NBER WORKING PAPER SERIES

CHALLENGES IN IDENTIFYING AND MEASURING SYSTEMIC RISK

Lars Peter Hansen

Working Paper 18505
http://www.nber.org/papers/w18505

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2012

I benefitted from helpful suggestions by Amy Boonstra, Gary Becker, John Heaton, Jim Heckman, Arvind Krishnamurthy, Monika Piazzesi, Toni Shears, Stephen Stigler and especially Markus Brunnermeier, Andy Lo, Tom Sargent and Grace Tsiang in writing this chapter. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Lars Peter Hansen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Challenges in Identifying and Measuring Systemic Risk
Lars Peter Hansen
NBER Working Paper No. 18505
November 2012
JEL No. E44

ABSTRACT

Sparked by the recent “great recession” and the role of financial markets, considerable interest exists among researchers within both the academic community and the public sector in modeling and measuring systemic risk. In this essay I draw on experiences with other measurement agendas to place in perspective the challenge of quantifying systemic risk, or more generally, of providing empirical constructs that can enhance our understanding of linkages between financial markets and the macroeconomy.

Lars Peter Hansen
Department of Economics
The University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
lhansen@uchicago.edu

1 Introduction

Discussions of public oversight of financial markets often make reference to "systemic risk" as a rationale for prudent policy making. For example, mitigating systemic risk is a common defense underlying the need for macro-prudential policy initiatives. The term has become a grab bag, and its lack of specificity could undermine the assessment of alternative policies. At the outset of this essay I ask, should systemic risk be an explicit target of measurement, or should it be relegated to being a buzz word, a slogan or a code word used to rationalize regulatory discretion?

I remind readers of the dictum attributed to Sir William Thompson, "Lord Kelvin":

> I often say that when you can measure something that you are speaking about, express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of the meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts advanced to the stage of *science*, whatever the matter might be.

While Lord Kelvin's scientific background was in mathematical physics, discussion of his dictum has pervaded the social sciences. An abbreviated version appears on the Social Science Research building at the University of Chicago and was the topic of a published piece of detective work by Merton et al. (1984). I will revisit this topic at the end of this essay. Right now I use this quote as a launching pad for discussing systemic risk by asking if we should use measurement or quantification as a barometer of our understanding of this concept.

One possibility is simply to concede that *systemic risk* is not something that is amenable to quantification. Instead it is something that becomes self evident under casual observation. Let me recall Justice Potter Stewart's famed discussion of pornography:

> I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description ["hard-core pornography"]; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it, and the motion picture involved in this case is not that.[1]

[1] Justice Potter Stewart, concurring opinion in Jacobellis v. Ohio 378 U.S. 184 (1964), regarding possible obscenity in The Lovers.

This is quite different from Kelvin's assertion about the importance of measurement as a precursor to some form of scientific understanding and discourse. Kelvin's view was that for measurement to have any meaning requires that 1) we formalize the concept that is to be measured and 2) we acquire data to support the measurement. Justice Stewart was not claiming to adopt a scientific perspective, but he did argue in support of the "know it when you see it" dictum as a matter of policy. If this approach works to enforce laws, perhaps it also works as a way to implement oversight or regulation of financial markets.

In the short run, we may be limited to some counterpart to Justice Stewart's argument about pornography. Perhaps we should defer and trust our governmental officials engaged in regulation and oversight to "know it when they see it." I have two concerns about leaving things vague, however. First, it opens the door to a substantial amount of regulatory discretion. In extreme circumstances that are not well guided by prior experience or supported by economic models that we have confidence in, some form of discretion may be necessary for prudent policy making. However, discretion can also lead to bad government policy, including the temptation to respond to political pressures. Second, it makes criticism of measurement and policy all the more challenging. When formal models are well constructed, they facilitate discussion and criticism. Delineating assumptions required to justify conclusions disciplines the communication and commentary necessary to nurture improvements in models, methods, and measurements. This leads me to be sympathetic to a longer-term objective of exploring the policy-relevant notions of the quantification of systemic risk. To embark on this ambitious agenda, we should do so with open eyes and a realistic perspective on the measurement challenges. In what follows, I explore these challenges in part by drawing on the experience from other such research agendas within economics and elsewhere.

The remainder of this essay :

i) explores some conceptual modeling and measurement challenges;

ii) examines these challenges as they relate to existing approaches to measuring systemic risk.

2 Measurement with and without Theory

Sparked in part by the ambition set out in the Dodd-Frank bill and similar measures in Europe, the Board of Governors of the Federal Reserve System and some of the constituent regional banks have assembled research groups charged with producing measurements of systemic risk. Such measurements are also part of the job of the newly created Office of Financial Research housed in the Treasury Department. Similar research groups have been assembled in Europe. While the need for legislative responses put pressure on research departments to produce quick "answers", I believe it is also critical to take a longer-term perspective so that we can do more than just respond to the last crisis. By now, a multitude of proposed measures exist and many of these are summarized in Bisias et al. (2012), where *thirty one* ways to measure systemic risk are identified. While the authors describe this catalog as an "embarrassment of riches", I find this plethora to be a bit disconcerting. In describing why, in the next section I will discuss briefly some of these measures without providing a full-blown critique. Moreover, I will not embark on a commentary of all thirty-one listed in their valuable and extensive summary. Prior to taking up that task, I consider some basic conceptual issues.

I am reminded of Koopmans's discussion of the Burns and Mitchell (1946) book on measuring business cycles. The Koopmans (1947) review has the famous title "Measurement without Theory". It provides an extensive discussion and sums things up saying:

> The book is unbendingly empiricist in outlook. ... But the decision not to use theories of man's economic behavior, even hypothetically, limits the value to economic science and to the maker of policies, of the results obtained or obtainable by the methods developed.

The measurements by Burns and Mitchell generated a lot of attention and renewed interest in quantifying business cycles. It served to motivate development of both formal economic and statistical models. An unabashedly empirical approach can most definitely be of considerable value, especially in the initial stages of a research agenda. What is less clear is how to use such an approach as a direct input into policy making without an economic model to provide guidance as to how this should be done. An important role for economic modeling is to provide an interpretable structure for using available data to explore the consequences of alternative policies in a meaningful way.

2.1 Systematic or systemic

Looking forward, a crucial challenge will be to distinguish "systemic" from "systematic" risk. In sharp contrast with the former concept, the latter one is well studied and supported by extensive modeling and measurement. Systematic risks are macroeconomic or aggregate risks that cannot be avoided through diversification. According to standard models of financial markets, investors who are exposed to these risks require compensation because there is no simple insurance scheme whereby exposure to these risks can be averaged out.[2] This compensation is typically expressed as a risk adjustment to expected returns.

Empirical macroeconomics aims to identify aggregate "shocks" in time series data and to measure their consequences. Exposure to these shocks is the source of systematic risk priced in security markets. These may include shocks induced by macroeconomic policy, and some policy analyses explore how to reduce the impact of these shocks to the macroeconomy through changes in monetary or fiscal policy. Often, but not always, as a separate research enterprise, empirical finance explores econometric challenges associated with measuring both the exposure to the components of systematic risk that require compensation and the associated compensations to investors.

"Systemic risk" is meant to be a different construct. It pertains to risks of breakdown or major dysfunction in financial markets. The potential for such risks provides a reason for financial market monitoring, intervention or regulation. The systemic risk research agenda aims to provide guidance about the consequences of alternative policies and to help anticipate potential breakdowns in financial markets. The formal definition of *systemic risk* is much less clear than its counterpart *systematic risk.*

Here are three possible notions of systemic risk that have been suggested. Some consider systemic risk to be a modern-day counterpart to a bank run triggered by liquidity concerns. Measurement of that risk could be an essential input to the role of central banks as "lenders of last resort" to prevent failure of large financial institutions or groups of financial institutions. Others use systemic risk to describe the vulnerability of a financial network in which adverse consequences of internal shocks can spread and even magnify within the network. Here the measurement challenge is to identify when a financial network is potentially vulnerable and the nature of the disruptions that can trigger a problem. Still others use the term to include the potential insolvency of a major player in or compo-

[2] A more precise statement would be that these are the risks that could require compensation. In equilibrium models there typically exist aggregate risks with exposures that do not require compensation. Diversification arguments narrow the pricing focus to the systematic or aggregate risks.

nent of the financial system. Thus systemic risk is basically a grab bag of scenarios that are supposed to rationalize intervention in financial markets. These interventions come under the heading of "macroprudential policies". Since the great recession was triggered by a financial crisis, it is not surprising that there were legislative calls for external monitoring, intervention or regulation to reduce systemic risk. The outcome is legislation such as the rather cumbersome and still incomplete 2,319 page Dodd-Frank Wall Street Reform and Consumer Protection Act. The sets of constructs for measurement to support prudent policy-making remains a challenge for future research.

Embracing Koopmans's call for models is appealing as a longer-term research agenda. Important aspects of his critique are just as relevant as a commentary on current systemic risk measurement as it was for Burns and Mitchell's business cycle measurement.[3] There are, however, important conceptual challenges that go along with the use of explicit dynamic economic models in formal ways. Paramount among these is how we confront risk and uncertainty. Economic models with explicit stochastic structures imply formal probability statements for a variety of questions related to implications and policy. In addition, uncertainty can come from limited data, unknown models and misspecification of those models. Policy discussions too often have a bias towards ignoring the full impact of uncertainty quantification. But abstracting from uncertainty measurement can result in flawed policy advice and implementation.

2.2 Systemic risk or uncertainty

There are various approaches to uncertainty quantification. While there is well known and extensive literature on using probability models to support statistical measurement, I expect special challenges to emerge when we impose dynamic economic structure onto the measurement challenge. The discussion that follows is motivated by this latter challenge. It reflects my own perspective, not necessarily one that is widely embraced. My perspective is consonant, however, with some of the views expressed by Haldane (2011) in his discussions of policy simplicity and robustness when applied to regulating financial institutions.

I find it useful to draw a distinction between risk and alternative concepts better designed to capture our struggles with constructing fully specified probability models. Mo-

[3] One way in which the systemic risk measurement agenda is more advanced than that of Burns and Mitchell is that there is a statistical theory that can be applied to many of the suggested measurements of systemic risk. The ability to use "modern methods of statistical inference" was one of the reasons featured by Koopmans for why formal probability models are valuable, but another part of the challenge is the formal integration with economic analysis.

tivated by the insights of Knight (1921), decision theorists use the terms *uncertainty* and *ambiguity* as distinguished from risk. See Gilboa and Schmeidler (1989) for an initial entrant to this literature and Gilboa et al. (2008) for a recent survey. Alternatively, we can think of statistical models as approximations and we use such models in sophisticated ways with conservative adjustments that reflect the potential for misspecification. This latter ambition is sometimes formulated as a *concern for robustness*. For instance, Petersen et al. (2000) and Hansen and Sargent (2001) confront a decision problem with a family of possible probability specifications and seek conservative responses.

To appreciate the consequences of Knight's distinction, consider the following. Suppose we happen to have full confidence in a model specification of the macroeconomy appropriately enriched with financial linkages needed to capture system-wide exposure to risk. Since the model specifies the underlying probabilities, we could use it both to quantify systemic risk and to compute so-called counterfactuals. While this would be an attractive situation, it seems not to fit many circumstances. As systemic risk remains a poorly understood concept, there is no "off the shelf" model that we can use to measure it. Any stab at building such models, at least in the near future, is likely to yield, at best, a coarse approximation. This leads directly to the question: how do we best express skepticism in our probabilistic measurement of systemic risk?

Continuing with a rather idealized approach, we could formally articulate an array of models and weight these models using historical inputs and subjective priors. This articulation appears to be overly ambitious in practice, but it is certainly a good aim. Subjective inputs may not be commonly agreed upon and historical evidence distinguishing models may be weak. To make this approach operational leads naturally to a sensitivity analysis for priors including priors over parameters and alternative models.

A model by its very nature is wrong because it simplifies and abstracts. Including a formal probabilistic structure enriches predictions from a model, but we should not expect such an addition to magically fix or repair the model. It is often useful to throw other models "into the mix" so to speak. The same limitations are likely to carry over to each model we envision. Perhaps we could be lucky enough to delineate a big enough list of possible models to fill gaps left by any specific model. In practice, I suspect we cannot achieve complete success and certainly not in the short term. In some special circumstances, the gaps may be negligible. Probabilistic reasoning in conjunction with the use of models is a very valuable tool. But too often, we suspect the remaining gaps are not trivial, and the challenge in using the models is capturing how to express the remaining skepticism. Simple models can contain powerful insights even if they are incomplete along some dimensions. As statisticians with incomplete knowledge, how do we embrace such models or collections of them while acknowledging skepticism that should justifiably go along with them? This is an enuring problem in the use of dynamic stochastic equilibrium models and it seems unavoidable as we confront the important task of building models designed to measure systemic risk. Even as we add modeling clarity, in my view we need to abandon the presumption that we can measure fully *systemic risk* and go after the conceptually more difficult notion of quantifying *systemic uncertainty*. See Haldane (2011) for a further discussion of this point.

What is at stake here is more than just a task for statisticians. Even though policy challenges may appear to be complicated, it does not follow that policy design should be complicated. Acknowledging or confronting gaps in modeling has long been conjectured to have important implications for economic policy. As an analogy, I recall Friedman (1960)'s argument for a simplified approach to the design of monetary policy. His policy prescription was premised on the notion of "long and variable lags" in a monetary transmission mechanism that was too poorly understood to exploit formally in the design of policy. His perspective was that the gaps in our knowledge of this mechanism were sufficient that premising activist monetary policy on incomplete models could be harmful. Relatedly Cogley et al. (2008) show how alternative misspecification in modeling can be expressed in terms of the design of policy rules. Hansen and Sargent (2012) explore challenges for monetary policy based on alternative specifications of incomplete knowledge on the part of a so-called "Ramsey planner". The task of this planner is to design formal rules for implementation. It is evident from their analyses that the potential source of misspecification can matter in the design of a *robust rule*. These contributions do not explore the policy ramifications for system-wide problems with the functioning of financial markets, but such challenges should be on the radar screen of financial regulation. In fact, implementation concerns and the need for simple rules underly some of the arguments for imposing equity requirements on banks. See, for instance, Admati et al. (2010). Part of policy implementation requires attaching numerical values to parameters in such rules. Thus concerns about systemic uncertainty would still seem to be a potential contributor to the implementation of even seemingly simple rules for financial regulation.

Even after we acknowledge that policy makers face challenges in forming systemic risk measures that could be direct and explicit tools for policy, there is another layer of uncertainty. Sophisticated decision-makers *inside* the models we build may face similar struggles with how to view their economic environments. Why might this be important? Let me draw on contributions from two distinct stands of literature to speculate about this.

Caballero and Simsek (2010) consider models of financial networks. In such models financial institutions care not only about the people that they interact with, say their neighbors, but also the neighbors of neighbors, and so forth. One possibility is that financial entities know well what is going on at all nodes in the financial network. Another is that while making probabilistic assessments about nearby neighbors in a network is straightforward, this task becomes considerably more difficult as we consider more indirect linkages, say neighbors of neighbors of neighbors .... . This view is made operational in the model of financial networks of Caballero and Simsek (2010).

In a rather different application Hansen (2007) and Hansen and Sargent (2010) consider models in which investors struggle with alternative models of long-term economic growth. While investors treat each such model as misspecified, they presume that the models serve as useful benchmarks in much the same way as in stochastic specifications of robust control theory. Historical evidence is informative, but finite data histories do not accurately reveal the best model. Important differences in models may entail subtle components of economic growth that can have long-term macroeconomic consequences. Concerns about model-misspecification become expressed more strongly in financial markets in some time periods than others. This has consequences for the valuation of capital in an uncertain environment and on the market tradeoffs confronted by investors who participate in financial markets. In the example economies considered by Hansen (2007) and Hansen and Sargent (2010), what they call uncertainty premia become larger after the occurrence of a sequence of bad macroeconomic outcomes.

In summary, the implications of systemic uncertainty whether in contrast or in conjunction with systemic risk are both important for providing policy advice and understanding market outcomes. External analysts, say statisticians, econometricians and policy advisors, confront specification uncertainty when they build dynamic stochastic models with explicit linkages to the financial markets. Within dynamic models with micro foundations are decision makers or agents that also confront uncertainty. Their resulting actions can have a big impact on the systemwide outcomes. Assessing both the analysts' and agents' uncertainties are critical components to a productive research agenda.

3 Current approaches

Let me turn now to some of the recent research related to systemic risk. Just the wide scope of Bisias et al. (2012) survey reminds us that there is not yet an agreed upon single approach to this measurement. To me, it suggests that what measurements will be the most fruitful to support our understanding of linkages between financial markets and the macroeconomy is an open issue. In a superficial way, the sheer number of approaches would seem to address the Kelvin dictum. The problem is complex and it has many dimensions to it and thus requires multiple measurements. But I am doubtful that this is a correct assessment of the situation. Alternative measures are supported implicitly by alternative modeling assumptions and it is hard to see how the full array of measurements provides a coherent set of tools for policy makers. Many of the measurements to date seem closer in spirit to the Burns and Mitchell approach and fall way short of the Koopmans standard. From a policy perspective, I fear that we remain too close to the Potter-Stewart "we know it when we see it" view of systemic risk.

What follows is a discussion of a few specific approaches for assessing systemic risk along with some modeling and data challenges going forward.

3.1 Tail measures

One approach measures co-dependence in the tails of equity returns to financial institutions. Some form of co-dependence is needed to distinguish the impact of the disturbances to the entire financial sector from firm-specific disturbances. Prominent examples of this include the work of Adrian and Brunnermeier (2008) and Brownlees and Engle (2011). Measuring tail dependence is particularly challenging because of limited historical data. To obtain estimates requires implicit extrapolations from the historical time series of returns because of the very limited number of extreme values of the magnitude of a financial crisis. While co-dependence helps to identify large aggregate shocks, all such shocks are in effect treated as a conglomerate when extracting information from historical evidence. The resulting measurements are interesting, but they put aside some critical questions that are needed to understand better policy advice. For example, while equity returns are used to identify an amalgam of aggregate shocks that could induce crises, how does the *mechanism* by which the disturbance is transmitted to the macroeconomy differ depending on the *source* of the disturbance? Not all financial market crises are macroeconomic crises. The big drops in equity markets on October 19, 1987 and April 14, 2000 did not trigger major macroeconomic declines. Was this because of the source of the shock or because of the macroeconomic policy responses? Understanding both the source and the mechanism of the disturbance would seem to be critical to the analysis of policy implications. Further empirical investigation of financial linkages with macroeconomic repercussions should be an important next step in this line of research.

It is wrong to say that this tail-based research is devoid of theory, and in fact Acharya et al. (2010) suggest how to use tail-risk measures as inputs into calculations about the solvency of the financial system. Their paper includes an explicit welfare calculation, and their use of measurements of tail dependence is driven in part by a particular policy perspective. Their theoretical supporting analysis is essentially static in nature, however. The macroeconomic consequences of crises events and how they unfold over time is largely put to the side. Instead, the focus is on providing a measure of the public cost of providing capital in order to exceed a specific threshold. This research does result in model-based measurements of what is called *marginal expected shortfall* and *systemic risk*. These measurements are updated regularly on the V-Lab web page at New York University. The use by Acharya et al. (2010) is an interesting illustration of how to model systemic risk and may well serve as a valuable platform for a more ambitious approach.

The focus on equity calculations limits the financial institutions that can be analyzed. The so-called shadow banking sector contains potentially important sectors or groups of firms that are not publicly traded. One could argue that if the monitoring targets are only SIFI's (so called systemically important financial institutions), then the focus on publicly-traded financial firms is appropriate. But system-wide policy concerns might be directed at the potential failure of collections of firms in the shadow banking sector including ones that are not publicly traded and hence omitted by calculations that rely on equity valuation measures.

3.2 Contingent claims analysis

In related research, Gray and Jobst (2011) apply what is known as contingent claims analysis. This approach features risk adjustments to sectoral balance sheets while featuring the distinct roles of debt and equity. It builds on the use of option pricing theory for firm financing where there is an underlying stochastic process for the value of the firm assets. Equity is a call option on these assets and debt is the corresponding put option. Gray and Jobst (2011) discuss examples of this approach extended to sectors of the economy including the government. In their applications, they measure sectoral balance sheets with a particular interest in financial crises. This approach neatly sidesteps statistical challenges by using "market expectations" and risk-adjusted probabilities in conjunction with equity-based measures of uncertainty and simplified models of debt obligations. Extending contingent claims analysis from the valuation of firms to systems of firms and governments is fruitful. Note however, if our aim is to make welfare assessments and direct linkages to the macroeconomy, then the statistical modeling and measurement challenges that are skirted will quickly resurface. Market expectations and risk-neutral probability assessments offer the advantage of not needing to distinguish actual probabilities from the marginal utilities of investors in financial markets, but this advantage can only be pushed so far. A more fundamental understanding of the market-based "appetite for risk" and a characterization of the macroeconomic implications of the shocks that command large risk prices require further modeling and a more prominent examination of historical evidence. Such an understanding is central when our ambition is to engage in the analysis of counterfactuals and hypothetical changes in policies.[4]

3.3 Network models

Network models of the financial system offer intriguing ways to summarize data because of its focus on interconnectedness. These models open the door to some potentially important policy questions, but there are some critical shortcomings in making these models fully useful for policy. A financial firm in a network may be highly connected, interacting with many firms. Perhaps these links are such that the firm is "too interconnected to fail". A critical input into a policy response is how quickly the networks structure will evolve when such a firm fails. As is well recognized, in a dynamic setting these communications links will be endogenous, but this endogeneity makes modeling in a tractable way much more difficult and refocuses some of the measurements needed to address policy concerns.

3.4 Dynamic, stochastic macroeconomic models

Linking financial market disruption to the macroeconomy requires more than just using off-the-shelf dynamic stochastic equilibrium models, say, of the type suggested by Christiano et al. (2005) and Smets and Wouters (2007). By design, models of this type are well suited

[4]The potential omission of firms not publicly traded limits this approach for the reasons described previously.

for econometric estimation and they measure the consequences of multiple shocks and model explicitly the transition mechanisms for those shocks. Identification in these multi-shock models is tenuous. More importantly they are "small shock" models. In order to handle a substantial number of state variables, they appeal to small noise approximations for analytical tractability. Since the financial crisis, there has been a rush to integrate financial market restrictions into these models. Crises are modeled as times when *ad hoc* financial constraints bind.[5] To use the existing methods of analysis, separate local approximations are made around crisis periods. See Gertler and Kiyotaki (2010) for a recent development and discussion of this literature.

Enriching dynamic stochastic equilibrium is a promising research agenda, but this literature has only scratched the surface on how to extend these models to improve our understanding of the macroeconomic consequence to upheaval in financial markets. It remains an open research question as to how best i) to model financial constraints, both in terms of theoretical grounding and empirical importance; ii) to characterize the macroeconomic consequences of crisis level shocks that are very large but infrequent; and iii) to model the origins of these shocks.[6]

3.5 Pitfalls in data dissemination and collection

Measurement requires data. Going forward, there is great opportunity for the Office of Financial Research in the United States and its counterparts elsewhere to provide new data for researchers. Some of the data in its most primitive form will be confidential. Concern for confidentiality will create challenges for sharing this information with external researchers. One approach is to restrict the use of such data to be "in house." This should be avoided. The best way to ensure the high quality of research within government agencies is to make important components of the data available to external researchers. This external access is necessary not only to allow for replication of results, but also to nurture innovative modeling and measurement.[7] Moreover, external analysis can provide a check against research with pre-ordained conclusions or inadvertent support for policies

[5] I use the term *ad hoc* in a less derogatory manner than many other economists. I remind readers of a dictionary definition: concerned or dealing with a specific subject, purpose, or end.

[6] For instance, the Macroeconomic Financial Modeling group funded by the Sloan Foundation explores the challenges to building quantitatively ambitious models that address these and other related challenges.

[7] Andy Lo has made the related point that potentially relevant sectors, such as the insurance sector, are not under the formal scrutiny of the federal government and hence there may be an important shortfall in the data available to the Office of Financial Research.

such as "too big (or too something) to fail." While external access will require that data be distributed in manners that respect individual confidentiality, the possibility of making such data available is a reality. The Census department has already confronted such challenges successfully.

There is an additional data issue that requires scrutiny. Distortions in the collection of publicly available data can hinder the measurement of aggregate risk exposures because of the temptation to disguise the problematic nature of policies in place. Brickell (2011) argues that this may have contributed to assessing housing market risk at the outset of the great recession. Concerns of this nature place an extra burden on empirical researchers to model the biases in data collection induced by both public and private incentives for distortion.

Given this state of econometric modeling and measurement, I see a big gap to fill between statistical analyses measuring co-movements in the tails of financial market equity returns and empirical analyses measuring the impact of shocks to the macroeconomy. This gap limits, at least temporarily, the scope of the analysis of systemic risk. Closing this gap provides an important opportunity for the future. The compendium of systemic risk measures identified in Bisias et al. (2012) should be viewed merely as an interesting start. We should not lose sight of the longer-term challenge to provide systemic risk quantification grounded in economic analysis and supported by evidence. The need for sound theoretical underpinnings for producing policy relevant research identified by Koopmans many decades ago still applies to the quantification of systemic risk. Policy analysis stemming from econometric models aims to push beyond the realm of historical evidence through the use of well-grounded economic models. It is meant to provide a framework for the conduct of hypothetical policies that did not occur during the historical observation period. To engage in this activity with the ambition to understand better how to monitor or regulate the financial sector to prevent major upheaval in the macroeconomy requires creative adjustments in both our modeling and our measurement.

4 Conclusion

The need to implement new laws with expanded regulation and oversight put pressures on public sector research groups to develop quick ways to provide useful measurements of systemic risk. This requires short cuts, but it also can proliferate superficial answers. These short-term research responses will be revealing along some dimensions by providing useful summaries from new data sources or at least data sources that have been largely ignored in the past. Stopping with short term or quick answers can lead to bad policy advice and should be avoided. It is important for researchers to take a broader and more ambitious attack on the problem of building quantitatively meaningful models with macroeconomic linkages to financial markets. Appropriately constructed, these models could provide a framework for the quantification of systemic risk.

We should not underestimate the difficulty of this challenge, but success offers the potential for valuable inputs into policy making. Wearing my econometrician's hat has led me to emphasize measurement challenges and the associated uncertainty caused by limited data or unknown statistical models used to generate the data. Of course clever econometricians can always invent challenges, and in many respects part of the econometrician's job is to provide credible ways to quantify measurement uncertainties. After all, quantitative research in economics grounded by empirical evidence should be more than just reporting a single number but instead ranges or distributions that include sensitivity to model specification. Good econometrics is supported simultaneously by good economics and good statistics. Exploring the consequences of potential model misspecification necessarily requires inputs from both economics and statistics. Economic models help us understand what statistical inputs are most consequential to economic outcomes and good statistics reveal where the measurements are least reliable. Moreover, such econometric explorations will benefit discussions of policy by providing repeated reminders of why gaps in our knowledge can have important implications.

Allow me to close by returning to the Kelvin dictum and drawing on some intellectual history of it as it relates to social science research. The decision to place this dictum on the Social Science Research building at the University of Chicago caught the attention of some distinguished scholars. This building housed the economics department for many years and the Cowles Commission for Research in Economics during the years 1939 to 1955 when many young scholars came there to explore linkages between economics, mathematics and statistics.[8] Two of the original pillars of the "Chicago school", Knight and Viner, had notable reactions to the use of the Kelvin quote and proposed amendments:[9]

> Knight: If you cannot measure a thing, go ahead and measure it anyway.
>
> Viner: ... and even when we can measure a thing, our knowledge will be meager and unsatisfactory.

Perhaps just as intriguing as Knight's and Viner's scepticism are the major challenges that were levied to Lord Kelvin's own calculations about the age of the sun. These challenges provide an object lesson in support of "model uncertainty." Kelvin argued that the upper bound of the sun's age was 20-40 million years, although his earlier estimates included the possibility of a much larger number, up to 100 million years. Kelvin's evidence and that provided by others were used to question the plausibility of the Darwinian theory of evolution. Darwin's own calculations suggested that much more time was needed to justify the evolutionary processes. In hindsight, Lord Kelvin's estimates were over one hundred times lower than the current estimate of 4.5 billion years. Kelvin's understatement was revised upward by substantive advances in our understanding of radioactivity as an energy source. This historical episode illustrates rather dramatically an impact of model uncertainty on the quality of measurement. While Knight's and Viner's words of caution were motivated by their perception of social science research several decades ago, their concerns extend to other research settings as well. It is difficult to fault Lord Kelvin for not anticipating the discovery of a new energy source. Nevertheless, I do not wish to conclude that the potential for model misspecification should induce us to abandon earnest attempts at quantification. Instead quantification should be a valued exercise, and part of this exercise should include a characterization of sensitivity to alternative model specifications. Unfortunately, there are no guarantees that we have captured the actual form of the misspecification among the possibilities that we consider, but at least we can avoid some of the pitfalls of using models in naive ways.

Quantitative ambitions have the virtue of providing clarity for what is to be measured. Models provide measurement frameworks and facilitate communication and criticism. While simple quantifications of systemic risk may be a naive hope, producing better models to support policy discussion and analysis is a worthy ambition. Building a single

[8]After moving to Yale in 1955, the Cowles Commission was renamed as the Cowles Foundation.
[9]See Merton et al. (1984).

consensus model is unrealistic in the near term, but even exploring formally the consequences of alternative models adds discipline to policy advice. Without such modeling pursuits, we are left with a heavy reliance on discretion in governmental course of action. Perhaps discretion is the best we can do in some extreme circumstances, but formal analysis should provide coherency and transparency to economic policy.

While systemic-risk modeling and measurement is a promising research agenda, caution should prevail about the impact of model misspecification on the measurements and the consequences of those measurements. A critical component to this venture should be to assess and guard against adverse impacts of the use of measurements from necessarily stylized models. Complete success along this dimension is asking too much, otherwise we would just “fix” our models. Nevertheless, confronting the various components of uncertainty with some formality will help us to use models in sensible and meaningful ways. As our knowledge and understanding advance over time, so will our comprehension and characterization of uncertainty in our model-based, quantitative assessments.

References

Acharya, Viral V., Christian Brownlees, Robert Engle, Farhang Farazmand, and Matthew Richardson. 2010. Measuring Systemic Risk. In *Regulating Wall Street*, edited by Viral V. Acharya, Thomas F. Cooley, Matthew Richardson, and Ingo Walter, 85–119. John Wiley and Sons, Inc.

Admati, Anat R., Peter M. DeMarzo, Martin F. Hellwig, and Paul Pfleiderer. 2010. Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity Is Not Expensive. Research Papers 2065, Stanford University, Graduate School of Business.

Adrian, Tobias and Markus K. Brunnermeier. 2008. CoVaR. Tech. rep., Federal Reserve Bank of New York, Staff Reports.

Bisias, Dimitrios, Mark Flood, Andrew W. Lo, and Stavros Valavanis. 2012. A Survey of Systemic Risk Analytics. Working Paper 0001, Office of Financial Research.

Brickell, Mark. 2011. Lessons Gleaned from Flawed Mortgage Risk Assessments. *Financial Times* April 13.

Brownlees, Christian and Robert Engle. 2011. Volatility, Correlation and Tails for Systemic Risk Measurement. Tech. rep., SSRN.

Burns, Arthur F. and Wesley C. Mitchell. 1946. *Measuring Business Cycles*. NBER Book Series Studies in Business Cycles.

Caballero, Ricardo and Alp Simsek. 2010. Fire Sales in a Model of Complexity. MIT.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy* 113 (1):pp. 1–45.

Cogley, Timothy, Riccardo Colacito, Lars Peter Hansen, and Thomas J. Sargent. 2008. Robustness and U.S. Monetary Policy Experimentation. *Journal of Money, Credit and Banking* 40 (8):1599–1623.

Friedman, Milton. 1960. *A Program for Monetary Stability*. New York: Fordham University Press.

Gertler, Mark and Nobuhiro Kiyotaki. 2010. Financial Intermediation and Credit Policy in Business Cycle Analysis. In *Handbook of Monetary Economics*, vol. 3 of *Handbook of Monetary Economics*, edited by Benjamin M. Friedman and Michael Woodford, chap. 11, 547–599. Elsevier.

Gilboa, Itzhak and David Schmeidler. 1989. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18 (2):141–153.

Gilboa, Itzhak, Andrew W. Postlewaite, and David Schmeidler. 2008. Probability and Uncertainty in Economic Modeling. *The Journal of Economic Perspectives* 22 (3 (Summer)):173–188.

Gray, Dale F. and Andreas A. Jobst. 2011. Modelling Systemic Financial Sector and Sovereign Risk. *Sveriges Riksbank Economic Review* 2:68 – 106.

Haldane, Andrew G. 2011. Capital Discipline. Tech. rep., Bank of England. Based on a speech given to the AEA meeting in Denver.

Hansen, Lars Peter. 2007. Beliefs, Doubts and Learning, Valuing Macroeconomic Risk. *American Economic Review* 97 (2).

Hansen, Lars Peter and Thomas J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.

———. 2010. Fragile Beliefs and the Price of Uncertainty. *Quantitative Economics* 1 (1):129–162.

———. 2012. Three types of ambiguity. *Journal of Monetary Economics* forthcoming:–.

Knight, Frank H. 1921. *Risk, Uncertainty, and Profit.* Boston: Hart, Schaffner and Marx; Houghton Mifflin Co.

Koopmans, Tjalling C. 1947. Measurement Without Theory. *The Review of Economics and Statistics* 29 (3):pp. 161–172.

Merton, Robert K., David L. Sills, and Stephen M. Stigler. 1984. The Kelvin Dictum and Social Science: An Excursion into the History of an Idea. *Journal of the History of Behavioral Sciences* 20:319–331.

Petersen, Ian R., Matthew R. James, and Paul Dupuis. 2000. Minimax Optimal Control of Stochastic Uncertain Systems with Relative Entropy Constraints. *IEEE Transactions on Automatic Control* 45:398–412.

Smets, Frank and Rafael Wouters. 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *The American Economic Review* 97 (3):pp. 586–606.