

Data and the Financial Crisis: How Might We Do Better Next Time?¹

Note: The remarks below are based on my discussion of material presented in the “Monitoring Recession and Recovery” session at the Brookings/Heritage Conference “Measuring Innovation and Change during Turbulent Economic Times.” These materials, as well as other information from the conference, are available at: <http://www.heritage.org/press/events/ev111709a.cfm>.

We have seen an excellent set of presentations this morning that covered a lot of ground—from Paul Smith’s and [Marshall Reinsdorf’s](#) discussions of ongoing improvements to our methods for estimating key household financial variables, to [Steve Landefeld’s](#) ideas for expanding and supplementing the national accounts in useful ways, to the policy data “wish lists” that Becky Blank and [Alan Krueger](#) laid out. In my remarks, I will focus on data needs related to the financial crisis. I will present four key points that draw off what the others have said but also reflect my own experience analyzing the crisis and developing and evaluating policy options. I have been engaged in these issues both as a Brookings scholar and prior to my recent arrival at Brookings, when I was at the Federal Reserve Board working on housing, household finance, and broader financial market issues.

To set the stage, let me define what I mean by financial-crisis-related data needs. I am referring to the need for data that speak to the following three questions:

1. How do we anticipate possible adverse financial developments?
2. How do we recognize and respond to such developments once they are underway?
3. How do we avoid or at least respond more effectively to the next (possible) crisis?

These three questions are clearly critical for policymakers, but they are also relevant to anyone who is making decisions that are contingent on the future path of the economy and the risks around it. Further, the task of providing answers to the questions does not fall singularly to the policymaking community but rather to a much broader group of scholars and analysts who are studying the issues, and there are important social gains to having them participate in the policy discussion.

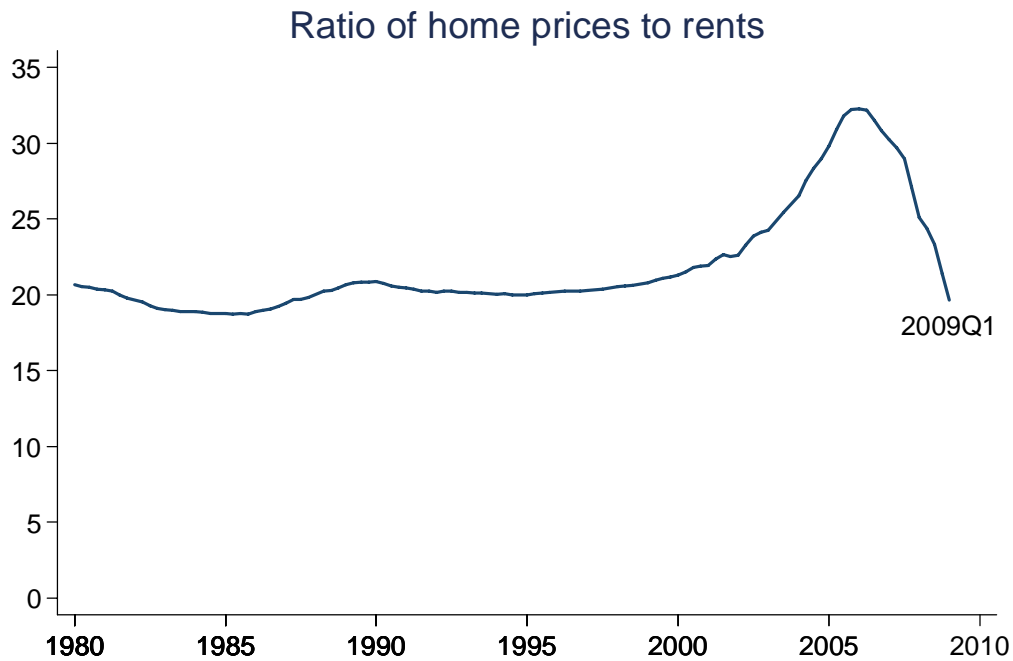
Point #1. Developing and publishing aggregate indicators of financial imbalances would be helpful but will take some work.

Such measures are also sometimes referred to as “sustainability” measures, as in [Steve Landefeld’s paper](#). They fit into a broader class of aggregates that more closely resemble things that analysts care about as opposed to aggregates that are of interest mainly because they are a component of some important top-line series. I agree with Steve that a lot of the pieces we need to create these measures are already available—and thus that the main work involved in publishing a set of such indicators is deciding what measures are of interest and then mixing and matching the right series to create them.

¹ I thank Howie Lempel for research assistance and helpful comments.

An example of one such indicator would be a series that captured the overvaluation of house prices. For example, many analysts have found the aggregate ratio of home prices to rents to be a useful gauge of imbalances in the housing market. Figure 1 shows a measure, based on data from a recent paper by Morris Davis, Andreas Lehnert, and Robert Martin.²

Figure 1



Note: Data from Davis, Lehnert, and Martin, 2009.
Available at <http://www.lincolnst.edu/subcenters/land-values/rent-price-ratio.asp>.

The measure speaks to the valuation of house prices because it shows whether, at the national level, home prices have gotten out of line with the flow of services that one might expect from these homes, as captured by average rents. The series is thus the housing market equivalent to the price-earnings ratios at which stock market analysts sometimes look to judge whether stocks are valued correctly. As can be seen, house prices rose dramatically above the level justified by rents in the early to middle part of this decade before correcting sharply. As of early this year, the ratio was back to a level that was fairly typical by pre-credit-boom standards.

A statistical agency such as the Bureau of Economic Analysis could create a set of indicators of financial imbalances and publish regular updates. The set might include

² See Davis, Morris A., Andreas Lehnert, and Robert F. Martin, 2008, "The Rent-Price Ratio for the Aggregate Stock of Owner-Occupied Housing," *Review of Income and Wealth*, vol. 54 (2), pp. 279-284. The authors used the decennial Census of Housing to calculate the average value of owner-occupied housing units and average annual rents for rental units in Census years and then interpolated between these years and extrapolated after 2000 using quarterly home price and rent indexes. The quarterly home price series used for this figure came from the Macromarkets LLC national house price index (formerly Case-Shiller-Weiss) and the quarterly rent series came from the Bureau of Labor Statistics.

both series that are constructed by the data provider like the housing valuation measure above as well as indicators of financial imbalances that are already constructed and published, such as the Federal Reserve's [debt service and financial obligations ratios](#). Such a collection would facilitate analysis of financial risks and conditions both within the policymaking community and elsewhere.

That said, the endeavor would not be completely straightforward for two reasons. First, if done by an agency primarily tasked with data publication, it would in part represent a new mission, as it would put the agency in the business of making recommendations as to what indicators should be watched. The agency would then bear some responsibility for whether the set of indicators actually captures imbalances in the financial system. Further responsibility would come with the very real possibility that the new information about imbalances could have important effects on market prices. All told, these considerations would move the mission of the data-producing institution closer to that of a policymaking institution.

Second (and relatedly), identifying what measures belong in the set would be challenging. Even once the broad concepts to include—such as asset valuation measures, saving measures, and leverage measures for households, businesses, and financial institutions—have been selected, there is considerable work in deciding exactly how to construct these measures. For example, there are many variants on the housing valuation measure presented above. As Davis, Lehnert, and Martin show, using a different measure of house prices in the numerator implies a considerably smaller degree of peak overvaluation than what is shown in the figure. Analysts also argue about whether rents or incomes should be in the denominator.

One way to identify what indicators should be included in the set would be to explore what measures show up repeatedly in reports by institutions whose work has been closely tied to the financial crisis, such as the Federal Reserve, the International Monetary Fund, and the Bank of International Settlements. One might also look at commentary by Wall Street analysts and others writing about the crisis. In addition, it would be sensible to solicit the views of academic experts who have done research using such measures. One could envision another conference focused exclusively on what measures would be useful and how best to deliver them to the public.

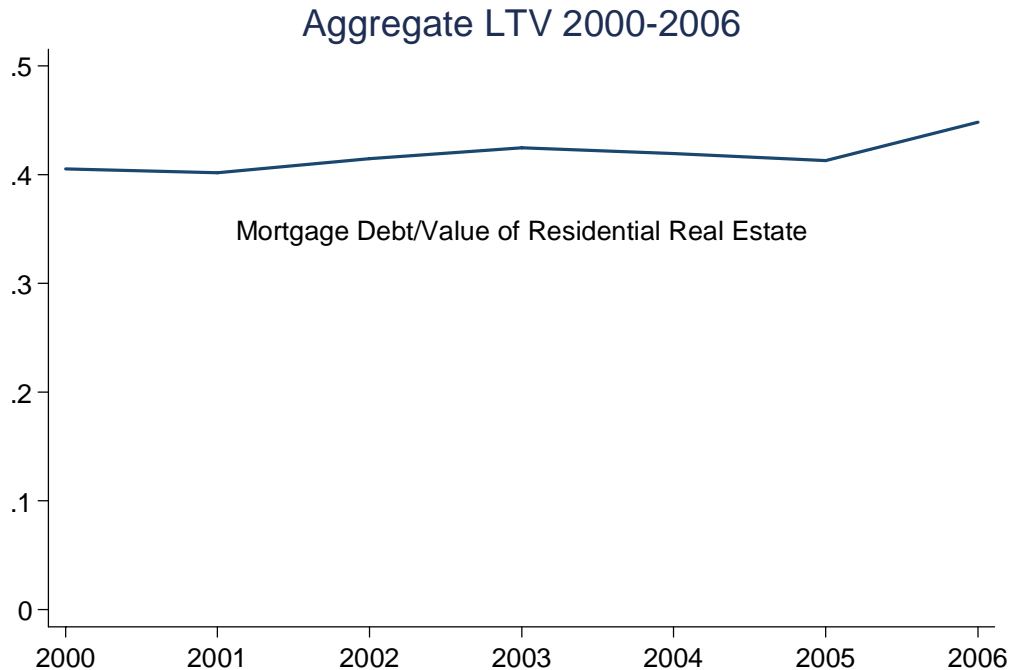
Point #2: To anticipate, recognize, and respond to this crisis, we needed more than macro data or data on the “average” experience—we needed timely data on the distribution and particularly the tails.

The best way to illustrate this point is with an example.

Figure 2 shows the aggregate value of home mortgage debt divided by the aggregate value of residential real estate owned by the household sector, both from the U.S. Flow of Funds accounts, in the period leading up to the crisis. The series essentially represents an aggregate loan-to-value ratio, which, in turn, reflects the *average* experience among U.S. homeowners. As can be seen, this measure barely budged during the credit boom early

this decade and thus was not setting off warning bells for those monitoring financial conditions.

Figure 2

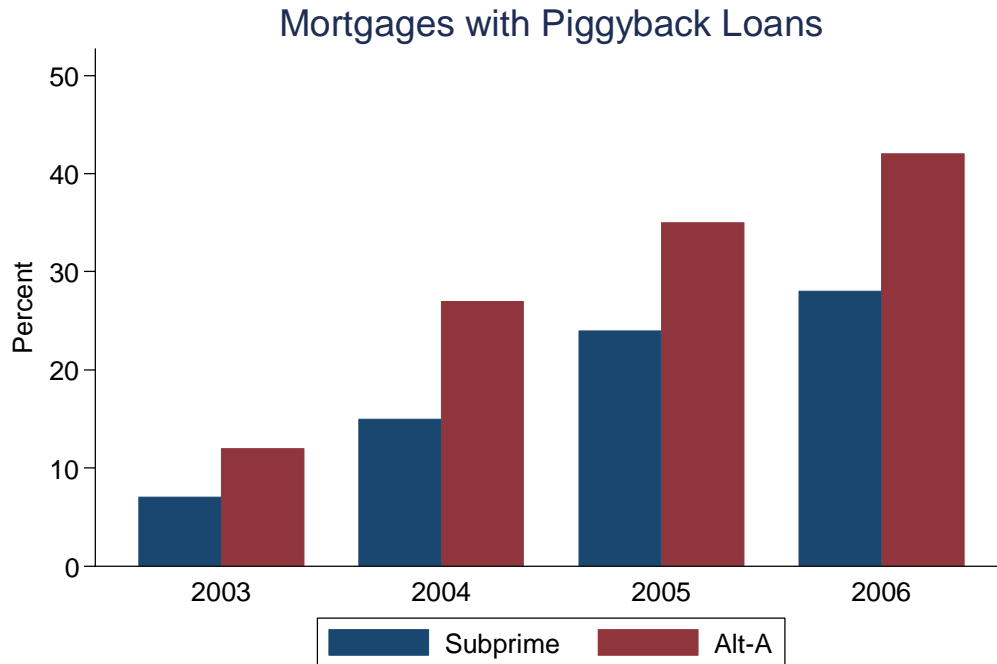


Source: U.S. Flow of Funds accounts.

A dramatically different picture emerges if we change our focus from the average homeowner to those at the upper tail of the distribution of mortgage leverage. Figure 3 shows data from a paper by Chris Mayer, Karen Pence, and Shane Sherlund that depicts the fraction of non-prime mortgages with so-called piggy back loans—second liens taken out at the time of origination of the mortgage, which allowed homeowners to take on much more leverage.³ This fraction showed an alarming rise from 2003 through 2006, capturing the trend toward excess mortgage leverage that most analysts now agree was a key precipitating factor for the crisis.

³ Mayer, Christopher, Karen Pence, and Shane M. Sherlund, 2009, “The Rise in Mortgage Defaults,” *Journal of Economic Perspectives*, vol. 23 (1), pp. 27-50.

Figure 3



Source: Mayer, Pence, and Sherlund, 2009.

Other series related to the distribution of credit would also have promoted more widespread recognition of the impending financial crisis. For example, media reports during the credit boom frequently expressed concern over the rising popularity of nontraditional mortgage products like “interest-only” and “negative amortization” loans, but hard numbers on their prevalence were rarely cited. Moreover, the rise in subprime “2/28” mortgages (loans for which payments were tied to a low teaser rate for two years before resetting to much higher rates) garnered very little attention until delinquencies among them began to skyrocket in the latter part of 2005. Information about trends in the fraction of households that had very high ratios of required debt payments to income would have provided another useful signal of the trouble to come.

How might such data be generated? Economists have traditionally turned to long-established large-scale surveys of households to study distributional aspects of household finances. However, these surveys are not a practical choice for detecting, monitoring, and responding to crises because of the infrequency at which they are done and the lags in data publication. For example, [as Alan Krueger mentioned](#), the results of the triennial *Survey of Consumer Finances* are not released to the public for more than a year after households are surveyed. Forcing a significant acceleration of the publication schedule for such data sets might increase their usefulness in some contexts but would likely come at considerable cost to data quality, as a great deal of time is needed simply to process and clean up the responses.

The good news is that we do have potentially excellent timely sources of distributional information from administrative records. The data in Figure 3 were derived from a large

continuously updated data set of loan-level mortgage records from LoanPerformance, a subsidiary of First American CoreLogic, Inc. Such records were originally compiled by this company and others so that investors could monitor the performance of mortgages that had been pooled into mortgage-backed-securities. However, the data have increasingly been used in recent years by researchers studying the crisis. The data sources tend to lack information about the current income and demographic characteristics of the borrower, but, as interest has increased in using these data for research purposes, analysts have made progress in devising ways to merge in data from other sources.

Another potentially valuable source of distributional information is consumer credit records. These records complement loan-level records, as they paint a full picture of an individual's credit use, and so reveal cases where the burden associated with any given loan is not large but the total burden on an individual is likely unsustainable. As with mortgage records, these data sets tend to be updated on a continuous basis and can, to some extent, be merged with information about other characteristics of the individual.

Point #3: We need to consider the appropriate role of the government with regard to proprietary sources of information about the state of household finances and financial markets.

This point builds off of my last point. The bad news about the good news above is that the loan-level mortgage records and broader credit records are, for the most part, proprietary and very expensive. So, a key part of the financial data discussion needs to concern whether the government should take a role in getting information to the public when timely useful data are available but hard to obtain privately for cost or other reasons.

One might argue that the most important need is for the relevant policymaking agencies to have such data. While this may be true, there are also arguments for making the information available more broadly. For example, some commentators have argued that regulatory agencies faced analytical limitations and tendencies towards "group think" that contributed importantly to their failure to recognize just how much risk was building up during the credit boom. To the degree that such arguments are valid, there is a case for providing greater information to the public so that there can be a more vigorous examination of and debate over financial trends.

Releasing loan-level or individual-level records from these data sources is probably not a practical option. The expense of doing so would be very large, as the data vendors would need to be compensated for essentially taking away their business. There are also important privacy concerns, as even if the most obvious personal information (such as names, specific addresses, and social securities numbers) were stripped from the records, it might still be possible to identify individuals through their credit usage patterns. Finally, using the loan-level data properly involves a considerable investment in time and data skills, raising the concern that impatient users might take shortcuts that resulted in inaccurate results.

That said, there is probably a happy medium for the government's role in providing information from these data sources. Instead of focusing on the release of loan-level mortgage data or individual credit records, public efforts could focus instead on releasing summary statistics for different parts of the distribution. For example, a statistical agency might generate information on the fractions of households in different risk buckets and publish updates on a regular basis.

Looking beyond distributional data on household finances, there are many other types of financial data that are difficult to obtain through current public sources. Examples include information about the outstanding levels and new issuance of various types of securities, as well as rates (or spreads) on these securities. The [Federal Reserve's daily commercial paper release](#) could serve as a model in this regard. To be sure, the costs of undertaking such a release for other types of securities would need to be weighed against the benefits, but the issue is certainly worth a serious discussion.

Point #4. Household survey data, while typically not useful for monitoring current financial conditions, are needed to understand underlying relationships, which can help both to prevent future crises and deal with them once they arise.

Thus far, I have primarily focused on the data needed to monitor financial conditions in real time. As I said earlier, the large-scale surveys of households are not especially useful in this regard because of the lags in publication of the results. However, that does not mean that these surveys are irrelevant to our ability to detect and respond to crises. Data sets such as the *Survey of Consumer Finances*, the *Panel Study on Income Dynamics*, and the *Consumer Expenditure Survey* have been the basis for longer-term research on key issues like the causes of household financial distress and differences in financial behavior across different types of households.⁴

Most of these surveys have been running for many years, but, given the strains associated with the federal budget outlook, we should not take their continued financing for granted. Probably most at risk are those surveys that receive large amounts of support from government grants. For example, funding just the "core" questions from the *Panel Study on Income Dynamics* (PSID) over the five years ending in 2011 has required \$25 million in grants, and the continuation of the survey depends on successful renewal of these grants.⁵

Further, there are arguments for increasing support of these surveys. An expansion of funding for household surveys would allow surveys to expand as new needs arise. For example, the PSID added questions on whether respondents had undergone a home foreclosure to its 2009 wave, but the section was limited in scope because of funding constraints. In addition, some surveys need more funding to maintain the usefulness of their results. For instance, researchers have voiced concerns that the *Consumer*

⁴ I mention these three data sets largely because I have used them in my own research, but there are a number of other sources that have yielded valuable results in this area such as the *Health and Retirement Survey*, the *Survey on Income and Program Participation*, the *National Longitudinal Surveys*, and the *Current Population Survey*.

⁵ In full disclosure I am currently on the Board of Overseers for this survey.

Expenditure Survey has become increasingly unrepresentative over time, based on (among other things) analyses that show that trends in measured saving amongst survey participants do not track trends in aggregate saving. A half-day session at a [National Bureau of Economic Research conference last summer](#) was dedicated to discussing the problem, with some participants arguing that the survey had not kept pace with changes in the way in which households track their spending and, accordingly, might be improved through use of electronic data such as credit card records.

Conclusion

My remarks today have stressed four points: (1) that developing and publishing a set of aggregate indicators of financial imbalances would be helpful, but will take some work to implement; (2) that to better detect and address the crisis, analysts needed more timely distribution data on household finances, particularly at the riskier tail; (3) that part of the discussion should concern what role the government should take in distribution information based on proprietary sources; and (4) that continued and, in some cases, expanded support of household surveys should yield information on how to better detect and address the next financial crisis.

Most of my suggestions involve committing more public resources to data production and distribution. An effort to create and publish a set of indicators of financial imbalances could require significant energy and money, particularly if it were to include measures based on distributional information. Likewise, initiatives to expand or significantly improve household survey data could be costly. But the evidence suggests that, as we debate steps that will reduce the likelihood and mitigate the costs of another financial crisis, we should focus not only on strengthening the financial regulatory system and policymaking institutions, but also on providing better information to the community of analysts, both within government and more broadly.