

See *important*
subscription
change inside

recent research

research review



Issue No. 11, January 2009 – June 2009

federal reserve bank of boston

Please note:

As a result of extraordinary budget pressures in these difficult times, we will be publishing future issues of *Research Review* online only. Like previous issues, these will be available at <http://www.bos.frb.org/economic/ResearchReview/index.htm>

If you would like to receive e-mail alerts when research department papers or publications, such as a new issue of *Research Review*, are posted on our website, please enroll at <http://www.bos.frb.org/genpubs/email/index.htm>

We appreciate your interest in this publication and your great response to our survey. (See the special feature in this issue for a summary of the results.) We are working to increase the research output of the department, and we anticipate posting many interesting summaries in coming issues. We hope you will stay with us during this transition and beyond.

Research Department

Jeffrey C. Fuhrer

Executive Vice President and
Director of Research

Geoffrey M. B. Tootell

Senior Vice President and
Deputy Director of Research

Economists

Giovanni P. Olivei, VP
Robert K. Triest, VP
Michelle L. Barnes
Katharine Bradbury
Mary A. Burke
Christopher L. Foote
Fabià Gumbau-Brisa
Julian Jamison
Yolanda K. Kodrzycki
Jane Sneddon Little
Scott Schuh
Joanna Stavins
J. Christina Wang
Paul S. Willen

Manager

Patricia Geagan, AVP

Editors

Suzanne Lorant
Elizabeth Murry

Graphic Designer

Heidi Furse

Research Review is a publication of the
Research Department of the Federal Reserve
Bank of Boston.

ISSN 1552-2814 (print)
ISSN 1552-2822 (online)

©Copyright 2009
Federal Reserve Bank of Boston

research review

Issue No. 11, January 2009 – June 2009

Research Review provides an overview of recent research by economists of the research department of the Federal Reserve Bank of Boston. Included are summaries of scholarly papers, staff briefings, and Bank-sponsored conferences.

Research Review is available without charge. To be placed on the mailing list, or for additional copies, please contact the Research Library:

Research Library—D
Federal Reserve Bank of Boston
600 Atlantic Avenue
Boston, MA 02210
Phone: 617.973.3397
Fax: 617.973.4221
E-mail: boston.library@bos.frb.org

Research Review is available on the web at

<http://www.bos.frb.org/economic/ResearchReview/index.htm>.

Views expressed in *Research Review* are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System. The authors appreciate receiving comments.

Research Department Papers Series of the Federal Reserve Bank of Boston

Public Policy Discussion Papers present research bearing on policy issues. They are generally written for policymakers, informed business people, and academics. Many of them present research intended for professional journals.

Working Papers present statistical or technical research. They are generally written for economists and others with strong technical backgrounds, and they are intended for publication in professional journals.

Public Policy Briefs present briefing materials prepared by Boston Fed research staff on topics of current interest concerning the economy.

Research department papers are available online only.

<http://www.bos.frb.org/economic/research.htm>

Executive Summaries in This Issue

Public Policy Discussion Papers

- p-09-1 Making Sense of the Subprime Crisis 4
**Kristopher S. Gerardi, Andreas Lehnert, Shane M. Sherlund,
and Paul S. Willen**
- p-09-2 Reducing Foreclosures 7
**Christopher L. Foote, Kristopher S. Gerardi, Lorenz Goette,
and Paul S. Willen**
- p-09-3 Reviving Mortgage Securitization: Lessons from the Brady Plan
and Duration Analysis 10
Fabià Gumbau-Brisa and Catherine L. Mann

Working Papers

- w-09-1 Why Are (Some) Consumers (Finally) Writing Fewer Checks? 13
The Role of Payment Characteristics
Scott Schuh and Joanna Stavins
- w-09-2 Another Hidden Cost of Incentives: The Detrimental Effect on 15
Norm Enforcement
Andreas Fuster and Stephan Meier
- w-09-3 Has Overweight Become the New Normal? Evidence of a Generational 19
Shift in Body Weight Norms
Mary A. Burke, Frank Heiland, and Carl Nadler
- w-09-4 Empirical Estimates of Changing Inflation Dynamics 22
Jeffrey C. Fuhrer, Giovanni P. Olivei, and Geoffrey M. B. Tootell
- w-09-5 Geographic Variations in a Model of Physician Treatment Choice 25
with Social Interactions
Mary A. Burke, Gary M. Fournier, and Kislaya Prasad
- w-09-6 The Optimal Level of Deposit Insurance Coverage 29
Michael Manz

Special Feature

- Results of the *Research Review* Survey 31

Contributing Authors

33



Public Policy Discussion Papers

p-09-1

Making Sense of the Subprime Crisis

by Kristopher S. Gerardi, Andreas Lehnert, Shane M. Sherlund, and Paul S. Willen

complete text: <http://www.bos.frb.org/economic/ppdp/2009/ppdp0901.htm>

e-mail: kristopher.gerardi@atl.frb.org, andreas.lehnert@frb.gov, shane.m.sherlund@frb.gov, paul.willen@bos.frb.org

Motivation for the Research

If market participants had anticipated the increase in defaults on subprime mortgages originated in 2005 and 2006, the nature and extent of the current financial market disruptions would be very different. Before the bursting of the housing bubble, investors in subprime mortgage-backed securities would have demanded higher returns and greater capital cushions. As a result, borrowers would not have found credit as cheap or as easy to obtain as it became during the subprime credit boom of 2005–2006. Rating agencies would have had a similar reaction, rating a much smaller fraction of each deal investment grade. After the downturn in house prices, the increase in foreclosures would have been significantly smaller, with fewer attendant disruptions to the housing market. In addition, investors would not have suffered such outsized, and unexpected, losses. To make sense of the subprime crisis, one needs to understand why, when accepting significant exposure to the creditworthiness of subprime borrowers, so many smart analysts, armed with advanced degrees, data on the past performance of subprime borrowers, and state-of-the-art modeling technology failed to anticipate that so many of the loans they were buying, either directly or indirectly, would go bad. This paper explores the question of whether market participants could have or should have anticipated the large increase in foreclosures that occurred in 2007 and 2008. Specifically, the authors explore whether market participants underestimated the likelihood of a fall in house prices or the sensitivity of foreclosures to house prices.

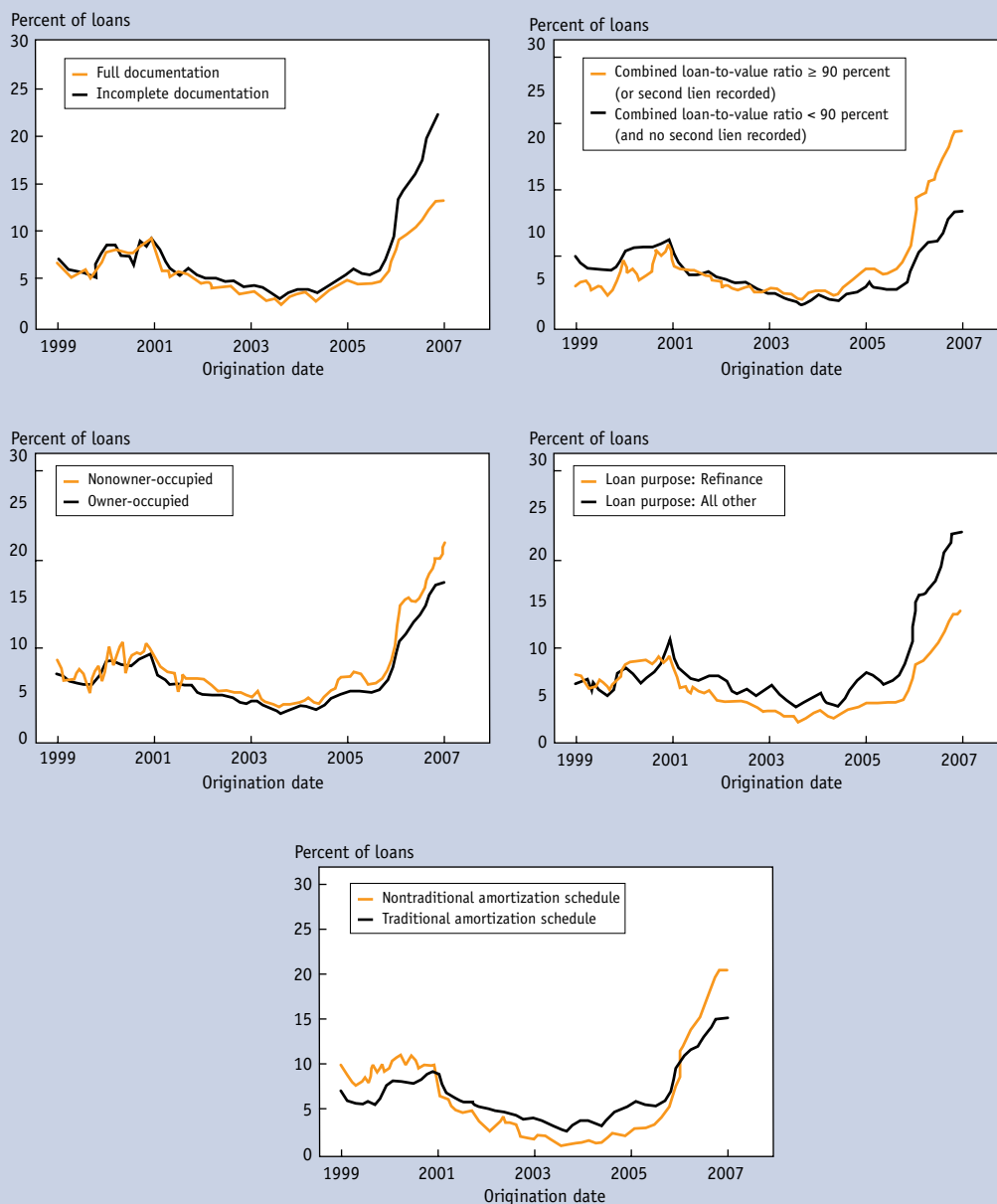
Research Approach

The authors use two different methods to put themselves in the place of analysts in 2005, when subprime lending was still thriving, and look forward. The first method forecasts performance using only data available in 2005; the second looks at what market participants wrote at the time. In the first approach, the authors estimate the type of hazard models commonly used in the industry to predict mortgage defaults. The estimates are based on two data sources: (1) a loan-level data set from First American Loan Performance used extensively in the industry to track the performance of mortgages in mortgage-backed securities and (2) an ownership-level data set from the Warren Group, which tracks the fates of homebuyers in Massachusetts from the late 1980s forward. Ownership-level data differ from loan-level data in that they consider the entire loan experience of an owner from purchase until sale or foreclosure, regardless of any instances of refinancing. Thus, a loan that terminates in refinancing within an ownership that ends in default appears as a happy ending in a loan-level data set but not in an ownership-level data set. These ownership-level data were not widely used in the industry but were available, at least in theory.

Key Findings

- A decline in underwriting standards alone cannot explain the dramatic rise in foreclosures, yet most of the foreclosures that occurred in 2007 and 2008 did stem from loans originated in 2005 and 2006, and loans originated in those years did carry extra risk factors—particularly increased leverage.
- Given available data, a researcher with perfect foresight about the trajectory of house prices from 2005 forward would have forecast a large increase in foreclosures starting in 2007, although models

Default Characteristics of Subprime Mortgages by Month of Origination

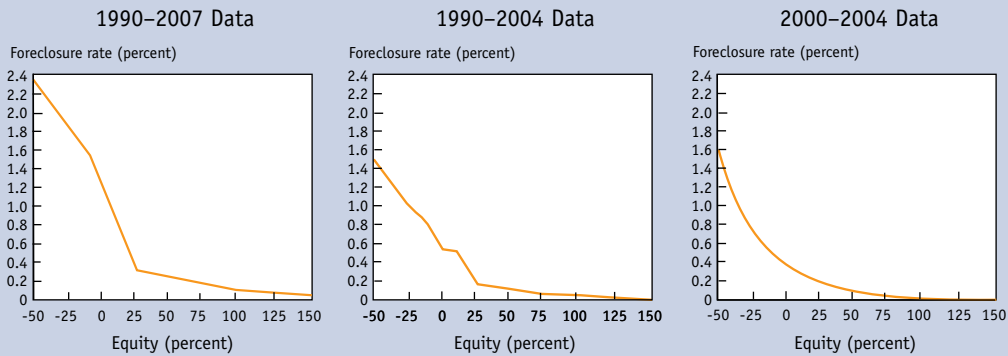


Source: First American Loan Performance ABS data.

estimated with loan-level (as opposed to ownership-level) data would have predicted a smaller rise than actually occurred.

- Loans made in 2005–2006 were not very different from loans made earlier, which had performed well despite carrying a variety of serious risk factors. Most of the misplaced optimism about losses stemmed from misjudgment concerning the evolution of house prices and not from misjudgment of the quality of the underwriting.

Estimated Effect of Equity on Foreclosure



Source: Authors' calculations, based on deed-registry data from the Warren Group.

- Many commentators have attributed the crisis to the payment shock associated with the first reset of subprime 2/28 mortgages. However, the evidence from loan-level data shows that resets cannot account for a significant portion of the increase in foreclosures. The overwhelming majority of defaults on subprime adjustable-rate mortgages occurred long before the first reset. (A 2/28 mortgage is a mortgage in which the contract rate is fixed at an initial “teaser” rate for two years, after which it adjusts to the six-month LIBOR rate plus a predetermined margin—often around 6 percentage points.)
- A review of analyst reports shows that investment analysts on the whole understood with remarkable accuracy how falling house price appreciation would affect the performance of subprime mortgages and the securities backed by these mortgages, but they assigned a low probability to such an outcome.
- Five main themes emerge from studying the analyst reports. First, the subprime market was viewed by market insiders as a great success in 1995. Second, subprime mortgages were viewed, in some sense correctly, as lower risk than prime mortgages because of their more stable prepayment behavior. Third, analysts used fairly sophisticated tools but were hampered by the absence of episodes of falling prices in their data. Fourth, many analysts anticipated the crisis in a qualitative way, laying out in various ways a roadmap of what could happen, but they never fleshed out the quantitative implications. Fifth, analysts were remarkably optimistic about house price appreciation.

Implications

This paper focuses almost exclusively on subprime mortgages, but many of the same arguments can be applied to prime mortgages.

Events with the potential to have catastrophic impact should not be discounted completely based on a judgment that they are unlikely to occur.

Reducing Foreclosures

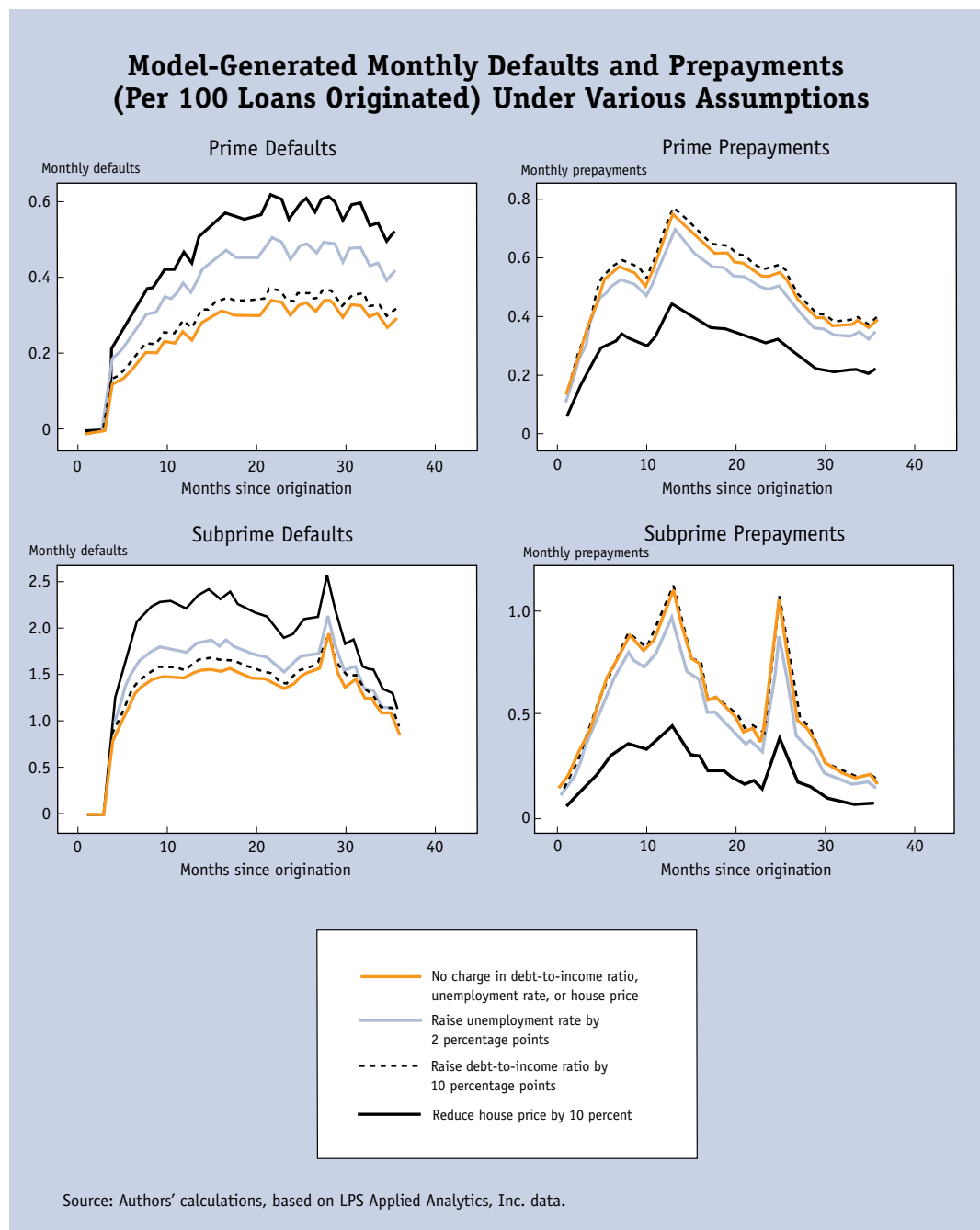
by Christopher L. Foote, Kristopher S. Gerardi, Lorenz Goette, and Paul S. Willen

complete text: <http://www.bos.frb.org/economic/ppdp/2009/ppdp0902.htm>

e-mail: chris.foote@bos.frb.org, kristopher.gerardi@atl.frb.org, lorenzfabian.goette@unil.ch, paul.willen@bos.frb.org

Motivation for the Research

One of the most important challenges now facing U.S. policymakers stems from the tide of foreclosures that is engulfing the country. There is no shortage of suggestions for how to attack the problem. One of the most influential strands of thought contends that the crisis can be attenuated by changing the terms of “unaffordable” mortgages. It is thought that modifying mortgages is beneficial not just



for borrowers in danger of losing their homes but also for lenders, who would recover more from modifications than from foreclosures. Proponents of this view, however, worry that without government intervention this win-win outcome will not occur. Their concern is that the securitization of mortgages has given rise to contract frictions that prevent lenders and their agents (loan servicers) from carrying out modifications that would benefit both borrowers and lenders. With respect to a borrower's decision to default, many commentators have argued that loose lending practices, including relaxed underwriting standards for the debt-to-income ratio (DTI) at origination and the issuance of adjustable rate mortgages with extremely low teaser rates, were the primary factor in subsequent defaults, by saddling borrowers with mortgages they could not sustain. The authors of this paper take a skeptical look at these arguments and investigate two economic decisions: the borrower's decision to default on a mortgage and the lender's choice to offer the borrower a loan modification or foreclose on a delinquent loan.

Research Approach

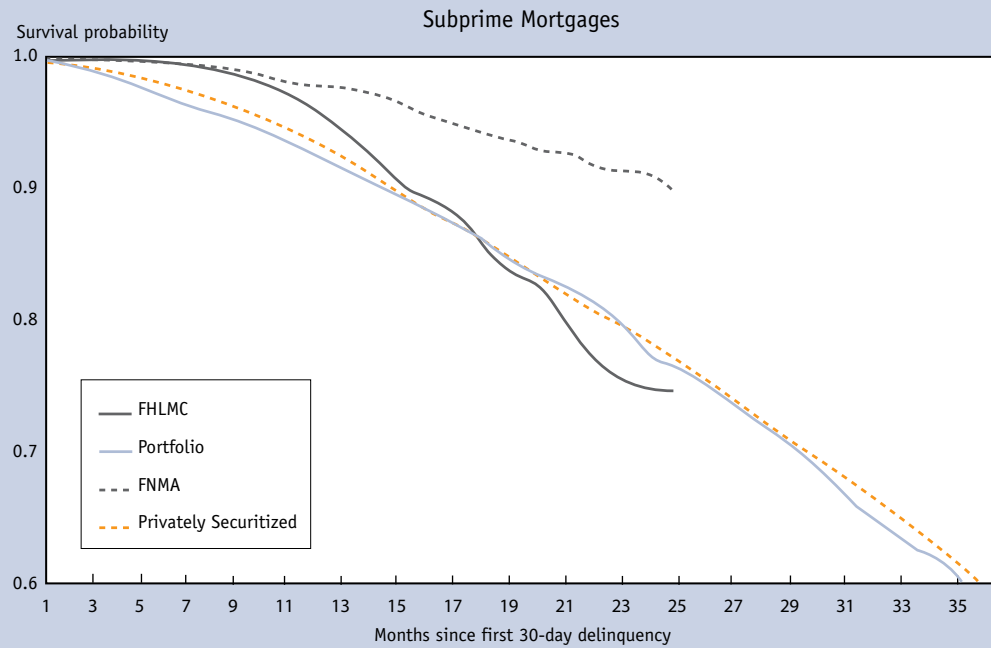
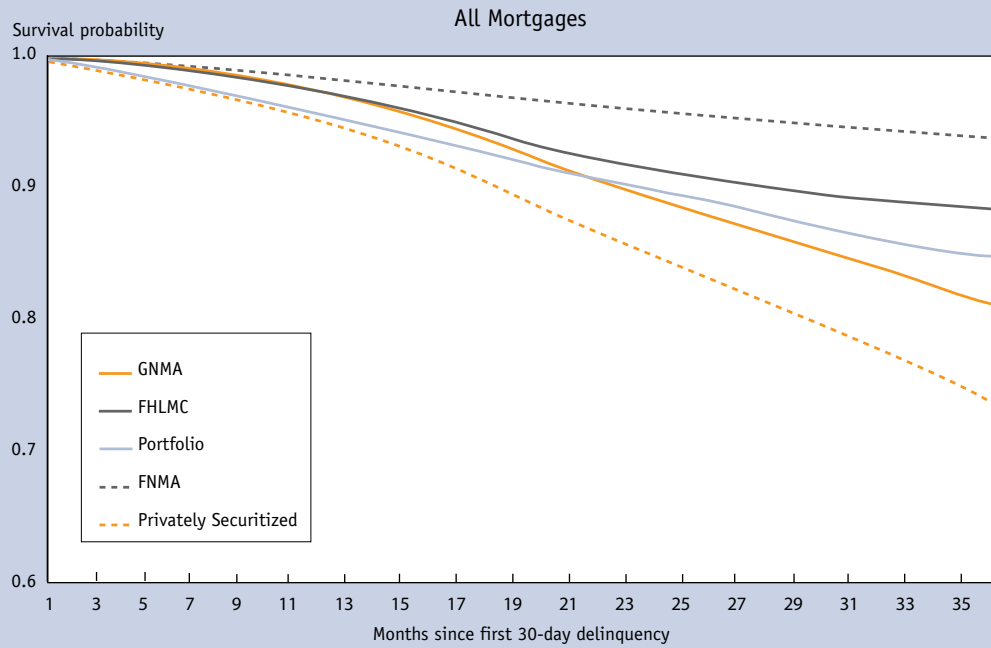
The authors perform two empirical tests, using a loan-level data set compiled from large loan-servicing organizations by LPS Applied Analytics, Inc., to examine the evidence regarding these two decisions. This data set, which covers more than half the U.S. mortgage market, contains information on key variables set at the time of each loan's origination, including the amount of the loan, the appraised value and location of the property that secures the loan, the borrower's DTI at the time of origination, whether the loan is classified as a prime or subprime mortgage, whether it is a first or second lien, and, importantly, whether the loan is held in the lender's portfolio or has been packaged into a mortgage-backed security (securitized). Information is included indicating whether the interest rate on the loan is fixed or adjustable; the manner in which the interest rate changes, if adjustable; whether a loan ended in payment, prepayment, or default; and, for each month that a given loan is in the data, the loan's outstanding balance, the current interest rate, and borrower's payment status (current, 30-, 60-, or 90-days delinquent, in foreclosure, etc.).

In the first empirical test, the authors run Cox proportional hazard models for both defaults and prepayments, using the LPS data, to learn how various risk characteristics, including the importance of affordability at origination, as measured by DTI, and macroeconomic variables, including the unemployment rate and house prices, affect the likelihood of default. The second empirical test asks whether securitized loans are less likely to be modified if they become delinquent than are loans held by the issuer in its portfolio. To examine this question, the authors plot the loans' survival functions, comparing those that have been securitized with those that have not.

Key Findings

- A 10-percentage-point increase in the debt-to-income ratio (DTI), defined here as mortgage payments relative to income (rather than the alternative definition of all debt to income), increases the probability of a 90-day delinquency by 7 to 11 percent, depending on the individual borrower. By contrast, a 1-percentage-point increase in the unemployment rate raises this probability by 10 to 20 percent, while a 10-percentage-point fall in house prices raises it by more than half. Thus, "unaffordable" loans, defined as those with a high DTI at origination, are unlikely to be the main reason that borrowers default.
- What really matters in a borrower's default decision is the monthly mortgage payment relative to the borrower's income in the present and future. Thus, the typical problem leading to default appears to be a combination of shocks to household income and an unprecedented fall in house prices that makes it difficult or impossible for the borrower to solve his unaffordable debt problem either by selling the house and paying off the mortgage with the proceeds or by refinancing. Resets of adjustable mortgage rates are of only limited importance. Many commentators have put resets at the heart of the crisis, but the simulations show that it is difficult to support this claim.

Survival Estimates: Transition from 30-Day Delinquency to Modification



Source: Author's calculations, based on LPS Applied Analytics, Inc. data.

Note: For all mortgages in the data set (upper panel), privately securitized loans and GNMA loans are more likely to have been modified than either loans held in the lender's portfolio or FNMA loans, over a fairly long horizon. Over a shorter horizon (less than one year), there is very little difference across various types of loans. when conditioning on a 30-day delinquency. For the sample of subprime and Alt-A mortgages, which are the loans most likely to become delinquent (lower panel), the incidence of modification is virtually the same over all horizons for portfolio-held and privately securitized loans.

- The evidence of contract frictions is weak, at least if these frictions result from the securitization of the loan: in this data set, which includes both securitized and nonsecuritized mortgages, the two types of loans were modified at about the same rate. The relatively small number of loan modifications to date is explained more plausibly by the greater efficiency of foreclosure over modification for investors than by contract frictions related to securitization agreements between servicers and investors.
- While it is true that lenders may lose a great deal of money with each individual foreclosure, it is also the case that the loan modifications may have negative net present value if they are sometimes extended to people who would be likely to pay on time anyway. And the benefits of modifications are uncertain if borrowers are unemployed, as jobless borrowers are more likely than other borrowers to default even after modification.

Implications

These findings suggest that policymakers intent on reducing foreclosures should focus on mitigating the effects of income volatility by helping people who lose their jobs get through difficult periods without having to lose their homes. For example, the government could replace a portion of lost income for a period of one or two years, through a program of loans or grants to individual homeowners. For more permanent and very large setbacks, foreclosure-reduction policy might help homeowners transition to renting, through short sales or other procedures. In any case, the results suggest that policies that encourage moderate, long-term reductions in DTIs face important hurdles in addressing the current foreclosure crisis.

The authors do not argue that the crisis is completely unrelated to looser lending standards, which saddled borrowers with high-DTI mortgages. Rather, they argue that foreclosure-prevention policy should focus on the most important source of defaults. In the data, this source appears to be the interaction of falling house prices with adverse life events.

p-09-3

Reviving Mortgage Securitization: Lessons from the Brady Plan and Duration Analysis

by Fabià Gumbau-Brisa and Catherine L. Mann

complete text: <http://www.bos.frb.org/economic/ppdp/2009/ppdp0903.htm>
e-mail: fabia.gumbau-brisa@bos.frb.org, clmann@brandeis.edu

Motivation for the Research

Since 2007, trading activity has declined dramatically in all U.S. credit securitization markets. Moreover, from a high of some \$4 trillion, new issuance of pooled mortgage securities and asset-backed securities has stalled. U.S. policymakers have implemented programs to inject liquidity into these markets and have even implemented direct purchases of some of these asset types, achieving some success—in the first months of 2009 mortgage rates reached record lows and mortgage-backed securities (MBS) issued by government-sponsored entities increased somewhat. Nonetheless, in broad terms, MBS still appear to be trading at large discounts, and private sector entities remain on the sidelines with respect to issuing these securities. Arguably, investors' excessive risk perception could explain the problem of issuance and the lack of secondary market trading for these important assets.

This paper draws analogies between the current MBS crisis and the Latin American debt crisis that extended from the early 1980s to the early 1990s, when the Brady Plan revived the issuance of debt and the trading of troubled legacy loans. At the heart of the problem in both instances, the “originate/syndicate” business model underpriced risk and spread the assets among many investors. In both

periods, both debtors and creditors faced solvency problems following macroeconomic stress. During the Latin American debt crisis, market liquidity disruptions gave rise to excessive risk premia; today the U.S. market for securitized credit faces the same situation.

Two approaches were tried to revive lending to Latin America. The Baker Plan, implemented between 1984 and 1988, focused on injecting liquidity into the market to solve the dual solvency problem. Yet the Baker strategy failed to revive trade in existing debt, did not revalue collapsed prices on existing debt, provided little new private money to borrowers, aided creditors' balance sheet problems only minimally, and by relying on official support, kept private investors on the sidelines. After this first systemic intervention failed, the Brady Plan, administered from 1989 to 1992/1994, successfully restored normal functioning in the Latin American sovereign debt markets and provided new policy tools to deal with sovereign debt crises thereafter. The key element of the Brady Plan was to collateralize specific cash flows of the renegotiated debt obligations.

The paper highlights key elements underpinning the success of the Brady Plan, and extracts lessons to formulate a similar policy for the current MBS crisis.

Research Approach

The authors examine critical elements of the Brady Plan and focus their attention on the collateralization of the principal repayment of cash flows at the 30-year maturity of the new (renegotiated) bonds. They use duration analysis to assess the debt's price risk, first using the Brady example and then applying the method to mortgage pass-through securities. Duration measures the semi-elasticity of a bond's price with respect to changes in yield—that is, it captures the sensitivity of the bond's price to changes in both riskless interest rates and any risk premia or spreads. Based on this analysis, the authors draw relevant lessons for fixing the problems in the MBS market related to the current credit crisis and argue that the current situation facing the U.S. market for pooled mortgage securities, given the interventions implemented to date, resembles the aftermath of the Baker Plan episode.

Extending the analysis to account for the main features of pooled mortgage securities today, the paper proposes a plan similar to the Brady Plan in that it targets cash flows key to pricing risk. Providing guarantees for these cash flows, which are payable near the mid-life of the underlying loans (some three to ten years hence), would have the greatest impact on market prices today at the lowest expected cost for the public institution acting as guarantor. Moreover, the risk for the guarantor of losing the collateral would only materialize years after the present period and its accompanying financial turmoil; thus, a successful intervention also protects the guarantor.

Key Findings

- The contrast in outcomes between the Baker Plan and the Brady Plan presents a cautionary tale on relying on liquidity provision alone to fix the solvency problem faced by both borrowers and creditors. Focusing on the dual solvency problem that the Baker Plan did not address, the Brady Plan recognized the need to write down the value of the existing debt that was not trading in an open market.
- The central feature of the Brady Plan was cash flow guarantees, along with offering a menu of alternatives to repackage the existing troubled debt into new, tradable 30-year bonds. Importantly, the principal of the bonds, payable at maturity, was fully collateralized with a zero coupon 30-year U.S. Treasury security. There were three reasons why this credit enhancement reduced the risk premium on the debt. First, the collateral improved the likelihood of repayment, which lenders accepted in exchange for some debt forbearance. Second, borrowers were more able to repay the lower levels of debt (even without considering the macroeconomic reforms they undertook, which also improved global macro conditions). Third, the enhancement promoted trade in the new assets in secondary markets, thus reducing the risk premia associated with a lack of market depth.

- The collateral enhancement had an important “bang for the buck.” Each \$1.00 invested in collateral resulted in the forgiveness of almost \$2.50 of outstanding debt—an instantaneous gain. Even in the worst-case scenario, involving eventual default and full loss of the collateral, the immediate social return was 150 percent. Thus, the expected cost of the program was small relative to its social benefit.
- Credit enhancements should focus on the cash flows that contribute the most to an asset’s duration, because these cash flows would require the smallest reduction in risk premia to achieve a given increase in asset price. In the case of the Brady bonds, the cash flow with the largest duration-contribution was the principal repayment at the 30-year maturity (the Brady bonds were 30-year bullet bonds). The financial instruments that are at the heart of today’s financial crisis are mortgages, which are amortizing loans with a prepayment option. Thus, the duration contribution profile for a mortgage pass-through security looks different from a bullet bond. Nevertheless, by assuming prepayment rates based on the PSA (Public Securities Association) benchmarks, the duration-contribution profile for a pool of 30-year fixed-rate mortgages suggests that targeting credit guarantees to some of the cash flows between years three and ten would have the lowest expected cost for a given target increase in the security’s price. The credit enhancement could be achieved by partial or full collateralization of the expected cash flow, for instance by pledging Treasury STRIPS. This collateral would provide a back-stop to potential losses, much as did the Brady bonds collateral.
- In the case of the Brady Plan, the authors show that the collateralization of the cash flow with the largest contribution to the asset’s duration was also the least costly approach to improving market functioning. Arguably, applying the Brady lessons and duration analysis to the market for MBS offers a cost-effective approach to complement strategies currently being employed to support this market.

Implications

Reviving the market for Latin American debt was a prolonged process, and delaying an effective workout program exacted costs on borrowers, creditors, and the global economy. The Brady Plan revived the normal functioning of the Latin American debt market, reduced the risk of borrower defaults, and improved creditors’ balance sheets. Applying lessons gleaned from the Brady Plan could revive the normal functioning of U.S. credit markets, particularly for assets with longer maturities such as mortgage pools, and could achieve this goal at a cost relatively low compared to its social benefits.

The authors contend that an optimally timed credit enhancement based on duration analysis would have the direct effect of increased valuations for MBS and reduced downside risk, which would boost secondary market activity for these assets and thereby also support new issuance. During the Latin American crisis, lenders were willing to reduce the value of the outstanding loans in exchange for greater certainty of repayment. The same principle could apply to current mortgage securitizations, since servicers should renegotiate mortgages as if they were the owners benefiting from this adjustment. More importantly, issuance of these securities would be cheaper, and part of these savings could be passed on to new borrowers. Making securitization more attractive could also reduce mortgage defaults, because renegotiating a troubled mortgage in order to include it in a pool that qualifies for this type of intervention would be more attractive for banks. Moreover, an important indirect benefit of the credit enhancement program is that if it successfully restores the market and contributes to the economic recovery, the actual probability of having to make any guaranteed payments is lowered.

Hence, refining the timing of the suggested enhancements has direct and indirect effects that reinforce each other. An optimally designed, large commitment can be more effective and have a lower final cost to the taxpayer than a smaller, ineffective one. Given the large public outlays to date, the authors recommend that policymakers consider the proposed cost-effective policy approach to mitigate the financial crisis and achieve market-improving outcomes.

Working Papers

w-09-1

Why Are (Some) Consumers (Finally) Writing Fewer Checks? The Role of Payment Characteristics

by Scott Schuh and Joanna Stavins

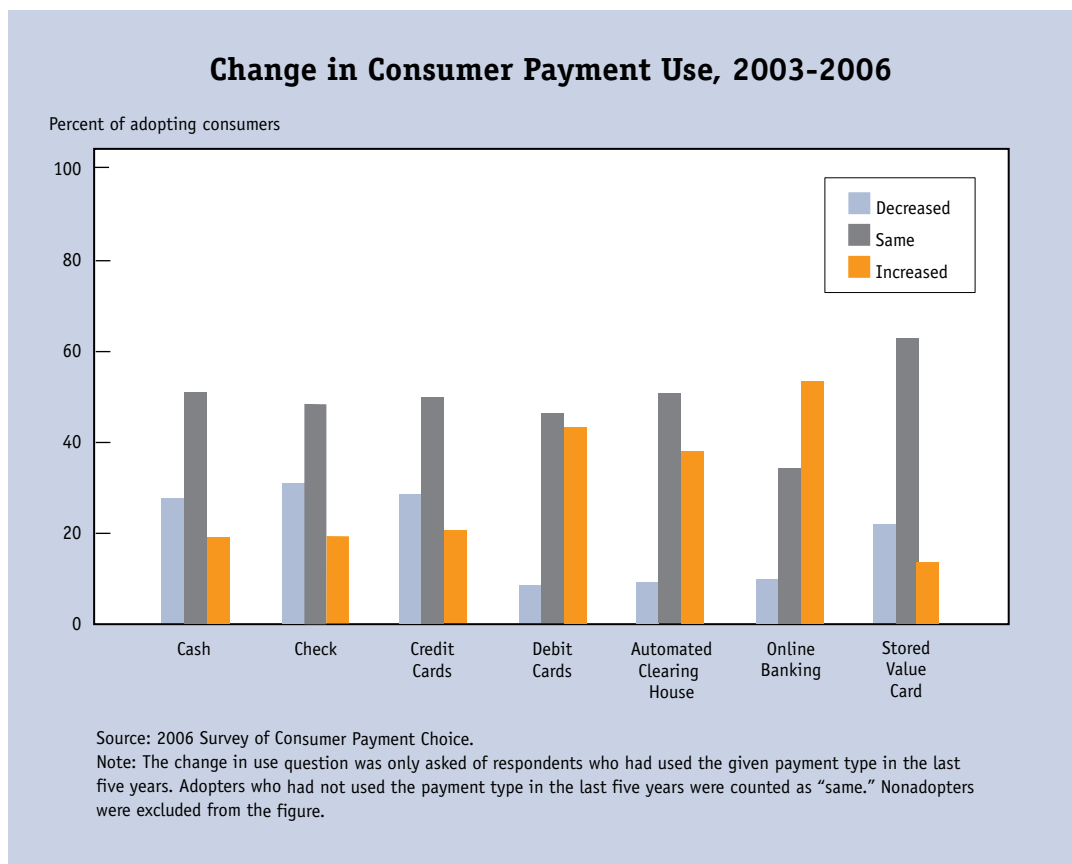
complete text: <http://www.bos.frb.org/economic/wp/wp2009/wp0901.htm>

e-mail: scott.schuh@bos.frb.org, joanna.stavins@bos.frb.org

Motivation for the Research

Since the mid-1990s, the U.S. payment system has been undergoing a transformation featuring a significant decline in the use of paper checks. Between 1995 and 2004, consumer adoption of checks and credit cards—traditional payment instruments—remained flat, while consumer adoption of debit cards, automated clearing house payments (ACH), and online bill payments jumped significantly. Meanwhile, the share of all noncash payments (by consumers, businesses, and government) made with checks fell from 77 percent to 33 percent. Although checks remain the single most common form of noncash retail payment, electronic payments together now account for the majority of non-cash payments.

This transformation has been quite uneven across consumers. Although aggregate check volume declined by 16 percent between 2003 and 2006, only 31 percent of U.S. consumers reduced their use of checks and only 0.4 percent stopped writing checks entirely. And during this era of supposed check demise, 19 percent of consumers actually reported increasing their use of checks. Nevertheless,



the consumer share of check use likely declined because most consumers reduced or held steady the number of checks they wrote and because the number of payments likely increased with personal income over time.

Many aspects of this payments transformation are not well understood. Why did it begin in the mid-1990s and not earlier or later? Why did paper check volume decline as much and as rapidly as it did? How much more will check volume decline, and what payment instruments will dominate in the coming years? And why are only one-third of U.S. consumers writing fewer paper checks, especially given that use of these checks is so costly relative to other payment instruments? Knowing the answers to these questions is vital for the Federal Reserve and the U.S. payment system. But answers have been elusive. This paper attempts to determine why some consumers are writing fewer checks and other are not.

Research Approach

Following the literature, the authors develop econometric models of payment adoption and payment use. The models extend the literature in two dimensions: (1) modeling both adoption and use—measured as the number of payments per month rather than a qualitative measure—of all payment instruments held by each consumer and (2) modeling adoption and use simultaneously, employing the Heckman (1976) selection model, which controls for potential selection bias in payment use.

The authors also extend the literature by using a comprehensive new data source with unique information on the characteristics of U.S. payment instruments. The 2006 Survey of Consumer Payment Choice (SCPC), sponsored by the AARP and the Boston Fed, contains data on adoption and monthly use of checks plus six other payment instruments, as well as many characteristics of adult U.S. consumers. The SCPC also includes consumers' assessments of seven characteristics of payment instruments: cost, convenience, safety, privacy, accuracy, timing, and recordkeeping. Together, these data are a more comprehensive source of information on consumer payment choice than was previously available.

Key Findings

- The number of payment instruments a consumer holds is the most important determinant of the cross-sectional variation in the use of checks and other payment instruments. Payment characteristics—especially convenience, cost, timing, and recordkeeping—are also important determinants of payment use. Although a commonly used term, “convenience” is not well defined and requires more research.
- Although some demographic and income-related variables also play a role in determining the adoption and use of payment instruments, these variables are economically and statistically less significant than the role of payment instrument characteristics.
- Observed changes in the primary determinants of check use can explain substantial portions of the actual 8.4 percentage point decline in the check share of consumer payments from 2003 to 2006. A decrease in the relative convenience of checks and an increase in the relative cost of checks can explain directly about 25 percent and 14 percent, respectively, of the decline in the check share of consumer payments.
- The largest part of the decline in check use (33 percent) occurred via an increase in the number of payment instruments per consumer, which likely was influenced by payment characteristics as well, but this indirect effect cannot be identified with available data.
- A key factor driving the success of payment characteristics in explaining the decline in check use is that payment demand is far more heterogeneous within narrow demographic groups than

across them and consumers' assessments of the characteristics explain a significant portion of the within-group variation.

- Three factors appear to explain why some consumers are finally writing fewer checks. First, some consumers use fewer checks because of the increased availability and acceptance of alternative payment instruments with presumably more appealing characteristics than checks—at least more appealing to these consumers. Other consumers do not find the characteristics of the newer payment instruments more appealing and do not use fewer checks. Second, convenience, cost, recordkeeping, and timing are the payment instrument characteristics that appear to have been most important to consumers in the payments transformation. Third, some demographic attributes of consumers are important, but these are more important in affecting the use of payment instruments that substitute for checks than in affecting the direct use of checks.

Implications

At least two caveats apply to these conclusions. First, no comparable data exist on the characteristics of U.S. payment instruments before 2006, so we cannot tell for certain whether or how much these characteristics may have changed. Therefore, we cannot distinguish between changes in check use caused by changes in the payment characteristics and changes in consumers' understanding (financial literacy) and attitudes (preferences) toward the payment characteristics. This distinction is crucial to understanding and especially to influencing consumer payment choice. Second, because our econometric models are reduced-form models, we cannot link the results back to the parameters of a structural model of demand for money and payments based on preferences (utility) and technology (production of payments services).

Both caveats suggest two future directions for research. First, more and better data on consumer payment choice are needed. Towards this end, the Federal Reserve Bank of Boston is sponsoring in 2008 and 2009 new versions of the Survey of Consumer Payment Choice conducted by the RAND Corporation using its American Life Panel. Other data development efforts, in the United States or elsewhere, would be helpful too. Second, structural models of consumer demand for money and payments that incorporate realistic features of actual U.S. payment instruments are needed to better guide econometric modeling.

w-09-2

Another Hidden Cost of Incentives: The Detrimental Effect on Norm Enforcement

by Andreas Fuster and Stephan Meier

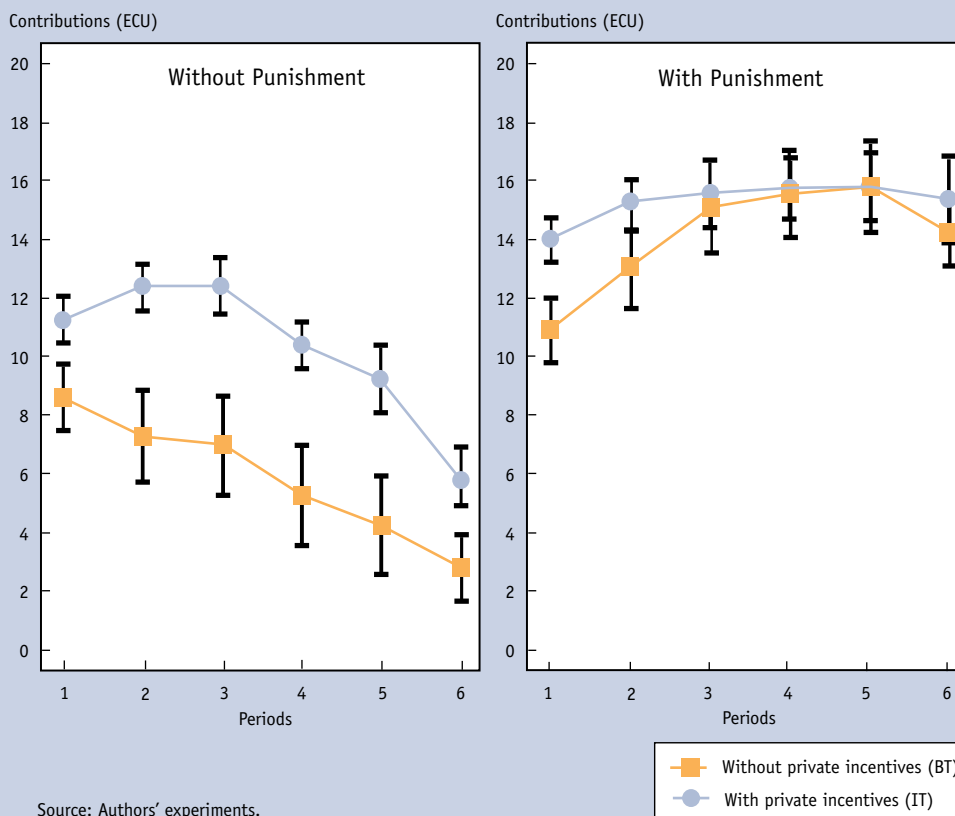
complete text: <http://www.bos.frb.org/economic/wp/wp2009/wp0902.htm>
e-mail: afuster@fas.harvard.edu, sm3087@columbia.edu

Motivation for the Research

Prosocial behavior, such as making private contributions to public goods, is crucial for the proper functioning of society and organizations. Various activities performed by humans, from hunting to holding a potluck party, require discretionary contributions by group members to be successful. Many common pool resources become exhausted if individuals do not refrain from consuming the privately optimal but socially suboptimal amount. Similarly, because it is difficult to fully control behavior through contracts, the success of organizations depends on members' willingness to take unselfish, efficiency-enhancing actions. However, the private incentives to free ride in such situations often make it difficult to sustain high levels of prosocial behavior—the well-known “tragedy of the commons.”

Recent research shows that peer punishment or norm enforcement, that is, people's willingness to incur a personal cost to punish a “free rider” (someone who contributes less than average), can be

Effect of Incentives on Mean Contributions to Public Good



effective in maintaining high levels of cooperation. The ability of a peer punishment mechanism to sustain or increase public good contributions depends largely on two factors: (1) how harshly free riders are punished, and (2) afterwards, how those free riders adapt their contributions in response to having been sanctioned by their peers.

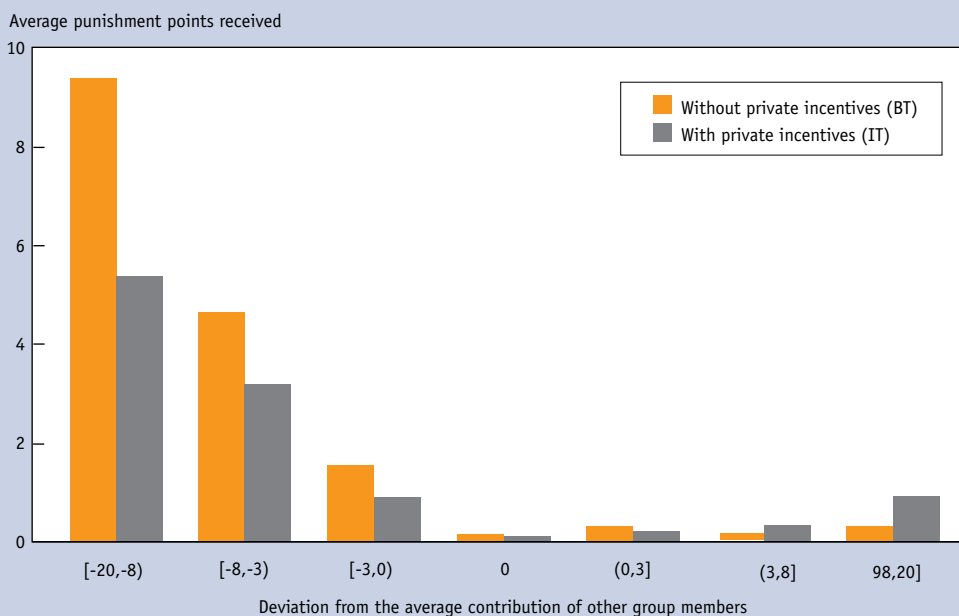
Previous research has also shown that monetary incentives, which are often used by policymakers and managers to foster prosocial behavior, may substantially increase contributions in some settings. This paper employs a laboratory experiment to examine how (centrally provided) private monetary incentives for contributing to a public good interact with (decentralized) peer-inflicted punishment, in particular with respect to the two factors mentioned above, and how this in turn affects contributions.

Research Approach

The authors had experimental subjects play a linear public good game with four treatment conditions. In the baseline treatment (BT), subjects participated in two six-period public good games, one with and one without the possibility of punishment. In the incentive treatment (IT), a private monetary inducement to contribute to the public good was added to the BT. Sixty students, the vast majority of them undergraduate students, participated in the BT and 76 in the IT. The experiments were conducted at the Computer Lab for Experimental Research at Harvard University, using the software z-Tree, developed at the University of Zurich for conducting economic experiments.

In the BT, participants first play six periods of a public good game in fixed groups of four. In each period every group member receives an endowment of 20 experimental currency units (ECU) and can

Received Punishment Points for Deviations from Others' Average Contribution



Source: Authors' experiments.

Note: Numbers above bars indicate the relative frequency of observations in a category.

contribute an integer (whole number, not fractional) amount of these units to a public good. All group members decide simultaneously on their contribution amount in a given period. Each individual's monetary payoff from the group project in a particular period is computed so that an individual's private cost of contributing 1 ECU to the public good is 0.6 ECU, while the total benefit to the fellow group members is 1.2 ECUs. This means that not contributing at all is the dominant action for each group member, while the total group payoff is maximized if all group members contribute their full endowment. After playing six periods without punishment, participants are informed that there will be a second game and are re-matched into new groups where they play for another six periods with the same parameter values as in the first six periods but with a peer punishment mechanism added. After participants make their contribution decision, they are informed about the contribution of each of the other group members and are allowed to assign an integer value of between 0 and 10 punishment points to the other group members. The punishment points (neutrally labeled "deduction points" in the instructions) are costly to both the sender and the receiver. Each punishment point costs the sender 1 ECU and the receiver 3 ECUs. However, the payoff-effective punishment costs imposed by the other group members on each subject cannot exceed the first-stage payoff.

The IT is identical to the BT except that the subjects are now given a private monetary incentive to contribute to the public good. To make the private incentive salient, participants receive a virtual "lottery ticket" for each ECU that they contribute to the public good. Each lottery ticket gives a 1 percent chance of winning an additional 20 ECUs at the end of the experiment, so it has an expected value of 0.2 ECU. For example, if a participant contributes 10 ECUs to the public good, he or she will on average win 2 ECUs. This incentive is in place for both parts of the treatment, with and without punishment. The monetary incentive is nontrivial, as, in expectation, it is equivalent to a reduction from 0.6 ECU to 0.4 ECU in the private cost of contributing an ECU to the public good. However, this incentive does not alter the benefits that the other members of the individual's group derive from the individual's contribution, and from a purely monetary perspective, it is still a dominant strategy for each subject to decline to contribute anything to the public good.

If private incentives do not interact with either the harshness of punishment or the response of free riders to being sanctioned for failing to contribute appropriately, we would expect the IT to yield higher contributions than the BT, with and without the punishment mechanism. However, the presence of private incentives may influence punishment behavior and the reaction of free riders—and as a result these interactions could yield higher or lower overall contribution levels than would be expected from the price effect alone. If high contributors interpret the low contributions of free riders as “more selfish” when additional incentives are in place, one may imagine that they would punish them more harshly in the IT than in the BT, a result that should lead to higher contributions. The competing hypothesis is that the extra incentive for prosocially inclined individuals mitigates their anger, and reduces their willingness to punish free riders, as they receive something that free riders do not. In this case, the effectiveness of the norm enforcement mechanism may be reduced as compared with the BT, and the private incentives may not lead to higher contributions.

Key Findings

- Without the punishment mechanism, contributions are significantly higher with incentives in place than without incentives, while with the punishment mechanism, there is no significant difference between the two treatments. This supports the hypothesis that incentives and norm enforcement are substitutes, meaning that one or the other in isolation is successful in raising contributions, while adding incentives in a setting with a peer punishment mechanism does not lead to higher contributions.
- Free riders (group members who contribute less than average) are punished significantly less severely when incentives are in place than when these are not. This is consistent with the hypothesis that the rewards received by the high contributors dampen their anger towards the free riders, resulting in less punishment being meted out.
- For a given level of punishment, free riders increase their subsequent contributions by less when incentives are present than when they are not. This result is consistent with the idea that the presence of contribution incentives mitigates the shame or guilt free riders may feel.

Implications

These findings indicate that policymakers and managers should be careful in using private incentives to foster prosocial behavior in settings where norm enforcement and social pressure are important. Incentives may affect prosocial behavior not only directly but also indirectly through their negative influence on the effectiveness of norm enforcement.

That said, it is not clear from these findings that the presence of incentives does not increase social welfare, even though contributions are not increased: achieving a similar level of prosocial behavior while incurring less (socially wasteful) norm enforcement is welfare-enhancing. On the other hand, while norm enforcement is costly in the short run, in the long run it is mainly the threat of punishment that maintains high levels of contributions, while providing incentives remains costly. Therefore, whether it is desirable for a policymaker or manager to introduce private incentives for prosocial behavior will depend on: (1) the weights the policymaker assigns to obtaining high contributions to the public good versus the costs of enforcing norms, (2) the cost of providing incentives, and (3) the time horizon over which the policy is considered.

Has Overweight Become the New Normal? Evidence of a General Shift in Body Weight Norms

by Mary A. Burke, Frank Heiland, and Carl Nadler

complete text: <http://www.bos.frb.org/economic/wp/wp2009/wp0903.htm>
e-mail: mary.burke@bos.frb.org, fheiland@fsu.edu, carl.nadler@bos.frb.org

Motivation for the Research

A large multidisciplinary literature has investigated self-perceptions of body size and body shape. Most of the existing studies have focused on the cross-sectional factors influencing the self-perception of body weight appropriateness, such as socioeconomic status, race/ethnicity, gender, and age. Consequently, little is known about the evolution over time of body weight perceptions, either within individual lifetimes or across generations.

A framework developed by Burke and Heiland (2007) in a theoretical paper posits that the social norm for body weight—defined as a reference point against which individuals judge the appropriateness of their own weight—varies directly with mean body weight in the population at large. This framework predicts that the social weight norm in the United States would have increased over the past 20–30 years in response to the increase in mean body weight observed over the same period. This paper tests that prediction by examining differences in self-perception of weight status between two survey cohorts observed 10 years apart on average.

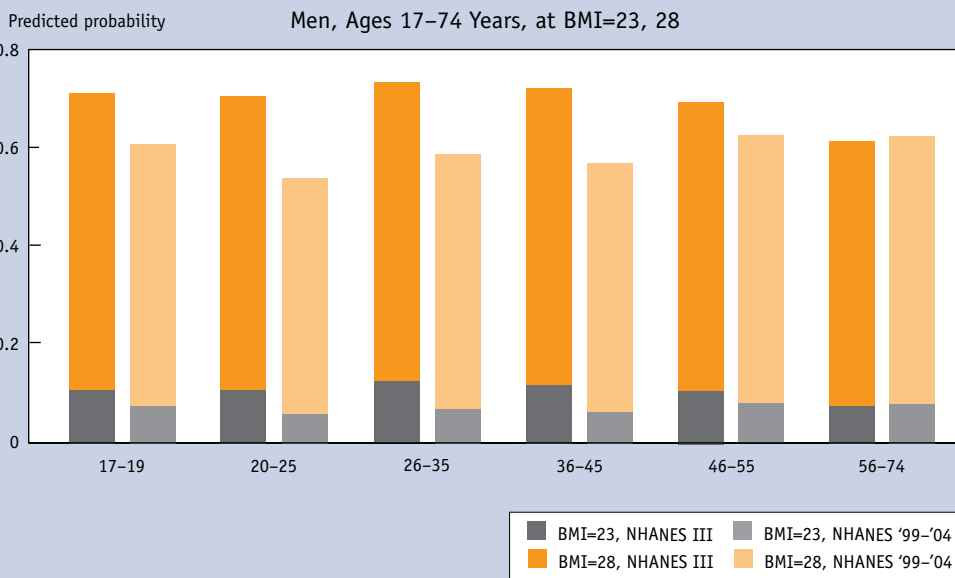
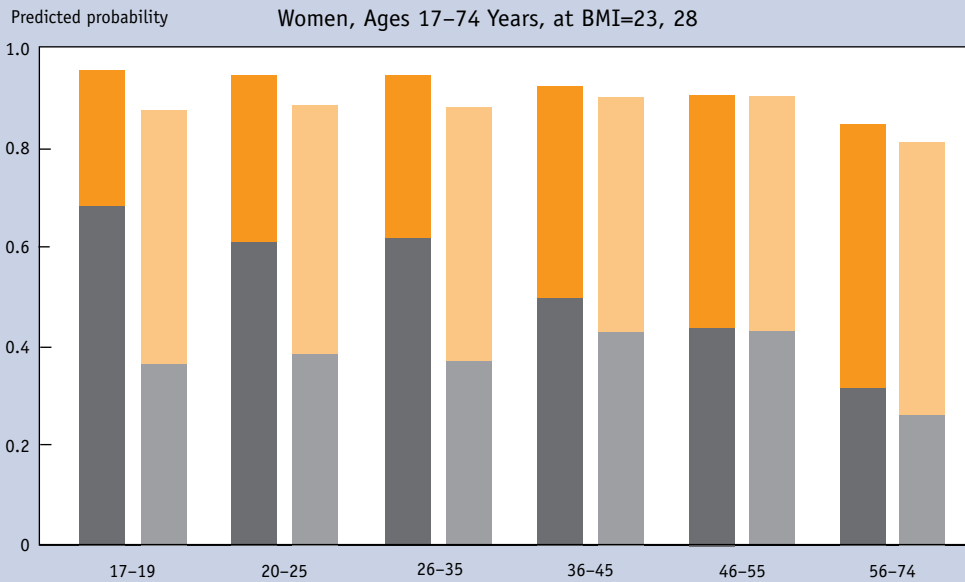
Research Approach

The authors examine data from the 1988–1994 and 1990–2004 National Health and Nutrition Examination Surveys (NHANES III and NHANES 1990–2004, respectively). These surveys provide data on the self-classification of body weight for two large, nationally representative samples of individuals, as well as extensive anthropometric and demographic information. In each survey, people ages 17 to 74 years were asked whether they consider themselves (at their current weight) to be “underweight,” “about right,” or “overweight.” The authors use descriptive analysis, correlation, and multivariate regression analysis to compare responses across different survey cohorts at the same age, separately by sex, controlling for various demographic and anthropometric factors.

Key Findings

- The probability of self-classifying as overweight is significantly lower, and—within some age ranges—the probability of self-classifying as underweight significantly higher, in the more recent survey, controlling for objective weight status and numerous socio-demographic variables, including race/ethnicity and age. The shift in overweight self-classification appears most pronounced between the earlier and later cohorts of young adults (ages 17–35 years), while the shift in underweight self-classification occurred exclusively across cohorts of adults ages 36–55 years.
- Members of groups (defined by age range, sex, and birth cohort) with higher average BMI are less likely to classify themselves as overweight, and more likely to classify themselves as underweight, than are members of groups with lower average BMI, controlling for own BMI and other factors. This relationship helps to account for the between-survey decline in the tendency to feel overweight.
- Changes in body fat percentage and/or waist circumference do not account for the changes in weight classification between the surveys, as both of these factors increased on average between the surveys, conditional on BMI, and would have increased the tendency to feel overweight, all else being equal. In the uncontrolled analysis, misperception relative to Center for Disease Control (CDC) standards declines among women, but no net change in misperception occurs among men.

Predicted Probability of Feeling Overweight, by Sex and Age



Source: NHANES III and NHANES 1994–2004.

- Individuals exposed to a higher childhood obesity rate (as children), as is the case, on average, in the later survey cohort, have a lower relative risk of self-classifying as overweight as adults, all else being equal.

Implications

These findings are consistent with theories of endogenous weight norms, and specifically with the argument of Burke and Heiland (2007) that the mean BMI in the population determines the social norm for body weight and that individuals assess their own weight in relation to the social norm. However, the authors do not claim to identify causality between the rightward shift of the weight distribution and the increase in the subjective threshold for overweight.

The authors acknowledge that a variety of social forces, such as imagery in the popular media, public health messages, the recalibration of women's dress sizes ("vanity sizing"), and scholastic content, may have changed between the survey periods, influencing self-perceptions of body weight and contributing to a more relaxed standard of overweight. However, the official CDC standard for overweight actually became stricter between the survey periods and the ideal female body in media imagery became progressively thinner between 1950 and 1980 before leveling off. With respect to males, previous research indicates that the ideal male figure became more muscular between 1973 and 1997, but it is difficult to determine whether trends in media portrayals of male physique played any role in the trends identified in this research.

In the past decade, numerous government programs have emerged that promote the development of a "healthy body image" and healthy eating behavior, particularly for young women. "Healthy body image" refers to positive self-esteem concerning physical size and shape. In the United States, such campaigns have sought to offset a cultural emphasis on thinness, celebrate diverse body types, and promote physical and mental health regardless of size. Programs of this type may have contributed to the finding that younger, normal-weight females experienced some of the greatest reductions of any demographic group in the probability of self-classifying as overweight. At the same time, in recent years several states have begun to require that children and adolescents in public schools be weighed and measured. Parents and children are informed of the child's official weight status, and overweight and obese children are advised to adopt healthier habits and target a healthier weight. Therefore, overweight students may be receiving mixed messages: on one hand, to lose weight, on the other, to have a healthy self-image. The net impact of such messages on self-assessments among children and adolescents remains unclear.

The welfare implications of a change in the weight standard are ambiguous. These results indicate that women's notion of overweight in the later survey shows stronger agreement with CDC standards than before, as the share of normal-weight women who consider themselves overweight declined significantly. However, a larger share of overweight and obese men and women now classify themselves as "about right," an indication that such individuals feel they do not need to lose weight. While departing from official classifications, such judgments may not have a negative net impact on health, especially for those individuals who are merely overweight and not obese. Recent research has shown that overweight (but not obese) individuals enjoy lower mortality rates than people in any of the other weight classes. Fat-acceptance advocates have pointed out that negative self-image results in significant mental health costs and that dieting is costly in numerous ways. More fundamentally, recent research has highlighted the limitations of using a one-dimensional index to define appropriate body size, resulting in increased calls for the use of alternative measures such as waist circumference and body fat percentage. Regardless of where one stands in these debates, there is likely to be broad agreement that public policy has, to date, focused too narrowly on promoting "healthy weight" rather than on promoting a more comprehensive definition of health that includes good nutrition, cardiovascular fitness, and sound mental health.

Empirical Estimates of Changing Inflation Dynamics

by Jeffrey C. Fuhrer, Giovanni P. Olivei, and Geoffrey M. B. Tootell

complete text: <http://www.bos.frb.org/economic/wp/wp2009/wp0904.htm>

e-mail: jeff.fuhrer@bos.frb.org, giovanni.olivei@bos.frb.org, geoff.tootell@bos.frb.org

Motivation for the Research

The Phillips curve, a controversial but widely used model of inflation, suggests that inflation is determined by expected inflation, resource utilization, and key relative price shocks, notably the relative price of energy. The backward-looking version of the Phillips curve, which incorporates sticky prices and inflation persistence, is widely used in policy circles.

Yet the literature on out-of-sample inflation forecasting casts doubt on the usefulness of backward-looking Phillips curves for forecasting U.S. inflation in the post-1984 period. While several studies have reached more nuanced conclusions, the essence of Atkeson and Ohanian's (2001) results about the inability of Phillips curve specifications to improve upon a simple univariate benchmark, whereby the forecast of inflation in the next four quarters is the value of four-quarter inflation today, has proved difficult to overturn. In the most comprehensive study to date, Stock and Watson (2008) corroborate Atkeson and Ohanian's findings, with the important qualification that when the unemployment rate is sufficiently different from the non-accelerating inflation rate of unemployment (NAIRU), Phillips curve specifications improve substantially upon a univariate benchmark.

This paper provides a battery of empirical tests to examine the stability of various aspects of the Phillips curve, including: (1) its slope (the coefficient(s) on the real activity measure), (2) the influence of key relative prices that shift the curve, (3) the influence of past inflation dynamics on the formation of inflation expectations, and (4) general equilibrium influences that arise from the conduct of monetary policy and the IS curve linking real activity to real interest rates. The paper uses both single-equation, reduced-form versions of the Phillips curve and multi-equation, constrained rational expectations models that incorporate a Phillips curve.

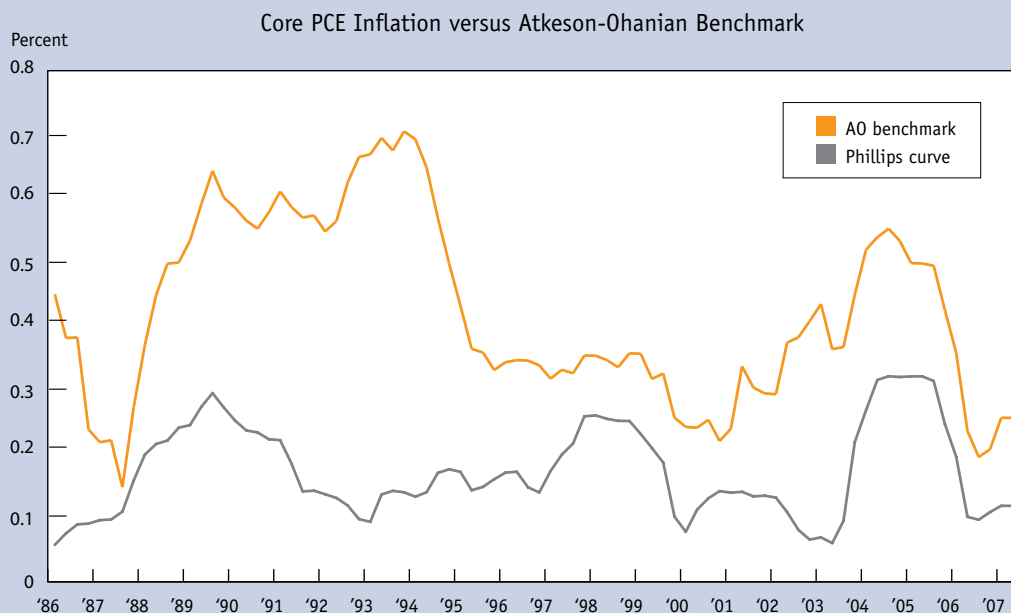
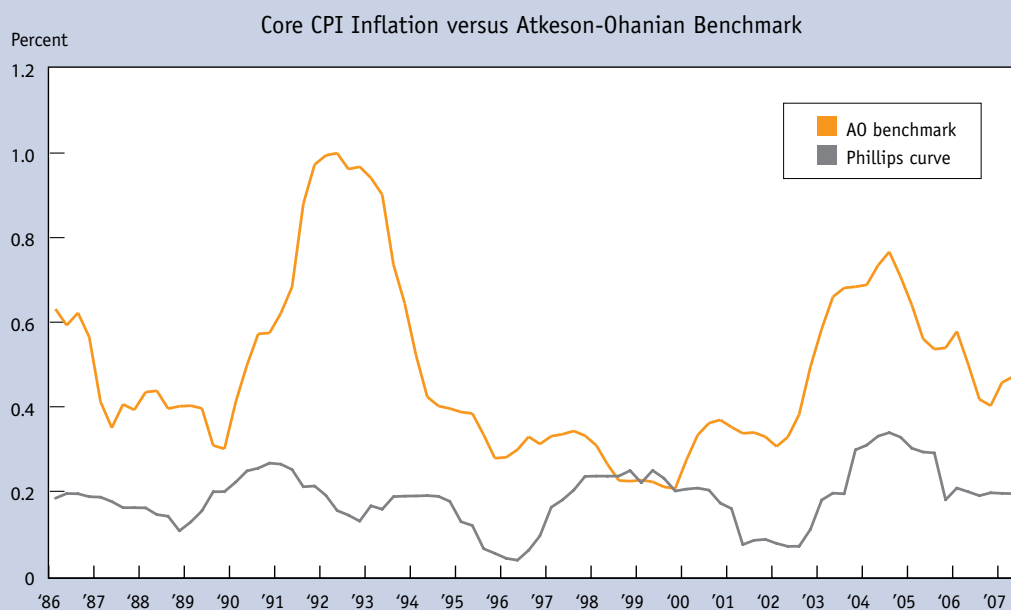
Research Approach

The authors discuss the Phillips curve forecasting performance, especially relative to the Atkeson-Ohanian benchmark, and examine the inflation-unemployment tradeoff implicit in recent Federal Reserve Greenbook forecasts. They then analyze reduced-form and structural estimates of shifts in key Phillips curve parameters. Because this analysis finds a clear shift in the influence the relative price of oil has on inflation expectations, they explore in more detail potential changes in the pass-through of key relative prices into core and headline inflation, and into wages.

To illustrate some important issues concerning the usefulness of traditional backward-looking Phillips curves in forecasting inflation, the authors first compare the in-sample versus the out-of-sample performance of the Phillips curve. The backward-looking Phillips curve they use is standard in the literature. Current inflation is a function of past inflation, the unemployment rate, and other control variables (supply shocks) that act as shifters in the inflation-unemployment rate relationship. The authors compare the four-quarter-ahead forecast of inflation in this standard specification with the Atkeson-Ohanian benchmark, whereby expected inflation over the next four quarters is equal to inflation over the previous four quarters, plus noise. The Atkeson-Ohanian benchmark can be thought of as an effective shortcut to a non-naïve, univariate representation of the inflation process.

Two measures of prices are considered, the core consumer price index and the core personal consumption expenditures index. Core measures of consumer prices are of particular interest to policymakers.

Core Inflation: Root Mean Square Error of Phillips Curve versus Benchmark



Source: Authors' calculations.

Note: The figures compare the root mean squared errors (RMSE) for inflation forecasts over a rolling window of eight quarters.

Also, by mitigating the role of energy and food price shocks, a Phillips curve relationship specified on core inflation measures should provide a better assessment of the role of economic activity in generating common price movements. The civilian unemployment rate for workers 16 years of age and older is specified either in levels or as a deviation from a time-varying NAIRU. As shock supply controls, the authors include lags of the change in the relative price of oil and lags of the change in the relative price of non-oil imports. The metric used to evaluate each forecast is the root mean squared forecast error (RMSE). The sample period is 1984 to 2007. Specifications with either a time-varying NAIRU

or a constant NAIRU, and with or without supply shocks are considered. The authors also distinguish between in-sample forecasts and (pseudo) out-of-sample forecasts. The out-of-sample forecasts simulate a real-time forecasting exercise.

Key Findings

- There have indeed been shifts in Phillips curve parameters. This conclusion is validated in all forms of the Phillips curve examined.
- The shifts are most pronounced for the effect of the relative price of oil on core inflation.
- The effect of the real activity variable may have diminished in recent decades. However, it is not zero, and the authors provide evidence suggesting that the Federal Reserve's Greenbook forecast may have underweighted the contribution of the unemployment rate in recent years.
- The shifts appear to be concentrated in the early 1980s. The Phillips curve has been relatively stable over the past two decades.
- Accounting for a change in parameters in the 1980s dramatically improves the out-of-sample forecasting performance of the Phillips curve. In fact, incorporating this shift in parameters overturns the results of Atkeson and Ohanian (2001).
- In a structural model of the New Keynesian variety, the source of the shifts in the more reduced-form Phillips curves is a combination of a smaller effect of the relative price of oil in the Phillips curve, a smaller Phillips slope coefficient, and a downward shift in the interest sensitivity of the IS curve, all of which occurred in the early 1980s. In more recent years, the share of backward-looking or indexing agents in the Phillips curve appears to have declined, although identifying that effect is difficult.

Implications

The authors document significant changes in the dynamics of U.S. inflation as described by an array of Phillips curve specifications. These specifications cover the conventional backward-looking Phillips curve and a hybrid New-Keynesian Phillips curve that is embedded in a simple general equilibrium representation of the macroeconomy. The different specifications yield similar results concerning the nature of the changes in inflation dynamics. They point to the impact of energy and food prices on core inflation measures having changed over time, with a much diminished pass-through of food and energy prices into core inflation. They also produce evidence of changes in the effect of past inflation on current inflation, and of potential changes in the number of unemployment point-years required to reduce the inflation rate by one percentage point, known as the sacrifice ratio.

Taking into account changing inflation dynamics produces more accurate out-of-sample forecasts. Indeed, the authors show that the Atkeson-Ohanian results concerning the inability of Phillips curves to produce better forecasts in the post-1984 sample than a naïve, random-walk-type forecast are overturned once time-varying coefficients are incorporated into an otherwise standard backward-looking Phillips curve. Provided changing inflation dynamics are taken into account, the Phillips curve framework continues to have, with a few notable exceptions, relevant economic content for explaining the dynamics of inflation and for inflation forecasting.

The estimated extent of changing inflation dynamics can differ according to the Phillips curve specification used—an issue that is even more pressing when the Phillips curve is estimated in the context of a general equilibrium framework. The uncertainty introduced by this variation complicates the forecasting exercise noticeably in real time. Here, the Federal Reserve Board's Greenbook inflation forecasts are instructive. Recent Greenbook inflation forecasts are predicated on a very high sacrifice

ratio. Some Phillips curve specifications do indeed point to an increase of the sacrifice ratio in recent years, but the authors show that the Greenbook inflation forecasts seem to have relied on a Phillips curve slope that is potentially too flat—that is, a Phillips curve that implies a severe tradeoff that requires many point-years of unemployment to lower inflation modestly.

Overall, while the evidence of changes in inflation dynamics is overwhelming, how to best model such changes is still an open issue. Traditional backward-looking Phillips curves provide flexibility in this regard, but their quasi-reduced form makes it difficult to attribute changes in inflation dynamics to particular structural features of the economy. Versions of the Phillips curve that are more structural can provide more economic content to the changes, especially when embedded into a general equilibrium framework. In this case, however, the risk is that the instability could be the result of a misspecified model.

The task of how to model changing inflation dynamics is somewhat simplified by the fact that the most notable changes seem to have occurred in the early 1980s. But the stability of the Phillips curve is still relative, with potentially important changes concerning inflation expectations formation and the sacrifice ratio having occurred in more recent years. At this point, we still have too few observations to be confident of these changes. Needless to say, the current recession should provide fertile ground for continued assessment of the economic content and forecasting power of the Phillips curve, and for better understanding more recent changes in inflation dynamics.

w-09-5

Geographic Variations in a Model of Physician Treatment Choice with Social Interactions

by Mary A. Burke, Gary M. Fournier, and Kislaya Prasad

complete text: <http://www.bos.frb.org/economic/wp/wp2009/wp0905.htm>

e-mail: mary.burke@bos.frb.org, gfourmier@fsu.edu, kprasad@umd.edu

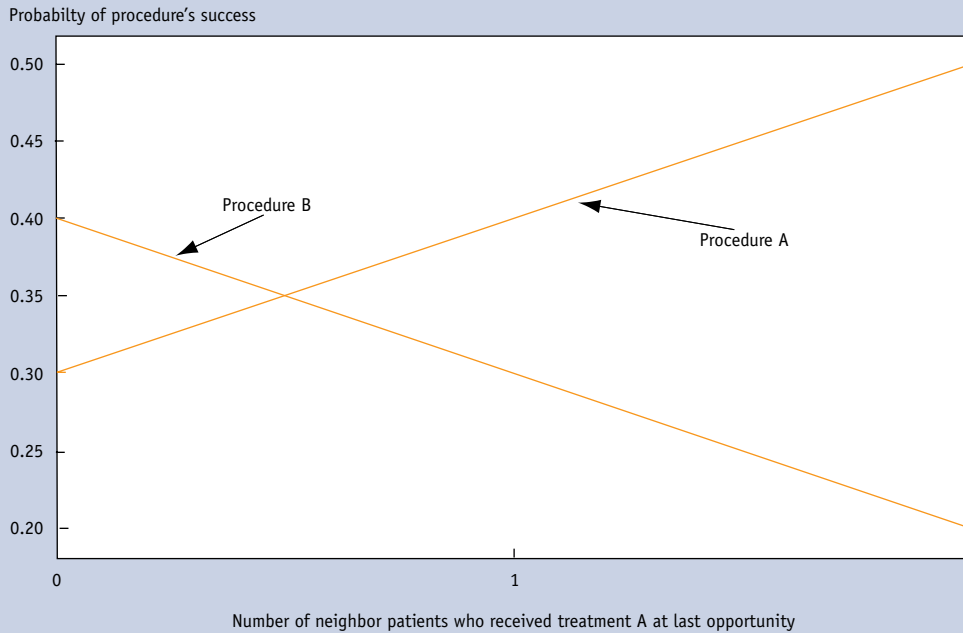
Motivation for the Research

Over the years, a number of studies have found that medical choices tend to be relatively uniform within geographic regions in the United States and yet quite diverse across regions. Surprisingly, the phenomenon persists today despite vast improvements in communications technology. For example, in a comparison of hospital referral regions across the country, rates of coronary artery bypass grafting among Medicare enrollees varied by a factor of more than 3.5, while the rates of coronary angioplasty ranged from 2.5 to 16.9 per 1,000 enrollees. Such high variation has been recorded for many procedures and treatments, spanning all areas of medicine and often persisting over time. The variations have proved remarkably robust to controls for incidence of illness and demographic and socioeconomic factors. This paper develops a general theory of the emergence of region-specific norms of behavior and uses it to suggest a possible resolution of a famous public health puzzle—the well-documented effect of geographic location on the choice of medical procedures in the United States.

Research Approach

The authors construct a theoretical model in which physicians' treatment choices depend on patients' characteristics and on the recent choices of nearby peer physicians—either because there are local knowledge spillovers or because physicians want to conform to local practice patterns. In the model, which is adapted from the mathematical theory of interacting particle systems, regional differences in the patient mix give rise to geographically divergent treatment patterns: the treatment a patient receives depends on where s/he lives.

Physician Payoffs for Procedures A and B for an Alpha Patient



Source: Authors' calculations.

Note: An alpha patient is the type for whom the preferred treatment is Procedure A.

In the model, two key features drive the emergence of regional treatment norms: local choice interactions among physicians and regional variation in patient characteristics. In the former, the authors assume that the choices of a physician's nearby colleagues exert an influence on her own choices, either because of local increasing returns or because of pure conformity effects. As an example of local increasing returns, a surgeon's own expertise in a given treatment or procedure may improve as her peers gain experience in administering the same treatment and share their insights. If so, a given treatment will yield better outcomes and so become increasingly favored within a group of interacting peers the more frequently it has been used in the past. Alternatively, choice spillovers among neighboring physicians could reflect pure conformity effects. Conformity of behavior within a group may arise because individuals have an innate preference for social approval or because of incentives in the institutional environment. For example, in the United States, malpractice claims are judged by comparing a doctor's actions to standard practices within the local community, thereby discouraging deviation from such practices. Either form of social influence contributes to the emergence of geographical variations in medical treatment, but the welfare implications are quite different in these two cases.

The model's second key feature is that patient characteristics differ on average across geographic regions, and these characteristics influence the treatment choice. Specifically, the authors assume that the "ideal" treatment for a given patient—the treatment that maximizes the probability of a successful outcome regardless of the local treatment history—depends on the patient's underlying traits, such as age, in addition to (unmodeled) transitory symptoms. This assumption accords with the observation, for example, that heart bypass surgery carries greater risks among the elderly than among the young and hence is less common among the former group. To see how this assumption contributes to the emergence of regional treatment variations, consider what happens when a young patient arrives for treatment in a region with a high proportion of elderly patients. The young patient's physician will take the patient's age into account, but will also be drawn toward the treatment she observes her neighbors performing most often, which will likely be the treatment that is ideal for elderly

patients. In the long run, the dynamic feedback between physicians' choices and patients' characteristics produces a stable treatment pattern in which the regional treatment norm represents the ideal treatment for the modal patient in the region. In this long-run equilibrium, patients of different types in the same region will receive the same treatment, and a patient of a given type will receive a different treatment depending on where s/he lives because the modal patient varies across regions.

The authors test the theoretical model's predictions using a census of Florida patient discharge records, focusing on patients over the age of 25 years with a primary diagnosis of either coronary atherosclerosis or acute myocardial infarction (AMI or "heart attack"). The usable sample covers over 500,000 inpatient stays during the 1995–2001 period.

Heart disease is an appropriate choice for this empirical analysis because its treatment mirrors a number of features of the theoretical model. First, the treatment options can be divided into two discrete categories: surgery and "medical management." The surgical options include coronary angioplasty and heart bypass surgery; medical management involves the use of drugs, such as beta blockers, calcium channel blockers, and ACE inhibitors, as well as other noninvasive therapies, such as diet and exercise modification. Second, consistent with the payoff structure in the model, there is evidence that the invasive treatments are, on average, less suitable for "older" patients (those 73 years of age and above) than for "younger" patients (those 62 years of age and younger, but above age 25). Third, the data agree with the presumption of age-dependence, in that the raw correlation between patient age and the probability of receiving surgical treatment for heart disease is negative and significant. Finally, the age distribution of patients admitted to hospitals for coronary care varies considerably across hospitals and regions in Florida, and older patients are well represented.

Key Findings

- Younger patients are less likely to be treated by invasive heart treatment—either coronary angioplasty or heart bypass surgery—for the same condition, the greater the share of older patients at the hospital during the same period, and conversely older patients are more likely to be treated invasively for the same condition, the greater the share of younger patients at the hospital during the same period. A one standard deviation increase in the proportion of younger heart patients would raise the expected number of angioplasties given to older patients by about 18 (from 330 per 1,000 to 348 per 1,000) and add six surgeries for older patients. Similarly, at the mean number of angioplasties for younger patients (450 per 1,000), a one standard deviation increase in the proportion of older heart patients at the hospital would reduce the expected number of angioplasties performed on younger patients at the hospital by about 17 (from 450 per 1,000 to 433 per 1,000) and decrease by 21 the expected number of surgeries performed on younger patients.
- The angioplasty rate (total surgery rate) increases for either younger or older patients with the hospital's total volume of angioplasties (volume of surgeries) during the quarter. This result most likely reflects specialization; for example, it has been observed that hospitals that perform relatively few invasive procedures have lower success rates for them. To the extent that this volume variable captures (time-varying) specialization at the hospital level, including this control helps to ensure that the effects of the patient-age composition are not merely proxying for the effects of specialization.
- The racial composition of the patient base also affects treatment systemically: surgery (or angioplasty alone) is less likely to be selected, for either younger or older patients, the greater the proportion of black patients treated at the hospital. Similar effects are observed for the share of Hispanic patients, but the significance is weaker on average.
- Hospitals with higher patient income (as of 1999) have a higher angioplasty rate for older patients, but income effects are otherwise only weakly significant. The share of insured patients exerts no significant effect on treatment tendencies.

- To the extent that the data provide information about patient outcomes, the authors find little evidence that patient welfare is adversely affected by the presence of regional treatment norms. While the welfare findings are not conclusive, they suggest that knowledge spillovers, rather than conformity effects, are the primary source of treatment variations in the data.

Implications

Based on the assumption that regional treatment variations across the United States imply net welfare losses, policies have been proposed that would seek more uniform compliance with guidelines for medical practice that are based solely on individual patient characteristics and symptoms. However, the authors' model implies that the welfare implications of treatment variations depend on the motivations underlying the social influences on treatment choice. In the presence of increasing returns, treatment variations have benefits, in the form of gains from specialization that accrue to the dominant patient type in a region, as well as costs that fall on the minority-type patients in a region. If the benefits outweigh the costs, treatment variations will dominate a system of strict treatment guidelines. On the other hand, if pressure to conform drives the choice spillovers, patients would be better off if physicians were allowed to make treatment decisions based solely on a patient's individual characteristics.

This welfare analysis holds important policy implications. Under conformity pressure, a tendency to overutilize invasive treatments may arise in some regions (and to underutilize them in other places), contributing to inefficient care and elevated costs. On the other hand, legislating strict treatment guidelines may be harmful in the presence of knowledge spillovers, because some patients will be deprived of gains from specialization. Regional treatment variations are not necessarily welfare-reducing, but further research along these lines is clearly warranted before policy conclusions can be drawn.

The authors' model makes significant contributions to the theoretical understanding of regional treatment variations in the United States. First, it provides an explanation for how regional norms emerge rather than just why norms might persist once in place. Second, the model shows that either productivity spillovers or conformity pressure can result in the emergence of treatment variations and demonstrates an important difference in welfare implications between the two mechanisms. In addition, the model implies that regional norms are grounded in relatively stable underlying fundamentals (in this case, patient characteristics), a result that contrasts with the multiple equilibria that often arise in settings involving conformity effects and other spillovers.

For simplicity, the authors assumed that patients do not respond to the emergence of treatment norms by choosing hospitals that best match their preferences and characteristics. In practice, patients (and their health care providers) exert some choice over the hospitals where they are treated, and they may even select residential locations based in part on proximity to a particular hospital. While such sorting would mitigate the welfare losses experienced by minority-type patients in a region, there are likely to be considerable constraints on optimal hospital-patient matching and hence the results should be largely robust.

w-09-6

The Optimal Level of Deposit Insurance Coverage

by Michael Manz

complete text: <http://www.bos.frb.org/economic/wp/wp2009/wp0906.htm>

e-mail: michael.manz@snb.ch

Motivation for the Research

In order to reduce the likelihood of bank runs, many countries have adopted explicit deposit insurance in recent decades. The major theoretical case for offering deposit protection relies on models in the style of Diamond and Dybvig (1983), where deposit insurance, by eliminating runs as the only source of bank failures, enables account holders to reap the full benefits of banking at no cost. However, there is now widespread agreement that most bank failures are driven by balance sheet weaknesses of individual banks rather than by self-fulfilling panics on the part of depositors. Thus, the adoption of deposit insurance may imply heavy losses and involves a policy tradeoff. Preventing runs is a welcome result if a bank is solvent, but it is less reasonable if depositors have good reason to withdraw their funds and force an insolvent institution to close. In a nutshell, deposit protection inhibits both inefficient and efficient bank runs and may encourage banks to engage in imprudent actions.

This paper focuses on the amount of coverage offered to depositors, which is a key feature governing the tradeoffs surrounding deposit insurance and is the centerpiece of any discussion of reforms. While partial deposit insurance (that is, any insurance that imposes any limit to the coverage) may fail to prevent runs, as the experience of U.K. bank Northern Rock in September 2007 well demonstrated, it may strengthen market discipline. This consideration gives rise to the vital question of whether there is an optimal degree of protection. In most countries we observe partial, rather than full, deposit insurance, although governments tend to extend protection in the wake of financial turmoil. Whether extending such protection is an appropriate measure remains open to debate.

While the basic arguments against or in favor of full deposit insurance seem to be reasonably well understood, existing theories offer little, if any, guidance on the optimality of offering partial insurance. A major difficulty resides in the fact that most models of bank runs, following Diamond and Dybvig (1983), exhibit multiple equilibria. This makes it hard to assess the difference between coverage of 1 percent and 99 percent of deposits. In either case, anything may happen, depending on which equilibrium is selected, which, in turn, is not explained by a model with multiple equilibria.

Although various studies investigate the impact of deposit insurance coverage empirically, very few theoretical papers deal with the amount of deposit coverage that should be offered. In light of its pivotal role, deposit coverage deserves more scrutiny. This paper offers the first rigorous analysis of the optimal amount of partial deposit insurance in a unique equilibrium context.

Research Approach

The paper considers a stylized model of a bank that is financed by a continuum of small depositors subject to (partial) deposit insurance, by a large uninsured lender, and by the bank owner. The bank is run by a self-interested manager who can influence the riskiness of bank assets. The performance of the bank depends on the value of a stochastic fundamental variable that represents the state of the economy, on how creditors make account withdrawals, and on the level of risk assumed by the manager.

To eliminate the indeterminacy introduced by multiple equilibria, the model is set in a global game framework by assuming that players privately observe noisy signals of the underlying state of

the world. With strategic complementarities and signals that are at least of a certain precision, the refinement technique of global games, originally initiated by Carlsson and van Damme (1993), ensures that the model has a unique equilibrium. In equilibrium, the behavior and the utilities of all involved agents are a function of the level of deposit insurance.

Key Findings

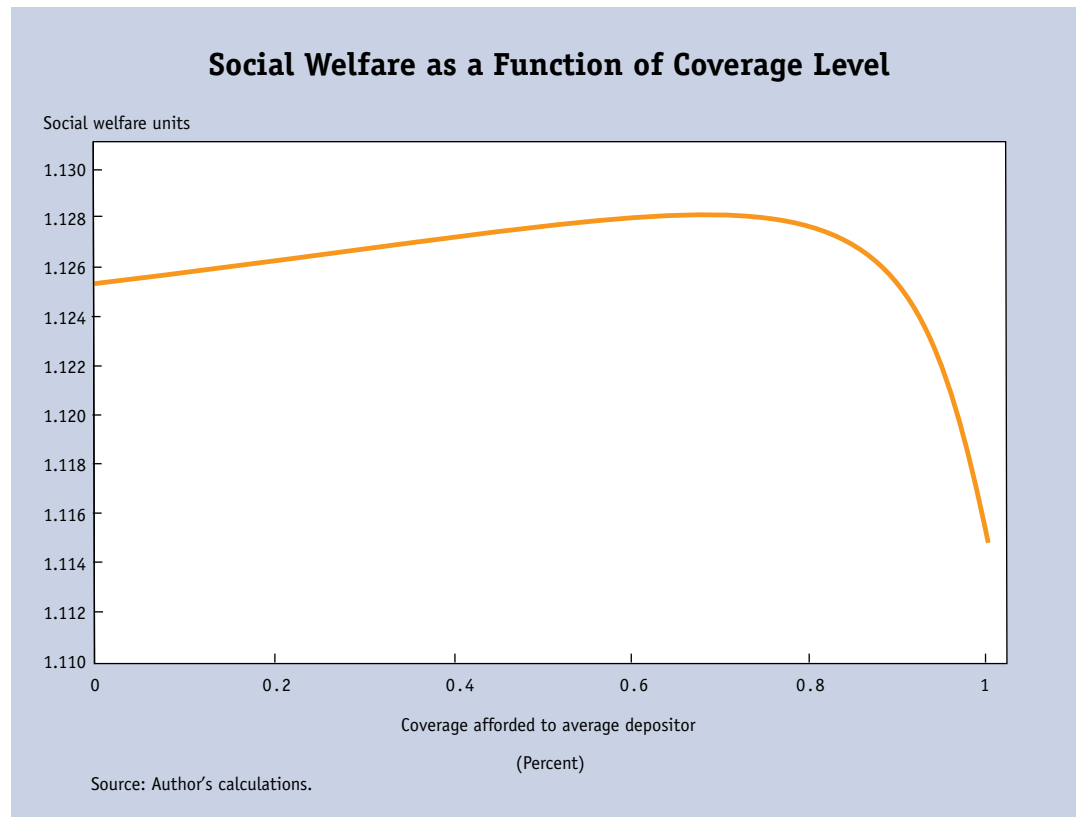
- The benefits of deposit insurance involve eliminating inefficient withdrawals and bank runs predicated upon noisy information and coordination failures, whereas the drawbacks consist of suppressing efficient withdrawals and inducing excessive risk taking. A hitherto hardly noticed conclusion is that a high level of coverage can even be detrimental if bank risk is exogenous, because it lowers the likelihood of efficient bank runs.
- A key contribution of the model is that it provides comparative statics on the optimal level of coverage. In particular, the optimal level of deposit insurance decreases with tighter liquidity requirements or with the liquidity of bank assets, increases with transparency (that is, with the quality of information available to depositors), decreases with the relevance of large creditors with uninsured claims, and increases with the weight that the deposit insurer assigns to the welfare of the uninsured creditors and the bank. This last consideration also reveals that offering a high amount of coverage is primarily in the interest of banks and uninsured lenders. Thus, the model is consistent with small banks being particularly active lobbyists in favor of extending deposit insurance. Moreover and perhaps surprisingly, the degree of deposit insurance should not vary with expectations regarding the development of the real economic sector.
- Systemic risk calls for a higher level of deposit insurance, albeit only for systemically relevant banks from which financial market contagion emanates, and not for the institutions that are potentially affected by such contagion.
- The model shows that a system of coinsurance (whereby the depositor bears a specified fraction of potential losses) outperforms a regime with caps (whereby the depositor bears the losses that exceed a given limit). Furthermore, in combination with full bailouts or optimal lending of last resort, the optimal level of deposit insurance is zero.

Implications

The results of comparative statics imply that while tightening liquidity requirements is a substitute for having deposit insurance, increasing transparency or imposing capital adequacy rules is not. In addition, deposit protection can be lowered in the presence of large investors with uninsured claims. The finding that the degree of deposit insurance should not vary with expectations concerning the real sector suggests that countries that in the past turned to increased, or even unlimited, deposit insurance as a reaction to a financial or economic crisis, such as Japan, Turkey, or the United States, would do well to pause and consider whether this is the right strategy to strengthen their banking systems.

Another key contribution of the model is its applicability to various policy issues. In practice, only a small, albeit growing, number of countries offering deposit insurance require bank customers to coinsure a portion of their deposits. Yet according to the model, optimally designed protection should build on coinsurance rather than on setting caps on insured deposits. The model further indicates that deposit insurance becomes redundant in combination with anticipated and complete bailouts or with an optimal lending of last resort policy. While an unconditional bailout policy is about as inefficient as can be, an optimal lender of last resort policy combined with perfect public disclosure comes closest to the first best outcome in terms of overall welfare. Yet such an optimal policy, which requires protection to be contingent on bank solvency, seems far more demanding and hence less realistic in practice than offering unconditional deposit insurance. If regulators or central banks cannot precisely

assess whether a bank is solvent, interventions are likely to be a mixture of the benchmark policies considered. Investigating these cases opens an interesting avenue for future research.



Special Feature

Results of the *Research Review* Survey

The editors would like to thank the many readers of *Research Review* who answered our survey. More than 1,050 readers responded, almost all via U.S. mail, for a response rate of over 14 percent. Many readers took the time to provide helpful suggestions in response to the open-ended questions.

Most respondents are happy with the publication as is, but we also received many thoughtful suggestions for improvement. While we are not able to implement every suggestion, we have been studying the responses carefully and we are taking the ideas into account as we plan our research department postings and publications.

Key Findings

- Respondents' professions fall mainly into three categories: banking/finance (28 percent), government (23 percent), and education (12 percent).
- Respondents overwhelmingly (92 percent) read the publication in print; a much smaller share (6 percent) read it about equally in print and online, while the rest (2 percent) read it mostly online.
- Of the two choices we offered for frequency of publication, 95 percent of respondents prefer to receive *Research Review* once every six months rather than once a year, with the remaining 5 percent preferring an annual frequency. However, many respondents wrote, as a suggestion for improvement,

that they would like to receive the publication more frequently. Most such requests were for quarterly publication.

- Respondents were allowed multiple responses to describe the value of the publication. Nine hundred and six responses, more than twice the number of responses for the “runner up” choice, were “saves time by helping me keep abreast of some interesting recent work in economics.” Other responses chosen by substantial numbers of respondents were: “reminds me of the work that the Boston Fed research department is doing” (424 responses), “saves time by helping me decide which papers to look at online or download” (344 responses), and “useful in my work as a publication I can share with others” (286 responses).
- Seventy-five percent of respondents find *Research Review* valuable (52 percent) or extremely valuable (23 percent), while another 23 percent find it moderately valuable.
- Respondents were allowed two responses to describe how they generally use *Research Review*. More than three quarters of respondents (77 percent) use *Research Review* as a focused summary of recent research department postings (806 responses). Forty-four percent use it for general information (464 responses), while 21 percent read it to decide which papers to access online (217 responses). Only 8 percent use *Research Review* to scan in print while offline (82 responses).
- Only 12 percent of respondents routinely check our web site for new additions, but 52 percent of respondents indicated that looking at *Research Review* has prompted them to go to the Boston Fed web site and look at our papers.
- Most respondents scan the entire issue and read selected summaries (57 percent). Another 23 percent scan the table of contents and read selected summaries, while 16 percent read it from cover to cover. The remaining 4 percent either scan selected summaries more or less at random (2 percent) or scan the table of contents only (2 percent).

Suggestions for Improvement

Ideas for improvement range widely. The following discussion touches on those that are most representative of the overall set of responses.

The largest group asked us to “stay the course,” as one respondent put it. Many suggestions bear on the subjects covered in the research, with by far the most requests in this category asking for more research and discussion focused on the current economic situation—particularly the financial crisis and its economic fallout, markets and bubbles, the policy challenges and responses involved, and similar topics. A number of respondents would like to see more focus on New England, while others requested a broader geographic focus and a stronger global perspective. There were requests for more applied economics research, more behavioral economics, more longitudinal studies, and more focus on industries. We will try to respond, subject to the limits imposed by our small staff and the demands on their time from real-time policy work, which have been unusually high of late.

A number of readers asked us to publish *Research Review* more frequently and/or to include more articles. We appreciate the appetite for more of what we produce and are striving to increase our research output within our capacity constraints.

Some readers requested links to the articles in the online version, links to the website, access to the authors, and/or online links to related data sets. In the online version of *Research Review* we already provide links to the articles and to the authors’ e-mail addresses (just below the authors’ names at the head of the articles), but this may not be apparent, and we are experimenting with using a different font that may make this more evident. With respect to links to related data sets, a few authors

provide such links in some of their articles, but this is not the usual practice. In some cases the data are proprietary, but we will explore the feasibility of providing such links in more cases.

A few readers commented on their preference for a print distribution instead of or in addition to an online distribution medium. Interestingly, these write-in responses were evenly divided in their preference for print (with or without online) versus online only. We note that the results of the survey show that the overwhelming majority of respondents read *Research Review* in print; however, we recognize that our sample may be biased in that print readers may be more likely than online readers to respond to a survey distributed via U.S. mail in a returnable, postpaid format. Although the survey was also posted online, only seven readers responded electronically, while over 1,050 readers mailed in responses. Print versus online has been an ongoing debate for us with respect to our entire portfolio of publications and postings, and while we had hoped to continue to publish *Research Review* in both media, we have regretfully decided, in the face of extraordinary budget pressures, to publish future issues online only. Like previous issues, these will be available for viewing, downloading, and printing at <http://www.bos.frb.org/economic/ResearchReview/index.htm>.

Contributing Authors

Mary A. Burke is a senior economist in the research department of the Federal Reserve Bank of Boston.

Christopher L. Foote is a senior economist and policy advisor in the research department at the Federal Reserve Bank of Boston.

Gary M. Fournier is a professor of economics at Florida State University.

Jeffrey C. Fuhrer is an executive vice president and the director of research at the Federal Reserve Bank of Boston.

Andreas Fuster is a graduate student in the economics department of Harvard University and a research associate in the research department of the Federal Reserve Bank of Boston.

Kristopher S. Gerardi is a research economist and assistant policy advisor in the research department of the Federal Reserve Bank of Atlanta and a visiting scholar at the Federal Reserve Bank of Boston.

Lorenz Goette is a professor of economics at the University of Lausanne. He was a visiting scholar in the research department of the Federal Reserve Bank of Boston when the paper summarized in this issue was written.

Fabià Gumbau-Brisa is a senior economist in the research department of the Federal Reserve Bank of Boston.

Frank Heiland is an assistant professor of economics at Florida State University.

Andreas Lehnert is chief of the household and real estate finance section in the division of research and statistics at the Board of Governors of the Federal Reserve System.

Catherine L. Mann is a professor of economics at Brandeis University and a visiting scholar in the research department of the Federal Reserve Bank of Boston.

Michael Manz is a senior economist at the Swiss National bank and a visiting scholar in the research department of the Federal Reserve Bank of Boston.

Stephan Meier is an assistant professor at the Graduate School of Business of Columbia University and a visiting scholar in the research department of the Federal Reserve Bank of Boston.

Carl Nadler is a senior research assistant in the research department of the Federal Reserve Bank of Boston.

Giovanni P. Olivei is a vice president and economist in the research department of the Federal Reserve Bank of Boston, where he oversees the macroeconomics/finance section.

Kislaya Prasad is the director of the Center for International Business Education and Research and a research professor in the decision, operations, and information department of the Robert H. Smith School of Business at the University of Maryland.

Scott Schuh is the director of the Consumer Payments Research Center and a senior economist and policy advisor in the research department of the Federal Reserve Bank of Boston.

Shane M. Sherlund is a senior economist in the household and real estate finance section in the division of research and statistics at the Board of Governors of the Federal Reserve System.

Joanna Stavins is a senior economist and policy advisor in the research department of the Federal Reserve Bank of Boston.

Geoffrey M. B. Tootell is a senior vice president and the deputy director of the research department at the Federal Reserve Bank of Boston.

Paul S. Willen is a senior economist and policy advisor in the research department of the Federal Reserve Bank of Boston and a faculty research fellow at the National Bureau of Economic Research.

This is
our last
printed
issue.

For the next issue visit us at:

<http://www.bos.frb.org/economic/ResearchReview/>

Research Department
Federal Reservebank of Boston
600 Atlantic Avenue
Boston, MA 02210

www.bos.frb.org/economic/index.htm
change service requested

PRSR STD
US POSTAGE
PAID
WILMINGTON, MA
PERMIT NO. 121

