now toxic assets and was about to collapse. The epidemic had escaped the mortgage market and infected the whole financial system. Nearly a year earlier, the global financial system had

entered into what Frederic Mishkin, a member of the Fed's Board of Governors, had called an "adverse feedback loop."¹¹ One failure induced another; a world-wide financial panic ensued.

Paulson, Bernanke, and Geithner threw in the towel and nationalized AIG. Their commitment to the free market had lasted one day; Congressman Barney Frank (D-Mass.) suggested we call it Free Market Day!¹² But while the sale of Bear Stearns had calmed the financial markets, the nationalization of AIG – arriving on the heels of Lehman Brothers' bankruptcy – only confused the market. The government had restated its ideals and then abandoned them in the twinkling of an eye. Investors could not predict what would come next.¹³

Barely functioning credit markets seized up completely. No one knew what the government policy was or if anyone was insured; no one wanted to purchase toxic assets. Economic activity and international trade came to a sudden halt. The brief reassertion of faith in the free market in 2008 was as counterproductive as fidelity to the gold standard had been in 1931. In both cases, the United States dragged the world down with it – doing so faster the second time than it had fifty years earlier.¹⁴

Fortunately, we are now in a Great Recession, not a repeat of the Great Depression. Ten percent unemployment and unemployment insurance compares favorably to 20 percent unemployment without a safety net. The primary reason for this divergence is the vagary of the American political cycle. Voters had to wait three years after the Great Depression began and a full year after the Fed turned a recession into a depression to vote on public policy; voters in 2008 had this opportunity just months after the financial crisis began. The similarity

between now and then is that it took a new group of leaders to change policy. The Obama administration has many holdovers – for example, Obama reappointed Bernanke as Fed Chairman – but there is no doubt that the theories underlying policy have changed since the last administration. An important difference between the past and current economic calamities is that because the present crisis is only a recession – not a depression – Obama does not have the opportunity for reform that Roosevelt did.

Roosevelt opened most banks quickly after their holiday; he took the United States off gold a month later. He introduced the National Industrial Recovery Act (NIRA) and the Agricultural Adjustment Act (AAA), pillars of the New Deal, shortly thereafter. These actions signaled a clear new direction in government policy, or what economists call a new policy regime. Investment rose and consumption began to recover; the long economic decline had ended.¹⁵

The continuation of high unemployment in the 1930s is commonly blamed on the high wages created by the NIRA and the subsequent growth of unions. This argument is inaccurate for several reasons. Economic growth progressed rapidly during Roosevelt's first term and may not have been able to occur any faster because of bottlenecks in the supply of raw materials and production. Faster growth, even if possible, likely would have led to inflation despite high unemployment.¹⁶ In fact, the recovery was so fast that both the Fed and the government decided to reverse policy and rein in demand through both monetary and fiscal policies. The result was the recession of 1937, which increased unemployment and delayed the return to full employment for several more years. It was policy, not gains by labor, that extended the Depression's length.

The growth of unions was only one result of the New Deal reforms. Not all of these reforms were consistent with each other, and not all of them lasted more than a few years. However, the enduring parts of the New Deal changed the economy in many ways. Labor and tax reforms preserved a stable income distribution in the economic expansion that followed World War II. Creation of the Food and Drug Administration helped expand the pharmaceutical industry that extended life for many people. Social Security improved the quality of life for many older people.

Reforms to the financial system produced a half-century free from financial crises.¹⁷ The Federal Deposit Insurance Corporation (FDIC) gave most people faith in the safety of their bank accounts. Deposit insurance was complemented by bank regulation to substitute for critical investors and depositors. The Glass-Steagall Act separated commercial and investment banks. The Federal Reserve System was restructured to empower its central office; the Securities and Exchange Commission (SEC) was created to regulate financial investments. Banking became a boring industry, and more people invested safely in the stock market. There was little excitement in the financial markets, and the economy grew rapidly and consistently after the war.

Nothing lasts forever, and prosperity generated a desire for more independent financial dealings. Economic turmoil in the 1970s hastened the transition, and the Washington Consensus arose in the 1980s. The Glass-Steagall Act was repealed, and the SEC's regulation relaxed. Americans urged the rest of the world to follow suit and deregulate both domestic and foreign capital movements. The distribution of income widened, the size of the financial sector rose, and a string of small-scale (at least to the United States) financial crises ensued.

This foreshadowing of our current problems was not seen as such at the time; even the failures of Long-Term Capital Management (LTCM) in 1998 and Enron in 2001 did not raise concerns. Most of the crises, like the Asian crises of 1997 (which spread to Russia, bringing down LTCM), were seen as problems of less developed countries, not mature economies like the United States. Economists and politicians alike pushed for less regulation at home and deregulation abroad. In particular, they sought to deregulate the international flow of capital and hailed the Washington Consensus as the way forward for all countries, developing and developed. Like Irving Fisher, a great economist of the early twentieth century who predicted continued prosperity just before the Great Depression, they too readily believed in the reigning economic model.¹⁸

Even Bernanke, chairman of the Federal Reserve and student of the Great Depression, did not see chaos ahead during most of 2008. Bernanke, to his credit, realized what was happening by the start of 2009. He resolved not to let the Fed duplicate its mistakes of the early 1930s, standing by as banks failed and supporting the gold standard instead of the domestic economy. He pulled all the strings – some of them on the outer edge of his authority – to loosen monetary policy and encourage economic activity. It was a bravura performance, but monetary policy lost its effectiveness as banks ran for cover even after the financial panic subsided. The banks used the Fed's services to rebuild their depleted reserves as the value of toxic assets went to zero, and they loaned only to the safest of customers.¹⁹

Obama, even before he took office, urged Congress to pass a stimulus bill – to create a fiscal expansion in addition to the hobbled monetary expansion.

Republican congressmen insisted he divert part of the stimulus to tax cuts, which went into savings as individuals – like banks – tried to build up their depleted reserves, limiting the size of the stimulus. This fidelity to the Washington Consensus reduced Obama’s ability to moderate the recession’s effects on ordinary people.

Expansive monetary and fiscal policies were effective enough to preclude a repetition of the Great Depression, and support for reforms on the order of the New Deal ebbed. Obama had campaigned on a program of bipartisan cooperation, and although he tried to bring Republicans along with his policies, they had not abandoned their belief in the Washington Consensus. Banks, moreover – newly prosperous from the government bailouts – resisted increased regulation. When Obama put extending health care to all Americans before reforming the financial system, resurgent banks blunted the impact of financial reforms.

There are two lessons to be drawn from this comparison. The first is that the open American economy is prone to collapse every once in a while. Favorable conditions – the New Deal and a vigorous post-war expansion – can eliminate “great” economic contractions for a generation

or so, but American exuberance appears to chafe under these conditions. As the memory of past economic difficulties fades, economic and political pressure for change rises to the fore. International economic imbalances are condoned until they have to be corrected, often painfully.

The second lesson is that there are strong pressures for unregulated capitalism that only abate in the face of sharp economic downturns like the Great Depression. We avoided another Great Depression by luck – the election cycle – and skill. Marx was correct when he argued that tragic history repeats itself as farce: we now have the oxymoronic Great Recession after all the fears of Great Depression II. Keynes was right, too; discredited economic theories – and the gold standard mentality – continue to dominate the actions of even “practical” men and women.²⁰ Recent policy initiatives have done little to reduce the underlying risk of another financial crisis. As of this writing, Jean-Claude Trichet, president of the European Central Bank, gave a striking illustration of the continuing gold standard mentality, calling for worldwide fiscal austerity in the early stages of a tentative recovery from our recent crisis.²¹

ENDNOTES

¹ John Maynard Keynes, *The General Theory of Employment, Income and Money* (New York: Harcourt, Brace, 1936), chap. 24, “Concluding Notes.”

² Karl Marx, *The Eighteenth Brumaire of Louis Bonaparte* (New York: International Publishers, 1964), 1.

³ Peter Temin, *Lessons from the Great Depression* (Cambridge, Mass.: MIT Press, 1989); Barry Eichengreen, *Golden Fetters: The Gold Standard and the Great Depression, 1919 – 1939* (New York: Oxford University Press, 1992).

⁴ John Williamson, “What Washington Means by Policy Reform,” in *Latin American Adjustment: How Much Has Happened?* ed. John Williamson (Washington, D.C.: Institute for International Economics, 1990).

- ⁵ Peter Temin, “Real Business Cycle Views of the Great Depression and Recent Events: A Review of Timothy J. Kehoe and Edward C. Prescott’s *Great Depressions of the Twentieth Century*,” *Journal of Economic Literature* 46 (September 2008): 669 – 684; Peter Temin, “Corrigendum,” *Journal of Economic Literature* 47 (March 2009): 3.
- ⁶ Hyman P. Minsky, *Can “It” Happen Again? Essays on Instability and Finance* (Armonk, N.Y.: M. E. Sharpe, 1982).
- ⁷ Carmen M. Reinhart and Kenneth S. Rogoff, *This Time Is Different: Eight Centuries of Financial Folly* (Princeton, N.J.: Princeton University Press, 2009).
- ⁸ Thomas Ferguson and Peter Temin, “Made in Germany: The German Currency Crisis of 1931,” *Research in Economic History* 21 (2003): 1 – 53; Peter Temin, “The German Crisis of 1931: Evidence and Tradition,” *Cliometrica* 2 (April 2008): 5 – 17.
- ⁹ Richard S. Grossman, “The Shoe That Didn’t Drop: Explaining Banking Stability During the Great Depression,” *Journal of Economic History* 54 (September 1994): 654 – 682.
- ¹⁰ Barrie A. Wigmore, *The Crash and Its Aftermath: A History of Securities Markets in the United States, 1929 – 33* (Westport, Conn.: Greenwood Press, 1985), 433 – 447.
- ¹¹ Frederic S. Mishkin, speech at the Risk USA Conference, New York, November 5, 2007, <http://www.federalreserve.gov/newsevents/speech/mishkin20071105a.htm>.
- ¹² David Wessel, *In Fed We Trust: Ben Bernanke’s War on the Great Panic* (New York: Crown Business, 2009), 26.
- ¹³ Thomas Ferguson and Robert Johnson, “Too Big to Bail: The ‘Paulson Put,’ Presidential Politics, and the Global Financial Meltdown,” parts 1 and 2, *International Journal of Political Economy* 38 (Spring 2009): 3 – 34 and 38 (Summer 2009): 5 – 45.
- ¹⁴ Christina D. Romer, “Back from the Brink,” speech to the Federal Reserve Bank of Chicago, September 24, 2009, http://www.whitehouse.gov/assets/documents/Back_from_the_Brink2.pdf.
- ¹⁵ Peter Temin and Barrie A. Wigmore, “The End of One Big Deflation,” *Explorations in Economic History* (October 1990): 483 – 502.
- ¹⁶ Christina D. Romer, “Why Did Prices Rise in the 1930s?” *Journal of Economic History* 59 (March 1999): 167 – 199.
- ¹⁷ Reinhart and Rogoff, *This Time Is Different*.
- ¹⁸ Irving Fisher, “Fisher Sees Stocks Permanently High,” *The New York Times*, October 16, 1929.
- ¹⁹ Richard C. Koo, *The Holy Grail of Macroeconomics: Lessons from Japan’s Great Recession* (Singapore: John Wiley, 2008).
- ²⁰ Paul Krugman, “How Did Economists Get It So Wrong?” *The New York Times*, September 6, 2009; Paul Krugman, “Misguided Monetary Mentalities,” *The New York Times*, October 12, 2009.
- ²¹ Jean-Claude Trichet, “Stimulate No More – It is Now Time for All to Tighten,” *Financial Times*, July 23, 2010.

From the writings of Alexander Hamilton

Editors' Note: Alexander Hamilton (1757 – 1804) was appointed the first Secretary of the Treasury of the United States in 1789. A major figure in the shaping of the new Republic, Hamilton fought in the Revolutionary War, eventually becoming Lieutenant Colonel in General Washington's Continental Army; served as a member of the Continental Congress; helped found the Bank of New York; and, with James Madison and John Jay, authored the Federalist Papers. He was elected a Fellow of the American Academy of Arts and Sciences in 1791.

In this first selection, from April 30, 1781, Hamilton writes to Robert Morris, a wealthy Pennsylvania merchant and an influential statesman who developed a plan for the first incorporated national bank, the Bank of North America. Hamilton, only twenty-four at the time, makes one of the earliest statements by anyone that banks and credit can strengthen a state and foster economic growth. These goals would become two of Hamilton's main objectives as a statesman.

The tendency of a national bank is to increase public and private credit. The former gives power to the state for the protection of its rights and interests, and the latter facilitates and extends the operations of commerce among individuals. Industry is increased, commodities are multiplied, agriculture and manufactures flourish, and herein consist the true wealth and prosperity of a state.

Most commercial nations have found it necessary to institute banks and they

have proved to be the happiest engines that ever were invented for advancing trade. Venice, Genoa, Hamburg, Holland and England are examples of their utility. They owe their riches, commerce and the figure they have made at different periods in a great degree to this source. Great Britain is indebted for the immense efforts she has been able to make in so many illustrious and successful wars essentially to that vast fabric of credit raised on this foundation. 'Tis by this alone she now menaces our independence.

She has indeed abused the advantage and now stands on a precipice. Her example should both persuade and warn us. 'Tis in republics where banks are most easily established and supported and where they are least liable to abuse. Our situation will not expose us to frequent wars, and the public will have no temptation to overstrain its credit.

Alexander Hamilton on the financial crisis & economic policy

The second selection is from Hamilton's December 1790 "Report on a National Bank." Now Treasury Secretary, Hamilton acknowledges that banks and borrowers can make mistakes, but that they are, on balance, beneficial.

That Banks furnish temptations to overtrading . . . must mean that by affording additional aids to mercantile enterprise, they induce the merchant sometimes to adventure beyond the prudent and salutary point. But the very statement of the thing shows that the subject of the charge is an occasional ill, incident to a general good.

[...]

If the abuses of a beneficial thing are to determine its condemnation, there is scarcely a source of public prosperity which will not be speedily closed. In

every case, the evil is to be compared with the good; and in the present case such a comparison will issue in this, that the new and increased energies derived to commercial enterprise from the aid of banks are a source of general profit and advantage, which greatly outweigh the partial ills of the overtrading of a few individuals at particular times, or of numbers in particular conjunctures.

Selections excerpted from "The Papers of Alexander Hamilton," vol. I – XXVII, ed. Harold C. Syrett (New York: Columbia University Press, 1961 – 1988), vol. II, 618 and vol. VII, 314 – 315, respectively. Special thanks to Richard Sylla, of New York University, for his assistance in choosing these selections and preparing the notes to accompany them.

Poem by Alice Notley

The Codex Eats Me

© 2010 by Alice Notley; reprinted with permission from the author

The dogs tell me it could be worse: I could be eaten by
oversubtlety rather than bold red or blue letters howling,

HOWLING! That word again. The dogs open their mouths
to word me.

Have you ever dreamed you didn't have a master. We dreamed
we ran down the gully to the river, but not without you; we
couldn't leave you. I entered this desert long past the middle

of my life, knowing that I could only have what I wanted
on paper. Everything's covered with dog hairs; shake them
off the illuminated pages, no they're painted on, ocher, gold,
and black filaments. I stepped outside the city of Paris,
and there was sun on water in clear air – surely a dream.
The past loved you, though it didn't know you, but

it projected itself towards my melancholy. Have I betrayed
the past?

Electricity and laughter, taunting, are inscribed on this page.
The old house full of creepy scholars, walking all over
colorful decades, squashing them with large, ignorant feet.
Can't you paint over their faces? Here is another

D for dog – a dog undermines anyone, whether it's the
dog of death, or the dog of dreams: more oscillating
hairs, white, shaded with grey and purple; or thinly gilded.
D persecutes you with its sweet disposition. Yet I dreamed
that a mean man announced, in veiled irritation – for
the dog had been part of his power – *The dog is dead. A new
dog will arise. Who volunteers to take charge of it?*
A healer raised her hand; but so did others. Who will be assigned

to this office? Who can call Death to order, now that the planet
has lost its blessing? No one.

Contributors

Lucian A. Bebchuk, a Fellow of the American Academy since 2000, is the William J. Friedman and Alicia Townsend Friedman Professor of Law, Economics, and Finance and Director of the Program on Corporate Governance at Harvard Law School. He is a Research Associate of the National Bureau of Economic Research and an Inaugural Fellow of the European Corporate Governance Institute. His publications include *Pay without Performance: The Unfulfilled Promise of Executive Compensation* (with Jesse Fried, 2004). Although Bebchuk served as a consultant to the Treasury Department Special Master for TARP Executive Compensation, the views expressed in his article should not be attributed to the Office of the Special Master.

Barry Eichengreen, a Fellow of the American Academy since 1997, is George C. Pardee and Helen N. Pardee Professor of Economics and Political Science at the University of California, Berkeley; Research Associate of the National Bureau of Economic Research; and Research Fellow of the Centre for Economic Policy Research. His book, *Exorbitant Privilege: The Decline of the Dollar and the Future of the International Monetary System*, will be published by Oxford University Press in January 2011.

Benjamin M. Friedman, a Fellow of the American Academy since 2009, is the William Joseph Maier Professor of Political Economy at Harvard University. He is author or editor of eleven books, including *The Moral Consequences of Economic Growth* (2005) and *Day of Reckoning: The Consequences of American Economic Policy Under Reagan and After* (1988), which received Columbia University's George S. Eccles Prize. He is a Director of the Council for Economic Education and the Encyclopedia Britannica.

Edward L. Glaeser, a Fellow of the American Academy since 2010, is the Fred and Eleanor Glimp Professor of Economics at Harvard University, where he also serves as Director of the Taubman Center for State and Local Government and the Rappaport Institute for Greater Boston. His recent publications include *Agglomeration Economics* (2010), *Housing Markets and the Economy: Risk, Regulation, and Policy* (with John M. Quigley, 2009), and *Rethinking Federal Housing Policy: How to Make Housing Plentiful and Affordable* (with Joseph Gyourko, 2008).

C.A.E. Goodhart is Professor Emeritus of Banking and Finance at the London School of Economics (LSE) and Program Director of Regulation & Financial Stability at the Financial Markets Group Research Centre at LSE. He served as an external member on the Bank of England's Monetary Policy Committee from 1997 to 2000 and was Chief Adviser at the Bank of England prior to coming to LSE. His publications include *Monetary Theory and Practice* (1984), *The Central Bank and The Financial System* (1995), and *The Evolution of Central Banks* (revised edition, 1988).

Robert E. Hall, a Fellow of the American Academy since 1985, is the Robert and Carole McNeil Joint Professor of Economics at Stanford University and Senior Fellow at the Hoover Institution. For 2010, he is President of the American Economic Association. He is the author of *Economics: Principles and Applications* (with Mark Lieberman; 5th edition, 2009). Other recent publications include *Forward-Looking Decision Making* (2010) and “The Burden of the Nondiversifiable Risk of Entrepreneurship” (with Susan Woodward), *American Economic Review* (2010).

Nolan McCarty, a Fellow of the American Academy since 2010, is the Susan Dod Brown Professor of Politics and Public Affairs and Associate Dean at the Woodrow Wilson School of Public and International Affairs at Princeton University. His publications include *Political Game Theory* (with Adam Meirowitz, 2006), *Polarized America: The Dance of Ideology and Unequal Riches* (with Keith T. Poole and Howard Rosenthal, 2006), and *The Realignment of National Politics and the Income Distribution* (with Keith T. Poole and Howard Rosenthal, 1997).

Alice Notley is a poet who has published more than twenty collections of verse. Her recent works include *Reason and Other Women* (2010), *In the Pines* (2007), *Grave of Light: New and Selected Poems, 1970 – 2005* (2006), *Disobedience* (2001), and *Mysteries of Small Houses* (1998). She is the recipient of the Lenore Marshall Poetry Prize from the Academy of American Poets, the International Griffin Poetry Prize, the Los Angeles Times Book Award for Poetry, the Academy Award in Literature from the American Academy of Arts and Letters, and the Poetry Society of America’s Shelley Memorial Award. “The Codex Eats Me” is from the forthcoming collection *Culture of One* (Penguin, March 2011).

Keith T. Poole, a Fellow of the American Academy since 2006, is the Philip H. Alston, Jr. Distinguished Chair in the Department of Political Science, School of Public and International Affairs, at the University of Georgia. His publications include *Ideology & Congress* (with Howard Rosenthal; second revised edition, 2007), *Polarized America: The Dance of Ideology and Unequal Riches* (with Nolan McCarty and Howard Rosenthal, 2006), and *Spatial Models of Parliamentary Voting* (2005).

Thomas Romer is Professor of Politics and Public Affairs and Director of the Research Program in Political Economy at the Woodrow Wilson School of Public and International Affairs at Princeton University. His publications include “An Empirical Investigation of the Dynamics of PAC Contributions” (with James Snyder), *American Journal of Political Science* (1994); “Interjurisdictional Sorting and Majority Rule: An Empirical Analysis” (with Dennis Epple and Holger Sieg), *Econometrica* (2001); and “Parties in Elections, Parties in Government, and Partisan Bias” (with K. Krehbiel and A. Meirowitz), *Political Analysis* (2005).

Howard Rosenthal, a Fellow of the American Academy since 1992, is Professor of Politics at New York University. He edited *What Do We Owe Each Other?: Rights and Obligations in Contemporary American Society* (with David J. Rothman, 2008) and is the author, most recently, of *Ideology & Congress* (with Keith T. Poole; second re-

vised edition, 2007) and *Polarized America: The Dance of Ideology and Unequal Riches* (with Nolan McCarty and Keith T. Poole, 2006).

Robert M. Solow, a Fellow of the American Academy since 1956, is Institute Professor Emeritus at the Massachusetts Institute of Technology. In 1987, he was awarded the Nobel Prize in Economics in recognition of his contributions to the theory of economic growth; in 1999 he received the National Medal of Science for his creation of the modern framework for analyzing the effects of investment and technological progress on economic growth. His publications include *Capital Theory and the Rate of Return* (1963); *The Nature and Sources of Unemployment in the United States* (1964); *Growth Theory* (1970); and *Inflation, Unemployment and Monetary Policy* (with John B. Taylor, 1998).

Jeremy C. Stein, a Fellow of the American Academy since 2008, is the Moise Y. Safra Professor of Economics at Harvard University. His recent publications include “Rethinking Capital Regulation” (with Anil Kashyap and Raghuram Rajan) in *Maintaining Stability in a Changing Financial System* (2008) and “A Gap-Filling Theory of Corporate Debt Maturity Choice” (with Robin Greenwood and Samuel Hanson), *Journal of Finance* (2010).

Peter Temin, a Fellow of the American Academy since 1986, is Elisha Gray II Professor Emeritus of Economics at the Massachusetts Institute of Technology. His books include *Did Monetary Forces Cause the Great Depression?* (1976), *Lessons from the Great Depression* (1989), *Elites, Minorities, and Economic Growth* (edited with Elise S. Brezis, 1999), and *The World Economy Between the Wars* (with Charles Feinstein and Gianni Toniolo, 2008). He served as President of the Eastern Economic Association (2001–2002) and the Economic History Association (1995–1996).

Luigi Zingales is the Robert C. McCormack Professor of Entrepreneurship and Finance and the David G. Booth Faculty Fellow at the Booth School of Business at the University of Chicago. He is the author of *Saving Capitalism from the Capitalists: Unleashing the Power of Financial Markets to Create Wealth and Spread Opportunity* (with Raghuram G. Rajan, 2003). He codeveloped the Financial Trust Index, designed to monitor the level of trust Americans have in the financial system.

AMERICAN ACADEMY
OF ARTS & SCIENCES

Chair of the Board
Louis W. Cabot

President & CEO,
William T. Golden Chair
Leslie Berlowitz

Treasurer
John S. Reed

Secretary
Jerrold Meinwald

Editor
Steven Marcus

Vice Chair, Midwest
John Katzenellenbogen

Vice Chair, West
Jesse H. Choper

Inside back cover : Unemployed men wait in long lines for bread and handouts in New York during the Great Depression. Photograph © Bettmann/Corbis.

Job-seekers wait in line for a job fair at a hotel in New York's Times Square, March 5, 2009. The job fair was sponsored by employment website Monster.com as part of its "Keep America Working" tour. Thousands attended the event featuring more than ninety employers. Photograph © REUTERS/Mike Segar.



coming up in Dædalus:

the meaning of
minority/majority

Gerald Early, Henry Louis Gates, Jr., Glenda R. Carpio, David A. Hollinger, Jeffrey B. Ferguson, Hua Hsu, Daniel Geary, Farah Griffin, Korina Jocson, Eric Sundquist, Waldo Martin, Werner Sollors, James Alan McPherson, Jeffrey B. Perry, Clarence Walker, Wilson Jeremiah Moses, Tommie Shelby, Amina Gautier, and others

race, inequality
& culture

Lawrence D. Bobo, William Julius Wilson, Michael Klarman, Rogers Smith, Douglas Massey, Jennifer Hochschild, Martha Biondi, Cathy Cohen, James Heckman, Taeku Lee, Pap Ndiaye, Alford Young, Marcyliena Morgan, Richard Nisbett, Jennifer Richeson, Daniel Sabbagh, Roger Waldinger, and others

the modern American
military

David Kennedy, Lawrence Freedman, David Segal, Lawrence Korb, Robert L. Goldich, Danielle Allen, Andrew Bacevich, James Sheehan, Brian Linn, Deborah Avant, Renée de Nevers, Errol Morris, Thomas Mahnken, Jonathan Shay, Charles J. Dunlap, Eugene Fidell, Martha McSally, William J. Perry, and others

plus protecting the Internet as a public commons, public opinion &c.

AMERICAN ACADEMY
OF ARTS & SCIENCES
Cherishing Knowledge · Shaping the Future