

SEPTEMBER 3, 2014 | NUMBER 754

Math Gone Mad

Regulatory Risk Modeling by the Federal Reserve

BY KEVIN DOWD

EXECUTIVE SUMMARY

The U.S. financial system faces a major, growing, and much under-appreciated threat from the Federal Reserve's risk modeling agenda—the “Fed stress tests.” These were intended to make the financial system safe but instead create the potential for a new systemic financial crisis.

The principal purpose of these models is to determine banks' regulatory capital requirements—the capital “buffers” to be set aside so banks can withstand adverse events and remain solvent.

Risk models are subject to a number of major weaknesses. They are usually based on poor assumptions and inadequate data, are vulnerable to gaming and often blind to major risks. They have difficulty handling market instability and tend to generate risk forecasts that fall as true risks build up. Most of all, they are based on the naïve belief that markets are mathematizable.

The Fed's regulatory stress tests are subject to all these problems and more. They:

- ignore well-established weaknesses in risk modeling

- and violate the core principles of good stress testing;
- are overly prescriptive and suppress innovation and diversity in bank risk management; in so doing, they expose the whole financial system to the weaknesses in the Fed's models and greatly increase systemic risk;
- impose a huge and growing regulatory burden;
- are undermined by political factors;
- fail to address major risks identified by independent experts; and
- fail to embody lessons to be learned from the failures of other regulatory stress tests.

The solution to these problems is legislation to prohibit risk modeling by financial regulators and establish a simple, conservative capital standard for banks based on reliable capital ratios instead of unreliable models.

The idea that the Fed, with no credible track record at forecasting, can be entrusted with the task of telling banks how to forecast their own financial risks, displacing banks' own risk systems in the process, is the ultimate in fatal conceits. Unless Congress intervenes, the United States is heading for a new systemic banking crisis.

Kevin Dowd is professor of finance and economics at Durham University in the United Kingdom and a partner in Cobden Partners based in London. He thanks John Allison, Roger Brown, Mark Calabria, Jim Dorn, Martin Hutchinson, Gordon Kerr, Jason Kuznicki, Harry Markowitz, Yalman Onaran, Antonio Pancorbo, Kevin E. Villani, Peter J. Wallison, Christopher Whalen, Margaret Woods, Basil Zafiriou and various anonymous readers for their helpful inputs. This Policy Analysis also involved interviews with U.S. bankers and regulators from a number of institutions; their participation is greatly appreciated. Any errors are the author's own.

“Faulty financial models were a big contributor to the recent financial crisis.”

INTRODUCTION

In the last two decades or so, there have been major problems with financial modeling, not least because faulty financial models were a big contributor to the recent financial crisis. A continuing threat comes from the regulatory risk modeling agenda—at the center of which is the use of the Federal Reserve’s mandated stress tests to determine banks’ regulatory capital requirements. The Fed is following a perfect recipe for a new systemic meltdown:

- It is stamping out the diversity in and competition among risk management practices of U.S. banks, on which financial stability depends;
- It is exposing the system as a whole to the blind spots in the Fed’s own risk models; and
- It is overwhelming the banks with a rapidly growing regulatory burden that is not just undermining recovery—greatly hindering the rebuilding of banks’ balance sheets and the resumption of bank lending to its normal levels—but crushing the financial system with ever more expensive and counterproductive regulations.

The Fed’s regulatory policy is part of the enormous growth in financial regulation generally. The Wall Street Reform and Consumer Financial Protection Act, best known as Dodd-Frank, is merely the culmination of a long process of ever-increasing regulatory pressure—which constitutes nothing less than a massive sustained regulatory onslaught against the banking system, greatly undermining its ability to function and service the real economy.

John Allison, former chairman and CEO of BB&T, notes that, “I don’t know a single time when federal regulators—primarily the FDIC—actually identified a significant bank failure in advance. Regulators are always the last ones to the party after everyone in the market (the other bankers) know something is going on. Thus, in that context, regulators

have a 100 percent failure rate.”¹

The Federal Reserve is not much better at peering into the future or even into the recent past. An example was the nonexistent “productivity miracle” in the 1990s. Puzzled by unexpectedly low inflation in the early to mid-1990s, Federal Reserve chairman Alan Greenspan had persuaded himself that the explanation must be that productivity was higher than estimated—and, indeed, it turned out that he was half-right. He then seized upon encouraging hints in the data that seemed to confirm his suspicions. As he announced to the Senate Banking Committee on July 22, 1997, in his usual cryptic way: “important pieces of information, while just suggestive at this point, could be read as indicating basic improvements in the long-term efficiency of our economy.”² The press seized on his comments as confirming the “productivity miracle” that everyone had been hoping for. It seemed plausible, too, not least because of the huge potential of the Internet and the fact that productivity figures for 1995–97 were unusually high. However, later revisions to the data were to show that these productivity gains had been spurious. What Greenspan overlooked was that lower-than-expected inflation had been due to a surge in consumer goods from China especially. There had been a productivity miracle, but it had occurred in Asia; the U.S. “productivity miracle” had just been noise in the data.

Federal Reserve officials were also caught off guard in the run-up to and early part of the global financial crisis. Consider the following transcript from a CNBC interview on July 1, 2005:

Interviewer: Tell me, what is the worst-case scenario, sir? We have so many economists coming on our air saying ‘Oh, this is a bubble, and it’s going to burst, and this is going to be a real issue for the economy.’ Some say it could even cause a recession at some point. What is the worst-case scenario if in fact we were to see [house] prices come down substantially across the country?

Ben Bernanke: Well, I guess I don't buy your premise. It's a pretty unlikely possibility. We've never had a decline in house prices on a nationwide basis. So, what I think what is more likely is that house prices will slow, maybe stabilize, might slow consumption spending a bit. I don't think it's gonna drive the economy too far from its full employment path, though.³

Now consider a few of Bernanke's other predictions:⁴

- **March 6, 2007.** At a bankers' conference in Honolulu, with subprime delinquencies on the rise, he anticipates that subprime will be no worse than prime: "The credit risks associated with an affordable-housing portfolio need not be any greater than mortgage portfolios generally . . ." Yet there was every reason to think they would be—that's what the "sub" in "subprime" means—and they were.
- **May 17, 2007.** At the annual Chicago Fed Conference on Bank Structure and Competition, he anticipated few spillovers from subprime: ". . . we believe the effect of the troubles in the subprime sector on the broader housing market will likely be limited, and we do not expect significant spillovers from the subprime market to the rest of the economy or to the financial system. The vast majority of mortgages, including even subprime mortgages, continue to perform well."⁶ The problem was that they didn't for much longer: subprime portfolios started to plunge a couple of months later.
- **January 10, 2008.** In response to a question after a speech in Washington, D.C., when the United States was already in recession: "The Federal Reserve is not currently forecasting a recession."⁷
- **February 27, 2008.** In testimony before the Senate Banking Committee, a

month before the Bear Stearns failure and seven months before the systemic meltdown of September 2008, he said: "I expect there will be some failures [among smaller regional banks]. . . . Among the largest banks, the capital ratios remain good and I don't anticipate any serious problems of that sort among the large, internationally active banks that make up a very substantial part of our banking system."⁸

- **July 16, 2008.** In testimony before the House Financial Services Committee, he said: "[Fannie Mae and Freddie Mac are] adequately capitalized. They are in no danger of failing . . ." Fannie and Freddie between them owned or guaranteed over \$5 trillion in mortgages and were placed into receivership less than two months later precisely because of fears of their default.⁹

I don't wish to single out Bernanke for particular criticism; most other officials were saying similar things, but he was the most senior over most of this period, and even he is no better able to predict the future than most of the rest of us.

The failure of the Fed's own forecasting models has not prevented it from dictating how the banks it regulates should construct their own risk-forecasting models. The idea that the Fed, with no credible track record at forecasting, can now be entrusted with the task of telling the banks how to forecast their own risks, displacing banks' own risk systems in the process, is the ultimate in fatal conceits.

WHY REGULATORY RISK MODELING DOES NOT WORK

A risk model is a computer algorithm that projects possible future financial outcomes and perhaps their associated probabilities. These models are used to manage risk—to guide investment decisions and give a sense of potential exposure to future losses, but their most important use is to determine capital re-

“The failure of the Fed's own forecasting models has not prevented it from dictating how the banks it regulates should construct their own risk-forecasting models.”

“If the risk model is wrong, the consequences can be serious: in particular, if risk estimates are too low, then the bank will be undercapitalized and more vulnerable to failure.”

quirements. For example, a bank might set out a target probability of insolvency and then use the risk model to determine how high its capital should be to achieve this target safety level. If the risk model is wrong, the consequences can be serious: in particular, if risk estimates are too low, then the bank will be undercapitalized and more vulnerable to failure.

It is helpful to think of a risk model as a black box based on calibrated data that spews out loss risk forecasts or loss projections, usually known as risk measures. Note the main elements:

- The black box or model itself;
- The model’s input, the data used to calibrate it; and
- The model’s output, the risk measure.

The Model

The first is the model itself. There are different types of risk models, but the most important for our purposes are probabilistic models and stress test models. The former forecast possible loss outcomes and their associated probabilities, while the latter seek to tell us the loss we would experience if a given scenario were to occur.

The big question is: How do we know if our model is correct? We don’t. We therefore have to rely on assumptions. Typically, those are derived from data on past losses on a bank portfolio. Analysts then try to find a statistical distribution that best fits the data. One then faces three major problems:¹⁰

- There will be many distributions that seem to fit the data, but no one will be able to confidently identify the “true” or “best” distribution.
- Most of the data will be “nickel and dime” loss data, with relatively little data on big losses, which means that the extremes or tails of the distribution will be particularly difficult to determine—and this is most unfortunate because it is the fitting of the high-losses data that we particularly want to get right.¹¹

- There is little statistical theory to help us select the “right” distribution.

Most risk modelers ignore these problems and assume the easiest and most commonly used distribution of all—namely the Gaussian or normal distribution. Unfortunately, this is almost the worst possible distribution to pick, as it provides a very bad fit to the tail data¹² and grossly underestimates how common and how big the very high losses can be.¹³ One might add that, to the extent statistical theory helps us here, it explicitly warns us *not* to choose this distribution for high-loss events.¹⁴

Model Calibration

The related problems of how to get good estimates of the model’s parameters and the lack of data on which to estimate them are most apparent in the calibration of subprime mortgage models in the years running up to the crisis. So what would we have needed, and what did we do?

- We would have needed estimates of default probabilities, but these instruments were new, and analysts only had a very short historical sample from which to estimate them. So they usually assumed something low, for example, 4 percent.
- We also would have needed estimates of the correlations between default probabilities, and these are even harder to estimate. In fact, analysts had so little data that they had no choice but to guess the correlation values. They went on to make the simplest assumption—that the correlations are zero—and they ignored the fact that they were targeting the same low-income people who they knew couldn’t pay off these mortgages, and who would soon default in droves.
- Finally, we would have needed data to help estimate the risk of a housing downturn, but analysts didn’t have any such data because the U.S. housing market had gone up and up, in nominal terms, since

at least the early 1930s. So they ignored this risk, as it wasn't in their data set.

I emphasize that these assumptions would all have been plausible at the time, as none was inconsistent with the historical data.

Suppose we now take 100 such mortgages, tranche them into 10 tranches ranked by seniority and sell them off to investors: the most junior tranche bears any losses until it is wiped out, then the second most junior tranche bears any further losses until it is wiped out, and so forth, so the more junior tranches "protect" the more senior ones. The probability that the most junior tranche loses all its investment is at most 0.68 percent. The probability the next most junior tranche loses all its investment is no more than 0.000003 percent—effectively zero—and the probabilities of the more senior tranches losing their investments are smaller still. It is no wonder that the originators of these securities described them as "super-senior," safer than U.S. government debt. They could then easily sell them off as AAA-rated securities. This was a perfect example of financial alchemy, in which a bunch of garbage is converted into instruments golder than gold.

Fool's gold, that is. Essentially, subprime mortgages are unsustainable. Once enough of them default, almost all of them do, and most of the investors are wiped out. But then the excuses begin: Everyone knew that subprime mortgages were a real risk—after all, who in their right minds would lend money to people who can't afford to repay their mortgages? Bankers might claim they were only lending to them to make commission and pass off the mortgages to mugs daft enough to buy them up. Investors, they might say, have only themselves to blame if they didn't do their due diligence properly. And who gives a damn about all those fools who took out subprime mortgages, and never mind that we targeted and often tricked them into signing up? Anyway, it is not our problem.

We all know the rest of the story. Wall Street built a doomsday machine and used it to manufacture the subprime catastrophe.

The models they used were badly built and poorly calibrated. They didn't just underestimate, but totally ignored, the biggest risk in the market—the risk of a market downturn. But this didn't matter. The simplest measure of sanity in the housing market is the ratio of median home price to income. Traditionally, the national average was 3:1, and bankers have long used the rule of thumb that a mortgage should not exceed 3 times the borrower's income. By late 2004, however, the ratio of national median house price to median income had risen to over 4:1 and began to cause some concern. Yet even this statistic masked two huge problems. First, there was a lot of variation across the country. In Miami it had risen to 7:1, and in Los Angeles it was nearly 10:1. Traditional bankers would have choked off such increases by refusing to continue lending. Second, one had to consider the quality of the borrowers—that is to say, they were often the worst possible. Traditional bankers would never have loaned to them in the first place. It was therefore obvious that the market was in a massive bubble and there would be a wave of defaults. Yet none of the models picked this up—and they *couldn't*, because it was not in the data sample on which they were calibrated.

This disaster was created not just under the noses of the regulators, but with active regulatory and governmental encouragement. Politicians and the media tut-tutted afterward about all the greed driving the subprime market, but they conveniently overlooked the fact that the government itself had started the subprime wagon going with its affordable-housing agenda. They also overlooked that the engine underlying subprime, mortgage-backed securitization was only made possible because of the GSEs, Fannie and Freddie. The crisis could not have occurred had the market been left free of government interference.¹⁵

The Risk Measure

The third element of the model is the risk measure. The most popular risk measure is the Value-at-Risk (or VaR)—in particular, the 99 percent VaR that purportedly gives us the

“Disaster was created not just under the noses of the regulators, but with active regulatory and governmental encouragement.”

“All that extra sophistication—the fortune invested in fancy risk modeling and all that that entails—had a *negative* value-added relative to much simpler approaches that would have cost virtually nothing.”

worst loss we can expect on the best 99 days out of 100. Its purpose is to provide the risk manager with some “risk comfort” by quantifying downside risks. However, the VaR is a very problematic risk measure:

- It tells us nothing about what we might lose on the remaining “bad” day out of 100: the VaR gives us the cutoff to the tail, but tells us nothing about what might happen in the tail itself. The VaR’s tail blindness is, to say the least, an unfortunate property in risk management, where it is the tail that matters.
- The VaR is very gameable: traders and other risk takers are very adept at gaming VaR-based risk management by “stuffing risk into the tail” where the model fails to detect it, exposing their banks to hidden risks that the VaR model can’t pick up.
- The combination of VaR and Gaussianity is particularly dangerous: as we have seen, the first is blind to the tail, and the second leaves the user much more exposed than they might think. The combination of the two is reckless and violates the principle of prudence that should permeate modern risk management, but which is now notable by its almost complete absence. It is therefore especially unfortunate that the Gaussian VaR model is still the most popular risk model in use.

Each of these weaknesses (and more¹⁶) has been known since the early days of VaR in the mid-1990s. The continued popularity of the VaR, and especially of the Gaussian VaR, would suggest that these serve the interests of those who want low risk numbers—as opposed to those who want effective risk management.

Besides these theoretical issues with the VaR, there is abundant evidence that VaR models don’t work well in practice either.

Back in 2002, a study by Berkowitz and O’Brien analyzed the performance of the daily VaR models of six large U.S. banks with large trading portfolios from January 1998 through

March 2000.¹⁷ Their results were striking: they found that the banks’ VaR forecasts worked well on most days, that is, when markets were fairly quiet, but during periods when the markets were volatile they often performed very badly, generating losses well beyond the VaR. The authors attributed the models’ poor performance in volatile market conditions to the difficulties of structural modeling when portfolios are large and complex. Large trading portfolios can have tens of thousands of individual positions. The models also employ many approximations to reduce computational burdens, have many imprecisely estimated parameters, and are undermined by regulatory constraints; these factors make them unwieldy and inaccurate, and also make it difficult for them to adjust to sudden large increases in market volatility.¹⁸

As a comparison, the authors also examined the performance of much simpler reduced-form Generalized Autoregressive Conditionally Heteroskedastic (GARCH) models that merely “fit” the observed profit-and-loss series. We can think of these as the modern equivalent of a back-of-the-envelope calculation: they can easily be fitted by a single modeler with a desktop computer and a standard econometrics package. In principle, the banks’ sophisticated models should (easily) be able to beat these GARCH models because the latter merely fit the aggregate data and cannot account for changes in positions, changes in underlying relationships, and so forth. However, they found that the opposite was the case. In particular, the GARCH models were able to adjust to increases in market volatility much better: they were much more nimble when it mattered. The bottom line: all that extra sophistication—the fortune invested in fancy risk modeling and all that that entails—had a *negative* value-added relative to much simpler approaches that would have cost virtually nothing—and the VaR models don’t work when we need them.

There are also many other cases where banks experienced losses *well in excess* of what their VaR models were indicating. Consider the following, out of many other possible examples:

“During the financial crisis, VaR models performed particularly badly on the trading book.”

- Long Term Capital Management’s risk model estimated its daily 99 percent VaR in August 1998 to be about \$30 million, implying that there was a 1 in 100 chance of a loss in excess of that amount, and yet LTCM’s *average* daily losses in that month were about *three times* that figure.
- At the end of 2011, the Structured Credit Portfolio at JP Morgan Chase in London had a value of around €157 billion. Its daily VaR was about \$120 million—suspiciously low. It was subsequently revised downwards to about half that. However, on a number of days the portfolio suffered losses well in excess of the VaR: \$319 million on March 30, 2012, and \$415 million on April 10. The eventual losses were \$6.2 billion, owing to a trader now known as the London Whale. Reported daily VaR turned out to be less than 1 percent of the actual losses from the positions to which it related.¹⁹
- During the financial crisis, VaR models performed particularly badly on the trading book, where they were most sophisticated: calculations performed by the Bank of England indicated that for the four biggest UK banks, cumulative trading losses over the height of the crisis were up to *six times* the value of the model-determined capital set aside to cover against trading losses.²⁰

Nor was this problem confined to the UK banks: almost all the big banks had VaR models that underestimated their market risk exposures to an almost unbelievable degree. Consider Merrill Lynch & Co., which was by no means atypical. By the end of 2007, Merrill had a market risk capital charge, based on its VaR models, of 0.23 percent of trading assets, implying that those assets were very safe indeed. Its trading book leverage, however, was about 427 percent, which would indicate the exact opposite. The ratio of its trading book to its total book was about 23 percent, indicating that the trading book amounted to a substantial proportion of its total expo-

sure, but the market risk capital charge was only 1.72 percent of its equity. One gets similar figures for the other big banks. Yet Merrill’s super-senior subprime portfolio—a delightful contradiction in terms, but really the posh end of its subprime portfolio—at the end of June 2007 was valued at \$32 billion and went belly-up soon after: by January 2008, Merrill had lost \$24.5 billion on its subprime exposures. As Pablo Triana Porela and Vasiliki Kosmidou conclude:

... By being unrealistically very low, VaR excused and afforded the accumulation of leverage and toxic assets that sank some of the world’s leading international banks, helping unleash a global financial and economic meltdown.

In sum, a VaR-based architecture led to a gargantuan leverage on huge trading portfolios filled with volatile and very toxic assets. . . . By not discriminating based on an asset’s fundamental nature [that is, by not taking account of its true risks!], VaR can categorize obviously risky assets as risk-lite, excusing and affording their vast accumulation in a very economical way capitalwise.²¹

Then there was Citi. By the end of 2007, Citi had a market risk capital charge of 0.30 percent of trading assets and a trading book leverage of about 336 percent. At the same time, the ratio of its trading book to its total book was over 24 percent, and its market risk capital charge was 1.41 percent of its equity. Triana and Kosmidou comment that Citi’s

internal models trading book regulatory architecture permitted vast undercapitalization and gearing. To add insult to injury, Citigroup was, by far, one of the firms to hold [the most] toxic subprime mortgage securities. If anyone needed to be soberly capitalized, it certainly was Citigroup; and yet the rules allowed it to run its market risks with *almost no cushioning capital*. [My italics]

“Citi also provides a uniquely interesting case study of how VaR modeling can go badly wrong.”

By early 2008, Citi had lost \$22.1 billion in sub-prime exposures.

Citi also provides a uniquely interesting case study of how VaR modeling can go badly wrong.²² Around 2003, the market and liquidity risk team at the Board of Governors began to collect daily profit-and-loss data and VaR data to help them identify trends and even perhaps get ahead of some of the risks. They were particularly concerned about Citi. Their concerns were well grounded: unknown to them, Citi's new CEO, Chuck Prince, was already in the process of weakening the bank's internal governance to facilitate greater risk-taking. The Board approached the New York Fed for these data—and remember that the New York Fed is the acknowledged leader amongst the Federal Reserve Banks, with a unique expertise in capital markets—but could not obtain them. It turned out later that the New York Fed didn't have the data to provide: all it had were three-month-old reports photocopied from originals provided to the Office of the Comptroller of the Currency; although the New York Fed had the authority to require the data from Citi, it hadn't bothered to do so. The Board team began to realize that the New York Fed seemed to lack both the expertise and the skepticism even to ask the right questions. A story then came out that a bank supervised by the New York Fed had lost between \$60 and \$80 million on the nascent market for carbon-emission credits. Although not a large loss, this loss was worrying because up to the moment it occurred, the bank's VaR had been only \$1 million: it had been taken for granted that the market was much more stable than it really was, and no one had challenged the model. This episode was a signal of more widespread modeling problems that could have been picked up by a rigorous inspection, but still the New York Fed refused. “Don't worry about that,” one of their examiners is reported to have said, “We just have to respond to these things when they happen. We can't get ahead of these problems.” They were too busy and didn't have enough staff.

Over the next few years, the market and liquidity risk team struggled to get the New

York Fed to tighten its supervision of Citi's modeling but met with little success. In the meantime, Citi added nearly a trillion dollars to its balance sheet and built up a \$43 billion super-senior collateralized debt obligation (CDO) exposure on its trading book—an exposure that was then *omitted* from the risk model. As Nick Dunbar explained,

Both the New York Fed and its watchdogs in Washington, D.C., failed to spot a fundamental breach of the thin blue line they created: recording the super-senior CDO as trading exposure and interrogating the bank's VaR model. A senior Federal Reserve official who is still angry about that screw-up says, “They didn't put them in their VaR [model]. . . . I mean this is just basic. You do not need to be a quant to catch this. They were supposed to mark to market. But the attitude seemed to be, ‘Why bother? They don't change in value. They're AAA.’ . . . You can stress test your heart out [but if] its not in the VaR [model], you're not going to get anything on it.”

The state of denial went all the way to the top. In an interview on July 9, 2007, Prince acknowledged there were problems on the horizon but insisted that Citi was still OK: “When the music stops, in terms of liquidity, things will be complicated. But as long as the music is playing, you've got to get up and dance. We're still dancing,” he told journalists. But not for much longer: the musical chairs stopped suddenly on October 31, when analyst Meredith Whitney publicly warned that Citi would go bust if it went ahead with its planned dividend distribution. Confidence in Citi then plunged: positions were rapidly written down and Chuck was chucked out of his job a few days later. As Keynes once said: “A sound banker, alas! is not one who foresees danger and avoids it, but one who, when he is ruined, is ruined in a conventional and orthodox way along with his fellows, so that no one can really blame him.”²³

Returning to the models, let me put it this way: all the risk models turned out to be woefully inadequate when the crisis occurred. In fact, I am not aware of a *single* model anywhere that correctly predicted the crisis. Those observers who *did* predict the crisis did so using old-fashioned methods—they identified the build-up looking at charts and key ratios signaling unsustainability. It is hard to conceive of a more total failure on the part of the models. Yet despite this most abject of track records, the conventional wisdom is as addicted to models as it ever was.

Market Instability

A greatly underrated problem with much financial modeling is the maintained (and often unquestioned) “scientific” belief that quantitative methods from the natural sciences, particularly physics, can be applied mechanically to social and economic problems. Hayek aptly observed in his Nobel acceptance speech that,

... the confidence in the unlimited power of science is only too often based on a false belief that the scientific method consists in the application of a ready-made technique, or in imitating the form rather than the substance of scientific procedure, as if one needed only to follow some cooking recipes to solve all social problems. It sometimes almost seems as if the techniques of science were more easily learnt than the thinking that shows us what the problems are and how to approach them.²⁴

This belief is naive for a number of reasons, but one of the most obvious is that the processes governing the operation of financial markets (and more generally, any social system) are not immutable “laws” comparable to the laws of physics.²⁵ Any social system is changing all the time. In social systems, including financial markets, time-invariant phenomena, if they exist at all, are the exception rather than the rule.²⁶

The comparative absence of stable laws is, in part, due to the fact that the broader environ-

ment in which markets operate is itself always changing. Pricing relationships fluctuate with supply and demand conditions, for instance, and there is a great danger of identifying spurious but superficially plausible patterns that are little more than accidental and have no serious predictive value. Other relationships are merely the temporary consequences of unusual conditions or the results of policy.

A recurring case is a fixed exchange rate that then suddenly changes. The problems involved here are nicely illustrated by the carry trade strategy on a fixed exchange rate, in which one goes long the currency with the higher interest rate and short the currency with the lower interest rate, so earning the difference. The less-intelligent market operators jump into the trade regarding it as risk-free, whereas the more intelligent ones appreciate that it is risky. However, it is ever so tempting, and the longer it goes on, the harder it is to resist. As one investor put it:

After ten successful years, everyone is doing the Thai baht carry trade. Why? Because even though you think it might be a risky trade, all your friends are getting rich doing it, and after a while it becomes difficult to resist the pull. You don’t want to be the only person at the hedge fund cocktail party who is not doing the trade *du jour*. Plus, the statistics show that it’s a risk-free trade. After eight years, it’s an immutable fact—Thailand doesn’t devalue. So you begin to look like a person who is not scientific—you’re a victim of your own unfounded insecurity, a man of the past. All your friends are getting wealthy. Why don’t you, too, take on these risk-free trades?²⁷

The risk models are useless in this context as they can’t pick up the riskiness of a variable that doesn’t change. Then the reckoning comes: one day, out of the blue, the exchange rate *does* change, and the underlying risk is revealed the hard way. In this particular case, it was by the Thai devaluation in 1997, which

“Those observers who *did* predict the crisis did so using old-fashioned methods—they identified the build-up looking at charts and key ratios signaling unsustainability.”

“The generalized use of risk models is itself destabilizing, as it lures the market as a whole into taking more risks when market participants think they are taking less.”

wiped out the hapless carry-traders and triggered the East Asia crisis. Those risk-free trades were not so risk-free after all! American banks have repeatedly got burnt the same way with the peso.

Yet one feature that one can confidently identify in financial markets is the apparently random oscillation between “normal” periods in which markets are stable and “crisis” periods in which markets are volatile. Most of the time, markets are fairly stable: volatilities and correlations are low, pricing relationships are steady, markets are liquid, credit is both cheap and easily available, and returns are good. However, once in a while, a crisis occurs and all the above phenomena disappear: volatilities rise, correlations radicalize, relationships break down, credit and liquidity dry up, risk management strategies unravel that had previously worked well, and financial institutions suffer large losses. Financial markets have fluctuated between these alternate states since their inception, and one cannot predict what will happen in the one state from what happened in the other. A good analogy is with fluid dynamics: markets generally follow a pattern of streamlined flow, obeying one set of equations with only local instances of turbulence where those equations break down, but in extreme circumstances, such as those of 2007–09, the turbulence spreads throughout markets, causing a general breakdown of relationships that had hitherto worked well.

This seesaw between alternating periods of stability and crisis causes havoc for risk modeling. Imagine a stable period, such as the Great Moderation. As the boom goes on, the models indicate that risks are decreasing, when in reality the bust is approaching, and true risks are actually increasing: the risk model gives an increasingly false risk comfort and lulls the risk manager to sleep as the danger approaches. The system looks strongest precisely when it is most vulnerable. In fact, just before the recent crisis, people were wondering where all the risk had gone: no one could find it. The crisis then hit, and the risk models were shown to have been useless just when they were most

needed. Had there been no model, on the other hand, the risk manager would have had to rely on other indicators—a sense of history, judgment, or rules of thumb—and would have been better off when the boom burst. In this sense, one can even say that the generalized use of risk models is itself destabilizing, as it lures the market as a whole into taking more risks when market participants think they are taking less.

There are also the usual psychological effects of a prolonged boom as past crises fade increasingly into the past, new people come to prominence who have little or no memory of past crises, and the “this time is different” mentality increasingly takes hold—all factors that are ignored by the models.²⁸ Even if the models *had* pointed to potential vulnerabilities, it would have been very difficult for participants to take them seriously, as hubris was at its peak and prudence at its low. As Claudio Borio, Mathias Drehmann, and Kostas Tsatsaronis pointed out: Who would have taken seriously a stress projection that the spread between interbank rates and the overnight index swap rate would rise to more than 300 basis points, as happened during the crisis, when it had generally fluctuated between 10 and 15 basis points?²⁹ The result is that only a few spot the warning signs, and they do so not by using risk models, but by a combination of charts, common sense, and a sense of market history.

In addition, the phenomena measured in physics typically do not change with the measurement itself or with the ways that the observer uses those measurements. The well-known exception, the Heisenberg Uncertainty Principle, is a feature of subatomic particle physics, but it does not affect the many problems where Newtonian physics gives good answers. By contrast, the financial equivalent of the Heisenberg principle is much more prevalent. The act of modeling a financial process over time—such as the movement of a stock price—will often lead observers to react in ways that affect the process itself by adopting a particular trading or risk management strategy. If enough people adopt the same strategy,

however, then that strategy will likely affect the dynamics of the stock price itself.

We then have a trading strategy that looks good on paper and even works well most of the time, when markets are stable. It then fails catastrophically when the market is stressed. The irony here is that it lets you down when you need it most: it protects you against the 2- and 3-sigma events that don't matter, at the price of leaving you wide open to the high-sigma events that can destroy you. And you have no idea of your true exposure because the model doesn't capture it. If you thought risk management was about managing risks, think again.

The same is the case with *any* risk management strategy. Say you have a VaR-based risk management strategy that calls on you to reduce your VaR by selling risky positions in the event of a crisis. This makes sense when you are the only person implementing this strategy and everyone else carries on as before. If everyone adopts the same strategy, however, then everyone will try to sell in a crisis. This collective reaction will itself exacerbate the fall in prices and create a positive feedback loop in which the crisis grows as it feeds off itself. Some initial trigger leads prices to fall and VaRs to rise. The increased VaRs then generate sales as risk managers struggle to get their VaRs back down, and the new sales cause further price falls and even higher VaRs. The collective attempt to get individual VaRs down destabilizes the market and inflicts the high losses that the risk management strategy was meant to avert. The problem is not with the risk management strategy as such—it does not matter whether this strategy is good, bad, or indifferent—but with the fact that everyone is doing it.

Remember that all this risk management stuff was meant to make the financial system more stable, not less. Market stability requires that players have *different* strategies so that some are willing to buy when others wish to sell. Thus the key to market stability is not some magic risk management strategy—those don't exist—but the presence of those willing to take contrary positions, preferably Warren

Buffetts with deep pockets and the patience to await their opportunity.

From this perspective, the *worst thing* that the regulators can do is endorse any particular model or risk management strategy as a regulatory standard for everyone to follow. Whatever its merits, the mere fact of one strategy being commonly adopted will undermine the stability of financial markets and leave those who use it more exposed than they imagine. Once again, you get the appearance of risk management (and scientific risk management, too!) but the reality of much greater risk exposure, all the more dangerous because it is hidden. Whether navigating risks or navigating icebergs, the last thing anyone needs is false risk comfort. But that is exactly what the models provide and the regulators endorse.

Implicit in much of the above are also the psychological impacts of model use. There is always a danger of a “black box” mentality: people start to think that a model—especially a sophisticated model—must cover everything; they then start to accept its results uncritically and lose sight of its weaknesses. The model crowds out critical thinking; people become overconfident and overextend themselves. In an institutional context, the model gets taken for granted and becomes difficult to challenge—especially when the models have been endorsed by the regulators and everyone is using much the same models anyway.

One should also remember that much market instability arises from the erratic monetary policies of the Federal Reserve itself. These policies are illustrated in Figure 1, which shows the Fed funds rate since the 1950s. By my count, there are 10 notable interest rate peaks. All but one—that of the mid '90s—were followed by sharp falls. Among the highlights were the massive, necessary-but-painful Volcker interest rate hike starting October 1979, which left much of the banking system insolvent in the early 1980s; the doubling of interest rates over 1994, which led to a wave of defaults (Orange County, etc.); Greenspan's warnings of “irrational exuberance” in 1996 followed by monetary *easing*, which stoked the tech bubble that burst

“The *worst thing* that the regulators can do is endorse any particular model or risk management strategy as a regulatory standard for everyone to follow.”

“On the one hand, the Fed endorses—in fact, requires—the use of risk models, but on the other, it undermines them by its own erratic monetary policies.”

Figure 1
Effective Federal Funds Rate (FEDFUNDS)



Source: Board of Governors of the Federal Reserve System (2014), <http://www.research.stlouisfed.org>.

in 2001; and after interest rates peaked again, another major volte-face occurred, in which interest rates became negative in real terms and stoked the subprime mortgage market. By 2007, interest rates had climbed again to just over 5 percent, but they were brought down to virtually zero in 2009 and have remained there since, well below inflation for nearly six years. If the past is anything to go by, these rates are stoking the mother of all booms and the mother of all busts as well.

So, on the one hand, the Fed endorses—in fact, requires—the use of risk models, but on the other, it undermines them by its own erratic monetary policies: the models cannot pick up the Fed’s sharp and unpredictable twists and turns. In fact, even the Fed itself can’t predict its own erratic twists and turns. President Clinton’s spin doctor James Carville is reputed to have said that “I used to think that if there was reincarnation, I wanted to come back as the President or the Pope or as a .400 baseball hitter. But now I would like to come back as the bond market. You can intimidate everybody.” At the time most people interpreted his comments as reflecting the power of the bond market vigilantes, especially their

presumed ability to “manipulate” government financing costs and hold the government to ransom. What this narrative overlooks is that these “vigilantes” were simply the market reaction to the volatile policies of the Federal Reserve—after all, it is the Fed that ultimately drives interest rates through its control of the money press, and the market merely reacts: the instability that everyone feared was created by the Fed itself. Mr. Carville should ask to be reincarnated as the Fed chairman.

Risk Modeling as Superstition

We must see risk modeling for what it really is. Given that it doesn’t work and is also highly counterproductive, perhaps the most charitable view one can take is to see it as an example of Coasean blackboard economics: policies that might work superbly on the blackboard, but which fail dismally in practice when confronted with the real world.

But perhaps the most helpful insights into the phenomenon of risk modeling come not from economics, but from a very different discipline: anthropology. From this perspective, most risk models, and regulatory risk models in particular, are textbook examples of the rit-

ualistic fetishes usually associated with primitive tribes. A fetish can be described as irrational attachment to an object—in this case, a risk model—regardless of its true usefulness. This is to treat the models as if they were ritual implements with magical properties and is the very essence of superstition. To quote the book *Alchemists of Loss*:

We overrate what we think we know. Nowhere is this more evident than in the fanciful belief systems we set up to comfort ourselves. The fertile minds of the ancient Egyptians imagined the sun god Ra in his daily cycle across the heavens; the Pharaohs performed daily devotions to Amun to preserve their precious harmonious order, *ma'at*, from the ever present threats of the evil god Set who would bring chaos in his wake. For his part, the modern risk manager tries to keep Set's financial equivalent at bay by building a VaR model. The difference is that the Pharaoh's daily devotions generally worked—ancient Egypt only experienced disaster every few hundred years—whereas the inadequacies of VaR were already very apparent early on, when the VaR on Nick Leeson's positions at Barings in February 1995 indicated zero risk two days before those same positions caused the bank to go bust, and in 1998, when LTCM's VaR model indicated that the firm was perfectly safe just weeks before it spiraled into collapse.³⁰

When we get the equivalent of a harvest failure, our response is not to reflect on the causes of that failure or see our quantitative risk management practices as the useless superstitious rituals that they are, but to do more of what we did before, only on a grander scale. Throughout the process, whatever the outcome, we continue to fool ourselves that what we are doing is scientific when it is anything but.

The bottom line is that markets are simply not mathematizable. To believe otherwise is to succumb to superstition.

The Failure of Risk Management

Going further, we can say that it is not just risk modeling that has failed, but risk management (and especially, quantitative risk management) generally. We can look at this failure from any of three different perspectives and still get the same answer.³¹

There is first the historical evidence, and what jumps out from the evidence is how the growth of risk management from the early 1990s onwards coincides with ever greater risk management problems that culminated in the meltdowns of 2007–2009. Call it the Great Risk Management Paradox: the growth of “Risk Management” coincides with the increasing instability and proneness to fail of the institutions practicing it—that is, with the very outcome that Risk Management was supposed to prevent—not exactly a ringing endorsement. Not only has risk management failed, but risk management is itself a major cause of the risk management failures it is meant to avert.

The failures of risk management can be explained from an engineering perspective. From this perspective, there are two principles at work: the first is the gap between theory and practice, as exemplified in the old engineering motto: in theory there is no difference between theory and practice, but in practice there is. The second is the old adage that any system is only as good as its weakest link.

In the case of risk management, theorists wrote down models without regard to the practicalities of implementation, with the ultimate (and unobtainable) goal of integrated risk management across the whole enterprise. Unfortunately, it all falls down in practice. Up to a point, we can integrate different types of market risk, such as equities and interest-rate risks. Integrating these with credit risks is more difficult, and integrating these with operational risks much more difficult still. Even then, the operational risks we can handle are the more trivial ones—such as the risks of running out of paperclips—rather than the ones that matter, such as the risks of uncontrolled rogue trading. Repeated scandals in recent years show that rogue-trading risk is still a ma-

“What jumps out from the evidence is how the growth of risk management from the early 1990s onwards coincides with ever greater risk management problems that culminated in the meltdowns of 2007–2009.”

“The net result was the appearance, but not the reality, of risk management—and many senior bankers were then *surprised* to discover that, when they really needed it, it wasn’t there.”

major problem that many institutions have failed to manage. Moving further out, we haven’t even begun to integrate these with other risks faced by a bank, such as its strategic risks, and it doesn’t help that the risk management and strategy functions typically live in mutually incomprehensible parallel universes. Most important of all, there are the broad governance and control risks, of which the main one is out-of-control senior executives—that is, the failure of corporate governance at its most fundamental level.

The net result was the appearance, but not the reality, of risk management—and many senior bankers were then *surprised* to discover that, when they really needed it, it wasn’t there.

The failure of risk management can also be explained from basic economics. There is a natural underlying tension between risk takers and risk managers: they are natural enemies. The former regard risk-taking as driving their bonuses: if they take risks, and those risks succeed, they get large bonuses; and if those risks fail, well, it’s only other people’s money, and they can always get another job. Heads they win, tails someone else loses. They therefore have an incentive to take excessive risks, and the job of the risk manager is to rein them in. However, in this warfare, traders are very good at undermining risk management in two fundamental ways.

First, they seek to find and exploit any weaknesses in the risk management systems set up to control them. They are good at this partly because they are incentivized to do so—rather like prisoners wishing to escape from prison, who are incentivized to study the system and find its weaknesses. They are also good at this because they are often vastly better paid: trading and similar activities such as financial engineering therefore attract the best talent. This is why so many of the top PhDs being produced in quantitative physics and similar subjects switch over to finance, where their skills can be almost incredibly lucrative—a career choice all the more attractive when the alternative (and I speak from experience) is a lifetime of penurious grind in academia. As

a result, the true risks faced by a financial institution will always be greater than the measured risks—and no one will know what the true risks really are. This is yet another reason why risk modeling does not work.

Second, traders and senior managers are natural allies. This is because it is the risk-taking that drives the bonus pool, and senior managers face similar asymmetric incentives as other risk takers down the line: in their case, if risk-taking succeeds, they get bigger bonuses (not to mention more stock options, greater pay based on measures of short-term performance, and so on), and if risk-taking fails, they can expect to be bailed out or, at worst, pensioned off in luxury. Traders then often find a sympathetic ear when they complain about risk managers preventing them from making profits—and everyone wants more profits. The risk manager’s problem is particularly stark when the main risk-taker is the CEO, and there are many reported cases of risk managers who pushed their luck too far and were then out of a job.

As Roger J. Brown aptly put it, “the risk manager is always unwelcome. When things are going well, he is the skunk at the garden party. When things are going badly, it is too late.”³²

These economics are reflected in the risk modeling process itself. If risk managers produce high-risk numbers, then these will lead to high regulatory capital charges. That would mean a drag on profits, so naturally there is pressure down the line—it makes no matter whether this is explicit or implicit—to generate acceptable risk numbers. In effect, risk modeling is often used not to manage risks, but to manage risk regulation.

REGULATORY RISK MODELING: A BRIEF OVERVIEW

We now consider a few generic problems with any form of regulatory risk modeling.

Standardization

One problem is that of standardization. In its literature on regulatory stress testing, the

Fed makes much of the fact that its approach allows for “consistency,” “horizontal comparisons” across the banks, and so on: in essence, a form of one-size-fits-all. Such an approach ignores the fact that banks are very heterogeneous. For example, the better regional banks are very different from the big zombies on state life support, such as Citi. It obviously makes no sense to run a risk analysis of the former using Citi’s loss data, or vice versa, or even to run a risk analysis of either using an average of their or the industry’s loss experience. Yet the regulatory process pushes for standardization and leads to exactly the latter outcome, giving results that are irrelevant to the specifics of the institutions concerned. Similarly, it makes no sense to apply a single stress-test scenario to such different banks, because they have very different business models and very different risk exposures, not least of which are very different geographical exposures.³³

There is a bigger problem. Any system of regulatory capital modeling inevitably means that the regulators have a preferred model of their own and then pressure regulated institutions to adopt similar models. Regulated banks’ risk models will then converge to something similar to the Fed’s own models. The end result is that banks will have much the same models and much the same risk management strategies. They will therefore take much the same risks and make much the same mistakes—dramatically magnifying systemic risk. This is exactly what happened with subprime.

The convergence to a single Fed-approved model standard also means that any weaknesses in the Fed’s models will be transmitted across the whole system, further increasing systemic risk—and all risk models have weaknesses. Indeed, there is already evidence that banks *are* becoming more focused on trying to mimic the Fed’s results to pass the Fed’s stress tests rather than projecting the risks they think most appropriate to their own institutions, with the result that innovation and diversity in risk modeling and management are being ground down across the system. As Til Schuermann, a former regulator and a noted

risk modeling expert, wrote in a recent article in the *Wall Street Journal*:

As the Fed’s models have become more and more important in deciding the fate of the biggest banks, those banks have focused more and more on trying to mimic the Fed’s results instead of tracing out their own risk profiles. This poses a real risk. . . .

The incentives to get close to the Fed’s numbers are powerful enough to stifle genuine creativity, imagination and innovation by risk managers and their modelers. Deviating from industry practice is now increasingly viewed with suspicion and often discouraged by bank regulators. . . .

But if everybody uses the same scenario (which they do) and works hard to get the same numbers (and they are trying), then we have a very narrowly specialized risk machine that is inflexible and unresponsive to unexpected shocks. That is, shocks that weren’t previously subject to a stress test.

The danger is the financial system and its regulators are moving to a narrow risk-model gene pool that is highly vulnerable to the next financial virus. By discouraging innovation in risk models, we risk sowing the seeds of our next systemic crisis.³⁴

Even if the Fed’s models were theoretically perfect, we would still have a major systemic risk, precisely because everyone would be doing the same thing: any risk management standard—especially a standardized regulatory risk model—is wrong on principle. To achieve systemic stability, it is essential that institutions have different risk management strategies and the freedom to choose them.

So should the Fed allow banks to use their own risk models for regulatory capital purposes? Unfortunately, no—or at least not for the zombie banks on state support. If the Fed is to impose a single capital regime, however—and it

“Any system of regulatory capital modeling inevitably means that the regulators have a preferred model of their own and then pressure regulated institutions to adopt similar models.”

“Because their corporate governance is broken and they exist on government support, senior executives of these banks have a clear incentive to produce the lowest risk numbers possible to allow distributions of capital to shareholders and, of course, to senior management.”

insists on doing so—then one that allowed too much freedom would leave the system open to gaming, and the big zombies are masters at this game. Because their corporate governance is broken and they exist on government support, senior executives of these banks have a clear incentive to produce the lowest risk numbers possible to allow distributions of capital to shareholders and, of course, to senior management. They are not concerned with building up their banks long-term, but with getting as much as possible out of them short term. Allowing these banks to use their own models to determine their regulatory capital charges is tantamount to allowing them to set their own capital requirements, which in the current context gives them license to loot their banks at the taxpayers’ expense.

The Fed has then cornered itself: given that it insists on a single regime to apply to both the sound banks and the zombies, it then has no choice but to prevent this looting by imposing controls on the risk models used to set regulatory capital charges—that is, the Fed *has* to impose standardization in some form. Yet we know that standardization is wrong too.

The Fed is then caught in a bind of its own making: too much freedom, and the zombie banks especially get to loot the system; too little, and we have a host of other problems—most notably the imposition of a hugely inefficient straightjacket on the banking system and the creation of a lot of systemic risk.

Procyclicality

A second problem is that of procyclicality. Ideally, capital requirements should be countercyclical: they should rise as the economy booms—and so help to counter the boom as it approaches its peak—and fall as the economy goes into recession. The problem is that all existing regulatory capital systems produce procyclical capital requirements, falling as the economy booms and rising in the recession. Thus, the regulatory capital requirements themselves serve to exacerbate the instability they are meant to counter. The root problem here is that the metrics on which capital

requirements are based all lead to procyclical capital ratios: the risk metrics, ratings, and model-based risk estimates indicate falling risk as the economy booms, and hence produce procyclicality; bank assets rise with the boom, on the other hand, but as they appear in the denominator of the capital ratios, they produce procyclicality as well. Despite various efforts³⁵ and a lot of ink expended on the subject, you cannot get countercyclical capital requirements out of metrics that can only sensibly produce procyclical requirements.³⁶ It is as simple as that.³⁷

We also have to take account of procyclicality in the regulatory process itself. One has to remember that the seeds of the crisis are sown during the good times that precede it. The regulators themselves have no special insights into the coming bust: in fact, they don’t really understand the business of banking anyway. Even if a regulator suspects a bank is taking excessive risks, there is nothing they can do. They only have a suspicion to work on and, in any case, the regulatory-sanctioned risk models are saying there is no problem. Bankers are well-connected and well-placed to fight off regulatory intervention. So even the best regulator will settle for an easy life and do nothing about it.

When the crisis comes, the regulatory process goes into reverse. There is then a lot of political heat to get tough with the banks, and the incentive of the regulator is to tighten up lending standards, even on banks that are well run. In this case, there is little cost to the regulator in overreacting, but a potentially a big cost in underreacting: the regulator does not want the blame if their bank subsequently gets into trouble. Unnecessarily high lending standards then put businesses out of business that would otherwise have survived. And when regulators take over a bank, they often botch the job because they don’t know how to fix banks or run them.

These considerations also prompt the usual glib regulatory “solution” that we are repeatedly told would achieve countercyclical capital requirements, that is, that we find some sup-

posedly “wise” regulator who would see the turning points in advance and adjust capital requirements accordingly, and have the courage as well as the foresight to be able run directly counter to prevailing opinion, not to mention the strength to prevail against the howls of anguish that such measures would provoke. If only such beings existed! Such “solutions” are literally a *deus ex machina*. Unfortunately, they do not work in the real world: far from being ahead of the market, regulators are typically well behind it. Thus, any proposed “solution” that relies on regulatory discretion—regulators applying discretionary countercyclical requirements and so on—is not only doomed to fail, but offends the intelligence of those to whom it is directed. That said, let me offer a compromise: if proponents of this solution can find their Superman, or even give us a credible plan to find him, then I am all in favor of handing the reins over to him.

In short, regulators have neither the knowledge, the skills, the models, nor the incentive to act in countercyclical ways. Instead, they simply exacerbate the fluctuations, and this pattern repeats itself again and again: it happened in the early 80s, the early 90s, and especially in the recent crisis. The pattern is no accident, but reflects the incentives that the regulators face. In any case, no regulator or regulatory system has the means to determine how much risk-taking there should be in the first place. Only a market can determine that, and only then in the absence of government distortions.³⁸

The Basel System of International Bank Capital Regulation

The original Basel treaty (now known as Basel I) was promulgated in 1988. It was 30 pages long and intended to provide a standardized regulatory capital regime for banks in the developed economies. At its heart was the notion of risk-weighted assets. Each asset would be bucketed into one of five risk classes and be assigned the risk weight corresponding to that class: commercial loans would be given a risk weight of 100 percent, residential mortgages

one of 50 percent, Organisation for Economic Co-operation and Development (OECD) government debt 0 percent, and so on. Capital requirements were then couched in terms of risk-weighted assets and required only a few calculations.³⁹ However, in the early 1990s there was increasing dissatisfaction with the bluntness of these risk weights, and Basel I was perceived as lacking risk sensitivity relative to the new wave of risk models then emerging.

After much lobbying by the banks, a Market Risk Amendment was agreed in 1996, which, for the first time, allowed banks to use their internal models to determine their regulatory capital requirements against market risk. The principle of risk-based capital regulation had now been conceded. A near decade-long process of renegotiation then followed, and a new treaty, Basel II, was agreed in 2004. Basel II allowed the (big) banks to use their internal models to determine their capital requirements against their credit and operational risks, while also setting out complex rules for derivatives positions. By design, Basel II served as an incentive for banks to upgrade their risk management technology and an explicit objective was to allow the bigger banks to obtain lower capital charges by using their risk models.⁴⁰ The move to more complicated regulation based on the banks’ own models was strongly promoted by the big banks themselves as it gave them more scope to “play the system”—indeed, the regulatory system itself was captured by them. Reflecting these changes, Basel II weighed in at 347 pages—an order of magnitude longer than its predecessor.

As an aside, note that an underlying (and unchallenged) principle here was that a risk model was a substitute for bank capital—and, in fact, a *very* good substitute, as a “good” risk model approved by Basel could allow a bank to dispense with most of its capital. This suited the banks, which could then cut down their capital and increase distributions via higher bonuses and dividend payments. Instead, the thinking was that everything would be fine if we just had more and better and more sophisticated models—and never mind that *actual*

“Any proposed ‘solution’ that relies on regulatory discretion is not only doomed to fail, but offends the intelligence of those to whom it is directed.”

“The parameter space of a large bank’s banking and trading books together can run to a considerable number of millions.”

capital levels were falling dangerously low. The focus on (so-called) risk-weighted capital was also helpful here, as it served to distract attention from the actual decline in capital as the banks were simultaneously loading up on risk. It is truly astonishing to note that this principle *still* reigns unchallenged in 2014, despite the fact that the manifestly obvious failure of the risk models has blown its evidentiary base right out of the water and broken much of the banking system—one thinks of Citi’s trading VaR model that had allowed it to operate on a wafer-thin capital cushion but overlooked the occasional \$43 billion in toxic subprime exposure.

Yet the length of the rulebook greatly understates its complexity. The move away from broad asset classes to internal models and individual loan exposures produced an explosion in the number of estimated risk weights, from single digits under Basel I to several million in recent years. The number of parameters for a big bank’s banking book might run to tens of thousands—three or four orders of magnitude greater than under Basel I. For the trading book, it is worse: a large bank might have several thousand risk factors in its VaR model, and the covariance matrix for all these risk factors might have several million parameters. Then there are the pricing models used to map from these risk factors to the valuations of individual instruments, each with more parameters of their own. Taken together, the parameter space of a large bank’s banking and trading books together can run to a considerable number of millions. These parameters have to be estimated from very limited samples—typically, less than 20–30 years for a credit model and less than 5 years for a market model—or simply guessed. Samples this short would encompass at most one extreme event, and perhaps not even that. Parameter estimates obtained from them are not only highly imprecise, but potentially very misleading, and—as we saw with the earlier subprime portfolio example—a recipe for trouble.

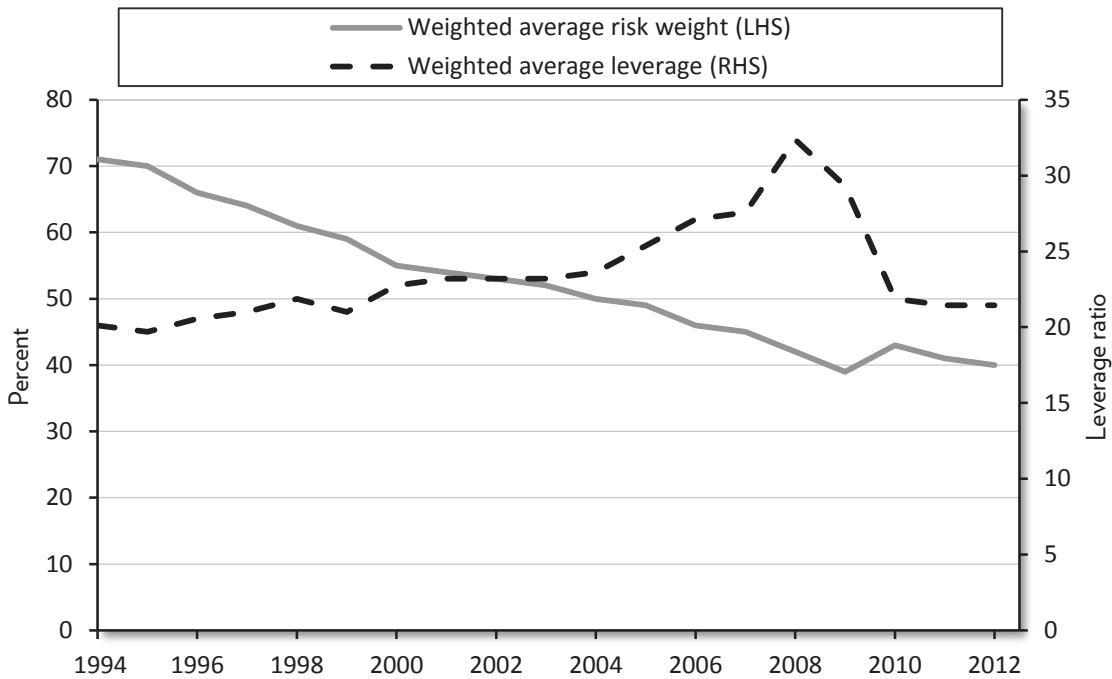
These models also involved an enormous amount of model risk—the prospect of differ-

ent models giving different answers to the same problem. The extent of this risk is illustrated by the results of some hypothetical portfolio exercises undertaken by UK regulators.⁴¹ The results indicated a huge range of variation, with estimated risk weights differing by factors of between three and five and estimated default probabilities differing by factors of between 5 and 10. The Basel Committee undertook similar exercises on a range of international banks and looked at the impact of model risk on various risk metrics. Even for simple metrics such as the VaR, the range of variability could vary by a factor of three to five. For more complicated measures and more complicated models, the range of variation was so large that it had to be reported on a log scale. For some, they differed by factors that ran into three and, in one case, four figures: in this case, one model suggested capital of \$1 and the other, \$1,000, for exactly the same exposure. These are astonishing variations: this evidence alone suggests that the models are practically useless. Moreover, even these results are likely to underestimate the true extent of the problem, because the exercise only considered a limited sample of banks and was biased toward relatively simple portfolios. Risk models are thus subject to a positively deafening volume of model noise.

There is also clear evidence that the banks were gaming the system. Perhaps the clearest comes from a chart put together by Andrew Haldane, which is reproduced as Figure 2.⁴²

This chart shows average Basel risk weights and leverage for a sample of international banks over the period 1994–2011. Over this period, average risk weights show a clear downward trend, falling from just over 70 percent to about 40 percent. At the same time, bank leverage—a simple measure of bank riskiness—moves in the opposite direction, rising from about 20 to well over 30 at the start of the crisis. The only difference is that while the latter then fell, the average risk weight continued to fall during the crisis, continuing its earlier trend. “While the risk traffic lights were flashing bright red for leverage [as the crisis

Figure 2
Average Risk Weights and Leverage(a)(b)



Source: The Banker, Bloomberg and Bank calculations.

(a) Sample consists of Deutsche Bank, HSBC, BNP Paribas, Barclays, Citigroup, UBS, BAML, BONY, Commerzbank, ING, JPM, LBG, RBS, Santander, State Street, UniCredit, Wells Fargo. Data is not available for the remaining G-SIBs.

(b) Leverage ratio defined as Total assets/Tier 1 capital.

approached], for risk weights they were signalling ever-deeper green,” as Haldane laconically puts it:⁴³ the risk weights were a *contrarian* indicator for risk, indicating that risk was falling when it was, in fact, increasing sharply.

The explanation is simple: the regulatory system was being gamed by banks engaged in risk-weight “optimization”—exploiting the loopholes and inconsistencies in the system via regulatory arbitrage, in effect gaming both the denominator⁴⁴ and the numerator⁴⁵ in the risk-weighted capital ratio. In so doing, they hijacked the system into a race to the bottom.

As an aside, we should dwell for a moment on the downward trend in the assets’ risk weights in Figure 2. Other things being equal, this means that a given *nominal* capital ratio will have *effectively* fallen, on average, by over 40 percent over this period: effective capital standards are eroding. Put another way, a bank with a capital ratio of 10 percent in 1995 would have had to increase its capital ratio to

over 16.7 percent by 2011 to maintain the same effective capital strength. If this trend were to continue, the risk-weights would hit zero by about 2030, by which point the effective capital ratios would have been watered down to nothing. However, my guess is that effective capital ratios are on course to nothingness even faster than this projection would suggest because the increasing complexity of the system gives ever more scope for the risk-weight arbitrage that the downward trend in Figure 2 so nicely illustrates.

To return to the main story: the banks’ risk models were all signed off as Basel-compliant—and barely was the ink dry on Basel II when the financial crisis struck and much of the banking system collapsed. A panicked regulatory response followed, and a new treaty, Basel III, was rushed out in the fall of 2010. And so Basel II, which took nearly 10 fairly quiet years to produce, was replaced with Basel III, which was the product of as many months, if that, and very unquiet they were too. Basel III

“The regulatory system was being gamed by banks engaged in risk-weight ‘optimization’.”

“A recurring theme in the recent history of bank-capital regulation is ever-growing complexity.”

inauspiciously weighed in at 616 pages—twice its predecessor—and involved some tightening up of capital standards (including higher standards and a tightening of capital definitions), but for the most part, the central principle was “more of the same, but better”: banks will hold more capital, and it will be of better quality; banks will have better and more extensive risk management; we will have more and better regulations and more and (naturally) better regulators—and rest assured, the new system really *will* work next time, trust us. In short, it is the same tired old response we have always had: a knee-jerk reaction, a rising regulatory burden, long on hot air and short on accountability, but no true reflection on what went wrong.

Complexity vs. Simplicity

A recurring theme in the recent history of bank-capital regulation is ever-growing complexity. The trend toward greater complexity is related in part to the desire to achieve greater risk sensitivity, or granularity. At first sight, it would seem obvious that more sensitive risk-weighting would produce greater risk sensitivity, but this intuition is mistaken:

- As a matter of empirics, it is mistaken because it ignores the risks that come from modeling itself. If model risk is sufficiently large—and the evidence suggests it is massive—then the noise in the model can drown out the signal; a much simpler approach might work better because it is unpolluted by model risk. This result might seem counterintuitive, but it is well understood in many areas outside of modern finance. It is well known that complex risk-weighted algorithms often perform poorly out of sample when making predictions about almost anything, whereas simpler approaches often perform better because they are free of noise pollution. A related example is the tax system: it is well-known that the more complex the tax code, the more loopholes it has and the more

opaque and the more gameable it becomes—and the less effective and more regressive is its impact: this is why Warren Buffett claimed to pay a lower tax rate than his secretary.

- It is mistaken theoretically because it ignores aggregation issues—that the risks across banks’ portfolios, taken as a whole, may bear little relation to the risks of each part. The focus has been on the micro risks, and there has been relatively little attention to the correlations among them that really determine aggregate risks. Risk modelers were so preoccupied with the grains in the wood that they overlooked the trees, let alone the forest. Yet under Basel III they are even more preoccupied with this atom-by-atom approach than they were before the crisis. I must confess to having the odd, ever so slight, twinge of doubt that using an ever stronger magnifying glass to focus ever more strongly on the wrinkles on the grains of the trees in the forest is the best way to give regulators a better sense of the big picture that they completely missed the last time.

Let’s consider some further evidence. The ultimate test is a horse race. So let’s ask the question: how would simple and complex models have performed predicting the bank problems that emerged in September 2008? To address this question, Haldane reports an experiment on a panel of 33 large international banks, distinguishing between those that got into trouble and those that did not.⁴⁶ His results were remarkable:

- If we take the Tier 1 capital ratios as a proxy for the risk forecasts of the complex models, he found that the Tier 1 capital ratios of the two sets of banks were essentially indistinguishable. These regulatory capital ratios were about as helpful predicting impending problems as a coin toss, that is, they were next to useless.

“Simple market-based metrics provided better predictors of bank failure than the complicated models.”

■ By contrast, if one used a variety of much simpler market-based metrics—including, in particular, the ratio of market bank capital to the book value of bank assets—we get a clear difference between the two: the simpler metrics give far timelier signals of impending stress. The market capital ratio, for instance, began to fall in April 2007, well over a year before the Lehman crisis.

Thus, the simple market-based metrics provided better predictors of bank failure than the complicated models. Other studies using a variety of techniques and samples come up with similar findings.⁴⁷ Moreover, studies found that the most sophisticated models performed the worst: Haldane and Madouros give an example where the simplest model in their horse race had 10 times the predictive explanatory power of the most sophisticated model.⁴⁸

The implications are profound and go against the training of both economists and regulators. Their natural instinct is to fight complexity with ever more complex solutions, but we have seen that these responses actually make matters worse. Haldane and Madouros give a nice analogy in their story about the dog and the Frisbee. Were a physicist to write down Frisbee-catching as an optimal control problem, they would need to apply Newton’s laws of gravity. Yet dogs are good at catching Frisbees without appearing to understand Newtonian mechanics:

So what is the secret of the dog’s success? The answer, as in many other areas of complex decision-making, is simple. Or rather, it is to keep it simple. For studies have shown that the [F]risbee-catching dog follows the simplest of rules of thumb: run at a speed so that the angle of gaze to the [F]risbee remains roughly constant. Humans follow an identical rule of thumb.

Yet financial regulators failed to anticipate the crisis.

So what is the secret of the watchdogs’ failure? The answer is simple. Or rather, it is complexity. . . . the type of complex regulation developed over recent decades might not just be more costly and cumbersome but sub-optimal for crisis control. In financial regulation, less may be more.⁴⁹

There are a number of reasons why complexity fails. One reason mentioned already relates to the vulnerability of complex models to model risk that more than drowns out the model’s risk signal. This problem is related to issues of small sample size, but it is also related to the problems of overfitting in statistics. An overfitted model leads noise to be mistaken for signal: the model sways with the slightest breeze in the data and produces fragile forecasts that are prone to fall apart. We see this outcome again and again in many different fields: simple models beat complicated ones.

In this context it is instructive to ask how much data would be needed under idealized conditions for a complex model to beat a simpler rival. Haldane and Madouros provide an interesting horse-race experiment that sheds some light on this question. They consider three three-asset VaR models in ascending order of complexity: the first uses a simple moving covariance matrix, the second a relatively simple exponentially weighted moving average (EWMA), and the third and most complex a multivariate GARCH. Note that this “complex” model is not by any means complex: it is just a little more involved than the others. Using monthly data from 1890, they then assess the out-of-sample forecast performance of the three models. Their results are striking: they find that the simplest models outperform the more complex for samples of 20–30 years; the performance of the latter gradually improves as the sample size is increased, but even with a sample size of 75 years, the simpler models perform at least as well as the complex one. They also find that as the portfolio becomes more complex—as more assets are added to it—the performance of the simpler models improves

“Under uncertainty, the appropriate response is not fine-tuned, but coarse-tuned.”

relative to that of the complex models. The message is both clear and consistent: for plausible sample sizes, the simpler models work at least as well as the complex ones. Furthermore, these tests implicitly assume that the world stands still, which it never does and definitely not over a long period. Putting these points together, we will *never, ever* realistically have a long enough sample to give the complicated models a chance to demonstrate their superiority; instead, *the simpler models will always win out*.

Another reason for the failure of complex models is that economists and finance quants are working with the wrong paradigm. They typically make overambitious assumptions about human calculation ability and the information available; most of all, their thinking is dominated by models of decisionmaking under risk—that is, by models of quantifiable uncertainty, not to mention their inbred obsession with optimal solutions, as opposed to merely good ones. This mindset leads them to complicated and often state-contingent optimal decision rules: policy responses are very finely tuned.

However, uncertainties—and certainly, the uncertainties that matter most—are unquantifiable, and for decision-making under true uncertainty, that logic is reversed: Complex uncertain environments instead call for *simple* rules. The reason is that these rules are more robust to ignorance, that is, to what we do not know. Under uncertainty, the appropriate response is not fine-tuned, but coarse-tuned. Indeed, this is exactly why simple rules of thumb—heuristics—are the rule, not the exception, in so many walks of life: the dog and the Frisbee again come to mind.

This message is also consistent with literature suggesting that relying on complex responses in a complex environment can be catastrophic: in nuclear power plants, oil rigs, aircraft navigation, and myriad other areas, disasters are more likely to occur when complex control mechanisms are used in complex, tightly coupled interdependent systems. This, in turn, reminds us of the limits of what the human mind can achieve and of the benefits

of decentralized, loosely coupled, systems. As Hayek wrote in *The Fatal Conceit*:

The curious task of economics is to demonstrate to men how little they really know about what they imagine they can design.

To the naive mind that can conceive of order only as the product of deliberate arrangement, it may seem absurd that in complex conditions order, and adaptation to the unknown, can be achieved more effectively by decentralizing decisions and that a division of authority will actually extend the possibility of overall order.⁵⁰

To paraphrase Haldane: to tackle complexity in banking through complexity in regulation is to fight fire with fire. Such an approach is unlikely to work in theory and does not work in practice. In banking regulation, less is more.⁵¹

REGULATORY STRESS TESTING

Before discussing the Fed’s stress tests, it is appropriate to offer a few remarks on regulatory stress testing in general. The most basic point to appreciate is that although the Fed often describes its stress tests as “forward-looking,” this description is misleading. Yes, a stress test projects into the future, but—and I really shouldn’t have to spell this out—no one can actually *foresee* the future; in fact, a stress test is merely a *guess* about what *might* happen. You pull a possible scenario or set of scenarios out of thin air as best you can and try to work out the loss or losses that would arise. In the case of the Fed’s stress test, there is only one scenario that matters, its “seriously adverse scenario.” Applied systemically, the Fed then implements a regulatory capital regime based on this guess scenario and declares we can be *confident* that the banking system is safe—a guess goes into the Fed’s magical mystery machine, and out comes systemic confidence—and that despite the fact that the regulatory regime itself creates a huge dose of systemic

instability, and never mind that the Fed's past guesses haven't been much good, either.

The reality is that a guess is still a guess. The guess might be wrong. You might imagine some adverse scenario—and perhaps even get the modeling right—and be covered if that scenario should occur. But then some other adverse scenario might occur that you hadn't thought of and that could be devastating: the world is uncertain. The whole notion of betting everything on a *single* scenario out of an infinite number of possible future scenarios is surely insane. Whatever you think is going to get you, something else will. In ancient Greece, the custom was to go to Delphi and ask the oracle for advice—to ask for the outcome to a scenario analysis—and empires rose and fell on the back of the oracle's answer.⁵² The Fed plays much the same role in the modern U.S. financial system. The difference is that the oracular priestesses high on volcanic gases had a better track record than the Fed with its army of “scientific” economists and risk modelers.

So how to be sure to get your guess “right”? Obviously—short of a crystal ball or an oracle that actually works—you can't. However, you can be prudent and not bet everything on a single scenario. If, for the sake of argument, regulators are going to do any regulatory capital stress testing at all, then I would suggest the following guiding principles:

1. Stress tests should be conservative, in the sense of being based on highly adverse possible scenarios: there is no point carrying out unstressful stress tests.
2. Note the plural: we wish to consider a range of plausible and heterogeneous adverse scenarios, not just one. We should then take the loss from each scenario and set the capital requirement equal to the maximum of these losses; that way, banks are covered against the losses from any of these scenarios.
3. Stress tests should be based on simple models, not highly sophisticated ones: the evidence indicates that the simpler

models work better anyway. Note also that we don't want huge armies of modelers.

4. The stress test models should be appropriate to the specifics of the institutions concerned: the stress tests don't make sense otherwise.
5. Any stress test system should be non-gameable by the parties involved.
6. Any system that involves a heavy regulatory burden or creates a large systemic risk exposure is destructive and counterproductive.
7. Any system should be transparent and accountable.

We shall see how the Fed's stress tests measure up to these guidelines presently.

It is also helpful in this context to look at what the expert literature on stress testing recommends. An example is the reference *Stress Testing for Financial Institutions*, edited by Daniel Rösch and Harald Scheuler.⁵³ In the book's epilogue, RiskMetrics' head of research Chris Finger summarizes the key principles of stress testing—I emphasize that we should regard these as authoritative—and two points in particular stand out:

The first is that stress testing should be used as a complement to other risk management tools. He makes a nice analogy with dietary supplements:

we see that in many guises, stress-testing practices are akin to a person taking dietary supplements rather than eating a healthier diet. As corrections to inferior risk models . . . stress tests do not solve our problem. We would be better served by eating more vegetables.

To a healthy diet, or a good risk model, some stress-testing practices do provide useful complements.⁵⁴

To put the point another way, we should never rely on stress tests on their own, or even as our primary risk metric: they are simply not reliable enough due to their Achilles heel—their

“Stress testing should be used as a complement to other risk management tools.”

“There is simply no point fine-tuning a single scenario and ignoring the infinite number of other possible outcomes, one of which will almost certainly occur, but we can’t tell which in advance.”

dependence on subjective scenario assumptions, which may or may not be any good.

The second point emphasized by Finger is essentially my second point above. As he puts it when discussing good practice in regulatory stress-testing:

a supervisory body specifies a (possibly large) set of prescriptive tests, and the party being tested is said to ‘pass’ if its worst loss across all the tests is less than a predetermined threshold. . . . *we do not look at any single scenario carefully, but rather hope that the set of scenarios covers the spectrum of risks we might face.*⁵⁵ (My italics)

Note that the tests should use a broad-brush approach across a possibly large set of scenarios. Precisely because the future is so uncertain, we don’t wish to bet everything on a single stress scenario. Underlying both these points is a deeper one, the principle of humility in the face of uncertainty. There is simply no point fine-tuning a single scenario and ignoring the infinite number of other possible outcomes, one of which will almost certainly occur, but we can’t tell which in advance. The chances of any single scenario actually occurring are vanishingly small, so we always need to consider a range of possibilities. We want to be approximately right, not exactly wrong.

You might have thought this was common sense, but as we shall see, modern financial regulators think otherwise: they think they can identify *the* relevant scenario—talk about hubris!—polish it ad infinitum and forget about the rest. We cannot expect such an approach to work, and the experience of other regulatory stress tests confirms that it doesn’t.

Lessons from Other Regulatory Stress Tests

So let’s look at the experiences of these other regulatory stress tests—in particular, the Fannie/Freddie and European stress tests.

THE FANNIE/FREDDIE STRESS TESTS. First, a bit of background: the Fannie-Freddie party line had always been that they were very safe.

Following the well-publicized bank difficulties of the late 1980s, however, there was concern about other potential bailout candidates, and the Treasury was preparing a report that would conclude that a rise in interest rates of 4 percent to 5 percent could make Fannie insolvent. This report would bolster the case of lawmakers like Jim Leach, who wanted Fannie and Freddie to hold more capital. However, in an audacious lobbying coup, Fannie pre-empted the Treasury report by commissioning ex-Fed chairman Paul Volcker to investigate the matter, and Volcker’s 1990 report concluded that Fannie was, in fact, adequately capitalized: the risk of any bailout was remote, he concluded. As Fannie’s chief executive David Maxwell said at a subsequent news conference, the company’s business wasn’t nearly as risky as that of the big banks with their dodgy lending. “There are no unpleasant surprises because of the nature of our assets,” he said. “We don’t have any see-through buildings, any Third World countries or any strip shopping malls. We just have those mortgages.”⁵⁶ And we all know that mortgages are as safe as houses.

Fannie and Freddie then had their way. A couple of years later, in 1992, the Federal Housing Enterprises Financial Safety and Soundness Act set up a new agency, the Office of Federal Housing Enterprise Oversight (OFHEO), and directed it to implement a stress-based capital standard to determine Fannie and Freddie’s regulatory capital requirements. The new capital regime was just what the GSEs wanted: it was very undemanding—the GSEs had capital requirements that were under half of those imposed on banks. It was no accident that the GSEs liked the new regime: as Jim Leach pointed out, it was their lawyers who had drafted the legislation. One of his colleagues, the Texas Democrat J. J. Pickle, was less impressed. “We should have done better,” he said, “We have, once again, left the public purse exposed to the risks of private greed and corporate misjudgment.” However, no one wanted to listen to the naysayers.

The new regime was state-of-the-art, ahead of anything in banking. Indeed, the

stress tests on which it was based were a regulatory first: it was to be nearly 20 years before banking regulation was to follow suit. The model and its risk-based rulebook then took a decade to become operational due to a combination of the complexity of the test and the politics of regulating two enormous, rapidly growing, and politically connected financial institutions.⁵⁷

The risks facing the GSEs were complicated. A decline in interest rates creates a risk of loss due to prepayments, as more homeowners refinance at lower interest rates. A rise in interest rates induces a decline in the price of mortgages and an increase in defaults. The stress test therefore incorporates shocks to both interest rates and default rates, and posits sustained stress over a 10-year period. At face value, the stress was a very severe one—a senior Fannie official described it as a “decade long ‘nuclear winter.’”⁵⁸ The required level of capital was then determined as the amount that would allow it to remain solvent through the stress-decade plus an additional 30 percent to cover other risks—again, at face value very conservative (and indeed, admirably so).⁵⁹

As the details were being finalized, Fannie commissioned a team of distinguished economists—including most notably, the recently en-Nobeled Joseph Stiglitz—to carry out some stress tests on Fannie and Freddie.⁶⁰ They made a series of assumptions relating to the precise form of the stress scenario, the GSE’s capital plans, and so on, and then obtained upper-bound estimates of the probability of default using historical data then available. Their results suggested estimates of “substantially less than one in 500,000” and maybe “smaller than one in three million.”⁶¹ The take-home was that the probability of default was “effectively zero.”⁶² They also reassuringly pointed out that even if their estimates were off by an order of magnitude, it really wouldn’t make that much difference. Such numbers were perhaps *too* reassuring and are reminiscent of NASA’s arguments in the aftermath of the *Challenger* disaster that the probability of *that* failure was only 1 in 100,000—prompting a fa-

mous rebuke from another Nobelist, the physicist Richard Feynman: “If a guy tells me the probability of failure is 1 in 10^5 , I know he’s full of crap,” he reputedly said.⁶³ Moreover, the failure probabilities cited by Stiglitz and his colleagues were between 5 and 30 times *smaller* than this number.⁶⁴

Despite the hype that greeted the OFHEO model, a recent re-evaluation by Federal Reserve economists identifies serious deficiencies.⁶⁵ To begin with, the model itself had very poor forecasting ability; they attribute this poor performance to the fact that the model was never updated. There was no effort to include missing variables or even recalibrate the model parameters throughout the seven-year period over which it was used, despite well-documented changes in mortgage underwriting practices during this time. There were also significant model risk problems, and the house price stress in the OFHEO model was significantly less stressful than the actual experience since the housing bust—in other words, the “nuclear winter” stress wasn’t particularly stressful. They also suggest that OFHEO faced “material challenges emanating from the statutory model disclosure requirements which, coupled with the political power of [the GSEs], would have made meaningful changes to the risk-based capital rule extremely costly.”⁶⁶ This inflexibility left the model open to gaming. As the authors continue: “By publishing model specification and parameter estimates, [the GSEs] were able to take on risks that were not well-captured by the models (e.g., loans to borrowers with weak credit histories or very high loan-to-book ratios). This disclosure, *by design*, allowed the GSEs to successfully manage to the stress test” (my italics). In short, the risk model itself encouraged the GSEs to take on risks that were hidden to the model, and this was a design feature of the system. And people were wondering afterwards why it didn’t work.⁶⁷

There was one other problem with the Fannie and Freddie stress tests, and a glaring one at that: they produced extremely low capital requirements. Indeed, Freddie’s estimated

“The risk model itself encouraged the GSEs to take on risks that were hidden to the model, and this was a design feature of the system.”

“The GSEs were taking on risk to the hilt, while the stress tests allowed them to operate on the thinnest possible wafer of risk capital.”

risk-based capital requirement was below 200 basis points for the entire period 2003–2007: it had taken the OFHEO rocket scientists a decade to come up with a model-based capital requirement of almost zero—and this over the very same period when the GSEs were loading up on toxic subprime. Those involved would have us believe that the obvious contradiction between the two didn’t occur to anyone—more likely, it did, but it didn’t suit anyone to point out the elephant in the room. So the GSEs were taking on risk to the hilt, while the stress tests allowed them to operate on the thinnest possible wafer of risk capital: the stress tests failed the sniff test.

Many of those involved were later to say that the GSEs’ subsequent problems could not have been anticipated. However, informed observers *had* been warning all along that the GSEs were heading for problems. One was Alan Greenspan. He regularly fulminated about the risks the GSEs were taking. Speaking to Congress in February 2004, he warned that they posed “very serious risks” to the U.S. financial system. However, as with his earlier “productivity miracle,” Greenspan was only half right and he missed the main risk factor that would subsequently undo the GSEs: the risk of a wave of mortgage defaults.

One of the few to be proven right all along, however, was Louisiana Congressman Richard Baker. On July 23, 2002, Baker was chairman of the House Subcommittee on Capital Markets, Insurance and Government Sponsored Enterprises. Addressing the subcommittee that day was the OFHEO director, Armando Falcon. Falcon was there to report on OFHEO’s recent stress tests while also requesting major increases in his office’s appropriations. “We have developed a strong, rigorous risk-based capital standard,” Falcon assured the Congressmen:

Fannie Mae and Freddie Mac are safe and sound and well capitalized. We all can see their financial health at the beginning of any discussion about them. . . . through our regulatory program we constantly probe for weaknesses and

vulnerabilities and assure that the enterprises maintain the highest standards of sound management.

We have a state-of-the-art risk-based capital rule in place. We are expanding our examination program, building our regulatory infrastructure and conducting valuable research. The OFHEO of 2002, which I am proud to direct, gets the job done.⁶⁸

Baker was not impressed. The point at issue was prompted by Baker’s earlier request to OFHEO to conduct another stress test after an earlier one had shown Fannie underwater by \$600 million. The new test result now put Fannie in the clear with an estimated net worth of \$2.4 billion. As Baker began:

I want to make it clear that my line of questioning today results from a great deal of frustration about the process and where I think we find ourselves, as of this moment.

It took eight-and-a-half years and considerable encouragement from the committee and others, including defense of appropriations process, to make sure that OFHEO had the resources to ultimately prepare the most professional analysis possible for this committee and for the benefit of taxpayers.

I find it troubling that at the end of that eight-and-a-half year period, we then had an approximately 60-day period, of which I am told by the agency it was a result of my request asking you to be thorough and thoughtful in the process, that it took as long as it did to make the subsequent modification, resulting in the post-amendment stress test, as we have it today.

Baker wanted to know why it had taken so long and, in particular, why the results were so different. There now followed a long drawn-out game of cat-and-mouse in which Baker tried and tried and tried to get an explanation

from Falcon, but could never get one. Former treasury secretary Lloyd Bentsen then got involved and passed Falcon a few softballs that allowed him to pat himself on the back again:

Every so often you might hear some criticism about the rules. Some think it is too lenient, some think it is too tough. What we have done is craft a rule which closely ties capital to risk. It is based on sound, historical analysis, historical data and our best judgment about the risk associated with all the different activities at the enterprises.

I think it is a very strong, robust rule. And I am proud of the work the agency did to put this together. It has never been done before by any regulator.

Baker then returned to the attack. To quote from the transcript:

What this [earlier test] says is that Fannie missed the mark by \$600 million, whereas if you use the post-amendment test, they were in excess of the requirement by \$2.4 billion. . . .

So it is a \$3 billion swing from one test to the other.

And then again:

And now whether or not the enterprise was in any financial duress or not is only something probably the GSE knows. But if the modifications to this test were made as a result of staff determination as to de-bugging—as to haircuts, as to multi-housing, as to funding costs and it results in a \$3 billion swing, you have got to ask one of two questions. Is the test valid? Or were the changes made as a result of some political involvement in the management of the structure of this test?

. . . the rule was built to depression-like standards of conduct.

Last quarter 2001 was a volatile financial quarter, but it was not depression-

like in its nature. There was nothing that changes in the business structure of Fannie from the last quarter of 2001 to the first quarter of 2002 that makes any structural difference.

I am just having a really difficult time in understanding how that test applied to Fannie Mae in the last quarter of 2001 could result in the analysis we get and the new test applied to the same quarter results in a \$2.4 billion capital surplus.

It does not make sense. . . .

The same data, the same quarter, two different tests, one gives one result, the other gives another.

Congressman Baker eventually gives a very frustrated summing up:

I guess I can best express my situation this way, Mr. Falcon, after eight-and-a-half years, we had a test. . . . [O]ne GSE passed; one GSE failed. Ninety days later, we have another test. Amazingly enough, both GSEs passed.

Imagine what our public school system would look like if we gave everybody in the class a test. And then 30 days later, we came back and gave them exactly the same test with prior announcement. I bet our test scores would go up.

Or if, before our children could go out to play in the afternoon, we asked them, “Did you clean up your room?” “Oh, yes, dad, sure. But could you come back about seven and check it out?”

That is my problem here. We have a test that was developed after eight-and-a-half years of very intense, very hard work with congressional oversight, people fighting to protect your budget, get you the staff you needed. And I feel, frankly, very let down. . . .

. . . there has got to be another place to land that can give the taxpayer of this country an accurate honest assessment of the true risk exposure they face in the indirect support of two enterprises

“The CEOs and boards were fired and the GSEs recapitalized at vast taxpayer expense. Fannie and Freddie’s stress tests had been a spectacular failure.”

“When it comes to screwing up stress testing, the Fannie/Freddie stress testers were rank amateurs next to their European counterparts.”

which have over \$3 trillion of exposure to this economy. That is enormous.

And in light of the financial conditions of the market we face today, giving all the [fraud] allegations that are floating [Enron had failed only months before]—thank goodness we are not hearing any of that about Fannie or Freddie. I can only imagine the consequences if one of these two corporations were to report a financial irregularity.

Those were to come to light a little later!

Fast forward to early July 2008 and the problems faced by the GSEs had become very apparent: their debt was sliding and their stock values falling sharply amongst widespread concern that they might default. Given their size—more than \$5 trillion in assets or guarantees—there was concern that a default by either could produce a systemic breakdown. Treasury Secretary Hank Paulson and Fed Chairman Bernanke made one last-ditch effort to reassure nervous investors. Fannie and Freddie are “adequately capitalized,” Bernanke told the House Financial Services Committee on July 16: “They are in no danger of failing.” By this point, however, both Paulson and Bernanke must have had some suspicion, at least, that the GSEs might *not* be adequately capitalized despite their reassurances to the contrary, and news had already come out that they were preparing more concrete measures. As Sen. Jim Bunning (R-KY) commented in disbelief:

When I picked up my newspaper yesterday, I thought I woke up in France. But, no, it turned out it was socialism here in the United States of America, and very well, going well. The Treasury Secretary is now asking for a blank check to buy as much Fannie and Freddie debt or equity as he wants. The Fed purchase of Bear Stearns assets was amateur socialism compared to this. And for this unprecedented intervention in our free markets, what assurance do we get that it will not happen again? Absolutely none.⁶⁹

Privately, the government was preparing a takeover to stabilize their debt. Paulson’s plan was to go in swiftly and seize them. The “first sound they [will] hear is their heads hitting the floor,” he told the president. On September 7, the government then made its move—one of dubious legality, too—in one of the most sweeping government interventions ever and placed the GSEs into a government conservatorship. The CEOs and boards were fired and the GSEs recapitalized at vast taxpayer expense. Fannie and Freddie’s stress tests had been a spectacular failure.

You ain’t seen nothing yet, however. As we shall see, when it comes to screwing up stress testing, the Fannie/Freddie stress testers were rank amateurs next to their European counterparts.

THE EUROPEAN STRESS TESTS. Consider Iceland as a warm-up. In 2004, the three largest Icelandic banks—Glitnir, Kaupthing, and Landsbanki—had assets equal to about 100 percent of GDP. They then began an enormous expansion financed through wholesale markets, and by the end of 2007 their assets amounted to almost 900 percent of GDP, a world record. By this point, their dependence on wholesale markets was a concern—the British bank Northern Rock had hit the rocks for this very reason in September 2007—and credit default swap spreads were strongly suggesting that the Icelandic banks were vulnerable. However, an International Monetary Fund (IMF) stress test in August 2008 gave results to “suggest that the system is resilient.”⁷⁰ Bank capital ratios had fallen a little—from 12 percent in 2007 to 11 percent in early 2008—but they were still well above the 8 percent Basel minimum. Liquidity ratios were also above minimum levels. Other stress tests carried out by the country’s central bank and its financial regulator, the Financial Supervisory Authority, gave similar results. The message was that although there were concerns in the troubled markets of the time, the system was essentially sound. However, the banks’ liquidity issues mounted, and then the banking sector collapsed suddenly in early October 2008: all three sets of stress tests had missed

the imminent collapse of the entire system. The result: Iceland was knocked back to an economic Ice Age.

Moving now to the European stress tests proper, a good starting point is the Europe-wide stress test implemented by the Committee of European Banking Supervisors (CEBS) in 2009. The objective was to see how the banks would fare on a common adverse scenario against a minimum target Tier 1 capital to risk-weighted assets ratio of 6 percent. The results announced on October 1, 2009, suggested that none of the 22 large banks in the sample would see their Tier 1 ratio fall below 6 percent for the adverse scenario: all banks passed with flying colors. The CEBS press release proudly talked of how the “resilience” of the banking system reflected the success of recent public-sector support to the banks: the banks were now in strong shape again, notwithstanding recent unfortunate difficulties. An alternative interpretation was that the stress test scenario was just too weak, and time would soon tell which interpretation was the correct one.

The next year, CEBS conducted another European-wide stress test: this exercise was expanded to 91 banks, and individual results were published for the first time. This sample of banks covered 65 percent of total European bank assets. In this case, the results reported on July 23, 2010, showed that only seven banks had failed to meet the 6 percent minimum Tier 1 ratio, and the aggregate shortfall to achieve this target was a mere €3.5 billion. This number alone was suspiciously low and should have set off red flags among the regulators, but they preferred to take the results at face value: after all, the stress test had confirmed what they wanted to believe—that is, that the banking system was fundamentally sound.

Critics pointed out that amongst other problems, the stress test had one very gaping hole: that the biggest stress facing the European banks, the risk of one or more sovereign defaults, was only covered in the trading book. Since most bonds were held in the banking book, this meant that most of the sovereign risk was not accounted for in the test. Regula-

tors defended the omission of most sovereign risks on the astonishing grounds that it wasn’t really a risk at all, as the EU would not allow any sovereign to fail! Thus, the stress test was compromised by wishful thinking on the part of the powers that be: it would never do to have a stress scenario that the European authorities had vowed to avoid at all costs, even though they would eventually fail, and the smart money knew it. One critic, Citi chief economist Willem Buiter, pointed out that at the end of 2009 there was more than €2.8 trillion worth of PIGS debt (Portugal, Ireland, Italy, Greece, and Spain) and there was not a single safe sovereign anywhere. He went on to argue that Europe’s then-current bailout facility of €860 billion might be insufficient by more than half. Other reports suggested even higher exposures, and one has to keep in mind that no one had a clue about the true value of Greek government debt. In response to these concerns, Christine Lagarde, the [then] French finance minister, reassured the doubters: “We will stress the system a little more to make the results more credible.” “For us, this is the definitive ‘credibility breaker,’” responded the pseudonymous Tyler Durden of Zero Hedge.⁷¹ He wrote off the whole exercise a few days later, even before the results were revealed:

Euro Area banks therefore need additional capital—a lot of it. This may not be apparent from their ratios of regulatory capital to risk-weighted assets but, in our view, both the numerator and the denominator of this ratio are deeply unreliable.

Tangible common equity is the only unconditional loss absorber, and the denominator is compromised by the deeply flawed risk weights. He then continued:

[T]he Stress Tests will be a joke as they will not take into account solvency haircuts, which, just so happens, is the primary concern in Europe. . . . In the absence of stress tests that include scenarios in-

“The stress test had one very gaping hole: that the biggest stress facing the European banks, the risk of one or more sovereign defaults, was only covered in the trading book.”

“It wasn’t just the huge potential for losses that was omitted from the stress test, but also the disruption to the wholesale finance on which many banks are very dependent.”

volving multiple sovereign defaults, we think it is difficult to view even the new exercise as a major confidence boosting event. Added to this is the problem that national implementation of these stress tests is left to the same national regulators and supervisors that failed in so many countries . . . to anticipate and prevent the excesses that undermined so many of the institutions they were responsible for regulating and supervising. . . . The results of stress tests performed by those who failed to prevent the last crisis are likely to convince few market participants.⁷²

However, it wasn’t just the huge potential for losses that was omitted from the stress test, but also the disruption to the wholesale finance on which many banks are very dependent. Two weeks later Durden wrote:

. . . by the time the stress tests are found to have been a joke, and the country hosting the bank blows up just because the bank’s assets are 3x the host nation’s GDP, and the country is forced to bankrupt, it will be far too late. So let’s get this straight—the very issue that is at the heart of the liquidity crisis in Europe, namely the fact that a bankrupt Greece has managed to destroy the interbank funding market in Portugal and Spain, and the other PIIGS, and has pushed EURIBOR and other money market metrics to one year stress highs, *and* forced the ECB to lend over \$1 trillion to various central and commercial banks, will not be *tested for?* Fair enough . . . But [the ECB] certainly should not be surprised if this latest show of idiocy by Trichet’s henchmen serves as the springboard for the latest round of spreads blowing up across Europe.⁷³ (his italics)

His warnings were soon realized, although it was not a country with a 300 percent ratio of bank assets to GDP that blew up, but one with

a ratio of 872 percent: Ireland. On November 21 2010, the Irish government requested assistance from the EU and the IMF. The problem was that having bailed out the banks in September 2008 and having thought the banks were now sound, the Irish government found itself unable to cover the banks’ wholesale financing requirements: the Irish banks had dragged down their own sovereign. So a mere four months after they passed the stress tests with flying colors, the Irish banks were now revealed to be in need of massive aid to stay afloat. The eventual bailout package was €85 billion, or nearly half of Ireland’s GDP. The stress test was utterly discredited. Ireland was already struggling, and this disaster knocked it into a depression from which it has still not recovered.

By this point, it was an open secret among everyone in the banking world that the stress tests were a parody, repeatedly attempting to restore systemic confidence, but repeatedly undermining it. But this being Europe, and European regulators being unable to think of anything else, they were soon proposing yet another round of stress tests. In late November, 2010, they announced that they were planning a new round of stress tests that they promised would be more rigorous than their now discredited predecessors. “Thank you for confirming [that] the prior stress test, the one which found that not one Irish bank was impaired, was a bunch of bullshit,” noted Durden gracefully.⁷⁴ The regulators also revealed that they were considering *not* publicly disclosing the results next time. “You can’t make this up,” wrote Durden: “Europe plans fresh round of ‘secret’ stress tests to ‘restore confidence.’”⁷⁵

The new test was to be carried out by new European Banking Authority that was to take over from CEBS at the start of 2011. Regulators were adamant that this next stress test would be more severe than its predecessors—it could hardly be less—and that important lessons had been learned and so on, the usual guff. As Enrico Enria, the nominee to head the European Banking Authority (EBA), reassured the European Parliament: “we need to do a stron-

ger, more credible and reliable exercise.” They also decided that it wasn’t such a clever idea to keep the results secret after all. However, it gradually became clear that the new stress tests would have *much the same* major design flaws as the previous stress tests, which even the regulators had had to admit had failed—in other words, *no* real lessons had been learned. It still wasn’t tough enough, it was still measuring the wrong things, it was still using unreliable data, and it *still* didn’t take proper account of the banks’ sovereign debt exposures. By this point, Durden had had enough:

The reason we have not been covering this year’s iteration of the European stress test closely . . . is because it was guaranteed a priori that it would be just as farcical as its original version, and result in glaringly failing institutions in the . . . bank sample tested as ‘passing.’ Sure enough, *The Independent* has just reported that all Irish banks have passed the test ‘comfortably’—a list that includes such horrors as Bank of Ireland, Allied Irish Banks and Irish Life and Permanent Plc . . . And with that we can close the book on this year’s stress test before it is even released.⁷⁶

When the results were released later that day, they further confirmed his suspicions. So we had a more severe stress test, one which involved a more severe minimum capital standard (5 percent core Tier 1 instead of 6 percent Tier 1, stronger because of the tighter capital definition); a slightly stronger sample (90 instead of 91: the weak German bank Helaba had withdrawn and observers expected it to fail soon); much more market volatility, in which the enormity of the sovereign debt problem had become much clearer, if still far from fully clear; and a desperate need on the part of the regulators for the new stress test to be much more credible than its predecessors. Yet what did the stress test conclude? It found that the aggregate shortfall to meet the higher standard was *even less* than it had been the year

before: €2.5 billion instead of the €3.5 billion produced by the previous stress test, which no one disputed had been a failure. You wouldn’t have needed any nose at all to tell that this test had failed the sniff test.⁷⁷

If anyone still had any doubt, the results were announced in the same week that the Italian sovereign debt crisis turned much worse—boy, these guys just can’t get anything right—when it was immediately obvious from that alone that the modest sovereign exposures in the stress test had been, well, way too modest. The 2011 stress test was thus discredited on publication, if not before.

And if anyone was brain dead enough to miss the signals provided by the noseless sniff test and the Italian crisis that was dominating the news, three months later the big Franco-Belgian bank Dexia suddenly went belly up. This bank had passed the test easily—under the stress scenario, it would have had a core Tier capital ratio of 10.4 percent, just over twice the minimum, and more than twice the ratios of the eight banks that failed the test.

Throughout the fall of 2011, the European situation continued to deteriorate and the ever-in-denial EBA found itself engaged in a rear-guard action to revise its estimates upward in a futile attempt to protect whatever credibility it imagined it might still have—the words “horses,” “stable door,” and “bolted” come to mind. This “revised” exercise involved a greater emphasis on capital shortfall, but even then, the EBA said that “these buffers are not explicitly designed to cover losses in sovereigns but to provide reassurance about the banks’ ability to withstand a range of shocks and still maintain adequate capital”—incredibly, *still* ignoring the main risk and making one wonder what on earth they thought the tests were designed *for*. The new estimate of capital needed was raised to €114.7 billion—over 45 times larger than their best estimate a few months before—and so once again, the EBA implicitly conceded that its predecessor exercise had produced results that were not worth the paper they were written on. Even this revised estimate was still well below other estimates that

“The new estimate of capital needed was raised to €114.7 billion—over 45 times larger than their best estimate a few months before.”

“The biggest problem with the European stress tests is illustrated by the case of Cyprus.”

were already circulating, not least the IMF estimate of €200 billion in its September *Global Financial Stability Report*; other estimates put the figure at €300 billion. The EBA's original shortfall estimate had been perhaps 1 percent of the estimates other people were getting.⁷⁸

Yet perhaps the biggest problem with the European stress tests is illustrated by the case of Cyprus. Out of the blue, the Cypriot banking system collapsed suddenly in March 2013. As Mark Grant observed in disbelief:

... where was the problem three weeks ago? There was not a mention, not a hint of anything that was wrong. All of the banks in Cyprus had passed each and every European bank stress test. The numbers reported out by the ECB and the Bank for International Settlements indicated nothing and everything reported by any official organization in the European Union pointed to a stable and sound fiscal and monetary policy and conditions. The IMF, who monitors these things as well, did not have Cyprus or her banks on any kind of watch list.

Let me assure you it was not some “Miracle on 34th Street” that took place overnight while everyone was in bed and counting sheep. I can also assure you it was not because some bean counter in Brussels or Frankfurt stumbled over some new bit of data and informed his superiors. Nothing of the sort. The culprit is what is counted and not counted in the European financial system and the quite real consequences of uncounted liabilities. . . .

In just two weeks' time we have gone from not a mention of Cyprus to a crisis in Cyprus because none of the official numbers were accurate. Without doubt, without question, if this can happen in Cyprus then it could happen in any other country in the Eurozone because the uncounted liabilities are systemic to the whole of Europe. The European Union does NOT count sovereign guarantees

of bank debt, sovereign guarantees of corporate debt, derivatives or many other types of contingent liabilities. They are all uncounted, but still there, no disappearing act, and as the bills roll in they have to be paid by someone. Dexia was fine, . . . Bankia was fine and overnight, “Snap, Crackle Pop!” *And Jack comes out of the Box.*

For investors here is the crux of the problem. *The Press is handed the official numbers and reports them out as if they were accurate. . . .* Some EU Finance Minister sends out a press release and the story is printed. . . . The ratings agencies take the word of Europe, make judgments based upon the faulty data and opine based upon the erroneous information. As the geeks among us would say: *Garbage in—Garbage out.*

Then many money managers rely upon the official data, rely upon the ratings agencies and make investment decisions based upon the data that they are given. The CDO issue again . . . where everything was rated inaccurately. So then we have Garbage in—Investments based upon Garbage. Therefore I will tell each and every one of you; if you are making decisions relying upon the official numbers of Europe and upon the ratings of European sovereign debt then you are going to get burned. You are relying upon falsified data and any ratio that you might run is wrong as the underlying numbers are inaccurate.⁷⁹

Grant's comments hit the nail bang on the head: underlying any other problems they may have (e.g., the conflict between the politics and credibility of the stress tests, the limitations of Basel risk-weights and so on) the root problem is that the stress tests can't work because they are inputted with meaningless data.

These cases are only illustrative, but they illustrate the general pattern with regulatory stress testing. In every case, the message they conveyed was simple and beguilingly reassur-

ing: there is no danger. Yet in almost every case, the stress test failed to anticipate the subsequent major stress event.⁸⁰ Instead, the stress tests provided a false comfort and blinded those involved to the real risks they were facing. To quote a recent Bank for International Settlements study:

To our knowledge, no macro stress test carried out ahead of the crisis identified the build-up of vulnerabilities. The relentless message was: ‘The system is sound.’ Rather than part of the solution, stress tests turned out to be part of the problem. They lulled policymakers and market participants into a false sense of security. There is a serious risk that, unless their limitations are fully understood, they will continue to do so in the future.⁸¹

THE FED’S STRESS TESTS

A Stress Test Alphabet Soup: SCAP, CCA, CapRP and DFAST

We turn now to bank capital regulation in the contemporary United States—and in particular, to the Fed’s stress tests. These started in early 2009 when the Federal Reserve introduced the Supervisory Capital Assessment Program (SCAP), a set of stress tests conducted by the Fed encompassing the 19 largest bank holding companies (BHCs) with assets of over \$100 billion, themselves representing about two-thirds of BHC assets. Its objective “was to ensure that these institutions have sufficient financial strength to absorb losses and to remain strongly capitalized, even in an economic environment more severe than that currently anticipated.”⁸² In this exercise, the Fed stipulated a hypothetical adverse economic scenario, under which growth, unemployment, and house prices were assumed to be somewhat worse than most forecasters anticipated. Using information provided by the firms, the Fed’s stress-test model projected that the 19 BHCs would lose about \$600 billion over the next two years under the as-

sumed adverse scenario. The Fed operated with a minimum capital adequacy standard of Tier 1 common capital at 4 percent or Tier 1 capital at 6 percent of risk-weighted assets. Where a bank’s capital fell below the minimum level, it was compelled to recapitalize up to at least the minimum standard. Fed officials repeatedly stressed that this exercise was “forward-looking” and allowed a consistent treatment across the different banks: these would be recurring themes in the later stress tests too.

The SCAP was followed in 2011 by a more ambitious program, the Comprehensive Capital Assessment Review (CCAR). This program centered on bank-run stress tests in which regulators specified the scenario but required each bank to model the stress event itself. Its purpose was to “assess the BHC’s capital planning processes to ensure that the BHCs have good capital plans in place.”⁸³ The minimum capital standard remained as before, but banks were also required to demonstrate that they could meet their Basel III and Dodd-Frank capital requirements under the stress scenario as well: any banks that failed to meet the minimum capital standard under the stress scenario would have their capital plans—their planned capital distributions, stock repurchases, and so forth—rejected.

The CCAR then became an annual exercise, each one becoming more extensive than its predecessor. There was now an annual CCAR cycle: details of the stress scenario and other changes were announced in late November, with submission of capital plans due in early January; the results were announced in March, when the Fed would accept or reject each bank’s capital plan. The 2012 CCAR maintained the same 5 percent Tier 1 minimum common capital ratio, then added a “seriously adverse scenario” to its earlier “adverse scenario”—displacing it as the most adverse scenario—and included a (fairly moderate) Eurozone downturn in this scenario. It was applied to the 19 largest BHCs, but it also stipulated a global stock market meltdown scenario for the six largest firms with large trading ac-

“In early 2009 the Federal Reserve introduced the Supervisory Capital Assessment Program (SCAP), a set of stress tests conducted by the Fed encompassing the 19 largest bank holding companies.”

“The 2013 planning cycle saw the introduction of regulatory stress tests mandated by the Dodd-Frank Act.”

tivities. The 2012 planning cycle introduced regulatory stress tests for banks with assets between \$50 billion and \$100 billion: this program, the Capital Plan Review (or CapPR), was somewhat less rigorous than CCAR because of the smaller banks’ more limited stress testing capabilities and experience. It also began the practice of publishing its results, identifying which banks had had their capital plans approved and which had not: this was another source of irritation to the increasingly beleaguered banking industry, not least because the Fed never deigned to explain why banks that failed had failed, leaving them and (especially) their stakeholders to guess, so undermining confidence—rather like a student being failed a capricious exam, being refused any reason why, *and* suffering the indignity of the result being announced to the whole world.

The 2013 planning cycle saw the introduction of regulatory stress tests mandated by the Dodd-Frank Act (DFA).⁸⁴ The DFA requires the Federal Reserve to conduct annual supervisory stress tests on all banks with \$50 billion or more in assets—so covering more banks in addition to the 18 now subject to the CCAR.⁸⁵ It further widened the scope of regulatory stress tests to nonbank financial companies designated by the Financial Stability Oversight Council (another creature of DFA) as being systemically important and therefore subject to supervision by the Fed.⁸⁶ In addition, these companies must conduct their own DFA stress tests—the DFAST tests—twice a year. Under the phase-in, the 18 original banks were subject to DFAST in the 2013 planning cycle, and the smaller ones were to be subject to it for the 2014 DFAST cycle. Note that the DFAST tests do not replace the existing CCAR and CapRP tests, but rather supplement them, adding yet another layer of regulatory stress testing—and a particularly useless one at that.

Regulators were at pains to point out that there are “important differences” between DFAST and CCAR. The principal difference comes down to the capital action assumptions involved, which affect the post-stress capital levels: under CCAR, the Fed inputs the banks’

own capital plans; under DFAST, the Fed inputs a set of *standardized* capital plans mandated by DFA to ensure consistency across the banks. Consequently, DFAST one-ups CCAR not just by standardizing the external stress projections, but by standardizing the banks’ capital action plans as well—regardless of what individual banks had actually planned to do(!)—and is therefore even less relevant to the circumstances of each individual bank. One can also say, with confidence, that the DFAST adds absolutely nothing to the CCAR stress-testing process except a much bigger regulatory burden: the supervisory stress tests under DFA add nothing to the bank-run CCAR stress tests, and the DFAST bank-run stress tests with their irrelevant standardized capital plans add nothing to them, either.

In summary: in 2009, the first Fed stress test, the SCAP, was carried out by the Fed itself involving the 19 biggest BHCs. The banks were then required to carry out their own regulatory stress tests by CCAR, and the scope and intensity of these stress tests increased markedly over time. Then came Dodd-Frank and additional regulatory stress tests alongside the CCAR ones: the big banks were now required to carry out DFAST tests twice a year in addition to their annual CCAR tests, and they were required to cooperate with new supervisory stress tests by the Fed; the stress testing remit was extended to other institutions, too. Meanwhile, banks are also subject to their Basel regulatory capital requirements as well. In September 2013, the Fed then announced that banks with assets over \$50 billion will be required to conduct *even more* stress tests under Basel 3, to overlap with the 2014 CCAR/DFAST planning cycles; and banks with assets over \$10 billion will be required to conduct Basel 3 stress tests for the year after. It is no wonder the banks are complaining that all this stress testing is stressing them out.

To add to the bankers’ regulatory burdens, with three weeks to go before their 2014 CCAR submission deadline, on December 16, 2013, the Fed circulated a letter to all banks indicating that it was changing the procedures

for its independent supervisory projections.⁸⁷ Explanation: banks had been working on their CCAR returns for almost a year, under intense pressure from the Fed to ensure they got the “correct” results in their models—that is, results similar to the Fed’s own. Then the Fed turned around and changed its own procedures at the last minute, pulling the rug from under the banks’ carefully prepared capital plans. Bankers were understandably furious. As if to rub salt into the wound, the Fed’s letter also reported projections that were compared to the banks’ previous CCAR projections, and the Fed’s projections were way out of line with the banks’. The Fed projected cumulative loan growth of between 1 percent and 2 percent, but the median BHC projection was minus 7.8 percent, and even the 25th percentile was minus 0.5 percent. Bankers were also highly skeptical of the Fed’s projections—as one said to me, “If bank assets are going to grow during a crisis, it is a strange financial crisis indeed” — but the bankers’ skepticism counted for nothing because the Fed is a law unto itself.⁸⁸

Assessing the Fed’s Approach to Stress Testing

We can evaluate the Fed’s approach against the principles of regulatory stress testing set out earlier. Let’s not stand on ceremony here: the Fed scores an obvious “F” on principles 4 (specificity), 6 (huge burden/counterproductive effects), and 7 (transparency/accountability), so there is no point dwelling on those any further. Let’s consider how well it scores on the other criteria:

PRINCIPLE 1: STRESS TESTS SHOULD BE CONSERVATIVE. At first sight, the Fed’s adverse scenarios seem to score well against this criterion. In his 2009 lecture on the SCAP, Chairman Bernanke stressed that the Fed’s adverse scenario was “appropriately conservative” because the loss estimates “significantly exceed those experienced in past recessions,” including the Great Depression. However, others disagreed. As Tyler Durden sardonically noted: “Here is where the Fed shows its disconnect with reality yet again, as the worst case scenario is already

the optimistic one for several parts of the country,” citing unemployment rates in a number of states as being already higher than the Fed’s worst case.⁸⁹ Another skeptic was Chris Whalen: he did his own independent analysis of the health of almost 7,000 banks across the country in 2009 and concluded that the stress tests might not have been tough enough.⁹⁰ There were also concerns about some of the “specifics.” It was often said, for instance, that the Fed’s scenario was meant to capture a repeat of 2007–2008, but a number of observers questioned whether the Fed had captured the high point of that period, the aftermath of Lehman’s failure. Yves Smith asked if the Fed was

really going to test for a repeat of the 2008 seize up in the repo market. . . . you couldn’t even repo 28-day Treasuries, which in trader land was tantamount to the opening of the sixth seal of the Apocalypse. No one wanted to deal with counterparties, period, even with the most pristine collateral. Yes, banks are widely reported to have much better liquidity buffers than last time round. But the Titanic was also believed to be unsinkable.⁹¹

A similar problem was apparent with the 2013 CCAR results: these projected trading losses for Goldman Sachs and JP Morgan Chase of only \$24.9 billion and \$23.5 billion.

PRINCIPLE 2: THERE SHOULD BE MULTIPLE SCENARIOS. The Fed scores *very* badly by this criterion: it places way too much reliance on a *single* macroeconomic scenario.⁹² Its vulnerability to other very different adverse scenarios—of which there is an infinite number—is then simply unknown. The Fed has blinded itself to these hidden risks, and its stress tests, in practice, pressure the banks to ignore them, too: most bankers will increasingly adopt a compliance mindset and go along with the Fed’s models simply as a compliance matter, and even bankers aware of these other risks will be distracted by the regulatory pressure and undermined by their weaker-willed colleagues. The

“Most bankers will increasingly adopt a compliance mindset and go along with the Fed’s models simply as a compliance matter.”

“We can’t confidently anticipate what the next crisis scenario will be like, but we *can* confidently predict that it will be *different* from what we anticipated.”

Fed thus creates a huge problem of risk blindness across the whole system—and without a shred of doubt, we can say that the Fed is not just overconfident in its models, but recklessly so.

If one lesson stands out from the experience of stress testing everywhere, it is simply this: we can’t confidently anticipate what the next crisis scenario will be like, but we *can* confidently predict that it will be *different* from what we anticipated and, when it hits, the outcome will be anyone’s guess. It will also almost certainly be much worse than the stress tests will have anticipated, not least because of this risk blindness and the inbuilt incentives toward greater systemic instability. Relying on a single scenario is not just bad practice, but a fatal flaw with potentially catastrophic consequences.

In this context we should also recall Chris Finger’s advice on the principles of good stressing: he wrote that stresses are like dietary supplements; they are at best a supplement to one’s diet, and certainly no substitute for it. So what does the Fed do? It relies on a *single* dietary supplement and doesn’t bother about any food! Were a doctor inflicting such quack medicine on a patient, they would be struck off and prosecuted for negligence.

PRINCIPLE 3: SIMPLE MODELS. FEW MODELERS.

One can only laugh or cry at this one: there *are* alternatives to armies of regulatory economic risk modelers—after all, we didn’t need them in the past when the banking system was more stable than it is now. If the Fed *absolutely has* to have regulatory stress testing, however, here is one suggestion:

Every year, each big bank sets up the following competition: it selects a small number of its best risk modelers to build simple desktop stress-risk models of the bank’s risk exposures. The modelers are given any basic parameters, access to any data they need (e.g., historical profit-and-loss for different positions), and details of the stress scenarios. Each modeler is given a week to produce a risk report—such a time frame is perfectly feasible for an experienced modeler working on a simple model. They are incentivized by the prospect of a

prize for the winner and are under orders not to communicate with each other. They are also told to build in explicit margins of prudence in their loss projections. The modelers are free to choose their own model: some might choose Monte Carlo, others some form of stressed historical simulation, and so forth: the main point is that they are free to experiment. They then present their results at a conference where an expert committee would select the “best” risk model. These experts would include experts from the bank and the Fed, but also outsiders, preferably from different disciplines (e.g., banking specialists as well as financial risk modelers, etc.) to give an interdisciplinary perspective. The winning model would then be used to determine the bank’s official risk projections to be reported to the Federal Reserve. These would then be published as they are—with no attempt by the Fed to impose consistency across the different banks’ models—along with the risk projections of the two runner-up models and comments by the judging panel and by independent referees from outside the panel. Publishing the results of the runner-up models and the associated comments would convey a sense of the uncertainty in the projections, allowing analysts to draw their own conclusions about their robustness, and hence the risks, in the risk projections.

This arrangement would beat the Fed system on every criterion I can think of:

- It would avoid the hugely expensive current system and its armies of risk modelers, regulatory modeling manuals, and other requirements; instead, it would be very inexpensive, have a light regulatory footprint, and greatly simplify risk reporting.
- It would produce simple risk models that the complexity literature suggests would be better than more sophisticated models, as well as being much less expensive and quicker to turn out.
- It would rely on competition to produce the best risk models, thus promoting innovation and improved modeling prac-

tice; this is much better than the Fed's command-and-control approach, which gets captured by groupthink and suppresses innovation and discovery.

- It would be transparent and open to outside scrutiny, unlike the current system, in which the Fed refuses to release details of its models that would allow them to be evaluated by outsiders.⁹³
- It would convey a sense of the uncertainty associated with the risk projections, and this is most important because all risk projections are inherently uncertain. The current system conveys no sense whatever of this uncertainty. In fact, the Fed's statements about its projections convey the exact opposite—they project a sense of hubris, a false confidence, even a 'false certainty' about the future—that is potentially fatal for good risk management.

The only limitation of this approach is that no risk modeling is likely to be of much use for banks that are already insolvent: any respectable stress test could only show them even further underwater than they already are. But the current Fed approach is even worse, because it can (and as we shall see, sometimes does) provide a fig leaf of apparent respectability to such banks, instead of highlighting their weakness for all to see.

PRINCIPLE 5: NON-GAMEABILITY. The current regulatory capital regime is riddled with opportunities to game the system. We have already touched on how banks can game both the numerator and the denominator in regulatory capital ratios: the former through the creation of non-equity Tier 1 instruments and the latter through risk-weight "optimization," i.e., reduction. Regulatory capital arbitrage (RCA) is, however, a vastly complicated subject—and one made all the more impenetrable by financial engineers' penchant for maddeningly obscure terminology that makes regulatory language look like the back of a cornflakes box. The underlying idea is to bring financial engineering to bear on any inconsistencies in regu-

latory treatment to achieve the minimum possible regulatory capital charges. This process is very similar to tax arbitrage, in which one arranges one's tax affairs to minimize the tax due. RCA often involves the channeling of bank risks through Special Purpose Vehicles (SPVs), entities created solely to unbundle and repack-age risks in ways that elicit preferential treatment under applicable accounting, regulatory, or tax standards. In essence, an SPV is set up that issues bonds backed by assets. The trick is to tranche the bonds into junior and senior claimants on the underlying assets. The senior tranche is protected by the junior one and can then obtain a very high credit rating. Typically, the senior tranche will have an AAA rating and account for maybe 80 percent of the total issue. The creation of the SPV will then often lead to much lower capital requirements—and you can repeat the process again and again, each time achieving further "optimization" of regulatory capital. Much RCA also revolves around the creative use of financial guarantees and quasi-guarantees to exploit differences in their regulatory treatment. A prescient article by David Jones in 2000 noted that RCA activity "is large and growing rapidly, especially amongst the biggest banks" and Jones observed that, even then, this activity amounted to a silent but very pronounced deterioration in effective capital standards. This was *before* the big explosion in credit derivatives, which greatly increased the scope of RCA even further. Credit derivatives—credit default swaps especially—are a financial engineer's dream come true: by allowing credit risks to be shorted, they create the potential to "complete the market" and arbitrage away virtually anything. The Holy Grail is then within reach: a decent financial engineer can obliterate any set of regulatory risk buckets and reduce the effective risk weights to virtually zero—and have the satisfaction of decapitating the capital regulation without the regulators even realizing it.⁹⁴

The Fed can also game the system itself. When carrying out stress tests, the Fed will be pressured to strike a balance in its model outputs and potentially massage results. If

“No risk modeling is likely to be of much use for banks that are already insolvent.”

“A regulatory stress test to promote confidence is an inherently self-contradictory exercise.”

its models delivered relatively low loss outcomes—for an individual bank or the system as a whole—then people in the Fed are going to say, “This can’t be right: no one will believe this” (assuming, that is, that they even notice the problem: see below). In this case the Fed has a credibility problem, and the temptation will be to massage the parameters of the system to get more plausible results. On the other hand, if its models delivered very high loss outcomes, then the Fed has a big political problem: releasing such loss outcomes would cause uproar in the banking system and potentially shatter confidence. In this case, the pressure would be on to fiddle the loss results downwards to produce something politically acceptable. So whatever its models might otherwise say, there is only a range of possible loss outcomes that the Fed can realistically live with. The danger, then, is that the Fed’s models merely become a vehicle to validate what is essentially the Fed’s own judgment—and one suspects that the Fed modelers are already massaging their models to ensure acceptable results. Again, it would be better to dispense with all the sophisticated modeling and go for some alternative that was much, much simpler and far less costly or counterproductive.

Underlying this latter point is a deeper and intractable problem, one intrinsic to any publicly released regulatory stress testing: the Fed is conflicted. Whereas a private institution might carry out a stress test, find that disaster is imminent, and publish anyway, regardless of the consequences, the Fed has to take account of how anything it says or does would affect confidence. After all, it has an obligation to promote confidence or at least not to undermine it. So even if the Fed *thinks* there is a major problem, it often can’t *admit* it. Everyone knows this, of course, and they discount Fed utterances accordingly: they seek to see through or second-guess what they are saying. One thinks of Paulson and Bernanke telling Congress in July 2008 that Fannie and Freddie were “capital adequate.” No one really believed any of this, but it would have been difficult for the individuals concerned to have said

much else, and people understood that, too. All commentators can really do in such situations is amuse themselves while they watch policymakers squirm around trying to suggest one thing but really meaning another: trying to reassure the markets while hiding their real concerns behind caveats, ambiguities and nit-picking distinctions that they hope won’t get picked up, but would give them the semblance of plausible deniability afterward. My main point, however, is simply this: the credibility of any scenarios produced by regulatory agencies, however good those scenarios might be, is undermined by those agencies’ obligations to protect or promote market confidence or, indeed, by any political agenda they may have or be subject to influence from. The integrity of the regulatory scenarios is undermined by the context in which they are produced and used. There is no escaping this point: regulatory stress scenarios will always have a credibility problem. A regulatory stress test to promote confidence is an inherently self-contradictory exercise.

I take as read all the other ways in which the Fed games the system. In particular, the way it uses the system to appropriate more power and control to itself along classic public-choice lines: we need more resources and more modelers, we need bigger and better models, we need ever more powers over the banking system, and we *don’t* need to be audited or otherwise held to account, as that would just get in the way of our doing our job. There are also the politics to consider, including not just the Fed’s own political agenda (such as a bias toward the big banks, etc.) but also its susceptibility to pressure from the government, all of which incentivizes the Fed to bury politically awkward problems. As we shall presently see, this pressure is one very big reason why the Fed’s stress scenarios are full of holes. Above all, we have to consider the enormous, if usually implicit, pressure on regulatory stress testers not to touch politically sensitive scenarios with a barge pole: as we saw earlier, the Fannie/Freddie stress testers were the last people to suggest that there might be anything wrong with

the subprime boom (while it was still a boom) and the European stress tests have been totally undermined by European regulators' refusal to face up to the scale of their sovereign debt exposures. Stress tests and politics don't mix: the latter inevitably destroys the credibility of the former.

EVALUATING THE FED'S OWN SCENARIOS. It is worth pausing to evaluate the Fed's own scenarios, and it is especially interesting to look at the risks that the Fed ignored or acknowledged late:

- A Eurozone collapse. It has been obvious to any observer since at least May 2010 that there was a serious risk of a blowup in the Eurozone. It was also obvious that a major crisis in Europe could place severe stress on the U.S. financial system and create the potential for another post-Lehman atmosphere or worse. So what did the Fed's scenario do? It *ignored* the risk until the 2012 CCAR. Had Europe blown up, the Fed's scenario would have left the U.S. wide open. One newspaper columnist applauding the Fed wrote that the inclusion of the Eurozone scenario "sent a clear signal that the Fed is not taking any chances with brewing troubles in Europe."⁹⁵ What she should have said was that the Fed had finally woken up and was beginning to take account of the possibility of more trouble in Europe. Well, guys: in stress testing, you are meant to anticipate the risks *in advance*. As the old saying goes: forecasting is a difficult business, especially forecasting the future.⁹⁶
- Counterparty credit failure. This has been an ongoing concern since the failure of AIG in 2008, the concern being whether a bank can survive the demise of an important counterparty. The Fed finally got around to including this risk in the 2014 CCAR. Ditto the risk of a major plunge in leveraged loans, another risk that had been obvious for some time.
- A rise in interest rates. Although the Fed effectively sets short-term interest rates and has managed to keep them down since '09, there has always been a danger that the Treasury bubble would burst, and then interest rates would rise sharply, bringing with them the prospect of the financial system collapsing again. A rising-interest-rate scenario—a risk that the Fed could hardly be unaware of as it had created it itself—was ignored until the 2014 CCAR.
- A rise in inflation. Still not included, although the Fed will doubtless include it when inflation has already started to rise again.
- Nothing on the possibility of a major downgrade or default by the United States, and also a scenario that would be, let us say, a little politically sensitive. The last thing the Fed would want to discover is that the U.S. banking system is highly exposed to Uncle Sam, even if we all knew that already: it is not just the European banks that are exposed to their sovereign.
- And nothing on the biggest threats of all: those posed by the practices of secret hypothecation,⁹⁷ rehypothecation,⁹⁸ and self-securitization:⁹⁹ These are not just huge, but positively eye-watering problems, and they are not even on the Fed economics PhD risk modelers' radar screen because they don't understand them. They are worrying because they allow risk-takers gambling with other people's money the possibility of stratospheric, even infinite, leverage—most of it perfectly legal too.

Meanwhile, the Fed pats itself on the back that its "forward-looking" projections have worked out so well—and assures itself they are credible, too. I would regard it as just lucky—and like any other gambler on a lucky roll, if it keeps going, at some point its luck will run out.

Independent analysts also expressed doubts about the Fed's scenarios. One has to bear in

“Stress tests and politics don't mix: the latter inevitably destroys the credibility of the former.”

“The last thing the Fed would want to discover is that the U.S. banking system is highly exposed to Uncle Sam, even if we all knew that already.”

mind here that the Fed’s approach is dominated by quantitative economics and risk modeling and is therefore subject to the blind spots to which practitioners of those disciplines are prone, including a limited understanding of banking, accounting, legal, and even many financial issues.

The views of independent analysts can then be illuminating. There was, to start with, the enormous accounting mess of the time that allowed banks to switch their assets back and forth between level 2 and level 3, using mark-to-model when it suited them, and resorting to mark-to-myth if that didn’t work. One accounting expert, Greg Hunter, dismissed the entire 2011 CCAR as a farce. As he explained:

Many of the 19 largest banks are sitting on possibly trillions of dollars of mortgage-backed securities (MBS) and underwater real estate. Because of an accounting rule change in April 2009 by the Financial Accounting Standards Board (FASB), “mark to market accounting” is largely ignored. Banks can hold diminished assets on their books at whatever value they think they can get for them in the future. . . . This accounting rule change makes banks look like they’re in much better shape than they really are. How many would be solvent, let alone be boosting dividends, if they had to write down toxic MBS and foreclosed property held at full value on their books? I am sure some banks would be insolvent if they used “mark to market accounting.”¹⁰⁰

Nor was he alone. In a Bloomberg article delightfully entitled “Class Dunce Passes the Fed’s Stress Test without a Sweat,” Jonathan Weil dismissed the 2012 CCAR exercise as “more about public relations and manufacturing confidence than they are about disseminating reliable information on banks’ health.”¹⁰¹ The root problem, he said, was that “to buy into the Fed’s conclusions, you would have to accept the theory that market values

for many types of market instruments don’t matter in a crisis. That would be foolhardy.” A case in point is the class dunce, Regions Financial. This bank passed the Fed’s stress test, and its CEO proudly announced that the Fed’s test “demonstrates the strength of our company.” Er, no it didn’t. As Weil pointed out, if you actually look at the bank’s accounts properly, the footnotes disclosed that the bank’s assets were worth \$8.1 billion less than what their balance sheet said. By comparison, the bank’s common equity, its net worth, was only \$7.6 billion. So the bank was actually insolvent, and the Fed hadn’t picked this up. The Fed had approved the capital plan of an insolvent bank: the stress test was a joke.¹⁰² The only thing keeping the bank going was the expectation that the government or the Fed would continue to prop it up.

Another prominent critic of the Fed’s approach was Chris Whalen. His preferred approach is financial analysis, that is, where you study the business model, turn up the volume, and ask how much capital the business needs to survive, say, a 3-sigma event. Whalen challenged both the methodology and the results of the Fed’s stress tests. As he explained:

The first rule of financial analysis is that you never mix apples and oranges, namely financial analysis and economic guesswork. The task of analytics is to stress the enterprise in an operational sense, then comes the job of trying to relate these observed limits with the outside world. Attempts to plumb econometric analysis into financial models are always bound to create more noise than anything else. But of course the Fed is run by economists and they cannot help themselves when it comes to building models. . . .

So when I look at the Fed stress tests [referring to the 2012 CCAR], which seem to be the result of a mountain of subjective inputs and assumptions, the overwhelming conclusion is that these tests are meant to justify past Fed policy

[with respect to allowing the big banks to pay dividends].¹⁰³

Whalen also suggested that there was an underlying political dimension in which the Obama administration wanted to play down real estate problems in advance of the forthcoming presidential election. He then highlighted some major weaknesses:

- Second lien mortgages: Since many households with two mortgages are underwater and half of residential homeowners cannot refinance, the true value of many second lien mortgages was close to zero. So the \$56 billion in total losses for a 20 percent fall in house prices was way too small. Try \$200 billion.
- First lien mortgages were assigned a \$62 billion loss, but since real estate is half the total \$13 trillion balance sheet of the U.S. banking system and more like three quarters of total exposure if you include residential mortgage-backed security, how does the Fed manage to keep total real estate losses below \$150 billion in the stressed scenario? Again, the Fed party line is that there is no problem with real estate.
- The Fed's projections ignored the large likely losses associated with foreclosure issues and the litigation cases that were already working their way through the court system.
- The Fed "almost completely ignores" the "vast financial risk" facing Bank of America and Ally, and to a lesser degree, Wells Fargo, JP Morgan Chase and, of course, Citi.

Whalen was even more scathing of the 2013 CCAR when its results were published in March 2013. In a devastating critique, he not only took issue with the CCAR, but with QE and Basel as well. His comments nicely pull together a number of the issues we have been talking about and are worth quoting at some length:

With the U.S. economy entering another speculative phase c/o quantitative easing, Wall Street is jubilant at the prospect of further capital and loan loss releases by the banks. But be careful what you wish for in this case. We have all seen this movie before.

The Fed and other regulators have made an enormous fuss about raising bank capital levels over the past several years. Yet the stated intent [of the CCAR and the CapPR] . . . is to allow increased return of capital to investors. This public evidence of schizophrenia on the part of regulators goes largely unnoticed in the press. . . .

Like utilities, the banks are anxious to increase the payout on common equity to appease unhappy investors. And the Fed is happy to facilitate . . .

The curtailment of OBS leverage focused on housing is significant for a number of reasons. First, driven by the idiotic Basel III framework, the large banks must now focus even more attention on OTC derivatives and structured products as they retreat from traditional business lines like residential and commercial mortgage lending.

Unless the loan is subsidized by Uncle Sam . . . the big banks don't want to know. The Basel III risk weights for mortgage lending are so severe that they will literally force the largest banks to withdraw from the bottom half of the U.S. mortgage sector. Nowhere in the CCAR capital adequacy scenario will you see any discussion of legacy mortgage risk and litigation, this even though the instructions for the tests specifically require banks to focus on risks 'not explicitly covered in information requested' . . .

Nor will there be any discussion of the possibility that a new, Fed-induced bubble in the stock market will result in higher losses to banks via derivative exposures over the time frame of the CCAR stress scenarios. Remember, the

“The Basel III risk weights for mortgage lending are so severe that they will literally force the largest banks to withdraw from the bottom half of the U.S. mortgage sector.”

“A key part of the problem with the Fed stress tests is that, despite the complexity of the process, they were still unable to address many of the risk concerns credibly identified by outside observers.”

net effect of Basel III and Dodd Frank is to make the TBTF [“too big to fail”] banks even more dependent upon derivatives, investment banking and capital markets business lines for profits than ever before.

The idea that the Fed will allow higher capital payouts by large banks illustrates the grotesque situation in Washington when it comes to bank regulation. Weak profitability and slow revenue growth should be the key areas of concern in the CCAR analysis, but there will be no discussion of these factors. Nor will banks be asked to model their forward capital needs in a ‘normal’ interest rate environment. The Fed is currently subsidizing the cost of funds for banks via QE to the tune of about \$100 billion per quarter, but we won’t talk about that either.

Instead, the Fed allows the banks to manipulate measures such as “risk weighted assets” to allow for greater short-term payouts to investors. . . . Even as regulators avert their eyes from issues like OTC derivatives, festering second lien exposures and the prospective re-default of modified mortgages, the Fed will allow the largest TBTF banks to pay out ever greater portions of their capital to yield-starved investors. . . .

. . . As Morpheus said to Neo in the film *The Matrix*: You still think that is air you are breathing?¹⁰⁴

In short, a key part of the problem with the Fed stress tests is that, despite the complexity of the process, they were still unable to address many of the risk concerns credibly identified by outside observers, especially those operating from outside the straightjacketed quantitative economic risk paradigm that dominates Federal Reserve thinking, with its fixation on macroeconomic scenarios and its inability to handle much else. However, the reality is much simpler: the Fed’s scenarios are *bound* to miss major risks that may or may not transpire but that are still real risks (or should

I say, uncertainties) *ex ante*. It would be unreasonable to expect anything else—and *this* is the main reason why we do not want to expose the whole banking system to the inadequacies of the Fed’s models.¹⁰⁵

THE FED’S STRESS TESTS: A SUMMARY. It is probably helpful if we now summarize the main problems with the Fed’s stress tests, any *one* of which, on its own, is enough to torpedo the whole exercise.

- The Fed’s stress tests ignore well-established and potentially fatal weaknesses in risk modeling, most notably their vulnerability to being gamed, their tendency to report diminishing risks as true risks are building up, and their tendency to blind risk managers to the true risks they face as the models crowd out critical thinking.
- The stress tests blatantly violate the core principles of good-practice stress testing: in particular, they ignore the principle that there should be multiple simple scenarios, not just one or two highly detailed scenarios. They also ignore the principle that stress tests should be used only as a complementary tool to other risk management methods, not as the principal tool, and certainly not the only tool, due to their inherent weaknesses.
- The Fed’s approach to stress testing is overly prescriptive and suppresses innovation and diversity in risk management practices; in so doing, it exposes the whole system to the weaknesses in the Fed’s approach and greatly exacerbates systemic risk.
- The Fed’s stress tests impose a huge and growing cost on banks and are counterproductive in their implementation, often undermining the objectives they seek to promote, not least the promotion of good risk management practices.
- The Fed’s approach is unreliable because it relies on data that are unreliable (e.g., mark-to-model valuations and other weaknesses in the accounting standards)

or inappropriate (e.g., applying industry-loss averages to individual banks).

- The Fed's approach is also unreliable because it uses unreliable metrics—in particular, it uses risk-weighted assets in the denominator of the regulatory capital ratios. Given that the risk weights are highly gameable, the denominator and hence the regulatory capital ratios, are at best highly misleading and at worst close to meaningless as indicators of banks' financial health. In this sense, the whole regulatory system, stress tests and all, is a very big edifice built on sand.
- The Fed's stress-test scenarios fail to address a considerable number of major risks credibly identified by independent experts, in large part because they are dominated by a quantitative macroeconomic mindset that has only a very limited understanding of other relevant considerations, for example, accounting and banking issues.
- The Fed's approach ignores the contradiction between the stress tests themselves and the use to which they are put—that is, to promote confidence. Promoting confidence is inevitably politicized, and this politicization undermines the tests' credibility: thus, you cannot use stress tests to credibly promote confidence. Put simply, politics and stress testing don't mix.
- The Fed's stress tests fail to embody any lessons learned from the astonishing failures of other regulatory stress tests—most notably, that all (or almost all) of them failed to anticipate the subsequent build-up of the stresses they were meant to anticipate, sometimes with catastrophic consequences.

THE REGULATORY ONSLAUGHT ON THE BANKING SYSTEM

We should also consider the enormous and rapidly growing cost of all the regulation under which banks are compelled to operate. The

Privacy Act requires hundreds of millions of useless notices sent out to clients that no one reads; Sarbanes-Oxley created a redundant and highly counterproductive system that forced banks, and especially senior management, to expend vast amounts of time and effort to cover their backs, lest paperwork errors land them in jail; then there is the USA PATRIOT Act, which never caught and convicted a single terrorist but costs the banking industry more than \$5 billion a year in a completely useless regulatory compliance while it tramples civil liberties underfoot as well; and there is the most oppressive financial legislation of all, Dodd-Frank, aptly described by Peter Wallison as “the most troubling—maybe even destructive—single piece of financial legislation ever adopted.”¹⁰⁶

It is helpful to look at the basic metrics—namely, the length of the rulebooks and the total cost, as much as the latter can be assessed at all. The original Federal Reserve Act was 31 pages long, Glass-Steagall was 37 pages, Sarbanes-Oxley spread out to 66 pages, and Dodd-Frank is a whopping 848 pages long—over 20 times longer than Glass-Steagall. Yet the size of Dodd-Frank is only the beginning: unlike its predecessor laws, which classically provide people with rules, Dodd-Frank is merely an outline directed at bureaucrats that instructs them to create still more regulators and even more bureaucracies. It is expected that the eventual Dodd-Frank rulebook will run to perhaps 30,000 pages long. As for the cost of all this regulation, especially of Dodd-Frank, the bottom line is that no one really knows, but a two-year-on interim assessment by S&P in 2012 suggested that the cost of DFA to the eight largest banks might be \$22 billion to \$34 billion a year, equivalent to a drop in pretax return on equity of 250 to 375 basis points for the biggest banks.¹⁰⁷ Given that these are merely interim assessments and most of the rules have still to be determined, it is fair to say that these cost estimates, enormous as they are, could still be below the eventual total cost.

This regulation creates serious distortion and associated resource misallocations. Con-

“Given that the risk weights are highly gameable, the denominator and hence the regulatory capital ratios, are at best highly misleading and at worst close to meaningless as indicators of banks' financial health.”

“Banks were not designed to be low-income lenders, but that is what they were forced to become.”

sider just two examples:

- The use of fair lending laws (and formulas based thereon, more models) forces banks to make loans to unqualified borrowers to increase home ownership, a political objective—and that one that backfired badly as we now know. Banks were not designed to be low-income lenders, but that is what they were forced to become. Subprime was then a moral imperative. However, the underlying driver was a political agenda. One also has to bear in mind the indirect damage as well, especially the way in which the meddling of the government and its GSEs enabled creation of mortgage-backed securities in the first place and then drove down standards across the whole industry, distorting everything.
- The regulatory treatment of small loan decisions has changed dramatically. In the old days, homeowners wanting to extend their kitchen or an entrepreneur with a business idea would ask for a loan, and an experienced loan officer with a lot of business and local-specific knowledge would make an informed decision—and most of these loans paid off. There wasn't much paperwork either. But now these decisions have to be put through a computer model to score the application, and there is much less scope for the exercise of judgment. The associated paperwork is huge, and any errors in the paperwork can lead regulators to challenge not just the loan, but the entire portfolio. As a result, loan decisions are much more centralized, there is a lot more paperwork, and many entrepreneurs with good venture capital ideas are being turned down because they do not meet the rigid lending standards of the models, even though loan officers might be happy to lend to them. Small business lending is now at a 40-year low.

There are also many other ways in which

the government meddles with the banking industry and has done for a long time. To quote Allison: “In theory, CEOs report to boards, who then report to shareholders. While that's true of most businesses, in the financial services industry, we only quasi-report to boards, quasi-report to shareholders, and definitely report to regulators.”¹⁰⁸ This requirement to report to the regulators is a key avenue by which the banks get pressured by the prevailing political winds: lend more to ethnic minorities, support affordable housing, and so on. Bankers have sometimes been subject to regulatory harassment, too. A case in point concerned a Boston Fed study of 1992 that concluded that there was a lot of racial discrimination still going on in mortgage lending.¹⁰⁹ Bill Clinton was elected president shortly afterward, and he and his cronies were convinced that the allegations were true. The government was then determined to “get the banks” on this issue, and a sustained regulatory attack followed. Most banks settled out of court—it was easier to admit the charges, pay a small fine, and move on, than to oppose them, even though they were unfounded. One bank refused, however, and the regulators responded by siege warfare, blocking approval for its mergers and acquisitions. The harassment only stopped when the balance of political power shifted back and the Republicans recovered control of Congress in 1994. It turned out that the study alleging race discrimination was discredited soon after it was published, but the damage had been done and no one was brought to account for it.¹¹⁰

This state of affairs deteriorated further in the Bush years with the extraordinary, indeed, extralegal regulatory powers doled out under Sarbanes-Oxley and the USA PATRIOT Act. To quote Allison again:

The intense focus from the regulators—especially on Sarbanes-Oxley and the Patriot Act—dramatically misdirected risk management in the financial industry. Regulators were threatening to put CEOs in jail and levy large fines on board members, which impacted our behavior

radically, and made us put a lot less focus on traditional risk management. I guarantee this happened across the whole industry.¹¹¹

One has to understand that the banks have no defense against this regulatory onslaught. There are so many tens or hundreds of thousands or maybe millions of rules that no one can even read them all, let alone comply with them all: even with armies of corporate lawyers to assist you, there are just too many, and they contradict each other, often at the most fundamental level. For example, the main intent of the Privacy Act was to promote privacy, but the main impact of the USA PATRIOT Act was to eviscerate it. This state of lawlessness gives ample scope for regulators to pursue their own or the government's agendas while allowing defendants no effective legal recourse. One also has to bear in mind the extraordinary criminal penalties to which senior bank officers are exposed. Government officials can then pick and choose which rules to apply and can always find technical infringements if they look for them; they can then legally blackmail bankers without ever being held to account themselves. The result is the suspension of the rule of law and a state of affairs reminiscent of the reign of Charles I, Star Chamber and all. Any doubt about this matter must surely have been settled with the Dodd-Frank Act, which doled out extralegal powers like confetti and allows the government to do anything it wishes with the banking system.

A perfect example of this governmental lawlessness was the "Uncle Scam" settlement in October 2013 of a case against JP Morgan Chase, in which the bank agreed to pay a \$13 billion fine relating to some real estate investments. This was the biggest ever payout asked of a single company by the government, and it didn't even protect the bank against the possibility of additional criminal prosecutions. What is astonishing is that some 80 percent of the banks' RMBS had been acquired at the *request* of the federal government when it bought Bear Stearns and WaMu in 2008, and now the

bank was being punished for *having* them. Leaving aside its inherent unpleasantness, this act of government plunder sets a very bad precedent: going forward, no sane bank will now buy a failing competitor without forcing it through Chapter 11. It's one thing to face an acquired institution's own problems, but it is quite another to face looting from the government for cooperating with the government itself.

The government takeover of the banks is now effectively complete: the banks are vassals of the state in all but name. Peter Wallison put this wonderfully in an op-ed immediately after the last presidential election.

The Reuters headline yesterday said it all: "Wall Street Left to Rebuild Obama Ties after Backing Romney." And so it begins. The government has become so powerful in the financial services field that private sector firms now have to "rebuild ties" after an election to avoid adverse rulings from their regulators.

If you are worried about crony capitalism, this is where it starts. Because of Dodd-Frank, Wall Street and the financial services industry generally will now be firmly in the control of the government. In the future, as long as the act remains in force, we can expect that Wall Street firms will be solid supporters of the administration in power. No CEO will risk the possibility that opposing administration policy will bring an adverse regulatory finding or an enforcement action.

To make matters even worse:

However, that isn't all. Under Dodd-Frank, if the secretary of the treasury believes that a financial firm in danger of failing could cause instability in the U.S. financial system, he has the power to seize the firm and turn it over to the FDIC for liquidation. If the company objects, the secretary can invoke the power of a court, but the court has only

“The government takeover of the banks is now effectively complete: the banks are vassals of the state in all but name.”

“Congress needs to pass a Prohibition of Regulatory Financial Modeling Act that would prohibit the use of any financial models by any financial regulators for any purposes of financial regulation.”

one day to decide whether the secretary’s act was reasonable. If the court does not act in that one day, the firm is turned over to the FDIC “by operation of law.” It does not take much political savvy for financial firms to realize that opposing the secretary of the treasury could be dangerous to their continued existence.

And worse again:

In addition, the Financial Stability Oversight Council, a new agency created by Dodd-Frank and headed by the secretary of the Treasury, is authorized to designate any financial firm as systemically important if the firm’s failure could cause instability in the U.S. financial system. The designation means that the firm is turned over for “stringent” regulation by the Federal Reserve, which has the power to control its capital, leverage, liquidity, and activities. . . .

What all this means is that in the future very few financial firms will be plaintiffs in actions against what they believe are illegitimate government regulations [witness the Uncle Scam scam], and when a regulator or, worse, the secretary of the treasury, calls to ask for support of an administration initiative he or she can be sure of a smart salute and a full-throated “aye-aye, sir.” Not only will that cooperation forestall an adverse regulatory action, but it will probably mean some administrative “flexibility” when the firm wants to make a controversial acquisition. And that, ladies and gentlemen, is crony capitalism to its core.¹¹²

All that is really missing at this point is for the government to apply the oath of fealty and rent out accommodation at the Tower of London.

WHAT SHOULD BE DONE?

The ideal solution to all these problems is to restore the rule of law and establish a genu-

ine free market in money and banking. We need to get rid of Dodd-Frank and other restrictive legislation holding down the banking system (including Sarbanes-Oxley, the Community Reinvestment Act and so on), abandon Too-Big-to-Fail, abolish bank capital regulation and federal deposit insurance, restore the gold standard, end the Fed and all the big regulatory agencies, rescind limited liability legislation, and remove all restrictions against private money or banking—and restore the rule of law. This is an ambitious and radical reform package, but it *would* work: Canada had a system that was close to this model during most of the 19th and part of the early 20th century, and it worked well. It was also much admired in the United States. It was unfortunate that the United States didn’t copy it; instead, it established the Federal Reserve and then one regulatory body after another over the course of the next century.

The banking sector’s immediately pressing problem, however, is that of carnage by computer: what to do about the Fed’s regulatory risk modeling, which is sleepwalking the banking system toward disaster and taking the country with it. The answer is simple: Congress needs to pass a Prohibition of Regulatory Financial Modeling Act. The act should be based on the following principles:

First, it would prohibit the use of any financial models by any financial regulators for any purposes of financial regulation, period. The Fed and other financial regulators would then be required to abandon the regulation of financial intermediaries based on any risk modeling. Since the Basel system is wedded to such modeling, this provision would compel the United States to withdraw from the Basel regime. In any case, there is no point haggling with other countries on the matter within the context of the Basel system; instead, if other countries wish to continue regulatory risk modeling, the United States should let them get on with it. Note also that this provision would apply not only to risk models, but to *any* financial models as defined above. This would put a stop to CRA modeling or dodgy regres-

sions looking for nonexistent race discrimination and, indeed, to any other regression-based models of this, that, or the other that regulators or interventionist social engineers think the banks should be attending to instead of attending to their core business. The banks would then be left free to make business decisions at their own risk, using their own judgment, free of outside meddling.

To any critics who might object that these provisions are so severe toward the Fed that they would force it to abandon all complex bank regulation, I would simply say: this is exactly the point. We need a *simple rule* that cannot be circumvented or filled full of holes: we do not need a repeat of the Volcker Rule fiasco with its 600-plus page rulebook and myriad gameable exemptions that render it totally ineffective.

Regulators' ability to perform calculations would then be very limited, which consideration takes us to the second principle: that if the United States is to have any financial regulation at all, then financial regulation should be extremely simple. It should be simple enough for a small number of regulators to do—preferably working by candle-light in a dark scriptorium and no central heating—armed only with pencil and paper and a few filing cabinets.

So what should the capital regime be? I would suggest that it should be as simple as possible, but also be conservative, robust to gaming, and have a light footprint. With these aims in mind, I would suggest a minimum capital ratio of 15 percent, which was approximately the banks' average operating capital ratio at the time the Fed was founded.¹¹³ This capital standard would be based on the simplest and least gameable metrics: the numerator, equity, would be common equity, not the gameable Tier 1 equity used in Basel II; the denominator, assets, would be GAAP (Generally Accepted Accounting Principles)-valued assets with no risk-weighting, so avoiding the gameability to which risk-weighting is prone.¹¹⁴ Calculating these capital ratios would involve a few calculations rather than a few million. The information used to determine the capital ratio would come from the banks' audited

accounts. All the regulator would have to do is find the two numbers from the accounts and divide the one by the other; the only challenge involved is long division. There would be an annual cycle based on the publication of the banks' accounts.

We then need an implementation process. The key points are automatic rules that leave no wiggle room for discretion and a strong incentive for bankers to look after their banks. Once the system had been phased in, I would suggest that any bank that failed to meet this standard would face automatic penalties. There would be no regulatory discretion in the system—either a bank meets its minimum standard or it does not. If it does, fine; if it does not, it would be prohibited from making any capital distributions or bonus payments, and its directors would go on minimum wage until the bank had met its minimum capital standard again: they would be put on the equivalent of water and gruel.

I would also extend directors' personal liability: we have to give them a strong incentive to look after their banks properly—and the minimum regulatory capital ratio would be supplemented with a solvency standard. I would suggest a 3 percent ratio of common equity to total assets. Any bank that fell below this ratio would be automatically placed into receivership, and bank directors would be first in line to pick up any losses involved. Of course, one could also envisage less draconian implementation mechanisms, but it is essential that they are simple and minimize scope for discretion and hence regulatory capture; otherwise, they lose their effectiveness.¹¹⁵ We should never underestimate the ability of Wall Street to game the slightest ambiguity or to hijack any consultation process until it has rendered useless any rules that constrain its ability to generate short-term profits: the fate of the Volcker Rule amply illustrates this point.

CONCLUSION

The Fed's obsession with math modeling is creating a huge and growing risk for the

“We need a *simple rule* that cannot be circumvented or filled full of holes: we do not need a repeat of the Volcker Rule fiasco.”

“The Fed asks us to believe that it alone, out of all the regulatory stress testers, will be the one to get it right—and this despite the Fed’s own disastrous forecasting record.”

U.S. financial system. Yet even though regulatory risk modeling was a key factor behind the weakening of the banking system that brought on the crisis, and even though regulatory stress testing has produced disaster after disaster in recent years, the Fed asks us to believe that it alone, out of all the regulatory stress testers, will be the one to get it right—and this despite the Fed’s own disastrous forecasting record, not least its total failure to anticipate the 2007–2009 financial crisis. The reality is that regulatory risk modeling doesn’t work in the long run. It didn’t work with Fannie and Freddie, it didn’t work in Iceland, it didn’t work in Ireland, it didn’t work in Cyprus, and it didn’t work across much of the rest of Europe, either. There is no reason to think that the Fed’s risk model will be any exception. A betting man would therefore describe the odds of the Federal Reserve leading the banking system into another disaster as a racing certainty. When the postmortem is over, the main damage will turn out to have been done by some risk that the Fed’s stress tests completely missed or, at best, greatly underestimated. That’s the way these things always are. But what clinches it for me is the fact that the Fed is so supremely *confident* that its stress tests will protect the banking system.

The best insights into the future come not from math modeling but from ancient Greek literature, which reiterates again and again the fates of those who were foolish enough to defy the gods. A case in point is an experienced sea captain who once said:

When anyone asks me how I can best describe my experiences in nearly 40 years at sea I merely say uneventful. Of course, there have been Winter gales and storms and fog and the like, but in all my experience I have never been in any accident of any sort worth speaking about. I have seen but one vessel in distress. . . . I never saw a wreck and have never been wrecked, nor was I ever in any predicament that threatened to end in disaster of any sort. . . .

I will go a bit further. I will say that I cannot imagine any condition which could cause a ship to founder. I cannot conceive of any vital disaster happening to this vessel. Modern shipbuilding has gone beyond that.¹¹⁶

These were the immortal words of E. J. Smith, the captain of the *Titanic*. However, one has to admit that the *Titanic* comparison is perhaps a little unfair—at least to Captain Smith. After all, he hadn’t ever encountered a crisis, and he did have good reason to be confident in his ship, whereas we have every good reason to believe that the Fed’s risk regulatory modeling agenda will end up underwater. So when the Federal Reserve *Titanic* eventually hits its iceberg, we can only hope that its captain will take a leaf out of Captain Smith’s book and have the decency to go down with her ship.

NOTES

1. John A. Allison, “Market Discipline Beats Regulatory Discipline,” *Cato Journal* 34, no. 2 (Spring/Summer 2014): 345.
2. Alan Greenspan, testimony given at the Federal Reserve’s Semiannual Monetary Policy Report Before the Committee on Banking, Housing, and Urban Affairs, U.S. Senate (July 22, 1997).
3. “Bernanke: Why Are We Still Listening to This Guy?” YouTube video, 5:07, posted by Paul Williams World, June 16, 2009, <http://www.youtube.com/watch?v=HQ79PtzGNJo>.
4. A nice compilation of these and other Bernanke gaffes is available at the Center for Economic and Policy Research, “Federal Reserve Board Chairman Ben Bernanke’s Greatest Hits,” <http://www.cepr.net/index.php/bernanke-greatest-hits>.
5. “Ben S. Bernanke: GSE Portfolios, System Risk, and Affordable Housing,” BIS Review 22 (2007), <http://www.bis.org/review/r070308a.pdf>.
6. Ben S. Bernanke, “The Subprime Mortgage

Market,” speech given at the Federal Reserve Bank of Chicago’s 43rd Annual Conference on Bank Structure and Competition, Chicago, Illinois (May 17, 2007), <http://www.federalreserve.gov/newsevents/speech/bernanke20070517a.htm>.

7. Jeannine Aversa, “Fed Chief to Slash Rates ‘As Needed’ to Avoid Recession,” Associated Press, January 11, 2008, <http://www.sfgate.com/business/article/Fed-chief-to-slash-rates-as-needed-to-avoid-3231862.php>.

8. “Fed Chairman: Some Small US Banks May Go Under,” CNBC.com, February 28, 2008, <http://www.cnbc.com/id/23390252>.

9. “Bernanke Tries to Calm Nerves,” July 16, 2008, WABC-TV, <http://abclocal.go.com/wabc/story?section=news/business&id=6266855>.

10. There is also a fourth difficulty—one is implicitly assuming that the “world stands still” and, in particular, that the statistical loss distributions we face do not change, and this assumption never holds true.

11. When dealing with financial risk, we are not interested in small losses—these don’t matter—but in the possibility of big losses that could put us out of business: we are therefore concerned with the downside risk represented by the more extreme tails of the loss distribution. However, we have relatively few observations in the tail region, and this makes fitting the tail highly problematic—and all the more so because we are often interested in the tail not just within the range of observations that we have, where observations are sparse, but in the tail well beyond it, where they are nonexistent. This problem is akin to extrapolating out into space from a relationship estimated on a few observations in the upper stratosphere.

12. Markets often experience very high sigma events—sigma being an estimate of the volatility or standard deviation of the market price—that are way too high to be compatible with Gaussianity. These include the crash of October 19, 1987, which was a 22-sigma event, and the laughable 25-sigma

events (and several days in a row, too!) that Goldman Sachs CFO David Viniar claimed their hedge funds experienced in August 2007. How unlikely are these? Under Gaussianity, a single daily 20-sigma event has a waiting period in years that greatly exceeds the higher NASA estimates of the number of particles in the entire universe—so pretty darn unlikely—and the waiting period for a 25-sigma event is a quinquicillion—that is, a trillion trillion trillion trillion times this number.

13. It is very easy to construct plausible examples showing how the Gaussian can grossly underestimate true risks. For instance, if we assume a familiar standard Gaussian when the world is actually governed by a standard Cauchy distribution, which has a much fatter tail, then the Gaussian Value at Risk (VaR) estimate could easily be 25 percent of the true VaR. If we made this same mistake in hydrology, we would be building seawalls that are only a quarter of the height they should be.

14. The distribution of extremes is *not* governed by the familiar Central Limit Theorem that is usually used to justify Gaussianity. The Central Limit Theorem says that if we draw from a sample of size n from an unknown distribution, then under general conditions, as n gets large, the sample mean approaches the Gaussian. However, the Gaussian only applies if we are concerned with the central mass of the distribution. By contrast, extremes are governed by the Extreme Value (EV) Theorem of Fisher and Tippett and are very non-Gaussian. See R. A. Fisher and L. H. C. Tippett, “Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample,” *Mathematical Proceedings of the Cambridge Philosophical Society* 24 (1928): 180–90.

15. This story of government involvement in this mess is well documented: see Johan Norberg, *Financial Fiasco: How America’s Infatuation with Homeownership and Easy Money Created the Economic Crisis* (Washington: Cato Institute, 2009); Kevin Dowd and Martin Hutchinson, *Alchemists of Loss: How Government Intervention and Modern Finance Crashed the Financial System* (Chichester, UK: Wiley, 2010); Patric H. Hendershott and Kevin E.

Villani, "The Subprime Lending Debacle: Competitive Private Markets are the Solution, Not the Problem," Cato Institute Policy Analysis no. 679, June 20, 2011; and John A. Allison, *The Financial Crisis and the Free Market Cure* (New York: McGraw Hill, 2013).

16. A major problem is that the VaR is not subadditive: subadditivity is the property whereby the risk of the sum of two or more risks is always no greater (and typically less) than the sum of those risks. Subadditivity is the property underlying portfolio diversification, that is, that risks put together typically offset each other, and it is the most basic property we would want of any "respectable" risk measure. The VaR's failure to be subadditive can then wreak havoc in situations where the VaR is used for delegated risk management because the delegated risk manager can exploit non-subadditivity to leave the delegating party exposed to the residual hidden risk. See Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath, "Coherent Measures of Risk," *Mathematical Finance* 9, no. 3 (1999): 203–28.

17. J. Berkowitz and J. O'Brien, "How Accurate Are Value-at-Risk Models at Commercial Banks?" *Journal of Finance* 57, no. 3 (April 2002): 1093–111.

18. At the centre of most of these models are huge covariance matrices that are not only difficult to calibrate, but inevitably slow to adjust to market developments. In order to produce their VaR estimates, they have to satisfy certain conditions, of which the most important is that the covariance matrix be negative semi-definite: this condition ensures that the portfolio variance is non-negative, and if it does not hold, the computational algorithms will stall. Major shortcuts are then required to ensure it holds: an example is the common use of highly restrictive Exponentially Weighted Moving Average (EWMA) processes applied across all risk factors, instead of more flexible processes such as multivariate GARCH processes. There is also a need to "map" individual positions to their equivalents in the space of underlying risk factors, which cannot be too large without exploding the covariance matrix. A sim-

ple example might be where a 15-day U.S. Treasury is mapped to an approximate equivalent in terms of 1-day and 30-day Treasuries. Mappings are inevitably inaccurate, especially so when dealing with derivatives such as options. For options, banks often use first- and second-order approximations, and the latter especially can be very inaccurate and far from robust. For more on these and other implementation issues, see Kevin Dowd, *Measuring Market Risk* (Chichester, UK: Wiley, 2005). The Berkowitz–O'Brien results would seem to suggest that all these implementation problems produce so much inaccuracy that they more than undermine any advantage of the sophisticated models over much simpler alternatives.

19. See Gordon Kerr, "The Case for Systemic Banking Reform," IREF Working Paper no. 201401 (March 2014), pp. 15–17.

20. Andrew G. Haldane, "Capital Discipline" (speech, American Economic Association, Denver, Colorado, January 9, 2011), chart 3, <http://www.bankofengland.co.uk/publications/Documents/speeches/2011/speech484.pdf>.

21. The figures quoted in the text are taken from Pablo Triana Portela and Vasiliki Kosmidou, "VaR and the Crisis: Empirical Analysis," *Corporate Finance Review* (September/October, 2012); and *Financial Times* (London), February 7, 2008.

22. This episode is taken from Nicholas Dunbar, "How Citi Sank Itself on the Fed's Watch," Reuters, November 30, 2011, <http://blogs.reuters.com/great-debate/2011/11/30/how-citi-sank-itself-on-the-feds-watch/>.

23. John Maynard Keynes, *Essays in Persuasion* (New York: W. W. Norton, 1963), p. 176.

24. Friedrich A. Hayek, "The Pretence of Knowledge," (Nobel Prize lecture, 1974), http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1974/hayek-lecture.html.

25. The response of the econo-physics crowd to these profound problems is to ignore them. A

prime example is the otherwise good book *Theory of Financial Risks*, by Jean-Philippe Bouchaud and Marc Potters, both PhD physicists (Cambridge: Cambridge University Press, 2000). The subtitle of their book says it all: “from statistical physics to risk management.” If only risk management were so simple!

26. A major difference between physical and social models is that the former ignore the ways in which thinking agents react to and try to anticipate each other. This points to perhaps the most important problem with the conventional risk modeling paradigm, which attempts to model market processes by treating them as if they followed “laws of motion” comparable to those in nature. The underlying assumption here is that one is in a “game against nature”: the intelligent human agent interacts with nature, and nature responds predictably and unintelligently in accordance with its own laws that the human understands. However, the physical “game against nature” is a poor analogy for many important problems in economics, finance, and society generally. Instead, many of these problems are characterized by strategic interaction or games and can only be understood if one takes account of how intelligent agents interact. Traders pit their wits against each other all the time, and it is these games of wits that determine market prices—not the statistical math processes that dominate the quantitative finance textbooks.

27. Stan Jonas, quoted in *Derivatives Strategy* 3, no. 4 (April 1998): 20.

28. These problems are compounded further when booms go hand-in-hand with rapid financial innovation. Financial innovation “holds out the promise of better risk management and, at the same time, stacks the deck against disproving this proposition,” as the authors succinctly put it; see Claudio Borio, Mathias Drehmann, and Kostas Tsatsaronis, “Stress Testing Macro Stress Testing: Does It Live up to Expectations?” BIS Working Paper 369 (November 29, 2011), p. 11. This is a similar mistake to the one that Alan Greenspan made in the mid-1990s, when his misreading of the tea leaves led him to infer that his nonexistent

productivity miracle was under way.

29. Borio et al., “Stress Testing Macro Stress Testing,” pp. 11–12.

30. Dowd and Hutchinson, *Alchemists of Loss*, p. 106.

31. In fact, we can also look at this from a fourth perspective and still get the same answer. From the perspective of organizational theory, there is a tension between tightly coupled and loosely coupled systems. Much risk management—one thinks in particular of Enterprise-wide Risk Management (ERM)—leads to more tight coupling, which means greater efficiency in stable conditions but greater vulnerability in unstable conditions—and this is exactly what we got. From this perspective, the ERM project is a doomed enterprise. Similarly, organisms that are highly adapted to the specifics of their environment tend to thrive when those environments are stable, but are prone to become extinct when the environment changes. Over time, the organisms that survive are less specialized but more robust.

32. Roger J. Brown, personal correspondence with author.

33. One of the bankers interviewed for this article gave me an amusing account of his bank’s experience with Federal Reserve standardization. His bank had little need for models because it took few risks. The risk models it then submitted to the Fed under the risk supervisory process used its own loss experience, which was much lower than the industry average. However, supervisors rejected their models and demanded that the bank use more sophisticated models and the industry loan loss experience instead of its own. This requirement made no sense at all. It saddled the bank with an unnecessary model development cost of hundreds of millions and a higher regulatory capital charge—and it will force the bank to take more risk to pay for the extra costs involved. In short, in the interests of promoting good risk management and discouraging excessive risk-taking, the Federal Reserve forced a well-run bank to

adopt highly expensive risk management technology that it neither wanted nor needed, imposed higher regulatory capital requirements that were not justified by the risks the bank wanted to take, and will then force the bank to take extra risks that it *didn't* want to take in order to recoup its higher costs!

34. Til Schuermann, "The Fed's Stress Tests Add Risk to the Financial System," *Wall Street Journal*, March 19, 2013.

35. The Bank of Spain introduced just such a measure in the late 1990s. This took the form of a capital requirement on loans that took account of expected losses at the time the loans are entered into. These expected losses were calibrated on past recession loss experience. The Spanish experiment was not successful—they took data from the past three wimpy recessions and the latest one blew through past parameters—but it was still better than what other regulatory bodies were doing.

36. Well, there is *one* way to achieve anticyclical regulatory capital requirements with procyclical models, but I imagine that regulators might have some reluctance to go that route: one could come up with a formula in which capital requirements move *inversely* with the risk numbers—and the obvious way to achieve that objective would be to put the risk number in the denominator rather than the numerator of the capital ratio! It would then be fun to see how the regulators might justify a capital regime that only worked thanks to its own perversity.

37. The only alternative to achieve anticyclical capital charges is to base those charges on a numerator variable that rises in the boom and falls in the bust. A natural choice is bank profits. In fact, this idea merely reflects the venerable principle articulated in the advice of the prophet Joseph to Pharaoh: store up in the good years to be ready for the lean years that will follow. This approach also has the attraction that it does not just allow for prudence: it is, indeed, the very embodiment of it. Another attraction is that it does not require any

risk model; all it requires is a reliable and not too gameable measure of profits. Having the regulatory capital regime based on profits rather than risk modeling is therefore not a bad idea at all. Needless to say, it never made it into Basel or any other regulatory capital regime.

38. One thing is for sure: if you didn't have regulation already, you wouldn't rationally introduce it. All that would do is create a vested interest group that would destabilize the financial system—and expose it to political manipulation. This might seem familiar.

39. There were also obvious problems with the arbitrariness of the risk weights. Perhaps the most severe was that it allows a zero weighting to be applied to government debt, and this feature survives intact into Basel III. Such a weighting implies that there is no risk attached to, say, GSEs like Fannie and Freddie or to Greek government debt. This weighting creates a big distortionary incentive to invest in such debt: in this sense, one can say that the Basel standardized weights were a major contributory factor behind both subprime (via Fannie and Freddie) and the Eurozone banking crisis, when banks' overexposures to financially weak European sovereigns came back to haunt them. The 50 percent weighting on residential mortgages relative to the 100 percent weighting on regular commercial loans didn't help either, as it misdirected lending from the latter to the former and further contributed to the U.S. mortgage crisis.

40. An obvious and quite intentional consequence of the model-based approach to regulation was to create an unlevel playing field: big banks could use their models to reduce their capital requirements in ways that small banks can't, because the latter were forced to use the simple Basel standardized risk-weight approaches that produce much higher charges. One can even say that this was a design feature of Basel II. The result was that a big bank and a small bank could have very different capital requirements against the same positions. As evidence, Haldane cites a UK exercise in which big and small banks were given the same portfolios

and had to work out their capital charges. On average, a small bank needed twice as much capital as a large bank; in some cases, the difference was as high as 9 to 1. See Andrew Haldane, “Containing Discretion in Bank Regulation” (speech given at the Atlanta Fed conference, “Maintaining Financial Stability: Holding a Tiger by the Tail(s),” April 9, 2013), p. 10, <http://www.bankofengland.co.uk/publications/Pages/speeches/2013/657.aspx>.

41. These and the other results cited here are reported in Haldane, “Containing Discretion,” pp. 5, 15–16.

42. Haldane, “Containing Discretion,” chart 2.

43. Haldane, “Containing Discretion,” p. 4.

44. A nice example is the “how to destroy the banking system” or “pig on pork” securitization. In 2002, Gordon Kerr and his colleagues working in the investment banking unit of a major UK bank came up with an ingenious scam to game the Basel capital rules. This arrangement produced immediate (but illusory) profits to the bank, created fictitious “virtual” capital and nice bonuses for all concerned, while leaving the bank’s real risk exposure almost unchanged. Here is how it worked: The bank had a bond portfolio worth over \$4 billion, which had been guaranteed by a U.S. insurer to ensure a AAA credit rating, and UK capital rules required that 8 percent of this value, or \$320 million, be set aside as capital. Kerr and his colleagues then persuaded the U.S. insurer to enter into a credit derivative contract with a European bank, which in turn wrote a matching credit derivative in favor of the UK bank. This transaction enabled the bond portfolio to be reclassified as a credit derivative and attract a regulatory capital requirement of 0.5 percent rather than 8 percent. The UK bank could then treat the transaction as raising \$300 million: 15/16 of the \$320 million was released. The transaction was engineered around the rule, and the release of capital was a chimera. A side effect was that profits could be brought forward and, under the lax accounting rules of the time, rather than recognizing the profit on the loans periodically as the loan margin was actually

paid, the bank booked as profit on execution the entire 30 years’ expected earnings for that part of the lending margin that now flowed through its derivative book. The other two parties were more or less unaffected. The U.S. bank earned an accommodation fee for the paperwork, but its economic position was unaltered because it was only guaranteeing the same assets again—and from its perspective, the second guarantee was redundant—hence pig on pork. For its part, the European bank earned an accommodation fee for taking very little exposure. The trick with credit derivatives is to define the circumstances under which they pay out: in this case, the financial engineers were careful to ensure that because the U.S. entity was guaranteeing the assets twice (via the guarantee and derivative), then the European bank’s derivative would not trigger until the UK bank had claimed under the guarantee. All very simple really! Senior management didn’t understand the scheme, but no one cared. This securitization was widely copied, and Mr. Kerr later found himself wondering why it took the banking system so long to collapse. See Gordon Kerr, “How to Destroy the British Banking System—Regulatory Arbitrage via ‘Pig on Pork’ Derivatives,” January 21, 2010, <http://www.cobdencentre.org/2010/01/how-to-destroy-the-british-banking-system>.

45. In addition to gaming the denominator in the Basel ratios, the banks were adept at gaming the numerator, bank capital. Basel I had a simple bare-bones definition of bank capital, common equity. However, Basel II based its capital ratios on more complex definitions: Tier 1, defined as share capital plus reserves minus various regulatory deductions and intangibles including especially goodwill; and Tier II, or supplementary bank capital, such as revaluation reserves, undisclosed reserves, and various forms of hybrid and subordinated debt. Minimum capital charges were 4 percent and 8 percent, respectively, of Tier 1 and Tier 2 against risk-weighted assets. So the banks developed new and less expensive but weaker forms of capital to meet these definitions, Tier 1 especially: gaming Tier 1 was a veritable cottage industry. The net result was that regulatory capital was artificially inflated with less stringent capital

instruments—especially various forms of Contingent Convertible or CoCo instruments—again making the banks appear stronger than they really were. To its credit, Basel III involves a move back toward more robust capital definitions, but even these are still gameable.

46. Haldane, “Capital Discipline,” pp. 6–8.

47. These include: Asli Demirgüç-Kunt, Enrica Detragiache, and Ouarda Merrouche, “Bank Capital: Lessons from the Financial Crisis,” World Bank Policy Research Working Paper Series No. 5473 (2010); David G. Mayes and Hanno Stremmel, “The Effectiveness of Capital Adequacy Measures in Predicting Bank Distress,” mimeo (2012); Allen N. Berger and Christa H. S. Bouwman, “How Does Capital Affect Bank Performance during Financial Crises?” *Journal of Financial Economics* 109 (2013): 146–76; Adrian Blundell-Wignall and Caroline Roulet, “Business Models of Banks, Leverage and the Distance-to-Default,” *OECD Journal: Financial Market Trends* 2012, no. 2 (2014); and Thomas L. Hogan, Neil Meredith and Xuhao Pan, “Evaluating Risk-Based Capital Regulation,” Mercatus Center Working Paper Series No. 13-02 (2013).

48. Andrew Haldane and Vasileios Madouros, “The Dog and the Frisbee” (speech given at the Kansas City Fed’s 36th economic policy symposium, “The Changing Policy Landscape,” Jackson Hole, WY, August 31, 2012).

49. Haldane and Madouros, “The Dog and the Frisbee,” p. 1.

50. F. A. Hayek, *The Fatal Conceit: The Errors of Socialism*, William W. Bartley III, ed. (London: Routledge, 1988), pp. 76–77.

51. As an aside, Haldane and Madouros cite the views of a former Bank of England deputy governor: “I confess that I fear we are in danger of becoming excessively complicated and that if so we may miss the wood from the trees,” he wrote. The official concerned was George Blunden writing in 1974 (!) in connection with the proposal to *introduce* regulatory returns by the UK banks. How

quaint. See Haldane and Madouros, “The Dog and the Frisbee,” pp. 1–12.

52. For a good overview of the prediction business, see Nate Silver, *The Signal and the Noise: Why Most Predictions Fail But Some Don’t* (New York: Penguin, 2012).

53. Daniel Rösch and Harald Scheuler, eds., *Stress Testing for Financial Institutions* (London: Risk Books, 2008).

54. *Ibid.*, p. 451.

55. *Ibid.*, p. 444 (my italics).

56. The quotes in this paragraph and the next are taken from James R. Hagerty, “How Paul Volcker Misread Fannie Mae,” *Wall Street Journal*, September 7, 2013.

57. More details and background are to be found in Joseph Stiglitz, Jonathan M. Orszag, and Peter R. Orszag, “Implications of the New Fannie Mae and Freddie Mac Risk-based Capital Standard,” *Fannie Mae Papers* 1, no. 2 (March 2002): 2–3; and W. Scott Frame, Kristopher Gerardi, and Paul S. Willen, “Supervisory Stress Tests, Model Risk, and Model Disclosure: Lessons from OFHEO,” Mimeo April 3, 2013, p. 5, https://www.frbatlanta.org/documents/news/conferences/13fmc_gerardi.pdf.

58. The quote comes from Fannie’s then-senior VP for regulatory policy, Arne Christenson, in his forward to Stiglitz, “Implications of the New Fannie Mae and Freddie Mac Risk-based Capital Standard.”

59. In addition to the stress-based capital requirement outlined in the text, the GSEs were also subject to a second capital standard: they had to hold capital equal to 2.5 percent of their on-balance sheet assets and 0.45 percent (!) of their off-balance sheet obligations and assets. However, the actual capital requirements turned out to be very low despite the superficially severe stress scenario explained in the text. The net results was that the

GSEs were subject to much weaker capital standards than the banks, while also having their own unique governance issues and the benefit of lower financing costs associated with the public perception that the government would never let them fail. As the joke goes, they would tell Congress and the news media, “Don’t worry, the government is not on the hook”—and then turn around and tell Wall Street, “Don’t worry, the government really is on the hook.” They were, in short, always a disaster waiting to happen.

60. One should also appreciate the context: the research was commissioned by Fannie and then published in a newly established series called *Fannie Mae Papers* set up by Arne Christenson, whose strategy was to pay eminent scholars who produced pro-Fannie results that Fannie officials could crow about. An innovative feature of *Fannie Mae Papers* was a helpful foreword: in this particular case, Christenson used the research study as a platform to boast of how Fannie’s risk management standards were well ahead of best market practice. Stiglitz and his colleagues had allowed themselves to be used in the most blatant manner and were to regret it. I am not suggesting, by the way, that the underlying research per se was compromised by the context in which it was commissioned and used—it was just bad research.

61. Stiglitz, “Implications of the New Fannie Mae and Freddie Mac Risk-based Capital Standard,” p. 6.

62. *Ibid.*, p. 5.

63. Feynman’s (more) detailed comments are available in Richard P. Feynman, *Personal Observations on Reliability of Shuttle of the Report of the Presidential Commission on the Space Shuttle Challenger Accident, volume 2: Appendix F* (Washington: NASA, 1986), <http://history.nasa.gov/rogersrep/v2appf.htm>.

64. Six years later—well before the first decade was up—the GSEs were taken into a government conservatorship to avert their failure and Stiglitz had altered his tune. The problem, he explained, was that he hadn’t anticipated the GSEs’ venture

into subprime lending—never mind that affordable housing was already well and truly baked into the cake by this point. Stiglitz’s explanation was less than convincing: He complained that people had been granted loans “beyond their ability to pay” [Wasn’t that the whole point of affordable housing?] and the GSEs’ management compensation was “designed to encourage excessive risk-taking . . . no one would have ever thought the regulators would allow them to get away with that. [Oh yeah?] We assumed that the regulator was doing its job.” Quoted in James R. Hagerty, “How Joseph Stiglitz Misread the Risks at Fannie Mae”, *Wall Street Journal*, October 1, 2012. If I understand his comments correctly, he seems to be suggesting that he and his colleagues overlooked subprime, overlooked the impact of compensation packages that encouraged excessive risk-taking, and overlooked the politics and assumed that the regulatory process would work as it was intended to, all of which were known issues even in 2002—but the modeling was otherwise fine.

65. Frame, “Supervisory Stress Tests.”

66. *Ibid.*, p. 4.

67. To be fair to OFHEO, they also faced significant political meddling. See, for example, Gretchen Morgenson and Joshua Rosner, *Reckless Endangerment: How Outsized Ambition, Greed, and Corruption Led to Economic Armageddon* (New York: Henry Holt, 2011), which would have considerably limited the willingness and ability of OFHEO officials to make changes to the rule or the model.

68. This and other quotes from this session are taken from OFHEO, “Risk-Based Capital Stress Test for Fannie Mae and Freddie Mac,” U.S. House of Representatives, Subcommittee on Capital Markets, Insurance and Government Sponsored Enterprises Committee on Financial Services (Tuesday, July 23, 2002), http://commdocs.house.gov/committees/bank/hba82684.000/hba82684_of.htm.

69. Hearing before the Committee on Banking,

Housing, and Urban Affairs, 110th Congress, 2d Sess., July 15, 2008.

70. “Iceland: Financial System Stability Assessment—Update,” IMF Country Report No. 8/368, p. 5.

71. Tyler Durden, “European Stress Tests Confirmed to Be a Farce: Will Not Include Discussion of Sovereign Risks,” June 6, 2010, <http://www.zerohedge.com/article/ecb-stress-test-confirmed-be-farce-will-not-include-discussion-sovereign-risks>.

72. Tyler Durden, “Second Straight Hungarian Bond Auction Failure as Citi’s Willem Buiters Calls for €2 Trillion European Rescue Facility, Ridicules Stress Tests,” June 24, 2010, <http://www.zerohedge.com/article/second-straight-hungarian-bond-auction-failure-citis-willem-buiters-calls-percentE2percent82percentAC2-trillion-rescue->

73. Tyler Durden, “Confirmed—Eurozone ‘Stress Tests’ Will Not Include any Default Scenario,” July 7 2010, <http://www.zerohedge.com/article/confirmed-eurozone-stress-tests-will-not-include-any-default-scenario>.

74. Tyler Durden, “You Can’t Make This Up: Europe Plans Fresh Round of ‘Secret’ Stress Tests to ‘Restore Confidence,’” November 30, 2010, <http://www.zerohedge.com/article/you-cant-make-europe-plans-fresh-round-stress-tests-restore-confidence>.

75. Ibid.

76. Tyler Durden, “Stress Test Part Two Discredited Before It Is Even Announced: All Irish Banks Pass ‘Comfortably,’” July 15, 2011, <http://www.zerohedge.com/article/stress-test-part-two-discredited-it-even-announced-all-irish-banks-pass-comfortably>.

77. A Moody’s report made an additional highly appropriate point: the 5 percent minimum common Tier 1 capital ratio used in the test was well below the 7 percent used in Basel III, which was

then forthcoming. They estimated that nearly half the banks would have failed at that threshold—yet another astonishing indication of how weak the European banking system really was, even granted the poor data being fed into the models (of which more below). They and other independent analysts painted a much bleaker picture than what the EBA considered in its stress tests. See Enam Ahmed, Andrea Appeddu, Melanie Bowler, Tomas Holinka, Juan Manuel Licari, Olga Loiseau-Aslanidi, and Zach Witton, “Europe Misses Again on Bank Stress Test,” Moody’s Analytics *Regional Financial Review* (July 2011): 11–15.

78. Months later, Dexia was followed by the failure in May 2012 of Spain’s fourth biggest bank, Bankia, which had also passed the 2011 stress test. A new EBA stress test exercise is being carried out over 2014. The new minimum hurdle will be raised to 8 percent, except for the simulated recession, where it be tweaked up half a point to 5.5 percent, and this despite fierce resistance from some European national regulators fighting for softer standards so their banks look better. One also assumes that the simulated recession will be different from the current, very real, recession in Europe, with unemployment at record highs. No mention either of how the European regulators are going to identify the hundreds of billions in bad debts, but my hopes were not high. The ECB will also get into the act with its own exercise covering 130 credit institutions across the EU, with the stated objectives of achieving transparency, repairing the banks, and, of course, promoting confidence. With objectives like these coming from any European governmental or regulatory institution, the outcome is a foregone conclusion.

79. Mark J. Grant, “Union: One Survived; One May Not,” March 22, 2013, <http://www.zerohedge.com/news/2013-03-22/union-one-survived-one-may-not>.

80. I say “almost” because there is one arguable exception where the regulators *did* identify a stress build-up years before it happened: the case of UK bank Northern Rock. Once a small, staid—that is to say, sound—regional building society, in the years

before the crisis it had transformed itself into an aggressive mortgage bank whose business model emphasized lower lending standards and a greater than usual reliance on wholesale markets—a British version of subprime. By the eve of the crisis it had become the UK's fifth-biggest mortgage provider. When the wholesale markets started freezing up in the fall of 2007 the bank then ran into trouble: on September 14 it experienced the first English bank run since Overend Gurney in 1866; it was bailed out and then nationalized. There was much criticism of way in which the UK authorities had mishandled the case. Government officials trotted out the usual lame excuses: no one could have anticipated the crisis and so on. However, it turned out that the UK authorities *had* anticipated it. It was later revealed that in 2004 they had conducted secret “war games” in which they correctly anticipated the Northern Rock crisis, including its systemic knock-on effects on HBOS, the biggest mortgage lender. When the war-game story came out, the excuses that came out after it were priceless: the regulators didn't have the authority to force the bank to change its business model, it was too hard to say that the bank really was taking on excessive risks, as other banks were doing the same and making good profits doing so, and—the best of all—that the focus of the exercise was not on predicting bank failure (which is *not*, by the way, what a stress test does: it is to identify vulnerabilities) but on uncovering weak regulatory practice (in which case, one has to concede that the exercise wasn't much of a success after all). There followed a nice little game of pass-the-parcel with the three British regulatory bodies—the Financial Services Agency, the Bank of England, and the Treasury—blaming each other for their collective failure to do anything. This case reminds us that we not only have to get the stress scenario right, but if it flags up a credible vulnerability, we should act on it too. See Norma Cohen and Chris Giles, “War Game Saw Run on Rock,” *Financial Times* (London), March 30, 2009.

81. Borio et al., “Stress Testing Macro Stress Testing,” pp. 8–9.

82. Ben S. Bernanke, “The Supervisory Capital

Assessment Program,” (speech at the Atlanta Fed 2009 Financial Markets Conference, Jekyll Island, Georgia, May 11, 2009), <http://www.federalreserve.gov/newsevents/speech/bernanke20090511a.htm>.

83. William C. Dudley, “U.S. Experience with Bank Stress Tests,” (speech at the Group of 30 plenary meeting, Bern, Switzerland, May 28, 2011), <http://www.newyorkfed.org/newsevents/speeches/2011/dud110627.html>.

84. The minimum capital ratios for the 2013 CCAR were: Tier common ratio = 5 percent, Tier 1 leverage ratio equal to 3 percent or 4 percent (3 percent for a BHC with a composite supervisor rating of “1” or that is subject to the Federal Reserve Board's market-risk rule) plus the standard Basel minimum capital ratios (Tier 1 risk-based capital ratio = 4 percent and Total risk-based capital ratio = 8 percent). In addition, the capital plan rule stipulates that a BHC must demonstrate the ability to maintain a Tier 1 common ratio above 5 percent: thus, the ‘headline’ minimum capital ratio was still 5 percent Tier 1 common. See “Comprehensive Capital Analysis and Review 2013: Assessment Framework and Results,” Board of Governors of the Federal Reserve System (March 2013), p. 11, <http://www.federalreserve.gov/bankinforeg/ccar-2013-results-20130314.pdf>.

85. There were now 18 rather than 19 banks subject to the CCAR. The missing one, MetLife, was really an insurance company and argued (with good reason) that the Fed's approach was inappropriate to its business. However, it had been included earlier because it had a banking license. It had had its capital plan rejected by the Fed in March 2012 and then surrendered its banking license to escape the 2013 CCAR.

86. The designation of an institution as “systemically important” effectively codifies “Too Big To Fail” and so creates a host of problems. Amongst these, it creates (or at least confirms) an enormous moral-hazard problem (how do we stop these institutions going on a gambling spree at the taxpayers' expense, again?) and it creates a very uneven

playing field. Of particular concern in the present context is that it brings non-banks under the Fed remit—one of the first institutions so designated was in fact AIG: this institution was designated “systemically important” not because it actually is, but in order to provide an *ex post* justification for its bailout by the Fed in 2008. The Fed’s one-size-fits-all approach is especially inappropriate for insurance companies that have very different asset liability structures than banks. So we know that the Fed’s stress tests are going to cause a lot of problems when it comes to AIG and other non-bank institutions that will be subjected to them. And, of course, the Fed knows even less about insurance than it does about banking.

87. Board of Governors of the Federal Reserve, (letter to bankers, December 16, 2013), <http://www.federalreserve.gov/bankinfo/reg/independent-projections-letter-20131216.pdf>.

88. The gentleman concerned also noted that the data used to justify the Fed’s projection that bank balance sheets would not shrink in a crisis was “totally misleading.” The main flaw is that the Fed data do not correct for mergers, that is, the merged assets of the surviving company were included in the growth statistics. Since the surveyed banks tend to acquire failed banks, the growth numbers then become grossly exaggerated. In short, the Fed’s approach confused true asset growth with merger and acquisition activity. Another distortion was the Troubled Asset Relief Program (TARP), which also artificially boosted banks’ balance sheets. As he pointed out, the only justification for including the TARP data in the Fed’s projections was if the Fed was planning to do TARP again!

89. Tyler Durden, “Stress Test Assumptions Nutshelled,” February 26, 2009, <http://www.zero hedge.com/article/stress-test-assumptions-nutshelled>.

90. Reported in Gretchen Morgenson, “What the Stress Tests Didn’t Predict,” *New York Times*, August 23, 2009.

91. Yves Smith, “The Fed Stress Tests While Eu-

rope Starts to Burn,” November 23, 2011, <http://www.nakedcapitalism.com/2011/11/the-fed-stress-tests-while-europe-starts-to-burn.html>.

92. To quote chairman Bernanke from his speech on the SCAP (Supervisory Capital Assessment Program) in 2009, “Projecting credit losses in an uncertain economic environment is difficult, to say the least, but the intensive, painstaking nature of this process gives *us* confidence in our results.” (My italics.) My response is that it shouldn’t: it doesn’t matter how much work you put into that single scenario, you should never be confident in its results—and *never ever* jump from that to confidence that the Fed has all risk bases covered. You also have to look at other (very different) scenarios, too, because the probability of any one scenario actually materializing is vanishingly close to zero. In any case, from the perspective of the complexity literature, the *more* effort goes into polishing the scenario, the *less* confident we should be of the outcome. There are the dangers of groupthink, too.

Fed apologists might point out that the system does not rely on one scenario, as the banks are meant to do their own stress tests on risks they think relevant, too. This argument is a cop-out. If the Fed really wanted them to do their own independent stress tests, they would incentivize them to do so instead of pressure them towards the Fed’s own risk model. Increasingly, we can expect to see the banks’ own risk projections merely become a sham, from the banks’ point of view an increasingly meaningless show-case exercise they go through to satisfy the Fed, and for no other reason—after all, since the Fed is increasingly taking over their own risk management, what is the point of fighting it?

The Fed now has recently moved to two core scenarios instead of one—its severely adverse scenario and the market-collapse scenario it imposes on the banks with the big trading activities—but this hardly addresses the underlying issue of the Fed’s excessive reliance on the former scenario. One also has to wonder what purpose the Fed’s baseline and adverse scenarios now serve if its seriously adverse scenario always produces higher losses than they do.

93. The Fed's reason for refusing to release more details is hinted at by William C. Dudley in his Bern speech of May 28, 2011 ("U.S. Experience with Bank Stress Tests"): "If all we ever did was run supervisory stress tests in which we instruct banks in detail how to perform the test, we would be in the position of a parent who shows his child how to solve each problem in her homework—and never discovers whether the child can do the work on her own or not." The parental analogy in this quote nicely captures the superior, even patronizing, attitude of the Fed toward the banks. The bankers, of course, are deeply frustrated: for them the process is like being forced to sit an exam with no way to work out the right answers—and, indeed, with the "correct" answers being changed at will. One can only sympathize. Releasing more details would also allow greater independent scrutiny of the Fed's models, but then the Fed consistently opposes any measure that would hold it to account. On the other hand, on this one occasion I confess to having a *little* bit of sympathy toward the Fed too: releasing more details would increase opportunities to game the system, and the OFHEO experience with open modeling is hardly encouraging. But on balance, a much greater degree of openness and accountability on this issue is much to be preferred.

94. For more on RCA—including plenty of how-to advice—see David Jones, "Emerging Problems with the Basel Capital Accord: Regulatory Capital Arbitrage and Related Issues," *Journal of Banking and Finance* 24 (2000): 35–58; Gordon Kerr, *The Law of Opposites: Illusory Profits in the Financial Sector* (London: Adam Smith Institute, 2011); and Gordon Kerr (2014), "The Case for Systemic Banking Reform."

95. Bonnie Kavoussi, "Federal Reserve Announces New Bank Stress Tests to Prepare for Europe Crisis," *Huffington Post*, November 22, 2011.

96. As Yves Smith noted in November 2011: "Somehow, the Fed has managed to wear blinkers since May 2010. As soon as Greece started going wobbly, a horde of analysts pointed out that the bond losses were likely to be in the 50 percent to

70 percent range, and contagion was a realistic possibility. Contagion not only arrived but metastasized while the Eurocrats engaged in policy shamanism in lieu of painful cures." Yet none of the Fed's risk modelers seemed to notice. As he continued, "There is every reason to expect the upcoming European crisis to be not as bad as the last one [2008–2009], as the Fed assumes, but worse. The banking systems in the U.S. and Europe are much more concentrated than before. Advanced economies are in much weaker shape [and so forth]. . . . The risk here is not a Lehman like disaster. It's a modern version of the Credit Anstalt [of 1931]: a major bank failure precipitating cascading collapses." This reminds us that it is one thing to *include* a particular kind of risk in the scenario, but we need to get the *nature and magnitude* of the risk event right as well.

97. This is a set of scams in which banks furtively pledge assets to counterparties while the assets ostensibly remain on their balance sheets. An example is a "failed sale," a repo-like transaction designed to secure finance by granting counterparties hidden hypothecations of prime bank assets. At first sight, it looks like a repo—a standard form of collateralized loan—but this particular beauty is accompanied by the sale of a repurchase option with the sole intent of hiding the preferment from other bank stakeholders. The trick is that under the rules, the repo'd assets never leave the borrowing bank's balance sheet—that is, the arrangement does not qualify as a true sale; hence the "failed sale" label. So, from the borrower's perspective, the bank's balance sheet is apparently unaffected. Since banks have substantial repo activity going on as normal derivative and hedging activity, the failed sale deals are very hard to spot. This type of transaction is damaging in at least two ways: First, it deceives other bank counterparties, who do not appreciate that they cannot recover the assets in question, even though they still remain on the bank's balance sheet. A failed sale transaction is thus essentially fraudulent. Should the bank then fail, creditors would lose almost everything when they found that they had taken possession of little more than an empty shell. Second, the fact that these practices are known to be going on means

regulator William Black who argued that some banks—such as Bank of America—were already insolvent and should be shut down; it had recently admitted that half of its 13.9 million mortgages were “bad.” To quote the article concerned:

This accounting scam produces enormous fictional “income” and “capital” at the banks. The fictional income produces real bonuses to the CEOs that make them even wealthier. The fictional bank capital allows the regulators to evade their statutory duties under the Prompt Corrective Action (PCA) law to close the insolvent and failing banks . . . We have made the CEOs of the largest financial firms—typically already among the 500 wealthiest Americans—even wealthier. We have rewarded fraud, incompetence and venality by our most powerful elites.

See William K. Black and L. Randall Wray, “Foreclose on the Foreclosure Fraudsters, Part 1: Put Bank of America in Receivership,” *Huffington Post*, October 22, 2010, http://www.huffingtonpost.com/william-k-black/foreclose-on-the-foreclos_b_772434.html.

101. Jonathan Weil, “Class Dunce Passes Fed’s Stress Test Without a Sweat,” March 15, 2012, <http://www.bloombergvew.com/articles/2012-03-15/stress-tests-pass-fed-s-flim-flam-standard-jonathan-weil>.

102. A bad joke it was, too. Instead of going through the farcical stress test that merely disguised the bank’s weakness, the regulators should simply have read the accounts properly and then obeyed the law. The bank was GAAP-insolvent and therefore fell under the stipulation of the Prompt Corrective Action statutes that *mandate* (note!) that banks that are severely undercapitalized, defined as banks with capital ratios below 2 percent, should be taken into receivership. The Fed’s passing of Regions Financial’s capital plan was illegal.

103. R. Christopher Whalen, “Bank Stress Tests

and Other Acts of Faith,” March 13, 2012, <http://www.zerohedge.com/contributed/2012-11-13/bank-stress-tests-and-other-acts-faith>.

104. R. Christopher Whalen, “CCAR|Stress Test Follies & Zombie Banks,” March 7, 2013, <http://www.zerohedge.com/contributed/2013-03-07/ccar-stress-test-follies-zombie-banks>.

105. There were also major implementation costs. The bankers I interviewed for this article mentioned many hundreds of millions in development costs required to feed the models demanded by the Fed, as opposed to the costs incurred for their own risk management purposes. One estimate put the costs of the Fed’s stress tests at 5 percent of that particular bank’s market capitalization. Much normal business activity was heavily disrupted in the process. For example, there were stories of major disruption to banks’ information technology innovation, as staff were overwhelmed with regulatory compliance. In addition, merger activity was drastically slowed as banks were unable to assess the regulatory risk in potential purchases; risk meetings became more frequent and longer, and regulatory reports could run to tens of thousands of pages.

106. See Peter J. Wallison, “The Dodd-Frank Act: Creative Destruction, Destroyed,” American Enterprise Institute for Public Policy Research *Financial Services Outlook* (June–August 2010), p. 1, <http://www.aei.org/files/2010/08/30/FSO-2010-July-August-g.pdf>. See also Peter J. Wallison, *Bad History, Worse Policy: How a False Narrative about the Financial Crisis Led to the Dodd-Frank Act* (Washington: American Enterprise Institute, 2012). Mr. Wallison’s book is highly recommended, but makes for a disturbing read.

107. Matthew B. Albrecht and Carmen Y. Manoyan, “Two Years On, Reassessing The Cost of Dodd-Frank for the Largest U.S. Banks,” August 9, 2012, <http://www.standardandpoors.com/ratings/articles/en/us/?assetID=1245338539029>.

108. John A. Allison, “The Real Causes of the Financial Crisis,” *Cato’s Letter* 12, no. 2 (Winter 2012), p. 1.

109. Alicia H. Munnell, Lynn E. Browne, James McEaney, and Geoffrey M. B. Tootell, "Mortgage Lending in Boston: Interpreting HMDA Data," Federal Reserve Bank of Boston Working Paper No. 92-7 (1992).
110. Stan J. Liebowitz, "A Study that Deserves No Credit," *Wall Street Journal*, September 1, 1993; Mark Zandi, "Boston Fed's Bias Study Was Deeply Flawed," *American Banker*, August 19, 1993.
111. John A. Allison, "The Fed's Fatal Conceit," *Cato Journal* 32, no. 2 (Spring/Summer 2012): 269.
112. Peter J. Wallison, "Dodd-Frank is the Midwife of Crony Capitalism," November 8, 2012, <http://www.aei-ideas.org/2012/11/dodd-frank-is-the-midwife-of-crony-capitalism/>.
113. One knock-on effect of this capital standard would be to force the mortgage GSEs to raise their capital levels: this pressure to deleverage would force them to cut back their activities and scale back the sizes of their balance sheets. With luck and the right political pressure (always the optimist!), it might also help them onto a path toward entering voluntary receiverships, so that they can be put down once and for all.
114. There are, of course, a host of problems with GAAP itself (see, e.g., John A. Allison, *The Financial Crisis and the Free Market Cure*, ch. 11, which has a good overview). This is a huge subject in itself but the answer, as ever, is to get the government out of the accounting standards business and go back to a much simpler principles-based system.
115. If Congress ever has the appetite, I would also recommend as much as possible of the following: (1) Deregulation—and lots of it: a bonfire of regulations, encompassing the repeal of all the restrictive bank regulation from Dodd-Frank all the way back to at least the Truth in Lending Act of 1968. Deregulation would free banks to seek out the most profitable lending opportunities and start to eliminate the many distortions created by decades of accumulated regulation—and give a big boost to bank lending and a soundly based economic recovery. Not least of these distortions are regulatory barriers to the continued consolidation of the American banking industry: deregulation would allow a natural consolidation process to develop and put an end to the current practice of mismatched regulatory shotgun marriages. (2) We should impose personal liability on regulators, too: given the way in which regulators now routinely ignore the law, we want to give those on the receiving end of this behavior effective legal redress. (3) It is important to protect the banking system by removing federal government interventions that create incentives for excessive risk-taking—get rid of federal deposit insurance and "Too Big to Fail" and prohibit any further Fed bailouts. Bankers need to have the fullest possible incentive to protect the safety and soundness of their own banks.
116. "Disaster at Last Befalls Capt. Smith," *New York Times*, April 16, 1912, <http://www.encyclopedia-titanica.org/disaster-at-last-befalls-capt-smith.html>.

RELATED PUBLICATIONS FROM THE CATO INSTITUTE

Run, Run, Run: Was Financial Crisis Panic over Institution Runs Justified? by Vern McKinley, Cato Policy Analysis no. 747 (April 10, 2014)

Liberalization or Litigation? Time to Rethink the International Investment Regime by Simon Lester, Cato Institute Policy Analysis no. 730 (July 8, 2013)

Central Banks: Reform or Abolish? by Gerald P. O'Driscoll Jr., Cato Institute Working Paper (October 15, 2012)

Regulation, Market Structure, and Role of the Credit Rating Agencies by Emily McClintock Ekins and Mark A. Calabria, Cato Institute Policy Analysis no. 704 (August 1, 2012)

Questioning Homeownership as a Public Policy Goal by Morris A. Davis, Cato Institute Policy Analysis no. 696 (May 15, 2012)

What Made the Financial Crisis Systemic? by Patric H. Hendershott and Kevin Villani, Cato Institute Policy Analysis no. 693 (March 6, 2012)

Fixing Mortgage Finance: What to Do with the Federal Housing Administration? by Mark A. Calabria, Cato Institute Briefing Paper no. 123 (February 6, 2012)

Reputation under Regulation: The Fair Credit Reporting Act at 40 and Lessons for the Internet Privacy Debate by Jim Harper, Cato Institute Policy Analysis no. 690 (December 8, 2011)

Capital Inadequacies: The Dismal Failure of the Basel Regime of Bank Capital Regulation by Kevin Dowd, Martin Hutchinson, Jimi M. Hinchliffe and Simon Ashby, Cato Institute Policy Analysis no. 681 (July 29, 2011)

Time to X Out the Ex-Im Bank by Sallie James, Cato Institute Trade Policy Analysis no. 47 (July 6, 2011)

The Subprime Lending Debacle: Competitive Private Markets Are the Solution, Not the Problem by Patric H. Hendershott and Kevin Villani, Cato Institute Policy Analysis no. 679 (June 20, 2011)

Fannie Mae, Freddie Mac, and the Future of Federal Housing Finance Policy: A Study of Regulatory Privilege by David Reiss, Cato Institute Policy Analysis no. 674 (April 18, 2011)

Fannie, Freddie, and the Subprime Mortgage Market by Mark A. Calabria, Cato Institute Briefing Paper no. 120 (March 7, 2011)

Has the Fed Been a Failure? by George Selgin, William D. Lastrapes and Lawrence H. White, Cato Institute Working Paper (November 9, 2010)

Lawless Policy: TARP as Congressional Failure by John Samples, Cato Institute Policy Analysis no. 660 (February 4, 2010)

Would a Stricter Fed Policy and Financial Regulation Have Averted the Financial Crisis? by Jagadeesh Gokhale and Peter Van Doren, Cato Institute Policy Analysis no. 648 (October 8, 2009)

Bright Lines and Bailouts: To Bail or Not To Bail, That Is the Question by Vern McKinley and Gary Gegenheimer, Cato Institute Policy Analysis no. 637 (April 20, 2009)

Financial Crisis and Public Policy by Jagadeesh Gokhale, Cato Institute Policy Analysis no. 634 (March 23, 2009)

Freddie Mac and Fannie Mae: An Exit Strategy for the Taxpayer by Arnold Kling, Cato Institute Briefing Paper no. 106 (September 8, 2008)

RECENT STUDIES IN THE CATO INSTITUTE POLICY ANALYSIS SERIES

- 753. **The Dead Hand of Socialism: State Ownership in the Arab World** by Dalibor Rohac (August 25, 2014)
- 752. **Rapid Bus: A Low-Cost, High-Capacity Transit System for Major Urban Areas** by Randal O'Toole (July 30, 2014)
- 751. **Libertarianism and Federalism** by Ilya Somin (June 30, 2014)
- 750. **The Worst of Both: The Rise of High-Cost, Low-Capacity Rail Transit** by Randal O'Toole (June 3, 2014)
- 749. **REAL ID: State-by-State Update** by Jim Harper (March 12, 2014)
- 748. **State-Based Visas: A Federalist Approach to Reforming U.S. Immigration Policy** by Brandon Fuller and Sean Rust (April 23, 2014)
- 747. **Run, Run, Run: Was the Financial Crisis Panic over Institution Runs Justified?** by Vern McKinley (April 10, 2014)
- 746. **State Education Trends: Academic Performance and Spending over the Past 40 Years** by Andrew J. Coulson (March 18, 2014)



Published by the Cato Institute, Policy Analysis is a regular series evaluating government policies and offering proposals for reform. Nothing in Policy Analysis should be construed as necessarily reflecting the views of the Cato Institute or as an attempt to aid or hinder the passage of any bill before Congress. Contact the Cato Institute for reprint permission. All policy studies can be viewed online at www.cato.org. Additional printed copies of Cato Institute Policy Analysis are \$6.00 each (\$3.00 each for five or more). To order, call toll free (800) 767-1241 or email catostore@cato.org.