Economic Review 2008

An annual publication of the Economic Research Department

Federal Reserve Bank of San Francisco

Table of Contents

Opinions expressed in the *Economic Review* do not necessarily reflect the views of the management of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

The Federal Reserve Bank of San Francisco's *Economic Review* is published annually by the Research Department, under the supervision of

John Judd, Executive Vice President and Director of Research;
Glenn Rudebusch, Senior Vice President and Associate Director of Research;
John C. Williams, Senior Vice President and Advisor;
Frederick Furlong, Group Vice President, Financial and Applied Microeconomic Research, and Executive Director of the Center for the Study of Innovation and Productivity;
Reuven Glick, Group Vice President, International Research;
Mary Daly, Vice President, Applied Microeconomic Research and Regional Developments, and Director of the Center for the Study of Innovation and Productivity;
John Fernald, Vice President, Macroeconomic Research;
Simon Kwan, Vice President, Financial Research;
Mark Spiegel, Vice President, International Research, and Director of the Center for Pacific Basin Studies.

The publication is edited, compiled, and produced by Anita Todd, Research Editor.

Please send editorial comments and requests for subscriptions, back copies, address changes, and reprint permission to:

Public Information Department
Federal Reserve Bank of San Francisco
PO Box 7702
San Francisco, CA 94120

phone (415) 974-2163
fax (415) 974-3341
e-mail Pubs.sf@sf.frb.org

The *Economic Review* and other publications and information from this Bank are available on our website

http://www.frbsf.org

See also *Fed in Print,* a searchable database of all research publications in the Federal Reserve System

http://www.frbsf.org/publications/fedinprint/index.html

Information and Communications Technology as a General Purpose Technology: Evidence from U.S. Industry Data*

Susanto Basu

Professor
Boston College and NBER

John G. Fernald

Vice President
Federal Reserve Bank of San Francisco

Many people point to information and communications technology (ICT) as the key for understanding the acceleration in productivity in the United States since the mid-1990s. Stories of ICT as a general purpose technology (GPT) suggest that measured total factor productivity (TFP) should rise in ICT-using sectors (reflecting either unobserved accumulation of intangible organizational capital, spillovers, or both), but with a long lag. Contemporaneously, however, investments in ICT may be associated with lower TFP as resources are diverted to reorganization and learning. We find that U.S. industry results are consistent with GPT stories: the acceleration after the mid-1990s was broad-based—located primarily in ICT-using industries rather than ICT-producing industries. Furthermore, industry TFP accelerations in the 2000s are positively correlated with (appropriately weighted) industry ICT capital growth in the 1990s. Indeed, as GPT stories would suggest, after controlling for past ICT investment, industry TFP accelerations are negatively correlated with increases in ICT usage in the 2000s.

1. Introduction

After the mid-1990s, both labor productivity and total factor productivity (TFP) accelerated in the United States. A large body of work has explored the sources and breadth of the U.S. acceleration. Much of this research focuses on the role of information and communications technology (ICT).[1]

In this paper, we undertake two tasks. First, we discuss industry-level TFP growth for data from 1987 to 2004. TFP is a relatively broad measure of productivity which, over time, largely reflects innovation and efficiency. Relative to labor productivity (output per hour worked), TFP also controls for capital deepening (increases in capital available per hour worked). Second, we use these results to show that the simple ICT explanation for the U.S. TFP acceleration is incomplete at best. In standard neoclassical growth theory, the use of ICT throughout the economy leads to capital deepening, which boosts *labor* productivity but not TFP in sectors that only *use* but do not produce ICT. TFP growth in *producing* ICT goods shows up directly in the economy's aggregate TFP growth. From the perspective of neoclassical economics, there is no reason to expect an acceleration in the pace of TFP growth outside of ICT production.

However, consistent with a growing body of literature, we find that the TFP acceleration was, in fact, broad-based—not narrowly located in ICT production. In an early study, Basu, Fernald and Shapiro (2001) found a quantitatively important acceleration outside of manufacturing. Triplett and Bosworth (2006, though the original working paper was from 2002) highlighted the finding that the TFP acceleration in the late-1990s was due, in a proximate sense, to the performance of the service sector.

Since these early studies, there have been several rounds of major data revisions by the Bureau of Economic Anal-

*We thank Robert Inklaar and conference participants for helpful comments, and David Thipphavong for excellent research assistance. This article was prepared for the conference on "The Determinants of Productivity Growth," held September 2006 in Vienna, Austria. It is reprinted (with minor edits and corrections) from *German Economic Review* 8(2), Special Issue on "Productivity Growth," pp. 146–173 (). This article draws heavily on and updates work reported in Basu, Fernald, Oulton, and Srinivasan (2003). Opinions expressed do not necessarily reflect the views of the management of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System.

1. Jorgenson (2001) and Oliner and Sichel (2000) provide early discussions of the role of information technology in the productivity acceleration. We discuss the literature in greater detail later. Since 2004, productivity growth has slowed relative to the preceding decade. In this paper, we do not take up the question of whether this slowdown will persist.

ysis (BEA) that changed the details of the size and timing of the measured acceleration in different sectors but did not affect the overall picture. Oliner and Sichel (2006) use aggregate data (plus data on the relative prices of various high-tech goods) and estimate that, in the 2000–2005 period, the acceleration in TFP is completely explained by non-ICT-producing sectors. Jorgenson, Ho, and Stiroh (2006) undertake a similar exercise and reach a similar conclusion. Indeed, both papers find that TFP growth in ICT production slowed from its rapid pace of the late 1990s. Using industry-level data, Corrado et al. (2006), and Bosworth and Triplett (2007) find that non-ICT-producing sectors saw a sizeable acceleration in TFP in the 2000s, whereas TFP growth slowed in ICT-producing sectors in the 2000s. In the data for the current paper, sectors such as ICT production, finance and insurance, and wholesale and retail trade accelerated after the mid-1990s; TFP growth in those sectors remained relatively strong in the 2000s, even as other sectors finally saw an acceleration.

The broad-based acceleration raises a puzzle. According to standard neoclassical production theory, which underlies almost all the recent discussions of this issue, factor prices do not shift production functions. Thus, if the availability of cheaper ICT capital has increased TFP in industries that use but do not produce ICT equipment, then it has done so via a channel that neoclassical economics does not understand well.

We discuss theories of ICT as a general purpose technology (GPT) in an effort to see if these theories can explain the puzzle of why measured TFP accelerated in ICT-using industries. The main feature of a GPT is that it leads to fundamental changes in the production process of those using the new invention (see, e.g., Helpman and Trajtenberg, 1998). For example, Chandler (1977) discusses how railroads transformed retailing by allowing nationwide catalog sales. David and Wright (2003) also discuss historical examples. Indeed, the availability of cheap ICT capital allows firms to deploy their other inputs in radically different and productivity-enhancing ways. In so doing, cheap computers and telecommunications equipment can foster an ever-expanding sequence of complementary inventions in industries using ICT. These complementary inventions cause the demand curve for ICT to shift further and further out, thereby offsetting the effects of diminishing returns.

As Basu, Fernald, Oulton, and Srinivasan (2003; henceforth BFOS) highlight, ICT itself may be able to explain the measured acceleration in TFP in sectors that are ICT users. In their model, reaping the full benefits of ICT requires firms to accumulate a stock of intangible knowledge capital. For example, faster information processing might allow firms to think of new ways of communicating with suppliers or arranging distribution systems. These intangible investments may include resources diverted from direct production to learning; they may involve purposeful innovation arising from research and development (R&D). The assumption that complementary investments are needed to derive the full benefits of ICT is supported both by GPT theory and by firm-level evidence.[2] Since (intangible) capital accumulation is a slow process, the full benefits of the ICT revolution show up in the ICT-using sectors with significant lags.

Note that the BFOS story hews as closely as possible to neoclassical assumptions while explaining the puzzle of TFP growth in ICT-using industries. From a firm's perspective, the story is essentially one of neoclassical capital accumulation. If growth accounting could include intangible capital as an input to production then it would show no technical change in ICT-using industries. (Of course, measuring intangible capital directly is very difficult at best; see Corrado, Hulten, and Sichel 2006.) But the story can easily be extended to include features that are not neoclassical and that would explain true technical progress in ICT-using industries via other mechanisms, such as spillovers. Indeed, to the extent that much of the intangible capital accumulated by ICT users is knowledge, which is a nonrival good, it would be natural to expect spillovers. For example, the innovations that have made Amazon.com and Wal-Mart market leaders could presumably be imitated, at least in the long run, at a fraction of the cost it took to develop these new ideas in the first place.

We assess whether the acceleration in measured TFP is related to the use of ICT. We develop a simple model to motivate our empirical work. The model predicts that observed investments in ICT are a proxy for unobserved investments in reorganization or other intangible knowledge. In this model, the productivity acceleration should be positively correlated with lagged ICT capital growth but *negatively* correlated with current ICT capital growth (with these growth rates "scaled" by the share of ICT capital in output). Note that the unconditional correlation between the productivity acceleration and either ICT capital growth or the ICT capital share can be positive, negative, or zero.

In the data, we find results that support the joint hypothesis that ICT is a GPT—i.e., that complementary investment is important for realizing the productivity benefits of ICT investment—and that, since these complementary investments are unmeasured, they can help explain the cross-industry and aggregate TFP growth experienced by the United States in the 1990s. Specifically, we find that industries that had high ICT capital growth rates in the 1987–2000 period (weighted by ICT revenue shares, as suggested by theory) also had a faster acceleration in TFP growth in the 2000s. Controlling for lagged capital growth, however, ICT capital growth in the 2000s was *negatively* correlated with contemporaneous TFP

2. For evidence, see Bresnahan, Brynjolfsson, and Hitt (2002).

growth. These results are consistent with—indeed, predicted by—the simple model that we present.

The paper is structured as follows. We present preliminary empirical results from industry-level growth accounting in Section 2 and document the puzzle we note above. We then present a simple model of intangible capital investment in Section 3 and show how measured inputs—especially ICT investment—can be used to derive a proxy for unmeasured investment in intangibles. We test the key empirical implications of the model in Section 4. Conclusions, caveats, and ideas for future research are collected in Section 5.

2. Data and Preliminary Empirical Results

We begin by establishing stylized facts from standard growth accounting. We focus on disaggregated, industry-level results for TFP. We first describe our data set briefly and then discuss results.

Our 40-industry data set updates that used in Basu, Fernald, and Shapiro (2001), Triplett and Bosworth (2006), and BFOS (2003). The data run from 1987 to 2004 on a North American Industry Classification System (NAICS) basis. For industry gross output and intermediate inputs, we use industry-level national accounts data from the BEA. For capital input—including detailed ICT data—we use Bureau of Labor Statistics (BLS) capital input data by disaggregated industry. For labor input, we use unpublished BLS data on hours worked by industry.[3]

Several comments are in order. First, there are potential differences in how the conversion from the earlier Standard Industrial Classification system to NAICS has been implemented across agencies; see Bosworth and Triplett (2006) and Corrado et al. (2006) for a discussion. Second, we do not have industry measures of labor quality, only raw hours, as estimated by the BLS. Third, we aggregate industries beyond what is strictly necessary, in part because of a concern that industry matches across data sources are not as good at lower levels of aggregation. (For example, in some cases, our BLS estimate of capital compensation shares in a sub-industry substantially exceeded the implied BEA figure, whereas in another sub-industry the share fell substantially short; once aggregated, the BLS figure was close to—i.e., only slightly smaller than—the BEA figure, as expected.)

Table 1 provides standard estimates of TFP for various aggregates, including the one-digit industry level. The first three columns show TFP growth, in value-added terms, averaged over different time periods. Since aggregate TFP is a value-added concept, we present industry TFP in value-added terms as well; by controlling for differences in intermediate input intensity, these figures are scaled to be comparable to the aggregate figures. The next two columns show the acceleration, first from 1987–1995 to 1995–2000; and then from 1995–2000 to 2000–2004. The final two columns show the average share of intermediate inputs in gross output and the sector's nominal share of aggregate value-added.[4]

The top line shows an acceleration of about ½ percentage point in the second half of the 1990s, and then a further acceleration of about ¾ percentage point in the 2000s. The other lines show various subaggregates, including the one-digit NAICS level. It is clear that in our data set, the acceleration was broad-based. First, suppose we focus on the non-ICT-producing sectors (fourth line from the bottom). They show a very small acceleration in the late 1990s (from 0.68 to 0.81 percent per year), but then a much larger acceleration in the 2000s (to an average of 1.98 percent per year). In contrast, ICT-producing industries saw a sharp acceleration in TFP in the late 1990s but then some deceleration in the 2000s.

A more detailed analysis of the non-ICT sectors shows more heterogeneity in the timing of the TFP acceleration. For example, trade and finance accelerated in the mid-1990s and growth then remained strong in the 2000s. Nondurable manufacturing, business services, and information slowed in the mid-1990s before accelerating in the 2000s. Nevertheless, by the 2000s, most sectors show an acceleration relative to the pre-1995 period (mining, utilities, and insurance are exceptions).

Griliches (1994) and Nordhaus (2002) argue that real output in many service industries are poorly measured—e.g., there are active debates on how, conceptually, to measure the nominal and "real output" of a bank;[5] in health care, the

3. The BEA data on gross product origination were downloaded from http://bea.gov/bea/dn2/gdpbyind_data.htm on March 15, 2006. The BLS capital data were downloaded from http://www.bls.gov/web/prod3.supp.toc.htm on March 21, 2006. We removed owner-occupied housing from the BEA data for the real estate industry. The BEA labor compensation data do not include proprietors or the self-employed, so we follow Triplett and Bosworth (2006) in using BLS data on total payments to capital that correct for this. We thank Steve Rosenthal at the BLS for sending us unpublished industry hours data, which adjusts for estimated hours worked by nonproduction and supervisory employees as well as the self-employed (received via email on June 27, 2006).

4. Aggregate TFP growth is a weighted average of industry gross-output TFP growth, where the so-called "Domar weights" equal nominal industry gross output divided by aggregate value added; the weights thus sum to more than one. See Hulten (1978) for an extensive discussion. In continuous time, this is equivalent to first converting gross-output residuals to value-added terms by dividing by (one minus the intermediate share) and then using shares in nominal value added. Hence, contributions to aggregate TFP growth are the same, using value-added weighted value-added TFP or using Domar-weighted gross-output TFP. (In discrete time, using average shares from adjacent periods, they are approximately equivalent.) Basu and Fernald (2001) discuss this aggregation and its extension to the case of imperfect competition; see also Oulton (2001).

5. See, for example, Wang, Basu, and Fernald (2004).

TABLE 1
U.S. TOTAL FACTOR PRODUCTIVITY BY INDUSTRY IN PRIVATE NONFARM BUSINESS, 1987–2004
(AVERAGE ANNUAL PERCENT CHANGES, EXCEPT WHERE NOTED)

	Productivity[a]			Acceleration		Intermediate input share of nominal VA	Industry share of nominal VA
	1987–1995	1995–2000	2000–2004	1995–2000	2000–2004	1987–2004	1987–2004
Private nonfarm economy (*not* adjusted for labor quality)	0.96	1.43	2.21	0.47	0.79	49.50	100.00
Mining	3.45	–0.37	–3.13	–3.82	–2.76	49.10	1.50
Manufacturing	2.07	3.81	4.09	1.74	0.28	66.36	19.12
Nondurables	0.70	–0.16	2.17	–0.86	2.33	69.77	7.98
Durables	3.06	6.51	5.50	3.45	–1.01	63.47	11.14
Construction	0.00	–2.89	–1.00	–2.90	1.89	49.31	5.36
Transportation	2.31	1.43	2.82	–0.88	1.39	51.51	3.51
Communication[b]	2.92	–0.47	5.75	–3.39	6.23	46.40	3.26
Utilities	2.04	1.04	1.15	–1.01	0.11	41.86	2.87
Information	1.95	0.60	6.33	–1.35	5.74	48.67	5.34
Wholesale trade	1.87	4.28	4.25	2.41	–0.03	36.75	7.66
Retail trade	2.48	4.52	4.48	2.03	–0.03	37.27	8.71
Finance & insurance	0.50	2.30	2.21	1.81	–0.09	43.64	8.62
Finance	0.27	2.84	3.88	2.57	1.04	39.51	5.78
Insurance	0.88	0.11	–1.63	–0.78	–1.74	50.40	2.83
Business services & real estate	0.42	–1.10	2.17	–1.52	3.27	40.41	12.82
Business services	0.19	–1.28	3.64	–1.47	4.93	34.93	5.12
Real estate	0.98	–0.41	1.05	–1.39	1.46	44.48	7.70
Other services[c]	–0.70	–0.03	0.16	0.67	0.20	40.15	26.29
ICT-producing[d]	6.84	14.85	9.39	8.00	–5.45	61.04	4.25
Non-ICT-producing	0.68	0.81	1.98	0.13	1.17	48.76	95.75
Well-measured industries[e]	2.18	3.51	3.58	1.33	0.07	56.71	43.66
Well-measured industries (excluding ICT-producing)	1.66	2.21	3.08	0.55	0.87	56.12	39.41
Poorly measured industries[f]	–0.07	–0.17	1.31	–0.10	1.48	42.23	56.34

a. Productivity is defined as (gross output TFP growth)/(1–share of intermediate inputs). Implicitly, this uses the Törnqvist index of value added for a sector.
b. Communication includes broadcasting and telecommunications from the information industry aggregate.
c. Other services includes NAICS codes 61–62, 71–72, and 81, as well as the noncommunications elements of "information" (NAICS code 51, excluding 513).
d. ICT-producing industries includes machinery, computer and electronic products, and electrical equipment, appliances, and components.
e. Well-measured industries include mining, manufacturing, transportation, utilities, and wholesale and retail trade.
f. Poorly measured industries include construction, information, finance and insurance, real estate and rental and leasing, professional and business services, etc.

hedonic issues are notoriously difficult. Nordhaus argues for focusing on "well-measured" (or at least, "better measured") sectors of the economy. The acceleration in TFP in well-measured industries (third line from the bottom) took place primarily in the 1990s with little further acceleration in the 2000s; but excluding ICT-producing sectors, the acceleration is spread out over the 1995–2004 period.

In the short term, nontechnological factors can change measured industry TFP. These factors include non-constant returns to scale and variations in factor utilization. Basu, Fernald, and Shapiro (BFS, 2001) argue that cyclical mismeasurement of inputs plays little if any role in the U.S. acceleration of the late 1990s. BFS also find little role in the productivity acceleration for deviations from constant returns and perfect competition.

In the early 2000s, some commentators suggested that, because of uncertainty, firms were hesitant to hire new workers; as a result, one might conjecture that firms might have worked their existing labor force more intensively in order to get more labor input. But typically, one would expect that firms would push their workers to work longer as well as harder; this is the basic intuition underlying the use of hours-per-worker as a utilization proxy in Basu and Kimball (1997), BFS, and Basu, Fernald, and Kimball (2006). In the 2000s, however, when productivity growth was particularly strong, hours per worker remained low.

BFS do find a noticeable role for traditional adjustment costs associated with investment. When investment rose sharply in the late 1990s, firms were, presumably, diverting an increasing amount of worker time to installing the new capital rather

than producing marketable output. This suggests that true technological progress was faster than measured. In contrast, investment generally was weak in the early 2000s, suggesting that there was less disruption associated with capital installation. Nevertheless, the magnitude of this effect appears small, for reasonable calibrations of adjustment costs. Applying the BFS correction would raise the U.S. technology acceleration from 1995 to 2000 by about 0.3 percentage points per year, but would have a negligible effect from 2000 to 2004. Hence, the investment reversal could potentially explain some portion of the second wave of acceleration, but not all of it.[6] These adjustment-cost considerations strengthen the conclusion that the technology acceleration was broad-based, since service and trade industries invested heavily in the late 1990s and, hence, paid a lot of investment adjustment costs.

3. Industry-Level Productivity Implications of ICT as a New GPT

The U.S. productivity acceleration in the late 1990s coincided with accelerated price declines for computers and semiconductors. But, as we just saw, much of the TFP acceleration appears to have taken place in the 2000s, and outside of ICT production. Can ICT somehow *explain* the measured TFP acceleration in industries using ICT? We first discuss broad theoretical considerations of treating ICT as a new general purpose technology and then present a simple model to clarify the issues and empirical implications.

3.1. General Purpose Technologies and Growth Accounting

Standard neoclassical growth theory suggests two direct channels through which ICT can affect aggregate labor and total factor productivity growth. First, faster TFP growth in *producing* ICT contributes directly to aggregate TFP growth. Second, by reducing the user cost of capital, falling ICT prices induce firms to increase their desired capital stock.[7] This *use* of ICT contributes directly to labor productivity through capital deepening.

Growth accounting itself does not take a stand on the deep causes of innovation and TFP. Neoclassical growth theory generally takes technology as exogenous, but this is clearly a modeling shortcut, appropriate for some but not all purposes. Endogenous growth theories, in contrast, generally presume that innovation results from purposeful investments in knowledge or human capital, possibly with externalities.

We interpret ICT's general purpose nature in the spirit of the neoclassical growth model, since the GPT arrives exogenously (i.e., technological progress in ICT production is exogenous). ICT users respond in a neoclassical way: Firms respond to faster, more powerful computers and software by reorganizing and accumulating intangible organizational capital. Measured TFP, which omits this intangible organizational investment as output and the service flow from organizational capital as an input, is also affected.

Our motivation for viewing ICT this way is the many microeconomic, firm-level, and anecdotal studies suggesting an important—but often indirect and hard to foresee—role for ICT to affect measured production and productivity in sectors using ICT. Conceptually, we separate these potential links into two categories: purposeful co-invention, which we interpret as the accumulation of complementary organizational capital and which leads to mismeasurement of true technology, and externalities of one sort or another. For example, Bresnahan and Trajtenberg (1995) and Helpman and Trajtenberg (1998) suggest that innovations in ICT cause unexpected ripples of co-invention and co-investment in sectors that seem almost arbitrarily far away.

First, firm-level studies suggest that benefiting from ICT investments requires substantial and costly co-investments in complementary capital, with long and variable lags.[8] For example, firms that use computers more intensively may reorganize production, thereby creating intangible capital in the form of organizational knowledge. Such investments include resources diverted to learning, or purposeful innovation arising from R&D. As Bresnahan (undated) argues, "advances in ICT shift the innovation possibility frontier of the economy rather than directly shifting the production frontier."

The resulting organizational capital is analogous to physical capital in that companies accumulate it in a purposeful way. Conceptually, we think of this unobserved complementary capital as an additional input into a standard neoclassical production function.[9]

6. These numbers are qualitatively the same but smaller than those reported in BFS (2001) for three reasons. The two main reasons are (a) data revisions have reduced the growth rate of investment in the second half of the 1990s, and (b) Jason Cummins and John Roberts pointed out a mistake in our conversion from Shapiro's (1986) framework to ours. This led us to reduce our estimate of the "disruption cost" per unit of investment growth (the BFS parameter ϕ from 0.048 in BFS to 0.035).

7. Tevlin and Whelan (2003) for the U.S. and Bakhshi, Oulton, and Thompson (2003) for the U.K. provide econometric evidence that falling relative prices of ICT equipment fueled the ICT investment boom.

8. See, for example, Brynjolfsson and Hitt (2000) and Bresnahan (undated) for a discussion of the kinds of complementary investments and co-invention that firms undertake to benefit from ICT, given its general purpose attributes. Bloom, Sadun, and Van Reenen (2005) use data on cross-country mergers to provide additional firm-level evidence for the importance of (partially transferable) intangible capital.

9. Much of Brynjolfsson's work tries to quantify the role of unobserved complementary capital. Macroeconomic studies of the effects of organizational capital include Greenwood and Yorukoglu (1997), Hornstein and Krusell (1996), Hall (2001), and Laitner and Stolyarov (2003).

Second, the GPT literature suggests the likelihood of sizeable externalities to ICT. For example, successful new managerial ideas—including those that take advantage of ICT, such as the use of a new business information system—seem likely to diffuse to other firms. Imitation may be easier and less costly than the initial co-invention of, say, a new organization change, because you learn by watching and analyzing the experimentation, the successes, and, importantly, the mistakes made by others.[10] Indeed, firms that *don't* use computers more intensively might also benefit from spillovers of intangible capital. For example, if there are sizeable spillovers to R&D, and if R&D is more productive with better computers, then even firms that don't use computers intensively may benefit from the knowledge created by computers.

The first set of considerations are completely consistent with the traditional growth accounting *framework* but suggest difficulties in implementation and interpretation. In particular, these considerations suggest that the production function is mismeasured because we don't observe all inputs (the service flow from complementary, intangible capital) or all outputs (the investment in complementary capital). Hence, TFP is mismeasured. The second set of ideas, related to externalities, suggest that ICT might also explain "true" technology.

Empirically, the challenge is to infer unobserved complementary investments. We now turn to a formal model that suggests variables that might proxy for these unobservables. Of course, our interpretation of the results will be clouded by our uncertainty about whether our proxies are capturing only neoclassical investment in unobserved organizational capital or whether the proxies are affecting TFP directly through spillovers.

3.2. Industry-Level Implications of ICT as a New GPT: A Simple Model[11]

Many papers modeling the effects of GPTs are motivated by the ICT revolution.[12] But it is difficult to derive industry-level empirical implications from this literature. For example, it is often unclear how to measure in practice some of the key variables, such as unobserved investment and capital; and even for observed variables, measurement conventions often depart from those used in national accounting.[13]

On the other hand, conventional industry-level growth-accounting studies of the sort reviewed and extended in Section 2 are typically hard to interpret in terms of GPT considerations because they generally lack a conceptual framework to interpret movements in TFP. Although some studies try to look for a "new economy" in which ICT has indirect effects on measured TFP in ICT-using industries, in the absence of clear theoretical guidance, it is not clear that many would know if they had, in fact, found it.

Finally, the empirical literature using firm-level data or case studies stresses the importance and costly nature of organizational change accompanying ICT investment. This literature is insightful but rarely makes contact with economy-wide productivity research. (An exception is Brynjolfsson and Yang 2002.) Our empirical work below is a tentative attempt to make that connection. The model below provides the bare bones of a theoretical framework to capture some of the key issues, focusing on cross-industry empirical implications. The model takes as given the arrival of a particular GPT, which here is taken to be the production of ICT capital at a continuously falling relative price. The distinguishing feature of a GPT is that its effects are general—going well beyond the industry of production—but require complementary investments by firms for them to benefit fully from its use. For empirical implementation, we focus on industries that use the GPT.

Value added, Q_{it}, in industries that use, but do not produce, ICT is given by

$$(1)\quad Q_{it} \equiv Y_{it} + A_{it} = F\left(Z_t G(K_{it}^{IT}, C_{it}), K_{it}^{NT}, L_{it}\right), i = 1, \ldots N,$$

where the production function F and the IT services function G are homogeneous of degree 1 in their arguments. Note that effective IT services depend on both ICT capital K^{IT} as well as complementary organizational capital C. Z is a technology term that each industry takes as exogenous. We discuss the distinction between A and Y shortly. For simplicity, we ignore intermediate inputs (though we incorporate them in our empirical work), imperfect competition, increasing returns, and capital adjustment costs. All could be added, at the cost of considerable notation.

10. Bresnahan (undated) provides a nice discussion of the channels for externalities to operate. Bresnahan and Trajtenberg (1995) highlight both "vertical" externalities (between GPT producers and each application sector) and "horizontal" externalities (across application sectors).

11. This section follows Basu, Fernald, Oulton, and Srinivasan (2003) fairly closely.

12. A very incomplete list is Caselli (1999), Greenwood and Yorukoglu (1997), the collection of papers edited by Helpman (1998), Hobijn and Jovanovic (2001), Jovanovic and Rousseau (2004), and Laitner and Stolyarov (2003).

13. For example, capital is typically measured as foregone consumption, which is sensible for an aggregative model but difficult to relate to industry-level capital accounts that deal with capital heterogeneity and quality change by (attempting to) measure capital input in efficiency units. Howitt (1998) attempts to bridge the two conventions.

Each industry hires labor L and rents ICT capital K^{IT} and non-ICT capital K^{NT} in competitive, economywide markets. The two types of capital depreciate at rates δ^{IT} and δ^{NT}, respectively. For given investment flows I^{IT} and I^{NT}, the aggregate stocks of the two types of capital (indexed by J) evolve as

$$K_t^{JT} = I_t^{IT} = (1 - \delta^{JT})K_{t-1}^{JT}, J = I, N. \tag{2}$$

Industries must individually accumulate complementary capital, C, representing business and organizational models or IT training. The investment flow A is the time and resource cost of training and creating new business structures.[14] Industries forego producing market output Y to accumulate this capital, which then depreciates at rate δ^C:

$$C_{it} = A_{it} + (1 - \delta^C)C_{it-1}. \tag{3}$$

Investment is irreversible. Since both A and NT investment goods cost the same to produce, the economic difference between the two types of capital is that they interact in different ways with ICT capital. The difference in terms of measurement is that Y is observable by national accountants but A is not.[15]

The main economic implication of the separability assumption built into equation (1) is that the marginal productivities of K^{IT} and C are closely tied to one another. We assume that the elasticity of substitution between the two inputs in the production of G is relatively small. We also assume standard conditions to the effect that the marginal productivity of each input is very low if the level of the other is close to zero. Thus, when the GPT arrives and ICT capital starts getting cheap, the incentive to also accumulate C is very strong.

Note that conceptually, "innovation" as traditionally construed can take two forms. First, we lump purposeful innovations into C (indeed, we have assumed that *all* purposeful innovation is closely linked to ICT). Second, we interpret Z as all exogenous increases in technology, including the component of organizational change that spills over as an externality from the sector of origin—for example, the idea of using individual electric motors at each workstation in a factory, rather than relying on the single drive train of a steam engine.

14. Chandler (1977) discusses innovations in inventory management made possible by railroads. The Wal-Mart inventory management system provides an example of innovations made possible by ICT.

15. Some fraction of A is probably measured: for example, consultant services and many forms of software. It is not clear how much of what is measured is properly capitalized, as required by equation (3).

3.3. TFP Measurement with Unobserved Inputs and Output

What are the implications of complementary capital accumulation for the measured TFP of ICT-using industries? Differentiating, we can write the production function in growth rates as

$$\begin{aligned}\Delta q &= \frac{Y}{Q}\Delta y + \frac{A}{Q}\Delta a \\ &= \frac{F_{K^{IT}}K^{IT}}{Q}\Delta k^{IT} + \frac{F_C C}{Q}\Delta c + \frac{F_{K^{NT}}K^{NT}}{Q}\Delta k^{NT} \\ &\quad + \frac{F_L L}{Q}\Delta l + \frac{F_Z Z}{Q}\Delta z.\end{aligned} \tag{4}$$

Lowercase letters are logs of their uppercase counterparts. Suppose P is the output price, W is the wage, and P_K^{IT} and P_K^{NT} are the rental prices for the two types of capital. Since we have assumed constant returns to scale and perfect competition, output elasticities equal factor shares in revenue. Hence,

$$\begin{aligned}&\frac{F_C C}{Q} + \frac{F_{K^{IT}}K^{IT}}{Q} + \frac{F_{K^{NT}}K^{NT}}{Q} + \frac{F_L L}{Q} \\ &= \frac{F_C C}{Q} + \frac{P_K^{IT}K^{IT}}{PQ} + \frac{P_K^{NT}K^{NT}}{PQ} + \frac{WL}{PQ} = 1.\end{aligned} \tag{5}$$

If we observed total output Q and knew the required rates of return to capital, we could back out the elasticity of output with respect to complementary capital, C:

$$\frac{F_C C}{Q} = 1 - \frac{WL}{PQ} - \frac{P_K^{IT}K^{IT}}{PQ} - \frac{P_K^{NT}K^{NT}}{PQ}. \tag{6}$$

Without independent information on the flow of A or the stock of C (perhaps from stock market valuations), one cannot implement this procedure using measured output, Y. We rewrite equation (6) as

$$\frac{F_C C}{Y} = \frac{Q}{Y} - \frac{WL}{PY} - \frac{P_K^{IT}K^{IT}}{PY} - \frac{P_K^{NT}K^{NT}}{PY}.$$

Since Q/Y is not observed, within broad limits we are free to believe that complementary capital is arbitrarily important in production by assuming that an arbitrarily large share of the true output that firms produce is not counted in the national accounts.

Some algebraic manipulations of equation (4) yield an expression for the measured Solow residual, i.e., measured TFP:[16]

16. The observed factor shares do not necessarily sum to one, even with perfect competition, as we assumed in our empirical work. In our experience, however, estimating the user costs rather than taking them as a residual makes little practical difference.

$$(7)\quad \Delta y - \frac{P_K^{IT}K^{IT}}{PY}\Delta k^{IT} - \frac{P_K^{NT}K^{IT}}{PY}\Delta k^{NT} - \frac{WL}{PY}\Delta l \equiv \Delta TFP = \frac{F_C C}{Y}\Delta c - \frac{A}{Y}\Delta a + s_G \Delta z,$$

where $s_G \equiv (F_Z Z/Y)$. Omitting complementary inputs can cause us to either overestimate or underestimate true TFP growth, $s_G\Delta z$. When unmeasured *output* is growing ($\Delta a > 0$), TFP growth is underestimated (the "1974" story) as resources are diverted to investment. When unmeasured *input* is growing ($\Delta c > 0$), TFP growth is overestimated. In the steady state, of course, the accumulation equation implies that $\Delta c = \Delta a \equiv g$. Hence, steady-state mismeasurement depends on r^*, the steady-state real interest rate, and g, the steady-state rate of growth:

$$\left[\frac{F_C C}{Y} - \frac{A}{Y}\right]g = \frac{C}{Y}\left[F_C - \frac{A}{C}\right]g = \frac{C}{Y}\left[(r^* + \delta^C) - \frac{g + \delta^C}{1+g}\right]g.$$

In a dynamically efficient economy the mismeasurement is necessarily positive: True steady-state TFP growth is *lower* than measured, not higher.[17]

This point is simple but important. Of course, if one corrects only output mismeasurement (Δa), then ICT will appear fantastically productive, far beyond what is ordinarily measured. But firms divert resources to unobserved investment Δa in order to create an intangible capital stock, which contributes to future production. The resulting unmeasured flow of capital services implies a bias in the other direction. The net bias may be either positive or negative at a point in time, but it is positive in the steady state.

We now seek an observable proxy for unobserved investment in, and growth in the stock of, complementary capital. Observed growth in ICT capital provides a reasonable proxy. Suppose G takes a constant elasticity of substitution (CES) form:

$$G = \left[\alpha K^{IT\frac{\sigma-1}{\sigma}} + (1-\alpha)C^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}.$$

σ is the elasticity of substitution between the two inputs; α gives their relative importance. Consider the optimization subproblem of producing G at minimum cost each period. Let P_K^{IT}/P_K^C be the relative rental rate of ICT capital to C capital. The solution of the subproblem is

$$(8)\quad \Delta c_t = \Delta k_t^{IT} + \sigma\Delta \ln(P_K^{IT}/P_K^C)_t.$$

This equation links growth in unobserved complementary capital and growth of observed ICT capital.

We can use the accumulation equation to express unobserved investment Δa in terms of current and lagged growth in unobserved capital Δc:

$$\Delta a_t = \frac{C}{A}\left[\Delta c_t - \frac{(1-\delta^C)}{(1+g)}\Delta c_{t-1}\right].$$

Substituting the last equation and equation (8) into (7), we have in principle an equation for TFP growth that indicates the importance of complementary capital accumulation:

$$(9)\quad \Delta TFP = \left[\frac{C}{Y}(F_c - 1)\right]\left[\Delta k_t^{IT} + \sigma\Delta\ln\left(\frac{P_K^{IT}}{P_K^C}\right)_t\right] + \left[\frac{C}{Y}\frac{(1-\delta^C)}{(1+g)}\right]\left[\Delta k_{t-1}^{IT} + \sigma\Delta\ln\left(\frac{P_K^{IT}}{P_K^C}\right)_{t-1}\right] + s_G\Delta z.$$

The first term is proportional to $(r^* + \delta^C - 1)$ where r^* is the steady-state interest rate; hence, under reasonable circumstances it is negative. The second term, on the other hand, is clearly positive. Hence, other things equal, industries that are making large ICT investments today will have low measured TFP growth, but those that made such investments in the *past* will have high measured TFP growth.

As an estimating equation, (9) has the difficulty that industries are likely to differ in their long-run *C/Y* ratios. Using the CES assumption for G, the cost-minimizing first-order condition implies that

$$\frac{C}{K^{IT}} = \left[\left(\frac{1-\alpha}{\alpha}\right)^\sigma\right]\left(\frac{P_K^{IT}}{P_K^C}\right)^\sigma.$$

If we define the observed share of payments to ICT capital in revenue as $s_{K,IT} = P_K^{IT}K^{IT}/PY$, then we can rearrange this condition as

$$(10)\quad \frac{C}{Y} = \left[\left(\frac{P}{P_K^C}\right)\left(\frac{1-\alpha}{\alpha}\right)^\sigma\left(\frac{P_K^{IT}}{P_K^C}\right)^{\sigma-1}\right]\left(\frac{P_K^{IT}K^{IT}}{PY}\right) \equiv \beta s_{K,IT}.$$

β is a composite parameter that depends on various things, including the share parameter α, the elasticity of substitution σ, and relative prices. In the Cobb-Douglas case, where σ equals one, *C/Y* is proportional to the observed ICT share. Hence, other things equal, the mismeasurement of complementary capital is more important in those industries where ICT capital is used to a greater extent. We can now substitute equation (10) into equation (9) to find

$$(11)\quad \Delta TFP = [F_C - 1]\beta\tilde{k}_t + \left[\frac{(1-\delta^C)}{(1+g)}\right]\beta\tilde{k}_{t-1} + s_G\Delta z,$$

17. Dynamic efficiency requires that the output elasticity equal or exceed the investment rate. In our discrete-time economy, one can show that dynamic efficiency requires that the marginal product of capital, which equals $r + \delta$, be greater than or equal to $(g + \delta)/(1 - g)$.

where $\tilde{k}_t = s_{K,IT}\left[\Delta k_t^{IT} + \sigma \Delta \ln\left(P_K^{IT}/P_K^C\right)_t\right]$. $\tilde{k}_t$ is a composite variable that, in essence, captures share-weighted ICT capital growth (with a relative-price adjustment).

This model has several general implications. First, one might find a link between ICT use and measured TFP even if there are no externalities to ICT use. Second, the correct proxy for ICT use involves the interaction of ICT intensity (the ICT share) and the growth rate. Intuitively, if ICT capital grows quickly but its share is small, then there probably isn't much complementary capital to cause mismeasurement. In contrast, if the share is large (implying complementary capital is likely important) but the growth of ICT capital is small, then the mismeasurement of true output and true investment is also small during the period at hand. Third, one needs to control for both current and lagged $\tilde{k}$. Since these values are correlated in the data, if one omits one of them, then the regression has an omitted variable problem. Indeed, the regression coefficient could have either sign, since the correlated omitted variable has a coefficient with the *opposite* sign.[18]

3.4. Extensions to the Basic Framework

The model above is, of course, stylized and imposes a lot of structure on the problem to derive an estimating equation. As a result, there are a number of challenges in implementing this framework empirically. First, it is unclear how long the lags are between ICT investment and complementary investment. In other words, the length of a period is a free parameter, and theory gives little guidance. The lagged $\tilde{k}$ may be last year's ICT capital accumulation, or the last decade's. Furthermore, equation (3) for the accumulation of complementary capital has no adjustment costs, or time-to-build or time-to-plan lags in the accumulation of *C*. But such frictions and lags are likely to be important in practice, making it even harder to uncover the link between ICT and measured TFP.

Second, suppose there are externalities captured in technology growth Δz and that they are a function of industry as well as aggregate *C*. Then one can no longer tell whether the $\tilde{k}$ terms represent accumulation of a private stock or intra-industry externalities that are internalized within the industry. Similarly, if we find that lagged $\tilde{k}$ is important for explaining current productivity growth, we do not know whether that finding supports the theory we have outlined or whether it indicates that the externality is a function of lagged capital.

Third, other variables might enter the production function for *A*, which we have not accounted for. We imposed the same production function for *A* and *Y*. But it is possible, as many have recognized, that the production of complementary capital is particularly intensive in skilled (i.e., college-educated) labor.[19] If true, the hypothesis implies that the relative price of accumulating complementary capital may differ significantly across industries (or across countries) in ways that we have ignored.

Fourth, even with the restrictions we've imposed, we need to make further assumptions about σ as well as the relative user costs for ICT and complementary capital. We made the strong assumption that the price of complementary investment is the same as that of output, so this relative price should largely reflect the trend decline in ICT prices. Nevertheless, that was clearly an assumption of convenience—reflecting our lack of knowledge—rather than something we want to rely on too strongly. In what follows, we ignore the relative price terms, but this needs to be explored further. (Suppose we assume that $\sigma = 0$. There is still a relative price effect which, if omitted, would imply a trend in the estimated coefficient over time; but in the cross-section, this relative price is close to common across firms, so its omission shouldn't matter much.)

Finally, given the difficulty of finding good instruments, we report ordinary least squares (OLS) regressions below. But current ICT capital growth is surely endogenous. Given the correlation between current and lagged share-weighted ICT capital growth, any endogeneity potentially biases both coefficients. The effect on estimates depends on the size of the true coefficient as well as the degree of endogeneity. The endogeneity bias might be positive or negative: Basu, Fernald, and Kimball (2006), for example, find that positive technology innovations tend to reduce inputs on impact. As is standard, one trades off bias against precision; indeed, weak instruments could lead to both bias and imprecision. In any case, one needs to interpret the results with caution.

4. Evidence for the GPT Hypothesis

Several studies explore whether TFP growth across industries is correlated with ICT intensity. In contrast to firm-level studies, these industry studies rarely find much correlation between ICT capital and TFP growth (e.g., Stiroh, 2002b). But as noted, our model implies that the contemporaneous correlation need not be positive—even if ICT is, in fact, an important contributor to measured TFP. BFOS found that the industry productivity acceleration in the second half of the 1990s was positively correlated with lagged ICT use but

18. Hence one needs to generalize the approach followed by, e.g., Stiroh (2002b) who looks for ICT spillovers by regressing TFP growth on only the current-year growth rate of IT capital. See Stiroh (2002a) and Stiroh and Botsch (2007) for related work. Brynjolfsson and Hitt (2003) also find significant lags in firm-level data, which nicely complements our more aggregative evidence.

19. Krueger and Kumar (2001) ask whether the different educational systems in the United States and Europe (especially Germany) may be responsible for their different growth experiences in the 1990s. See also Lynch and Nickell (2001).

TABLE 2
U.S. COMPUTER, SOFTWARE, AND COMMUNICATION SHARES OF VALUE-ADDED REVENUE (%)

	1990				2004			
	Computer	Software	Communication	Total ICT	Computer	Software	Communication	Total ICT
Private nonfarm economy	1.22	1.00	1.29	4.91	1.00	2.18	1.34	5.55
Mining	0.25	0.22	0.29	1.62	0.23	1.15	0.13	2.58
Manufacturing	1.08	0.93	0.44	3.60	0.62	2.73	0.29	4.83
Nondurables	0.88	0.75	0.33	3.70	0.71	2.67	0.32	5.54
Durables	1.23	1.07	0.52	3.52	0.56	2.78	0.27	4.32
Construction	0.06	0.05	0.02	0.20	0.27	1.22	0.55	2.38
Transportation	0.11	0.10	1.13	1.44	0.56	1.39	2.58	4.73
Communication[a]	1.27	1.12	24.03	26.89	2.56	4.56	20.60	28.05
Utilities	1.30	1.13	1.87	6.91	0.77	3.54	0.73	6.61
Information	2.33	2.19	15.62	21.16	1.95	5.24	13.05	20.62
Wholesale trade	1.47	0.44	0.63	4.21	1.39	1.06	1.15	4.20
Retail trade	0.94	0.34	0.36	2.10	0.60	0.56	0.52	1.94
Finance & insurance	3.14	3.14	0.73	8.71	1.89	2.29	0.35	5.07
Finance	4.56	3.74	0.97	11.63	2.31	2.08	0.40	5.44
Insurance	0.51	2.02	0.31	3.35	0.88	2.81	0.24	4.19
Business services & real estate	2.03	1.02	0.68	4.28	1.75	2.06	1.07	5.21
Business services	1.24	2.33	0.85	5.36	1.51	4.25	1.66	7.94
Real estate	2.51	0.20	0.57	3.62	1.92	0.49	0.65	3.25
Other services[b]	0.86	1.18	0.50	5.17	0.64	2.63	0.59	5.93
ICT-producing[c]	1.41	1.29	0.91	4.52	0.61	3.30	0.26	4.87
Non-ICT-producing	1.21	0.99	1.31	4.93	1.01	2.14	1.38	5.57
Well-measured industries[d]	1.02	0.65	0.60	3.38	0.75	1.79	0.72	4.07
Well-measured industries (excluding ICT-producing)	0.98	0.58	0.56	3.25	0.76	1.67	0.75	4.00
Poorly measured industries[e]	1.39	1.30	1.88	6.20	1.16	2.42	1.74	6.49

a. Communication includes broadcasting and telecommunications from the information industry aggregate.
b. Other services includes NAICS codes 61–62, 71–72, and 81, as well as the noncommunications elements of "information" (NAICS code 51, excluding 513).
c. ICT-producing industries includes machinery, computer and electronic products, and electrical equipment, appliances, and components.
d. Well-measured industries includes mining, manufacturing, transportation, utilities, and wholesale and retail trade.
e. Poorly measured industries includes construction, information, finance and insurance, real estate and rental and leasing, professional and business services, etc.

Source: Authors' calculations using payments to ICT capital from BLS and nominal value added from BEA.

negatively correlated with current ICT use.[20] This section updates their analysis.

Table 2 shows ICT shares in value added. For the entire private nonfarm economy, ICT accounts for about 5.6 percent of value added. Communications is a substantial outlier. Business services are ICT intensive. Finance was a substantial outlier in 1990, but returned to near average by 2004.

There is considerable uncertainty about how long it takes to build complementary capital and how long it takes for any spillovers to occur. The time lags depend on factors such as the time it takes to learn, innovate, and reorganize, which depends in turn on the adjustment costs associated with that complementary capital investment. Brynjolfsson and Hitt (2002) find evidence of long lags in firm-level data; Howitt (1998) calibrates a model to U.S. data, and finds that the beneficial effects of a new GPT will not show up in national accounts data for more than 20 years. Thus, our regressions need to be interpreted with a high degree of caution, and should be interpreted in the spirit of data exploration.

To capture the idea behind equation (11) in a loose way, Table 3 considers the following regression:

$$\Delta p_i^{00-04} - \Delta p_i^{90-00} = c + a\tilde{k}_i^{87-00} + b\tilde{k}_i^{00-04} + \varepsilon_i,$$

where Δp is gross-output TFP growth; we report results using $\tilde{k}$ for computers, software, and communications equip-

20. In subsequent work, Oulton and Srinivasan (2005) applied the BFOS framework to later U.K. data.

TABLE 3
ICT REGRESSIONS WITH CURRENT AND LAGGED ICT CAPITAL GROWTH

	All industries	Excluding outliers	Mfg.	Nonmfg.	Well-measured
C	–0.58	–0.60	0.47	–1.53	–0.83
	(0.35)	(0.41)	(0.32)	(0.63)	(0.47)
$\tilde{k}_{1987-2000}$	7.15	8.87	7.42	7.75	9.51
	(2.27)	(2.96)	(4.77)	(2.66)	(2.36)
$\tilde{k}_{2000-2004}$	–5.12	–7.72	–13.10	–3.62	–7.83
	(3.03)	(3.43)	(6.52)	(3.04)	(2.90)
R^2	0.22	0.20	0.34	0.35	0.27
Observations	37	34	15	22	29

Notes: Table shows regressions of $\Delta p_i^{00-04} - \Delta p_i^{90-00} = c + a\tilde{k}_i^{87-00} + b\tilde{k}_i^{00-04} + \varepsilon_i$, where Δp_i is the average industry TFP growth over the period specified and where $\tilde{k} = s_{K,IT}\, dk_{IT}$ uses computers, software, and communications equipment as the measure of ICT capital; $s_{K,IT}$ is the share in gross output. Robust standard errors are in parentheses. We omit ICT-producing industries.

ment (in all cases, the variables are averaged over the period shown):

$$\tilde{k} = s_{K,IT}\Delta \ln k_{K,IT},$$

where $s_{K,IT}$ is the share of computers and software in gross output (using value-added shares instead makes little qualitative difference to the results). This measure of $\tilde{k}$ drops the relative price terms from the alternatives discussed; BFOS found that results appeared more stable with this measure than when the relative price terms appeared, although qualitative results were generally similar.

Thus, we regress the acceleration in the 2000s, relative to the late 1990s, on $\tilde{k}$ for the 1990s, and on $\tilde{k}$ for the 2000s. Using the acceleration partially controls for differences in underlying growth rates across sectors.[21] (Separating out $\tilde{k}$ for the early and late 1990s led to multicollinearity problems—i.e., coefficients on the 1990s regressors were statistically insignificant, although in magnitude they remained similar to the combined coefficient a; the coefficient b was not (much) affected.) A virtue of the specification is that it imposes relatively minimal restrictions on lags and on coefficient stability, since the period averaging smoothes through some of that. Nevertheless, there is no reason to expect that the relationship will work equally well when applied to all time periods, since the link between ICT and unobserved intangible investment that we highlight in the model could easily vary over time and over industries. We also throw out ICT-producing industries, since they are such enormous outliers in the productivity dimension. (When we include the ICT-producing industries, results overall and within manufacturing are qualitatively similar but are very sensitive to outliers. Nonmanufacturing is not, of course, affected.)

Table 3 shows that, with long lags, ICT capital growth is positively associated with the industry TFP acceleration; but after controlling for past values, contemporaneous ICT capital growth is negatively associated with the acceleration. That is, $\tilde{k}_i^{87-00}$ enters positively but $\tilde{k}_i^{00-04}$ enters negatively. Thus, the productivity acceleration in the 2000s is somewhat consistent with the predictions of the theory section.

These results are robust to outliers. We identified influential observations (based on the Belsley-Kuh-Welsch "hat matrix" test).[22] When we omit those industries, in the second column, the standard errors go up, but so do the coefficients. We run the regression separately for manufacturing and nonmanufacturing, and the qualitative conclusions remain. In manufacturing, the current growth rate is particularly negative; in nonmanufacturing, the lagged growth rate is a bit more positive. When we look only at "well-measured" industries (basically, manufacturing plus a subset of nonmanufacturing, as discussed in the appendix), the magnitude and significance of the coefficients remains substantial.

These regression results are certainly preliminary and tentative. But they suggest that we can relate productivity growth to relatively current as well as lagged ICT investment in the cross section. Given that we are running an OLS regression, we cannot, of course, necessarily infer causation from the results. But the results are broadly consistent with the notion that ICT investments affect measured productivity growth with a long (and possibly variable) lag. Contemporaneously, they are correlated with a lot of diverted resources towards unmeasured complementary investment, and hence—once one controls for lagged growth rates—they are negatively correlated with output.

5. Conclusions

Even though ICT seems to be the major locus of innovation in the past decade, the TFP acceleration in the United States since the mid-1990s has been broad-based. We reconcile these observations by emphasizing the role of the complementary investments and innovations that ICT induces

21. A better way to do so would be a panel regression, since the model is inherently driven by time-series considerations. The panel setup would impose, however, that the dynamics were constant over time, which might not be the case. In any case, we have so far done only preliminary explorations in the panel dimension.

22. One standard statistical test is to look at the diagonal of the "hat" matrix, $X'(X'X)^{-1}X'$. For a regression with k coefficients and n observations, Belsley, Kuh, and Welsch (1980) identify influential observations as those where the diagonal element of the hat matrix exceeds $2k/n$.

in firms that use it. We thus link the literature on ICT as a general purpose technology with the literature on intangible capital. To the extent that there is unmeasured intangible output and unmeasured intangible capital, conventional TFP growth is a biased measure of true technical change. This GPT view suggests that productivity slowdowns and speedups might reflect the dynamics associated with complementary investment.

A fundamental difficulty, of course, is that complementary investment and capital are unmeasured. We present a simple theoretical framework in which observed ICT capital intensity and growth should serve as reasonable proxies. In line with this GPT view, the U.S. industry data suggest that ICT capital growth is associated with industry TFP accelerations with long lags of 5 to 15 years. Indeed, controlling for past growth in ICT capital, contemporaneous growth in ICT capital is negatively associated with the recent TFP acceleration across industries. More work remains to be done to explore the robustness of the theoretical framework—for example, allowing for different production functions for intangible investment and for final output—and for extending the empirical work. For example, we have not exploited the panel nature of the theory, nor have we explored the importance of the relative price of intangible capital. But we are encouraged by the preliminary results that link aggregate and industry-level U.S. TFP performance in the 2000s to both the persuasive macro models of GPTs and to the stimulating micro empirical work that supports the GPT hypothesis.

Appendix

Further Description of Industries[23]

The BEA publishes GDP-by-industry data for 61 private industries shown in the first column of the appendix table. To focus on private nonfarm industries, we first remove Farms (NAICS 111–112) and Forestry, Fishing, and Related Activities (NAICS 113–115).

Next, we sum up the BEA data for Motor Vehicles, Bodies and Trailers, and Parts (NAICS 3361–3363) and Other Transportation Equipment (NAICS 3364–3366, 3369) because the BLS capital data that we use has data only on their aggregate.

When we calculated TFP by disaggregated industry, we discovered that many subcategories of larger service industries had quite large fluctuations from year-to-year—with offsetting fluctuations in other, closely related subcategories. This could reflect differences in NAICS industry classifications across our data sources. We chain-aggregated these industries into broader aggregates, as indicated in the table. This 40-industry data set comprises three ICT-producing industries and 37 non-ICT-producing industries.

The table also shows other classifications used in the paper.

23. David Thipphavong contributed to the writing of this appendix.

TABLE A1
INDUSTRY LISTS USED IN SUBAGGREGATES AND REGRESSIONS

Industry name	NAICS 1997 Code	ICT prod.	Non-ICT prod.	Well-measured	Poorly measured	Mfg.	Non-mfg.
Farms	111-112						
Forestry, Fishing, and Related Activities	113-115						
Oil and Gas Extraction	211		X	X			X
Mining, except Oil and Gas	212		X	X			X
Support Activities for Mining	213		X	X			X
Utilities	22		X	X			X
Construction	23		X		X		X
Wood Products	321		X	X		X	
Nonmetallic Mineral Products	327		X	X		X	
Primary Metal Products	331		X	X		X	
Fabricated Metal Products	332		X	X		X	
Machinery	333	X					
Computer and Electronic Products	334	X					
Electrical Equipment, Appliances, and Components	335	X					
Transportation Equipment	336		X	X		X	
Motor Vehicles, Bodies and Trailers, and Parts	*3361-3363*						
Other Transportation Equipment	*3364-3366, 3369*						
Furniture and Related Products	337		X	X		X	
Miscellaneous Manufacturing	339		X	X		X	
Food, Beverage, and Tobacco Products	311-312		X	X		X	
Textile Mills and Textile Product Mills	313-314		X	X		X	
Apparel and Leather and Allied Products	315-316		X	X		X	
Paper Products	322		X	X		X	
Printing and Related Support Activities	323		X	X		X	
Petroleum and Coal Products	324		X	X		X	
Chemical Products	325		X	X		X	
Plastic and Rubber Products	326		X	X		X	
Wholesale Trade	42		X	X			X
Retail Trade	44, 45		X	X			X
Air Transportation	481		X	X			X
Rail Transportation	482		X	X			X
Water Transportation	483		X	X			X
Truck Transportation	484		X	X			X
Transit and Ground Passenger Transportation	485		X	X			X
Pipeline Transportation	486		X	X			X
Other Transportation and Support Activities	487, 488, 492		X	X			X
Warehousing and Storage	493		X	X			X
Information	51		X		X		X
Publishing Industries	*511*						
Motion Picture and Sound Recording Industries	*512*						
Broadcasting and Telecommunications	*513*						
Information and Data Processing Services	*514*						
Finance and Insurance	52		X		X		X
Fed. Res. Banks, Credit Intermed., and Related Activities	*521-522*						
Securities, Commodity Contracts, and Investments	*523*						
Insurance Carriers and Related Activities	*524*						
Funds, Trusts, and Other Financial Vehicles	*525*						
Real Estate and Rental and Leasing	53		X		X		X
Real Estate	*531*						
Rental and Leasing Svcs. and Lessors of Intang. Assets	*532-533*						
Professional and Business Services	54-56		X		X		X
Legal Services	*5411*						
Computer Systems Design and Related Services	*5415*						
Misc. Professional, Scientific, and Technical Services	*5412-5414, 5416-5419*						
Management of Companies and Enterprises	*55*						
Administrative and Support Services	*561*						
Waste Management and Remediation Services	*562*						
Educational Services, Health Care, and Social Assistance	61-62		X		X		X
Educational Services	*61*						
Ambulatory Health Care Services	*621*						
Hospitals and Nursing and Residential Care Facilities	*622-623*						
Social Assistance	*624*						
Arts, Entertainment, Recreation, Accommodation, and Food Services	71-72		X		X		X
Perf. Arts, Spectator Sports, Museums, and Related Activities	*711-712*						
Amusements, Gambling, and Recreation Industries	*713*						
Accommodation	*721*						
Food Services and Drinking Places	*722*						
Other Services, except Government	81		X		X		X

References

Bakhshi, Hasan, Nick Oulton, and Jamie Thompson. 2003. "Modeling Investment When Relative Prices Are Trending: Theory and Evidence for the United Kingdom." Bank of England Working Paper 189.

Basu, Susanto, and John G. Fernald. 2001. "Why Is Productivity Procyclical? Why Do We Care?" In *New Developments in Productivity Analysis,* eds. C. Hulten, E. Dean, and M. Harper. Cambridge, MA: National Bureau of Economic Research.

Basu, Susanto, John G. Fernald, and Miles Kimball. 2006. "Are Technology Improvements Contractionary?" *American Economic Review* 96(5) pp. 1,418–1,448.

Basu, Susanto, John G. Fernald, Nicholas Oulton, and Sylaja Srinivasan. 2003. "The Case of the Missing Productivity Growth." In *NBER Macroeconomics Annual 2003,* eds. M. Gertler and K. Rogoff. Cambridge, MA: MIT Press.

Basu, Susanto, John. G. Fernald, and Matthew D. Shapiro. 2001. "Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?" *Carnegie-Rochester Conference Series on Public Policy* 55, pp. 117–165.

Basu, Susanto, and Miles Kimball. 1997. "Cyclical Productivity with Unobserved Input Variation." NBER Working Paper 5915 (December).

Belsley, David A., Edwin Kuh, and Roy E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: John Wiley and Sons.

Bloom, Nicholas, Raffaella Sadun, and John Michael Van Reenen. 2005. "It Ain't What You Do It's the Way That You Do I.T.—Testing Explanations of Productivity Growth Using U.S. Transplants." Unpublished manuscript, London School of Economics.

Bosworth, Barry P., and Jack E. Triplett. 2007. "The 21st Century Productivity Expansion Still in Services?" *International Productivity Monitor* 14 (Spring) pp. 3–19.

Bresnahan, Timothy F. Undated. "The Mechanisms of Information Technology's Contribution to Economic Growth." Prepared for presentation at the Saint-Gobain Centre for Economic Research.

Bresnahan, Timothy, Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117(1) pp. 339–376.

Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. "General Purpose Technologies: 'Engines of Growth'?" *Journal of Econometrics* 65(1) pp. 83–108.

Brynjolfsson, Erik, and Lorin M. Hitt. 2000. "Beyond Computation: Information Technology, Organizational Transformation, and Business Performance." *Journal of Economic Perspectives* 14(4) pp. 23–48.

Brynjolfsson, Erik, and Lorin M. Hitt. 2003. "Computing Productivity: Firm-Level Evidence." *Review of Economics and Statistics* 85(4) (November), pp. 793–808.

Brynjolfsson, Erik, and Shinkyu Yang. 2002. "Intangible Assets and Growth Accounting: Evidence from Computer Investments." Manuscript.

Caselli, Francesco. 1999. "Technological Revolutions." *American Economic Review* 89(1) (March) pp. 78–102.

Chandler, Alfred Dupont, Jr. 1977. *The Visible Hand.* Cambridge, MA: Harvard University Press.

Corrado, Carol, Charles Hulten, and Daniel Sichel. 2006. "Intangible Capital and Economic Growth." Finance and Economics Discussion Series 2006-24 (April). Board of Governors of the Federal Reserve System. http://www.federalreserve.gov/pubs/feds/2006/200624/200624abs.html

Corrado, Carol, Paul Lengermann, Eric J. Bartelsman, and J. Joseph Beaulieu. 2006. "Modeling Aggregate Productivity at a Disaggregate Level: New Results for U.S. Sectors and Industries." Presented at the 2006 Summer Institute sponsored by the NBER and the Conference on Research in Income and Wealth, Cambridge, MA, July 17.

David, Paul A., and Gavin Wright. 2003. "General Purpose Technologies and Surges in Productivity: Historical Reflections on the Future of the ICT Revolution." In *The Economic Future in Historical Perspective,* eds. P.A. David and M. Thomas. Oxford: Oxford University Press.

Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. "1974." *Carnegie-Rochester Conference Series on Public Policy* 46 (June), pp. 49–95.

Griliches, Zvi. 1994. "Productivity, R&D, and the Data Constraint." *American Economic Review* 84(1) pp. 1–23.

Hall, Robert E. 2001. "The Stock Market and Capital Accumulation." *American Economic Review* 91(December) pp. 1,185–1,202.

Helpman, Elhanan (ed). 1998. *General Purpose Technologies and Economic Growth.* Cambridge, MA: MIT Press, 1998.

Helpman, Elhanan, and Manuel Trajtenberg. 1998. "Diffusion of General Purpose Technologies." In *General Purpose Technologies and Economic Growth,* ed. E. Helpman. Cambridge, MA: MIT Press.

Hobijn, Bart, and Boyan Jovanovic. 2001. "The Information-Technology Revolution and the Stock Market: Evidence." *American Economic Review* 91(December) pp. 1,203–1,220.

Hornstein, Andreas, and Per Krusell. 1996. "Can Technology Improvements Cause Productivity Slowdowns?" In *NBER Macroeconomics Annual,* eds. B. Bernanke and J. Rotemberg. Cambridge, MA: National Bureau of Economic Research.

Howitt, Peter. 1998. "Measurement, Obsolescence, and General Purpose Technologies." In *General Purpose Technologies and Economic Growth,* ed. E. Helpman. Cambridge and London: MIT Press.

Hulten, Charles. 1978. "Growth Accounting with Intermediate Inputs." *Review of Economic Studies* 45 (October) pp. 511–518.

Jorgenson, Dale W. 2001. "Information Technology and the U.S. Economy," *American Economic Review* 91 (March) pp. 1–32.

Jorgenson, Dale W., Mun Ho, and Kevin Stiroh. 2006. "The Sources of the Second Surge of U.S. Productivity and Implications for the Future." Manuscript.

Jovanovic, Boyan, and Peter L. Rousseau. 2004. "Mergers as Reallocation." Unpublished manuscript, New York University (June).

Krueger, Dirk, and Krishna B. Kumar. 2003. "U.S.-Europe Differences in Technology Adoption and Growth: The Role of Education and Other Policies." Manuscript prepared for the *Carnegie-Rochester Conference,* April.

Laitner, John, and Dmitriy Stolyarov. 2003. "Technological Change and the Stock Market." *American Economic Review* 93(4) (September) pp. 1,240–1,267.

Lynch, Lisa M., and Stephen Nickell. 2001. "Rising Productivity and Falling Unemployment: Can the U.S. Experience Be Sustained and Replicated?" In *The Roaring Nineties,* eds. A. Krueger and R. Solow. New York: Russell Sage Foundation.

Nordhaus, William D. 2002. "Productivity Growth and the New Economy." *Brookings Papers on Economic Activity* 2.

Oliner, Stephen D., and Daniel E. Sichel. 2000. "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives* 14(4) (Fall) pp. 3–22.

Oliner, Stephen D., and Daniel E. Sichel. 2006. Unpublished update to Oliner and Sichel (2000). September.

Oulton, Nicholas. 2001. "Must the Growth Rate Decline? Baumol's Unbalanced Growth Revisited." *Oxford Economic Papers* 53(4) (October) pp. 605–627.

Oulton, Nicholas, and Sylaja Srinivasan. 2005. "Productivity Growth and the Role of ICT in the United Kingdom: An Industry View, 1970–2000." Centre for Economic Performance discussion paper 0681, London School of Economics.

Shapiro, Matthew D. 1986. "The Dynamic Demand for Capital and Labor." *Quarterly Journal of Economics* 101(3) (August) pp. 513–542.

Stiroh, Kevin J. 2002a. "Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?" *American Economic Review* 92(5) pp. 1,559–1,576.

Stiroh, Kevin J. 2002b. "Are ICT Spillovers Driving the New Economy?" *Review of Income and Wealth* 48(1) (March) pp. 33–57.

Stiroh, Kevin J., and Matthew Botsch. 2007. "Information Technology and Productivity Growth in the 2000s." *German Economic Review* 8(2) (May) pp. 255–280.

Tevlin, Stacey, and Karl Whelan. 2003. "Explaining the Investment Boom of the 1990s." *Journal of Money, Credit, and Banking* 35(1) pp. 1–22.

Triplett, Jack E., and Barry P. Bosworth. 2006. "'Baumol's Disease' Has Been Cured: IT and Multifactor Productivity in U.S. Services Industries." In *The New Economy and Beyond: Past, Present, and Future,* ed. Dennis W. Jansen. London: Edward Elgar Publishing.

Wang, J. Christiana, Susanto Basu, and John G. Fernald. 2004. "A General-Equilibrium Asset-Pricing Approach to the Measurement of Nominal and Real Bank Output." FRB Boston Working Paper 04-7. http://www.bos.frb.org/economic/wp/wp2004/wp047.pdf

The Role of Relative Performance in Bank Closure Decisions*

Kenneth Kasa

Professor
Simon Fraser University

Mark M. Spiegel

Vice President
Federal Reserve Bank of San Francisco

This paper studies a banking industry subject to common and idiosyncratic shocks. We compare two types of regulatory closure rules: (1) an "absolute closure rule," which closes banks when their asset–liability ratios fall below a given threshold, and (2) a "relative closure rule," which closes banks when their asset–liability ratios fall sufficiently below the industry average. There are two main results: First, relative closure rules imply forbearance during "bad times," defined as adverse realizations of the common shock. This forbearance occurs for incentive reasons, not because of irreversibilities or political economy considerations. Second, relative closure rules are less costly to taxpayers, and these savings increase with the relative variance of the common shock. To evaluate the model, we estimate a panel-logit regression using a sample of U.S. commercial banks. We find strong evidence that U.S. bank closures are based on relative performance. Individual and average asset–liability ratios are both significant predictors of bank closure.

1. Introduction

A critical component of any bank regulatory package concerns the timing of bank closures, i.e., when exactly should regulators close or forcibly merge a bank? While other policies, like auditing, capital requirements, and deposit insurance are designed to contain risks, nevertheless bank failures do occur. By the principle of backward induction, how and under what circumstances banks expect closure has important consequences for how they behave before they are closed. An efficient closure policy should account for these incentive effects.

The goal of this paper is to study the incentive effects of bank closure policy. In doing so, we abstract from all other aspects of bank regulation, not only for analytical convenience but also because our goal is rather modest. We do not attempt to formulate a set of incentive-compatible policies that implement an informationally constrained Pareto optimum. We merely want to compare the cost effectiveness, in terms of expected taxpayer liability, of alternative closure rules. We can do this without taking a stand on exactly what banks do, or should be doing. Instead, we just consider two general types of rules which are simple, transparent, and pragmatic. Both rules are designed to elicit the same response (i.e., risk) by banks. Given this, we can then ask the following question: For any desired level of bank risk, which closure rule is less costly?[1]

There are two key inputs to our analysis. The first is the assumption that banks are subject to both common and idiosyncratic shocks. Interest rate fluctuations are one example of a common banking shock. The second key input is the assumption that regulators are unable to monitor bank portfolio decisions perfectly.

Since bank actions are unobserved, closure policy must be based on ex post realized outcomes. This confronts the regulator with a signal extraction problem. For incentive reasons, an efficient policy should attempt to distinguish between banks that are in trouble as a result of their own actions (i.e., moral hazard), and banks that were simply unlucky. While a policy of "prompt corrective action" can indeed discourage moral hazard and save taxpayer money, it can also cause banks to be unduly cautious in the presence of idiosyncratic shocks. Alternatively, separating moral hazard from bad luck

*We thank Simon Kwan and Mark Levonian for helpful comments. Laura Haworth and Hiroshi Kokame provided excellent research assistance. Opinions expressed do not necessarily reflect the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System.

1. There are studies that explore the interaction between bank closure policy and other policy instruments. For example, Acharya and Dreyfus (1989) study the potential complementarities between deposit insurance pricing and bank closure policy. However, they assume symmetric information and focus their analysis on dynamics and timing issues, while we focus on moral hazard and incentives.

can achieve the same overall level of banking industry risk at a lower (expected) cost to the taxpayer.[2]

We show that the key to separating moral hazard from bad luck is to base closure decisions on relative performance. With a large number of ex ante identical banks, relative performance is a good indicator of relative "effort."[3] Consequently, a rule that closes banks whenever their asset–liability ratios fall below the average of a cross-section of comparable banks by a given amount is superior to one based solely on each individual bank's asset–liability ratio.[4] An interesting implication of a relative closure rule is that it leads to forbearance during "bad times," defined as adverse realizations of the common shock. It is important to realize, however, that this forbearance occurs solely for ex ante incentive reasons, not because of irreversibilities or political economy considerations.[5]

In fact, if he could, our regulator would like to renege ex post on the announced relative closure rule once the bank's portfolio decision was made. The regulator would like to close unlucky banks in order to keep them from "gambling for resurrection." However, this finite-horizon time consistency problem would be mitigated in a repeated framework. In a repeated game, regulators would have an incentive to follow through on their announced closure policies if failing to do so would lead to loss of future credibility.[6] Moreover, legal and institutional constraints (e.g., the Federal Deposit Insurance Corporation Improvement Act of 1991 or FDICIA), even when they contain generous opt-out agreements, undoubtedly provide some degree of commitment.

We are not the first to point out the potential incentive benefits of a relative closure rule for moral hazard reasons. Nagarajan and Sealey (1995) also make this point. Our contribution is to formulate the problem in a way that leads to empirically testable predictions. We do this by explicitly modeling a large number of banks subject to continuously distributed shocks and by basing our closure rule on the cross-sectional average asset–liability ratio, as opposed to some notion of a "market return." One way to think of the difference between these two approaches is that our regulator is more concerned with accounting information, while the regulator in Nagarajan and Sealey's paper is more concerned with stock prices.[7]

Other arguments against constant regulatory rules have been made in the literature. Blum and Hellwig (1995) argue that capital adequacy regulations can reinforce macroeconomic fluctuations. In their model, economic downturns reduce the quality of bank balance sheets. Because their regulatory constraints are now more binding, banks respond by reducing their lending, exacerbating the economic downturn. A relative closure rule would mitigate this problem by easing the regulatory constraints faced by banks in aggregate downturns.

Other arguments have been made on the basis of reducing resolution costs. Acharya and Yorulmazer (2007, 2008) generate countercyclical forbearance based on the ex post cost of resolution of failed banks. In their model, acquisition of failed banks by surviving banks is posited to result in lower resolution costs than liquidating the failed bank's assets. As it becomes more difficult to find surviving banks to acquire failed banks in environments with higher numbers of bank failures, they find that the cost of closing a failed bank increases with the number of bank failures, a phenomenon they describe as "too-many-to-fail."

Arguments have also been made that there may be "contagion" in failures across the banking system. Diamond and Rajan (2005) describe an environment in which bank closures may lead to adverse spillovers as they reduce the liquidity of the banking system. Their model also suggests that proper time-varying regulatory rules can mitigate the costs of resolving bank failures. Lang and Stulz (1992) demonstrate that bankruptcy announcements lead to adverse valuations of industry competitors, suggesting that closures of failed banks could lead to reduced confidence in other banks in the system.

More generally, our results relate to the literature on optimal bankruptcy procedures and the evaluation of Chapter 11 proceedings. For example, Baird (1991) argues that Chapter 11 protection can encourage managers to initiate bankruptcy procedures, and Mooradian (1994) argues that Chapter 11 protection may serve as a mechanism for achieving a sepa-

2. One issue that has attracted attention in the literature is the limits to government information. Some have argued (e.g., Berger, Davies, and Flannery 2000, Flannery 1998) that private sector information could be superior in certain situations. To the extent that this is true, regulators would also want to incorporate the private information concerning relative performance in the manner modeled below.

3. The advantages of relative performance contracts were first discussed in the labor literature. See, e.g., Lazear and Rosen (1981) or Nalebuff and Stiglitz (1983). It should be noted that, while our model presumes ex ante identical banks, our empirical work attempts to control for one potentially important source of heterogeneity, namely, size.

4. One might wonder whether such a rule would be consistent with the dictates of FDICIA, which calls for "prompt corrective action" and contains no explicit reference to relative performance. However, as discussed in more detail by Mailath and Mester (1994), FDICIA also directs regulators to resolve troubled banks in the *least costly way* and grants regulators a large degree of discretion in deciding how to do this.

5. Kane (1989) discusses forbearance based on regulatory malfeasance. Demirgüç-Kunt (1991) and Fries, Mella-Barral, and Perraudin (1997) analyze forbearance based on irreversibility and the resulting option value of waiting. Boot and Thakor (1993) base forbearance on a principal–agent problem between the bank regulator and the taxpayer.

6. See Mailath and Mester (1994) for a detailed analysis of the time consistency problem in bank closure policy.

7. In a later paper, Nagarajan and Sealey (1998) extend their framework to a setting of adverse selection as well as moral hazard, although in it they only focus on the pricing of deposit insurance, not on bank closure policy.

rating equilibrium, by making it prohibitively costly for inefficient firms to mimic efficient ones in debt restructurings, while Aghion, Bolton, and Fries (1999) argue that strict bank closure policies can discourage managers from truthfully disclosing their bank's asset position.[8]

There is relatively little empirical evidence concerning whether time-varying closure rules have been used by regulators. Thomson (1991) examines the determinants of bank closure during the 1980s. He finds that various measures of macroeconomic conditions are usually significant predictors of bank failure. Brown and Dinç (2006), examine bank failures among 21 emerging market economies in the 1990s. Their results suggest that conditions in the rest of the banking sector are incorporated into the regulatory closure decision in a manner consistent with the predictions of the model in this paper.

Below, we estimate a panel-logit regression using a sample of annual data for over 12,000 U.S. commercial banks during the period 1992 through 1997, which corresponds to the post-FDICIA period in which failures were relatively prevalent. We find strong evidence that U.S. bank closures are based on relative performance. Our results demonstrate that both individual and average asset–liability ratios for a bank's home state are significant predictors of bank closure, and the coefficient estimates are consistent with this theory. Moreover, the results are robust to the exclusion of small banks from the sample, as well as to the inclusion of other controlling variables. Overall, we conclude that relative performance has been an input to bank closure decisions.

Our paper is organized as follows. Section 2 discusses our model. Section 3 describes the controlling variables we use and our results. Section 4 concludes. An appendix provides the proofs of our propositions.

2. A Simple Model of Bank Closure

2.1. Bank Investment Decision

We begin with a very simple model of bank closure. We assume that there are an infinite number of homogenous banks of measure zero. We model a representative bank i, which decides the amount of "effort," μ_i, to invest in enhancing the quality of its asset portfolio, where possible chosen values of μ range between the positive finite values $\min(\mu)$ and $\max(\mu)$.[9] The cost of supplying an amount of effort equal to μ is assumed to satisfy the function $V(\mu)$, where $V_\mu > 0$ and $V_{\mu\mu} > 0$. For simplicity, we assume that effort costs are borne up front. This simplifies the analysis by making this cost independent of the probability of bankruptcy, but this assumption drives none of our results. Finally, we assume that the optimal effort choice μ^* is always interior.

There are two shocks: a common shock, θ, which affects all banks, and an idiosyncratic shock, ε_i, which falls on bank i alone.[10] We assume that ε_i and θ are distributed on the intervals $\left[\underline{\varepsilon}, \overline{\varepsilon}\right]$ and $\left[\underline{\theta}, \overline{\theta}\right]$, respectively.

The model has one period, although our analysis extends to the repeated case if shocks are independently and identically distributed (i.i.d.). The timing of the model is as follows: First, the regulator announces a closure rule. Next, the bank chooses its effort level, μ_i. At the end of the period, the shocks are realized and the value of bank assets minus liabilities are determined, which we define as A_i. We assume that A_i satisfies

$$A_i = \mu_i + \theta + \varepsilon_i. \tag{1}$$

Finally, the regulator makes its closure decision consistent with its announced rule.

To make the model interesting, we assume the regulator only observes the total value of A_i, not the values of its components. We therefore limit the regulator to closure rules conditional on A_i. Moreover, as we noted in the introduction, we assume that the regulator can commit to a closure rule. Later we discuss the implications of constraining the regulator to time-consistent rules.

Banks are assumed to have limited liability, having zero value under bankruptcy. As in Marcus (1984), we assume that if the bank is allowed to continue, it has a charter value. We allow the charter value, $C(\mu_i)$ to be increasing in current bank effort. The charter value represents the expected future profits from continued banking operations.

Define ε^* as the minimum realization of ε_i under which the regulator chooses to allow the bank to continue in operation. Clearly, ε^* will depend on the regulator's closure rule. Because regulators are constrained to follow closure rules based on A_i, their observable indicator of bank financial health, ε^* will in practice be the level of ε_i that yields the minimum value of A_i that does not result in closure. For now,

8. The literature on the merits of Chapter 11 proceedings is mixed. Strict closure rules can provide managerial discipline, so that Chapter 11 protection of borrowers may exacerbate moral hazard problems (Gertner and Scharfstein 1991, Weiss and Wruck 1998). However, Chapter 11 protection may yield benefits when contracting technology is limited (Aghion and Bolton 1992, Hart and Moore 1998, Berkovich, Israel, and Zender 1998, and Harris and Raviv 1995).

9. Similar frameworks for studying bank regulation can be found in Dewatripont and Tirole (1993) and Giammarino, Lewis, and Sappington (1993).

10. We do not model the lending choices of banks directly. In particular, we do not allow banks to manipulate the relative importance of common and idiosyncratic shocks to their portfolios. The literature has expressed some concern that banks will respond to countercyclical forbearance policies by adjusting their portfolios to give less weight to idiosyncratic shocks (e.g., Goodhart, et al. 1998 and Acharya and Yorulmazer 2007).

we note that for all the closure rules we entertain below, ε^* is a decreasing function of both bank effort μ_i and the common shock θ, since A_i is increasing in both of these arguments.

The representative bank's investment decision is to choose μ_i to maximize expected bank value net of effort cost, which is

$$(2) \qquad \int_{\underline{\theta}}^{\overline{\theta}} \left[\int_{\varepsilon^*}^{\overline{\varepsilon}} \left[A_i + C(\mu_i) \right] f(\varepsilon) d\varepsilon \right] g(\theta) d\theta - V(\mu),$$

where $f(\cdot)$ is the density of ε and $g(\cdot)$ is the density of θ.

The bank's first-order condition satisfies

$$(3) \qquad \int_{\underline{\theta}}^{\overline{\theta}} \left[\int_{\varepsilon^*}^{\overline{\varepsilon}} (1 + C_\mu) f(\varepsilon) d\varepsilon - \left(\frac{\partial \varepsilon^*}{\partial \mu_i} \right) (\mu_i + \theta + \varepsilon^* + C) f(\varepsilon^*) \right] g(\theta) d\theta = V_\mu.$$

The two arguments on the left-hand side of equation (3) represent the marginal benefits of additional effort. The first term reflects the increased expected payoff in nonbankruptcy states, holding the probability of bankruptcy constant. The second term reflects the value of the change in the probability of bankruptcy that results from a marginal change in effort.

2.2. Case 1: Regulatory Standard Based on Absolute Performance

We first consider a closure rule based solely on absolute bank performance. Suppose that a bank is closed if

$$(4) \qquad A_i \leq m,$$

where $m = 0$ is obviously a special case where banks are closed on insolvency.

Define μ_i^p as the privately optimal choice of effort made by an individual bank given the absolute closure rule. Alternatively, define μ_i^s as the socially optimal level of effort, the one that maximizes the expected stream of revenues from the bank plus bank charter value, net of effort costs, and includes the expected regulatory liabilities under insolvency. We show in the appendix that $\mu_i^p \leq \mu_i^s$. This is our first result:

PROPOSITION 1: *With a closure rule based on insolvency, the level of privately chosen bank effort is below that consistent with maximizing the total "social revenue stream."*

Proposition 1 derives from the standard moral hazard result with limited liability: Since its losses are bounded from below, the private bank chooses a lower level of effort because it does not share in the gains to returns in bankruptcy states. These instead accrue to the regulator as a reduction in tax liabilities.

2.3. Case 2: Regulatory Standard Based on Relative Performance

Next, we assume the regulator bases closure on relative performance. Before showing how this can enhance efficiency, we should emphasize that for simplicity we allow the bank to alter the mean of its net asset position, but not its variance. If banks can also (independently) choose the variance of their net asset positions then relative performance schemes can produce bad equilibria, in which contestants choose very risky actions and low effort levels.

For example, Hvide (2002) shows that if effort is costly but risk-taking is not, then it will be optimal for the contestants to choose maximally risky outcome distributions with very low effort. Intuitively, injecting a lot of variance into the outcome reduces the marginal benefit of effort, since the noise is so great, which then permits low effort levels to be sustained in equilibrium. This is optimal if effort is costly but risk-taking isn't. Alternatively, the literature has expressed some concern that banks will respond to tournaments by adjusting their portfolios to give less weight to idiosyncratic shocks (e.g., Goodhart, et al. 1998). In both of these cases, the importance of adjustment of overall variance or the weight on idiosyncratic vs. common shocks will be dependent on the costliness to the bank of making such an adjustment. When such adjustments are costly, it is unlikely that much adjustment would take place in an uncoordinated environment.

We assume that there are a large number of banks, so that the law of large numbers yields,

$$(5) \qquad \theta = \overline{A} - \overline{\mu},$$

where $\overline{A}$ and $\overline{\mu}$ are the cross-sectional average levels of bank asset positions and efforts, respectively. By equations (1) and (5), and since $\mathrm{E}(\varepsilon_i) = 0$,

$$(6) \qquad \mathrm{E}(\mu_i - \overline{\mu}) = \mathrm{A}_i - \overline{\mathrm{A}}.$$

By incorporating relative performance, then, the regulator can infer relative effort. We therefore posit a relative closure rule which satisfies[11]

$$(7) \qquad A_i - \overline{A} \leq n,$$

where $\underline{\varepsilon} \leq n \leq \overline{\varepsilon}$.

We solve for the equilibrium level of bank effort under this closure rule in the appendix.

Note that our model does not allow banks to manipulate the relative importance of common and idiosyncratic shocks to their portfolios. While the literature has expressed some

11. The fact that the benchmark can be taken as the mean, rather than some more general weighting, depends on our homogeneity assumption. If banks differed by size or idiosyncratic risk, then it would no longer be optimal to use the mean as a benchmark. See, e.g., Holmstrom (1982, p. 337).

concern that banks will respond to tournaments by adjusting their portfolios to give less weight to idiosyncratic shocks (e.g., Goodhart, et al. 1998), it is difficult to see how such coordinated action could be sustained in a competitive banking system.

To allow for a common basis of comparison, we first find the relative closure rule which elicits the same level of effort as the absolute closure rule. We then compare the expected liability of the regulatory institution under the two closure rules. We designate as preferable the rule that delivers a given level of bank effort with the lowest expected regulatory liability.

To obtain analytic solutions for the regulator's expected liability, we must put more structure on the distribution of ε_i. Accordingly, without essential loss of generality we assume from here on that ε_i is distributed uniformly on the interval $\left[\underline{\varepsilon},\overline{\varepsilon}\right]$ with density f.

Define A^m and A^n as the minimum realizations of A_i necessary to avoid closure under the absolute and relative closure rules, respectively. We demonstrate in the appendix that

$$A^n - A^m = \theta - \mathrm{E}(\theta). \tag{8}$$

This leads to our second result:

PROPOSITION 2: *For a given level of bank effort, closure takes place at higher (lower) levels of A_i under the relative closure rule than under the absolute closure rule when θ exceeds (falls short of) its expected value.*

Intuitively, the proposition states that the relative closure rule will be more stringent in good times, i.e., when the common shock θ is above its mean, and more lenient in bad times.

Note that the implied "forbearance" has nothing to do with the opportunity cost of irreversibly shutting down banks, or with regulatory malfeasance. Rather, forbearance is advantageous here solely for ex ante incentive reasons. Basing closure on relative performance allows the regulator to more accurately separate those banks choosing low effort levels from those banks that are unlucky. If a bank knows its effort level is likely to be detected and incorporated in the regulator's closure decision, it will choose a higher level of effort.

Finally, we turn to the relative liability of the bank regulator. Define L_m as the expected liability of the regulatory institution under the absolute closure policy that elicits level of effort $\widehat{\mu}$. L_m satisfies[12]

$$L_m = -\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\varepsilon}}^{\varepsilon^*} A_i(\widehat{\mu},\theta,\varepsilon_i) f(\varepsilon_i) d\varepsilon_i g(\theta) d\theta. \tag{9}$$

12. Note that we do not consider the loss of bank charter value as part of the closure cost. The inclusion of charter loss would not change the results systematically with either closure rule.

Define L_n as the expected liability of the regulatory institution under the relative closure policy that elicits the same level of effort $(\widehat{\mu})$. Substituting for ε^* as above, L_n satisfies

$$L_n = -\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\varepsilon}}^{n} A_i(\widehat{\mu},\theta,\varepsilon_i) f(\varepsilon_i) d\varepsilon_i g(\theta) d\theta. \tag{10}$$

Assuming that ε_i is distributed uniformly, we demonstrate in the appendix that

$$L_m - L_n = \frac{1}{2}\left[\frac{Var(\theta)}{\overline{\varepsilon} - \underline{\varepsilon}}\right]. \tag{11}$$

This leads to our third result:

PROPOSITION 3: *For closure rules that elicit the same level of bank effort, the relative closure rule has a smaller expected liability to the bank regulator than the absolute closure rule. Moreover, the cost advantage of the relative closure rule is increasing in the variance of the common shock and decreasing in the variance of the idiosyncratic shock.*

Note that this cost advantage implies that sustaining a relative closure rule is valuable to the regulator. In a repeated context, a standard trigger strategy argument can be used to show that concerns about losing these benefits in the future can induce a discretionary regulator to comply with the relative closure rule ex post.

3. Empirical Results

3.1. Estimation Method

In this section, we investigate whether relative performance matters for bank closure decisions in the United States. Based on our theoretical model above, we formulate a binary choice model in which the regulator chooses at each point in time either failure or continuation of operations.

The definitions and sources for all variables used in this study are listed in Table 1. We represent the regulator's binary choice as a random variable *FAIL*, which takes the value one if the regulator chooses failure and the value zero if the bank is allowed to continue. Failure is defined as the end of a bank's existence whose resolution is arranged by the FDIC or other regulatory agency.

Our base specifications come directly from the theoretical model above. As a benchmark for relative performance, we consider the average financial position of banks in the home state of bank i in period t, $\overline{A}_{jt}$, where j denotes the home state of bank i. The base absolute closure rule then specifies closure decisions as depending solely on a bank's current asset position, A_{ijt}, while our base relative closure rule specifies that closure decisions also depend on the average financial position of banks in the home state of bank i in period t, $\overline{A}_{jt}$.

Table 1
Variable Definitions

FAIL	Binary variable which takes the value 1 when a bank fails and the value 0 when a bank is allowed to continue. Failure occurs when an entity ceases to exist and its resolution was arranged by the Federal Deposit Insurance Corporation, Resolution Trust Corporation, National Credit Union Administration, or state or other regulatory agency.
A_{ijt}	Book value of total assets divided by book value of total liabilities of bank i in state j at time t. Total assets exclude loan loss reserves. Total liabilities exclude subordinated debt.
$\overline{A}_{jt}$	Average value of A_{ijt} for all entities in a state in a given year.
SIZE	Book value of total assets excluding loan loss reserves.
COMMERCIAL	Commercial and industrial loans divided by total assets.
AGRICULTURAL	Loans to finance agricultural production and other loans to farmers divided by total assets.
REALESTATE	Loans secured by real estate divided by total assets.
NON-INTEREST	Total non-interest expense divided by total assets.
CD	Total time deposits of $100,000 or more divided by total assets.
90DAYSLATE	Total loans and lease financing receivables: 90 days or more past due and still accruing divided by total assets.

Source: All data are from the Federal Reserve Bank of Chicago's Bank Condition and Income Database.

In addition to the base specifications, we add a number of conditioning variables commonly used in the literature to forecast bank closures (e.g., Wheelock and Wilson 2000). First, we introduce a variable to measure relative bank size. $SIZE_{ijt}$ is proxied by the book value of bank i in state j in period t. It is widely believed that regulators might be more hesitant to close large banks in poor financial condition because of the potential for adverse systemic results from large bank closures. Second, we introduce three proxies for sectoral exposure: $COMMERCIAL_{ijt}$ represents the share of commercial and industrial loans to total assets. $AGRICULTURE_{ijt}$ represents the share of agricultural loans as a share of total assets. $REALESTATE_{ijt}$ represents the share of total assets in the real estate sector. We also introduce $NON\text{-}INTEREST_{ijt}$, the ratio of non-interest expenses to total assets, as an indicator of bank efficiency. As an indicator of the composition of bank liabilities, we introduce CD_{ijt}, the ratio of time deposits exceeding $100,000 as a share of total assets. This indicates the share of uninsured deposits. Finally, as an indicator of asset quality, we introduce $90DAYSLATE_{ijt}$, total loans and receivables 90 days or more past due as a share of total assets.

We measure A_{ijt} as the book value of the asset-to-liability ratio of bank i in period t. The use of book values is consistent with the maintained hypothesis that the bank regulator has imperfect information about individual banks' financial health.[13] Bank equity values would partially reflect the regulatory environment in which the bank operates and hence would raise simultaneity problems in our specification. Finally, asset book values are actually used by regulators in making closure decisions. The average financial position of banks in period t is represented by $\overline{A}_{jt}$, the cross-sectional mean value of the book asset–liability ratios of banks in the home state of bank i in period t.

The following binary model then nests both the absolute and relative closure rules, as well as the conditioning variables discussed above:

$$\Pr(F = 1)_{ijt} = \gamma_t + \beta_1 A_{ijt} + \beta_2 \overline{A}_{jt} + \beta_3 \xi_{ijt} + e_{ijt}, \quad (12)$$

where γ_t represents a time dummy for period t, ξ_{ijt} is the vector containing the conditioning variables listed above, and e_{ijt} represents an i.i.d. disturbance term.[14] A prediction of the structural model above is that $\beta_1 = -\beta_2$.

3.2. Data

The data set used in this study consists of a panel of annual data for 12,303 U.S. commercial banks from 1992 through 1997. We chose the starting date because, following FDICIA, a relatively homogenous regulatory environment has existed.[15] We choose the ending date because subsequent to 1997 years exist in which no failures take place. All data for individual banks were acquired from the Federal Reserve Bank of Chicago's Bank Condition and Income Database.

Because banks both fail and come into existence over the course of our sample, the panel is not balanced. However, this should not lead to biases in the data because the missing

13. Book values are likely to measure the financial positions of banks with errors. However, market measures were unavailable due to the extremely small number of bank failures among banks that issue equity. Indeed, it appears clear that a sample of equity-issuing banks would have a selection bias towards healthy banks.

Concerning the issue of errors in variables, both A_{ijt} and $\overline{A}_{jt}$ are likely to be measured with error. De Varo and Lacker (1995) demonstrate that the net effect in this case is still some attenuation towards zero.

14. As no failures occurred in 1996 or 1997, to avoid perfect multicollinearity we do not include time dummies for these years.

15. While FDICIA was only formally passed by the United States Congress in December of 1992, it is clear that these reforms were already being incorporated in the closure decisions of bank regulators throughout the year. Indeed, the 1991 data also seem to reflect the stricter regulatory activity called for under FDICIA, although we left this year out of our reported sample to limit ourselves to the post-FDICIA period.

variables due to entry or random exit (as in the case of an unassisted merger) are likely to be uncorrelated with the error term in our model. In the case where observations are missing because of bank failure, the reason for the missing data is precisely what we are attempting to identify in our model specification.

Summary statistics for the data are shown in Table 2. Our data set includes 113 bank failures over the 1992–1997 period. Because the number of failures in our sample is very small relative to the number of nonfailures, we use a logit specification in all our analysis. The logit specification is insensitive to uneven sampling frequency problems (Maddala 1983).

Two patterns stand out in the data. First, the average asset-to-liability ratio of the banking sector increases over the sample, implying an increase in the overall health of the banking system. Unsurprisingly, the number of bank failures diminishes over the panel, reflecting this increase in the financial system's overall health. The year 1992 is particularly active for bank failures, primarily reflecting closures associated with the new tighter regulatory policies under FDICIA. However, even excluding 1992 it is clear that the number of bank failures diminishes over the sample. To rule out time-specific effects in the data stemming from these trends, we include time dummies, γ_t, in our specifications.[16]

We conduct a preliminary investigation of the relationship between bank closure and local conditions in Table 3. We divide the observations in our sample into six subsamples, based on the value of local conditions at time t, or $\overline{A}_{jt}$. In particular, we separate the observations into six groups between 0 and 1, 1 and 2, and greater than 2 standard deviations from either side of the sample mean value of $\overline{A}_{jt}$. We then investigate the probability of bank closures given a potentially closed bank. For our purposes, we consider banks with values of A_{ijt} between 1.00 and 1.05.[17] This range includes all of the failed banks in our sample.

For the subsample with below-average local conditions, the relationship between local conditions and the probability that a bank within the critical range was closed is monotonic: 7.84 percent of banks in the critical range within the subsample with the poorest local conditions (more than two standard deviations below the mean) were closed, 13.89 percent of the banks in the critical range in the next higher subsample were closed, and 16.06 percent of the banks in the next higher subsample were closed. However, the monotonicity breaks down for banks experiencing above-average conditions, as only 11.34 percent of the banks with local conditions between the sample mean value and one standard deviation above the mean were closed. The next higher category obtains a 25 percent figure, resuming the monotonic pattern. However, as there are only four failures within that subsample with local conditions more than one standard deviation above average, one would not draw much of an inference from this observation.

Overall, the results in Table 3 do suggest that regulators are more likely to close a bank of a given asset position the stronger are local conditions, although the results appear to be strongest for substandard local conditions. Nevertheless, it is when times are bad that we would expect regulatory forbearance to be most prominent. In the following section, we formally test this hypothesis for the full sample.

Table 2
Summary Statistics

Year	$\overline{A}_{jt}$	# of bank failures	Avg. value of A_{ijt} for failed banks
1992	1.103	70	1.023
1993	1.109	26	1.018
1994	1.110	9	1.039
1995	1.119	4	1.007
1996	1.123	4	1.013

Source: Federal Reserve Bank of Chicago, Bank Condition and Income Database.

3.3. Results

The results for logit estimation of the entire sample are listed in Table 4. The first and second columns report the results for the base absolute and relative closure rule specifications, respectively. Absolute bank performance, A_{ijt}, enters significantly with its predicted negative sign in both specifications. However, the coefficient estimate on absolute bank performance is sensitive to the inclusion or exclusion of a relative performance measure. In the specification including relative performance, its value almost doubles.

The mean industry performance measure included in the second column, $\overline{A}_{jt}$ is also highly significant. Moreover, its value is of the opposite sign and of the same order of magnitude as the coefficient estimate on A_{ijt}. The formal theory above predicts that these coefficients would be of equal and opposite sign, but we do not find that to be the case. We conducted likelihood ratio and Wald tests of this restriction, and both were strongly rejected. Nevertheless, the similarity

16. Because there are no failures in 1997, we are forced to drop two of the time dummies, one of which must be 1997, to allow for estimation. We include dummies for 1992 through 1995 in the specifications, which yielded the results reported in Tables 3 and 4. Our results were not sensitive to which time dummies were included. Estimates of the coefficients on these time dummies, as well as those for specifications including alternative time dummies, are available from the authors upon request.

17. The qualitative results were robust to minor perturbations around this range.

TABLE 3
CLOSURES OF BANKS AND LOCAL CONDITIONS

Subsample	(1)	(2)	(3)	(4)	(5)	(6)
Percent of failures	7.84	13.89	16.06	11.34	25.00	—
With $1.00 < A_{ijt} < 1.05$						
# of banks	2,216	3,596	23,207	34,359	5,979	1,977
# of failures	11	6	76	16	4	0
Values of $\overline{A}_{jt}$ in subsample range	<1.093	1.093–1.102	1.102–1.110	1.110–1.119	1.119–1.127	>1.127

Note: Subsamples are based on home-state conditions of bank A_{ijt} at time t. Subsample 1 contains observations with $\overline{A}_{jt}$ more than 2 standard deviations lower than the mean of the entire sample. Subsample 2 contains observations with $\overline{A}_{jt}$ between 1 and 2 standard deviations below the mean, subsample 3 between 0 and 1 standard deviations below the mean, subsample 4 between 0 and 1 standard deviations above the mean, subsample 5 between 1 and 2 standard deviations above the mean, and subsample 6 more than 2 standard deviations above the mean.

in the magnitudes of these coefficients is supportive of the model above.

Comparing the base specifications, all of the regression diagnostics strongly favor the relative closure rule specification. Adding $\overline{A}_{jt}$ to the specification reduces the Akaike Information Criteria (AIC) statistic from 1,253.6 to 951.3. Similarly, the second specification lowers the Schwartz criterion from 1,298.9 to 1,005.6 and the –2 log-likelihood from 1,243.6 to 939.3. Likelihood ratio tests strongly reject the restriction that the coefficient on $\overline{A}_{jt}$ is equal to zero at a 1-percent confidence level.

The relative rule specification also does a much better job of predicting bank failures. Under the rule that a bank failure is predicted for probability values greater than or equal to 50 percent, the absolute specification with all of the conditioning variables predicts only six of the 113 bank failures in the sample. In contrast, the relative rule predicts 13 of the 113 bank failures correctly.

The third and fourth columns add the $SIZE_{ijt}$ variable to both specifications. A "too big to fail" theory of bank closure policy would suggest a negative coefficient on this variable, as regulators would resist closing large banks due to systemic concerns. While size does have the predicted negative coefficient estimate, it fails to achieve statistical significance in either specification, a disappointing performance in such a large sample. It may be that the impact of too-big-to-fail protection is nonlinear, such that bank size is only relevant after banks become large enough that their failures would threaten the stability of the payments system.

More importantly for our purposes, our base specification results are robust to the consideration of bank size. A_{ijt} and $\overline{A}_{jt}$ enter in the presence of a bank size variable with quite similar coefficient estimates as they obtained in the base specifications. Both are again highly significant and consistent with the prediction of the theory. Again, the diagnostic and classification statistics strongly support the relative closure rule specification over a simple absolute closure rule, although there is little improvement from the inclusion of the $SIZE_{ijt}$ variable. For the relative closure rule specification (models 2 and 4), likelihood ratio tests fail to reject the restriction that the coefficient on the $SIZE_{ijt}$ variable is equal to zero, although the restriction is rejected when comparing the absolute closure rule specifications (models 1 and 3).

The fifth and sixth columns add the other conditioning variables to the specification. Of the sectoral exposure measures, the $COMMERCIAL_{ijt}$ and $REALESTATE_{ijt}$ variables are robustly positive and significant, suggesting that exposure to these sectors increases the probability of bank closure. In contrast, the $AGRICULTURE_{ijt}$ variable is insignificant. Of the remaining conditioning variables, the CD_{ijt} and $90DAYSLATE_{ijt}$ variables both enter significantly with their predicted positive coefficients. The $NON\text{-}INTEREST_{ijt}$ variable is insignificant.

Again, the base specification results are robust to the inclusion of these conditioning variables. Again, A_{ijt} and $\overline{A}_{jt}$ enter significantly with quite similar coefficient estimates to those that they obtained in the base specifications. Both are again highly significant. Finally, the diagnostic and classification statistics strongly support the relative closure rule specification over a simple absolute closure rule. Likelihood ratio tests do reject the restrictions that the coefficients on the additional conditioning variables are jointly equal to zero, although there is again little improvement from the inclusion of the $SIZE_{ijt}$ variable.

To investigate whether our results were driven by the large number of small banks in our sample, we reran the specification excluding banks that had less than $50 million in book value of total assets during the sample period. This truncation reduced the number of both banking entities and bank failures in our specification roughly in half, from 12,303 to 6,052 and from 113 to 66, respectively. The results for this truncated sample are reported in Table 5.

The results are quite similar to those for the entire sample. The coefficient estimates are all highly significant and enter

TABLE 4
LOGIT ANALYSIS RESULTS: ENTIRE SAMPLE, 1992–1997
DEPENDENT VARIABLE: *FAIL*

Variables	Absolute closure rule (1)	Relative closure rule (2)	Absolute closure rule (3)	Relative closure rule (4)	Absolute closure rule (5)	Relative closure rule (6)
A_{ijt}	–43.42**	–71.81**	–43.50**	–73.10**	–32.59**	–69.74**
	(3.04)	(3.38)	(3.03)	(3.45)	(3.50)	(3.83)
$\overline{A}_{jt}$		55.24**		56.38**		52.26**
		(2.68)		(2.74)		(2.99)
SIZE			–2.19 E–7	–5.14 E–7*	–2.32 E–7	–3.92 E–7
			(1.88 E–7)	(2.40 E–7)	(2.0 E–7)	(2.43 E–7)
D92	41.76**	10.89**	41.90**	11.10**	28.02**	10.07**
	(3.23)	(1.04)	(3.22)	(1.05)	(3.79)	(1.08)
D93	41.18**	10.01**	41.90**	10.25**	27.36**	9.27**
	(3.27)	(1.06)	(3.22)	(1.07)	(3.82)	(1.09)
D94	40.26**	9.29**	40.40**	9.53**	26.44**	8.51**
	(3.29)	(1.11)	(3.28)	(1.11)	(3.85)	(1.13)
D95	39.69**	7.04**	39.85**	7.30**	25.63**	6.32**
	(3.33)	(1.12)	(3.32)	(1.13)	(3.88)	(1.14)
COMMERCIAL					6.71**	5.30**
					(1.06)	(1.18)
AGRICULTURAL					–1.12	–0.28
					(2.14)	(2.15)
REALESTATE					2.55**	2.21**
					(0.78)	(0.78)
NON-INTEREST					1.16	0.36
					(4.38)	(4.37)
CD					3.71**	5.29**
					(1.49)	(1.57)
90DAYSLATE					35.91**	26.66**
					(6.15)	(6.95)
Diagnostics						
AIC	1,253.63	951.28	1,252.43	943.04	1,152.86	886.31
Schwartz	1,298.89	1,005.60	1,306.75	1,006.41	1,261.50	1,004.00
–2 log-likelihood	1,243.63	939.28	1,240.43	929.04	1,128.86	860.31
# observations	63,135	63,135	63,135	63,135	63,135	63,135
Pseudo R^2 classification	.249	.433	.251	.439	.318	.480
Type I error	107/113	100/113	107/113	99/113	104/113	89/113
Type II error	0	6	0	8	6	13
Total correct	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%

Note: See Table 1 for variable definitions. Standard errors are in parentheses. * and ** indicate Wald chi-squared statistical significance at 5-percent and 1-percent levels, respectively. Time dummies for years 1992 through 1995 were included in the specification. Dummy coefficient estimates are available upon request from authors. Type II error figure represents the number of non-events incorrectly designated as events.

with their predicted signs. $\overline{A}_{jt}$ enters significantly positively with a coefficient of opposite sign and a similar magnitude as the absolute performance measure, A_{ijt}.[18] Moreover, the diagnostic statistics strongly suggest a role for relative performance in regulatory closure decisions, as specifications including relative measures continue to outperform those excluding relative performance. The inclusion of the relative performance measure strongly enhances sample fit and reduces Type I error.

18. However, the two variables again fail to enter with equal and opposite coefficient estimates, which would satisfy a strong restriction implied by the formal model.

TABLE 5
LOGIT ANALYSIS RESULTS: SMALL BANKS EXCLUDED, 1992–1997
DEPENDENT VARIABLE: *FAIL*

Variables	Absolute closure rule (1)	Relative closure rule (2)	Absolute closure rule (3)	Relative closure rule (4)	Absolute closure rule (5)	Relative closure rule (6)
A_{ijt}	–42.96**	–73.25**	–43.25**	–74.18**	–35.14**	–71.76**
	(4.27)	(4.88)	(4.26)	(4.89)	(4.81)	(5.39)
$\overline{A}_{jt}$		53.44**		54.26**		50.95**
		(3.68)		(3.70)		(4.01)
SIZE			–2.75E–7	–4.63E–7	–2.68E–7	–4.14E–7
			(2.16E–7)	(2.64E–7)	(2.17E–7)	(2.67E–7)
D92	41.36**	14.90**	41.78**	15.18**	31.24**	14.47**
	(4.55)	(1.50)	(4.53)	(1.50)	(5.20)	(1.61)
D93	40.45**	13.84**	40.87**	14.13**	30.23**	13.36**
	(4.61)	(1.53)	(4.59)	(1.54)	(5.25)	(1.64)
D94	40.25**	13.89**	40.67**	14.15**	30.06**	13.36**
	(4.62)	(1.58)	(4.61)	(1.58)	(5.26)	(1.68)
D95	39.11**	12.49**	39.55**	12.78**	28.68**	11.82**
	(4.68)	(1.69)	(4.66)	(1.69)	(5.34)	(1.80)
COMMERCIAL					5.56**	4.87**
					(1.37)	(1.56)
AGRICULTURAL					–0.93	1.66
					(4.01)	(3.87)
REALESTATE					1.31	1.26
					(0.99)	(1.02)
NON-INTEREST					6.15	6.55
					(5.32)	(5.63)
CD					4.22**	4.78*
					(1.76)	(2.02)
90DAYSLATE					30.83**	12.82
					(8.93)	(10.06)
Diagnostics						
AIC	728.17	581.94	726.18	576.29	695.30	564.79
Schwartz	769.90	632.02	776.25	634.72	795.45	673.30
–2 log-likelihood	718.17	569.94	714.18	562.29	671.30	538.79
# observations	31,143	31,143	31,143	31,143	31,143	31,143
Pseduo R^2 classification	.277	.426	.281	.434	.324	.458
Type I error	61/66	59/66	61/66	56/66	60/66	55/66
Type II error	0	3	0	9	2	3
Total correct	99.8%	99.8%	99.8%	99.8%	99.8%	99.8%

Notes: Analysis excludes banks with total assets below $50 million at any time during the sample period. * and ** indicate Wald chi-squared statistical significance at 5-percent and 1-percent levels, respectively. Time dummies for years 1992 through 1995 were included in the specification. Dummy coefficient estimates are available upon request from authors. Type II error figure represents the number of non-events incorrectly designated as events.

Finally, the conditioning variables perform similarly to the results for the entire sample. There is again little evidence that bank size is a useful predictor of bank closure. Bank size fails to enter significantly, and both specifications appear to be insensitive to its inclusion. Among the other conditioning variables, the $COMMERCIAL_{ijt}$ and CD_{ijt} variables are again robustly significant, while the $AGRICULTURE_{ijt}$, and $NON\text{-}INTEREST_{ijt}$ variables again fail to enter significantly. The notable changes are in the $REALESTATE_{ijt}$ and $90DAYSLATE_{ijt}$ variables, which now fail to enter significantly under the relative closure rule specification (model 6). This discrepancy probably reflects some degree of collinearity between these variables, which provide information about loan quality and a bank's relative performance.

Our empirical results give a strong indication that U.S. regulators considered relative performance in their closure decisions during the post-FDICIA period. This finding is consistent with the desirable policy in the theoretical model above.[19] Moreover, the results are robust to the inclusion of the conditioning variables we consider, as well as the exclusion of small banks from the sample.

4. Conclusion

This paper examines the role of relative performance in bank closure decisions. We show that when banks are subject to common shocks, a closure rule that incorporates relative performance will be less costly than one based solely on absolute performance. Our empirical results provide robust evidence that relative performance has indeed been considered in bank closure decisions in the United States during the post-FDICIA period.

As we note earlier, neither the relative performance rule nor the absolute performance rule is time consistent in a static one-shot game. Instead, a regulator whose loss function solely involves minimizing expected taxpayer liability would always choose prompt closure when regulatory rules allow such behavior. As such, our empirical test should be viewed as a test of the joint hypothesis that the regulator would choose to pursue a relative closure policy and that he has the commitment capacity to do so. Our empirical results suggest that relative performance is incorporated in closure decisions and, therefore, that some form of commitment is achieved. The source of this commitment poses interesting questions beyond the scope of this paper. An interesting extension of this paper would be to endogenize the commitment power of the regulator as a function of its closure strategy. One might conjecture that this would strengthen the superiority of a relative closure rule, because the regulator could more easily commit to the pursuit of a less costly closure strategy.

19. However, relative performance might also be important for considerations outside of our model, such as the ex post political-economy considerations discussed by Kane (1989).

Appendix

Proof of Proposition 1

Under the absolute closure rule in equation (4), ε^* satisfies

$$\varepsilon^* = m - \mu_i - \theta \tag{13}$$

and

$$\frac{\partial \varepsilon^*}{\partial \mu_i} = -1. \tag{14}$$

Substituting these into the bank's first-order condition, we obtain

$$\begin{aligned}(1 + C_\mu)\left[1 - \int_{\underline{\theta}}^{\overline{\theta}} F(m - \mu_i - \theta)g(\theta)d\theta\right] \\ = V_\mu - (m + C)\mathrm{E}\left[f(\varepsilon^*)\right].\end{aligned} \tag{15}$$

To ensure an interior solution for ε^*, we require the parameter restriction

$$\overline{\theta} + \max(\mu) + \underline{\varepsilon} < m < \underline{\theta} + \min(\mu) + \overline{\varepsilon}, \tag{16}$$

which we adopt. Note that feasibility of this condition requires $\max(\mu) - \min(\mu) < \overline{\varepsilon} - \underline{\varepsilon}$.

Consider the special case $m = 0$, i.e., the closure rule is to close all banks on the loss of solvency. In this case, the bank's first-order condition becomes

$$\begin{aligned}\mu_i^p = V_\mu^{-1}\left\{(1 + C_\mu)\left(1 - \mathrm{E}\left[F(-\mu_i^p - \theta)\right]\right)\right. \\ \left. + C\mathrm{E}\left[f(-\mu_i^p - \theta)\right]\right\},\end{aligned} \tag{17}$$

where μ_i^p is the privately optimal choice of effort.

As discussed in the text, we contrast the privately chosen level of effort with μ^s, the socially optimal effort level. The expected social stream of revenues includes expected bank revenues plus bank charter value, net of effort costs, plus expected regulatory liabilities under insolvency. The revenue stream therefore satisfies

$$\int_{\underline{\theta}}^{\overline{\theta}}\left[\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} A_i f(\varepsilon)d\varepsilon + \int_{\varepsilon^*}^{\overline{\varepsilon}} Cf(\varepsilon)d\varepsilon\right]g(\theta)d\theta - V(\mu). \tag{18}$$

The first-order condition for μ^s then satisfies

$$\begin{aligned}\mu^s = V_\mu^{-1}\left\{1 + C\mathrm{E}\left[f(-\mu_i^p - \theta)\right]\right. \\ \left. + C_\mu\left(1 - \mathrm{E}\left[F(-\mu_i^p - \theta)\right]\right)\right\}.\end{aligned} \tag{19}$$

A comparison of equations (17) and (19) leads to Proposition 1. The proof follows directly from the fact that $V_{\mu\mu} > 0$, since $f(-\mu_i^p - \theta) < 1$ and $\mathrm{E}\left[F(-\mu_i^p - \theta)\right] > 0$. As discussed in the text, this is the standard moral hazard result with limited liability: Since its losses are bounded from be-

low, the private bank chooses a lower level of effort because it does not share in the gains to returns in bankruptcy states. These are instead completely enjoyed by the regulator as a reduction in liabilities.

Also, note that when the level of effort is lower, the expected probability of bankruptcy, and hence the regulator's expected liability, will be higher.

Proof of Proposition 2

Under the relative closure rule in equation (7), ε^* satisfies

$$\varepsilon^* = n + \overline{\mu} - \mu_i \tag{20}$$

and

$$\frac{\partial \varepsilon^*}{\partial \mu_i} = -1. \tag{21}$$

Substituting these into the first-order condition yields

$$\int_{n+\overline{\mu}-\mu_i}^{\overline{\varepsilon}} (1 + C_\mu) f(\varepsilon) d\varepsilon + \left[\overline{\mu} + \mathrm{E}(\theta) + n + C\right] \mathrm{E}\left[f(\varepsilon^*)\right] = V_\mu. \tag{22}$$

In equilibrium, since banks are homogenous, all banks make the same effort decision and the first-order condition will satisfy

$$\int_{n}^{\overline{\varepsilon}} (1 + C_\mu) f(\varepsilon) d\varepsilon + \left[\mu_i + \mathrm{E}(\theta) + n + C\right] \mathrm{E}\left[f(\varepsilon^*)\right] = V_\mu. \tag{23}$$

As discussed in the text, we assume that ε_i is distributed uniformly on the interval $\left[\underline{\varepsilon}, \overline{\varepsilon}\right]$ with density f.

Define $\widehat{\mu}$ as the level of effort which satisfies equation (15), i.e., the equilibrium level of effort implied by the absolute closure rule in equation (4). When ε_i is uniformly distributed, equation (15) can be simplified to yield the following relationship between m and $\widehat{\mu}$:

$$m = \frac{\left[V_{\widehat{\mu}} - Cf\right]\left(\overline{\varepsilon} - \underline{\varepsilon}\right) - \left[\overline{\varepsilon} + \widehat{\mu} + \mathrm{E}(\theta)\right]\left(1 + C_{\widehat{\mu}}\right)}{f\left(\overline{\varepsilon} - \underline{\varepsilon}\right) - \left(1 + C_{\widehat{\mu}}\right)}. \tag{24}$$

Next, substituting into the solution above for the level of effort under the relative closure rule, equation (23), the value of n which results in banks choosing effort level $\widehat{\mu}$ satisfies

$$n = \frac{\left(\overline{\varepsilon} - \underline{\varepsilon}\right)\left[V_{\widehat{\mu}} - \left[\widehat{\mu} + \mathrm{E}(\theta) + C\right]f\right] - \overline{\varepsilon}\left(1 + C_{\widehat{\mu}}\right)}{f\left(\overline{\varepsilon} - \underline{\varepsilon}\right) - \left(1 + C_{\widehat{\mu}}\right)}. \tag{25}$$

Combining, $m - n$ satisfies

$$m - n = \widehat{\mu} + \mathrm{E}(\theta). \tag{26}$$

Recall that A^m and A^n are the minimum realizations of A_i necessary to avoid closure under the absolute and relative closure rules. By equations (4) and (7), it is clear that

$$A^m = m \tag{27}$$

and

$$A^n = n + \overline{A}. \tag{28}$$

Substituting from equations (26) and (5), and using the fact that in equilibrium $\overline{\mu} = \widehat{\mu}$,

$$A^n - A^m = \theta - \mathrm{E}(\theta), \tag{29}$$

which directly proves Proposition 2.

Proof of Proposition 3

Substituting into equations (7) and (9) for ε^*, and using the relationship between m and n and the fact that ε_i is uniformly distributed, L_m satisfies

$$L_m = -\int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\varepsilon}}^{n-\theta+\mathrm{E}(\theta)} A_i(\widehat{\mu}, \theta, \varepsilon_i) f(\varepsilon_i) d\varepsilon_i g(\theta) d\theta. \tag{30}$$

By equations (9) and (10),

$$L_m - L_n = \int_{\underline{\theta}}^{\overline{\theta}} \int_{n-\theta+\mathrm{E}(\theta)}^{n} A_i(\widehat{\mu}, \theta, \varepsilon_i) f(\varepsilon_i) d\varepsilon_i g(\theta) d\theta. \tag{31}$$

Assuming that ε_i is distributed uniformly, this simplifies to

$$L_m - L_n = \frac{1}{2}\left[\frac{Var(\theta)}{\overline{\varepsilon} - \underline{\varepsilon}}\right], \tag{32}$$

which gives us Proposition 3.

References

Acharya, Sankarshan, and Jean-Francois Dreyfus. 1989. "Optimal Bank Reorganization Policies and the Pricing of Federal Deposit Insurance." *Journal of Finance* 44(5) (December) pp. 1,313–1,333.

Acharya, Viral V., and Tanju Yorulmazer. 2007. "Too Many to Fail—An Analysis of Time-Inconsistency in Bank Closure Regulation." *Journal of Financial Intermediation* 16, pp. 1–31.

Acharya, Viral V., and Tanju Yorulmazer. 2008. "Cash-in-the-Market Pricing and Optimal Bank Bailout Policy." *Review of Financial Studies* (forthcoming).

Aghion, Philippe, and Patrick Bolton. 1992. "An 'Incomplete Contracts' Approach to Financial Contracting." *Review of Economic Studies* 59, pp. 473–494.

Aghion, Philippe, Patrick Bolton, and Steven Fries. 1999. "Optimal Design of Bank Bailouts: The Case of Transition Economies." *Journal of Institutional and Theoretical Economics* 155, pp. 51–70.

Baird, Douglas. 1991. "The Initiation Problem in Bankruptcy." *International Review of Law and Economics* 11, pp. 223–232.

Berger, Allen N., Sally M. Davies, and Mark J. Flannery. 2000. "Comparing Market and Supervisory Assessments of Bank Performance: Who Knows What When?" *Journal of Money, Credit, and Banking* 32(3), pp. 641–667.

Berkovitch, Elazar, Ronen Israel, and Jaime F. Zender. 1998. "The Design of Bankruptcy Law: A Case for Management Bias in Bankruptcy Reorganizations." *Journal of Financial and Quantitative Analysis* 33(4) (December) pp. 441–464.

Blum, Jurg, and Martin Hellwig. 1995. "The Macroeconomic Implications of Capital Adequacy Requirements for Banks." *European Economic Review* 39, pp. 739–749.

Boot, Arnoud W., and Anjan V. Thakor. 1993. "Self-Interested Bank Regulation." *American Economic Review* 83, pp. 206–212.

Brown, Craig O., and I. Serdar Dinç. 2006. "Too Many to Fail? Evidence of Regulatory Reluctance in Bank Failures When the Banking Sector Is Weak." Manuscript.

Demirgüç-Kunt, Asli. 1991. "Principal-Agent Problems in Commercial-Bank Failure Decisions." FRB Cleveland Working Paper 9106 (April). http://www.clevelandfed.org/research/Workpaper/1991/wp9106.pdf

De Varo, Jed L., and Jeffrey M. Lacker. 1995. "Errors in Variables and Lending Discrimination." FRB Richmond *Economic Quarterly* 81(3) (Summer) pp. 19–32. http://www.richmondfed.org/publications/economic_research/economic_quarterly/pdfs/summer1995/lacker.pdf

Dewatripont, Mathias, and Jean Tirole. 1993. *The Prudential Regulation of Banks.* Cambridge, MA: MIT Press.

Diamond, Douglas W., and Raghuram G. Rajan. 2005. "Liquidity Shortages and Banking Crises." *Journal of Finance* 60(2) pp. 615–647.

Flannery, Mark J. 1998. "Using Market Information in Prudential Bank Supervision: A Review of the U.S. Empirical Evidence." *Journal of Money, Credit, and Banking* 30(3), pp. 273–305.

Fries, Steven, Pierre Mella-Barral, and William Perraudin. 1997. "Optimal Bank Reorganization and the Fair Pricing of Deposit Guarantees." *Journal of Banking and Finance* 21, pp. 441–468.

Gertner, Robert, and David Scharfstein. 1991. "A Theory of Workouts and the Effects of Reorganization Law." *Journal of Finance* 46, pp. 1,189–1,222.

Giammarino, Ronald M., Tracy R. Lewis, and David E.M. Sappington. 1993. "An Incentive Approach to Banking Regulation." *Journal of Finance* 48(4) (September) pp. 1,523–1,542.

Goodhart, Charles, Philipp Hartman, David Llewellyn, Liliana Rojas-Suarez, and Stephen Weisbrod. 1998. *Financial Regulation: Why, How, and Where Now?* New York: Routledge.

Harris, Milton, and Artur Raviv. 1995. "The Role of Games in Security Design." *The Review of Financial Studies* 8(2) (Summer) pp. 327–367.

Hart, Oliver, and John Moore. 1998. "Default and Renegotiation: A Dynamic Model of Debt." *Quarterly Journal of Economics* 113, pp. 1–42.

Holmstrom, Bengt. 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13, pp. 324–340.

Hvide, Hans K. 2002. "Tournament Rewards and Risk Taking." *Journal of Labor Economics* 20(4) pp. 877–898.

Kane, Edward J. 1989. *The S&L Insurance Mess: How Did It Happen?* Cambridge, MA: MIT Press.

Lang, Larry H.P., and René M. Stulz. 1992. "Contagion and Competitive Intra-Industry Effects of Bankruptcy Announcements." *Journal of Financial Economics* 32, pp. 45–60.

Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89(5) (October) pp. 841–864.

Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Mailath, George J., and Loretta J. Mester. 1994. "A Positive Analysis of Bank Closure." *Journal of Financial Intermediation* 3, pp. 272–299.

Marcus, Alan. 1984. "Deregulation and Bank Financial Policy." *Journal of Banking and Finance* 8, pp. 557–565.

Mooradian, Robert M. 1994. "The Effect of Bankruptcy Protection on Investment: Chapter 11 as a Screening Device." *Journal of Finance* 49, pp. 1,403–1,430.

Nagarajan, S., and C.W. Sealey. 1995. "Forbearance, Deposit Insurance Pricing, and Incentive Compatible Bank Regulation." *Journal of Banking and Finance* 19(6) pp. 1,109–1,130.

Nagarajan, S., and C.W. Sealey. 1998. "State-Contingent Regulatory Mechanisms and Fairly Priced Deposit Insurance." *Journal of Banking and Finance* 22, pp. 1,139–1,156.

Nalebuff, Barry J., and Joseph E. Stiglitz. 1983. "Prizes and Incentives: Towards a General Theory of Compensation and Competition." *Bell Journal of Economics* 14(1) (Spring) pp. 21–43.

Thomson, James B. 1991. "Predicting Bank Failures in the 1980s." FRB Cleveland *Economic Review* 27(1) pp. 9–20. http://www.clevelandfed.org/research/Review/1991/index.cfm#Q1

Weiss, Lawrence A., and Karen H. Wruck. 1998. "Information Problems, Conflicts of Interest, and Asset Stripping: Chapter 11's Failure in the Case of Eastern Airlines." *Journal of Financial Economics* 48, pp. 55–97.

Wheelock, David C., and Paul W. Wilson. 2000. "Why Do Banks Disappear? The Determinants of U.S. Bank Failures and Acquisitions." *Review of Economics and Statistics* 82(1) pp. 127–138.

The Quantity and Character of Out-of-Market Small Business Lending*

Elizabeth S. Laderman

Economist
Federal Reserve Bank of San Francisco

Most small business lending from banks originates with institutions that have a local branch, but "out-of-market" lending does not. Supporting the view that proximity is conducive to lending, I find that only about 10 percent of small business lending is from banks with no branch in the local market. About half of this appears to be from banks with a branch in the same state, further supporting the role of proximity, while, at the same time, supporting the current regulatory practice of considering out-of-market loans when assessing local competitive conditions. I also find that out-of-market and in-market loans are of similar average size and are about equally likely to be secured by commercial real estate.

1. Introduction

Small businesses play an important role in the U.S. economy, accounting for roughly half of all private employment and more than half of output.[1] These small businesses need financing in order to operate and grow, and bank lending is an important source of this financing.[2] A key issue is whether geographic proximity of banks to small business borrowers is important for the establishment of credit relationships. In other words, how significant is a bank's physical presence in a local market to the provision of credit to small businesses in that market?

This paper discusses the quantity and type of small business loans in an area that are made by banks that do not have a physical presence in that area and the implications of those characteristics for defining small business loan markets. In addition to assessing the role of out-of-market lenders, the analysis explores the appropriate geographic scope and measurement of the level of competition among banks in providing small business financing. The latter is important to public policy since competition in banking can affect the quantity and price of banking services, including credit services to small businesses.

The paper is structured as follows. Section 2 discusses related literature. Section 3 provides background related to small business lending markets. Section 4 discusses the data used in this analysis, and Section 5 outlines the results of the analysis. Section 6 concludes.

2. Related Literature

Broadly speaking, this paper fits into the existing literature regarding the relationship between bank small business lend-

*I would like to thank Jin Oh and Armando Franco for valuable research assistance. Opinions expressed do not necessarily reflect the views of the management of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System.

1. According to the U.S. Small Business Administration, nonfarm businesses with fewer than 500 employees employ about half of all nonfarm private-sector workers, create more than half of nonfarm private-sector gross domestic product, and have generated 60 to 80 percent of net new nonfarm jobs annually over the last decade. See http://app1.sba.gov/faqs/faqIndexAll.cfm?areaid=24 for related data.

2. About 46 percent of the nonfarm businesses with fewer than 500 employees that participated in The Federal Reserve Board's 2003 Survey of Small Business Finances stated that they had a credit line, loan, or capital lease from a commercial bank, savings and loan, or savings bank in 2003. In comparison, about 22 percent said they had used a finance or factoring company for at least one of these types of credit, 6 percent had used family or individuals, 4 percent had used a leasing company, 4 percent had used a credit union, and 2 percent had used an insurance or mortgage company in 2003. None of these figures include financing through credit card borrowing or borrowing from the owner of the firm, even if, for example, a commercial bank issued the credit card. About 47 percent of the small businesses surveyed had used a personal credit card in 2003, and about 48 percent had used a business credit card. Of the small businesses that could have received loans from owners (those organized as corporations or partnerships), about 30 percent had obtained such a loan. Finally, about 60 percent of the small businesses used trade credit in 2003. (Mach and Wolken 2006)

ing and the proximity of lenders and business borrowers. This literature has generated several widely accepted findings. First, historically there has been a strong negative relationship between small business lending and distance. For example, examining the distance between the center of the census tract in which a borrower is located and the nearest office of the lender, Brevoort and Hannan (2006) find that distance operates as a statistically and economically significant deterrent to lending within local markets. In addition, Wolken and Rohde (2002) show that, in 1998, the median distance between a small business's headquarters and the financial institution making the loan was only ten miles.

Second, both the mean and the median distance between small business borrowers and their lenders have been increasing. For example, using data on small business borrower-lender relationships that existed in 1993, Petersen and Rajan (2002) find that the median borrower-lender distance increased from two miles for relationships that began in the 1970s to five miles for relationships that began in the 1990s.

Third, banks began adopting a new lending technology, small business credit scoring, in the early 1990s. In credit scoring, banks assess borrowers' creditworthiness using computer-generated models based mainly on information about the owner's credit quality from consumer credit bureaus and information about the small business's credit quality from commercial credit bureaus. Scoring models in essence automate the credit underwriting process. Credit scoring has the potential to reduce the cost of small business lending, at least for certain types of small business loans, and therefore has the potential to increase the distance over which loans are made. Small business credit scoring likely entails a relatively sizable fixed cost, which would give large banks a comparative advantage over small banks in using this technology.

Some papers have further explored the relationship between small business credit scoring and small business lending. Petersen and Rajan (2002) attribute at least part of the increase in median borrower-lender distance to the adoption of credit scoring by some banks. Frame, Srinivasan, and Woosley (2001) find a positive relationship between credit scoring and small business lending for a sample of large banks. Frame, Padhi, and Woosley (2004) find that banks that use credit scoring have a higher ratio of loans outside their local markets to total loans than do banks that do not use credit scoring. DeYoung, Glennon, and Nigro (2006) find that credit scoring is a relatively more efficient lending technology for more distant borrowers and lenders.

Two other papers also discuss the quantity of out-of-market small business loans. Krainer and Beauchamp (1999) find that, in 1997, for California, most of the small business loans in terms of number were from outside the local market. In small markets, most of the out-of-market lenders were either large banks that were relatively near the market or national credit card banks. For the San Francisco Bay Area, Laderman (2006) finds that, in 2005, after excluding credit card banks, the out-of-market share of small business lending by dollar volume was very minimal.

3. Background

For urban areas, the Federal Reserve currently defines small business lending markets to be about the size of metropolitan statistical areas (MSAs).[3] The Federal Reserve includes in these markets the small business loans of all banks that make small business loans in the MSA. Many of the banks that make loans in the MSA also have physical branches within the MSA, but some do not. In this paper, I refer to the banks that make loans in the market but do not have branches in the market as "out-of-market" banks and their loans in the market as "out-of-market" loans. In contrast, "in-market" loans are made by banks with a physical presence in the MSA.

The very existence of out-of-market small business loans raises the natural question of whether the size of small business lending markets is too small and whether, despite previous evidence suggesting that small business lending markets are very local, the geographic boundaries of these market definitions ought to be expanded, or whether a geographically based market definition even makes sense at all. These questions are especially compelling given the increase over time in the distance between borrowers and lenders.

I begin to address these issues by examining the share of small business lending within MSAs that is coming from out-of-market lenders. Intuitively, if out-of-market shares for MSAs overall are substantial, then MSA-based small business loan markets may be too small. But, even if out-of-market shares are small, if a great majority of those out-of-market loans are from lenders with a physical presence near the MSA, then MSA-based small business loan markets still may be too small.

I proxy the degree to which out-of-market small business loans are coming from near the market with the difference between the out-of-market share for MSA-based markets and the out-of-market share for state-based "markets."[4] I find that the out-of-market shares for both types of markets are quite small and that out-of-market lending from outside the

3. For simplicity, throughout this paper, I refer to "MSA markets" or "MSA-based markets," even though actual Federal Reserve urban banking markets differ somewhat from MSAs. For a discussion of the differences between the two and a conclusion that, for research, MSAs are reasonable approximations of urban Federal Reserve banking markets, see Laderman and Pilloff (2007).

4. I refer to state-based "markets," rather than, say, "areas," for simplicity. However, as explained above, actual Federal Reserve banking markets are comparable to MSAs, not states.

MSA but within the state accounts for about half of the out-of-market lending for MSA-based markets.

The relatively greater role of within-state banks compared to out-of-state banks in providing out-of-market small business loans is consistent with other evidence indicating that distance (the proximity of borrower and lender) affects the likelihood of a credit relationship between a bank and a business.[5] And while in-state banks' share of out-of-market lending is not large enough to compel a shift away from markets based on MSAs to markets based on a larger geographic area, it is large enough to suggest that the geographic borders of small business lending markets are not finely demarcated and therefore that it makes sense to consider out-of-market lenders in determining market competition. And it makes a difference: competition as measured by market competition excluding out-of-market lenders is notably weaker than competition including out-of-market lenders.

Who are the out-of-market small business lenders, and what are some characteristics of the lenders and their loans? I find that nearly 1,600 banks do some out-of-market lending. However, the dollar volume of out-of-market lending is quite concentrated in the biggest out-of-market lenders, while the number of out-of-market loans is even more so.

The calculation of small business loan concentration that includes both in-market and out-of-market loans assumes that out-of-market loans are good substitutes for in-market loans. Is this assumption warranted? While I do not provide an in-depth answer to this question, I do provide relative information to begin a comparison between out-of-market and in-market lending along a few basic dimensions: the average sizes of the loans, the sizes of the lenders, the lenders' small business loan-to-asset ratios, the size distribution of the loans, and the share of small business loans that are secured by commercial real estate.

Except for lender size, the differences between out-of-market and in-market characteristics, although almost always statistically significant, are relatively minor. For example, out-of-market loans tend to be a bit smaller than in-market loans, but both average less than $100,000. Out-of-market lenders as a whole tend to be markedly larger than in-market lenders. Consistent with this size difference and prior findings in the literature regarding the relative propensity of large banks and small banks to make small business loans, I find that the ratio of small business loans to assets is smaller for out-of-market lenders than for in-market lenders. However, the difference is minimal, and the ratio for out-of-market lenders is larger than for large banks in general.

I also compare the shares of the number of out-of-market business loans under $1 million that are under $100,000, between $100,000 and $250,000, and between $250,000 and $1 million, to those for in-market business loans. Akhavein, Frame, and White (2005) report survey data indicating that small business credit scoring is most likely to be used for loans under $100,000, less likely to be used for loans between $100,000 and $250,000, and least likely to be used for loans between $250,000 and $1 million. I find that the proportion of out-of-market small business loans that is under $100,000 is modestly higher than the proportion of in-market small business loans that is under $100,000. But the great majority of both in-market and out-of-market small business loans is under $100,000; this remains true whether or not I count the out-of-market loans of the banks that dominated out-of-market lending in the sample year—Wells Fargo Bank Northwest and Wells Fargo Bank.

I also compare the shares of loans of out-of-market lenders that are secured by commercial real estate to the shares of loans of in-market lenders that are secured by commercial real estate. I find a statistically significant but minor difference between out-of-market lenders' and in-market lenders' shares of small business loans that are backed by commercial real estate.

4. Data

I use data on the flow of small business lending in 2004, gathered from reports that roughly 5,000 "banks" (commercial banks, savings banks, and savings and loans) submitted in compliance with the Community Reinvestment Act (CRA). I use CRA reports of loans under $1 million to businesses with revenues under $1 million, thereby focusing on small loans to small businesses.[6]

I define out-of-market small business lending as cases when a bank lends to a borrower in either a state or an MSA, as the case may be, where the bank does not have a physical branch. Banks report loan totals by the census tract of the borrower's headquarters or by the census tract where the majority of the funds are being used. I aggregate from the census tract level to the MSA or state level. Commercial and

5. For any given state, the pool of potential out-of-market lenders from outside the state is far larger than the pool of potential out-of-market lenders from within the state.

6. During 2004, only those banks with assets of at least $250 million and banks that were in a holding company with at least $1 billion in assets were required to complete CRA compliance reports. (The cutoff was changed effective September 2005 to $1 billion in assets, with no holding company criterion.) Therefore, following previous research, for banks that do not meet the CRA reporting requirement criteria, I have estimated small business lending by MSA by using Call Report and Thrift Financial Report data. Specifically, I have allocated total small business lending as reported on the Call Report or Thrift Financial Report to different MSAs in proportion to the bank's share of the bank's total deposits in that MSA.

industrial loans (loans for a business purpose that are not secured by real estate), commercial real estate loans (loans that are secured by commercial real estate), and loans through a business credit card all are considered business loans for CRA reporting purposes.[7] Consistent with prior research, I exclude the loans of credit card banks from my sample.

5. Results

5.1. In- and Out-of-Market Loans and Loan Sizes

Table 1 presents some introductory sample statistics for 362 MSAs and the 50 states plus the District of Columbia. The number of small businesses demarcates the state and MSA size categories. It is apparent even from these aggregations that out-of-market lending accounts for a relatively small share of total small business lending for states and MSAs, as well as for size subcategories within those groups.

However, the quantity and distribution of out-of-market lending still may have a meaningful effect on competition in small business lending. I will discuss this further below. In addition to the quantity and distribution of out-of-market lending, one might want to consider whether out-of-market loans are similar enough to in-market loans. One aspect of this comparison is the size of the loan. Table 1 indicates that the mean of the average loan size, where the average loan size is the ratio of the total dollar volume of loans to the total number of loans for each of the geographic areas within the indicated category, is almost always statistically significantly smaller for out-of-market loans than for the comparable group of in-market loans.[8] The one exception is for small MSAs. However, all average loan size means are less than $100,000 and fall within a relatively narrow range of about $60,000 to about $90,000.

5.2. In-Market Shares

Table 2 shows further that, for MSAs, the great majority of small business loans, whether measured by dollar volume or number of loans, consists of in-market loans. Although a considerable number of lenders are from outside of the market, they appear to be making relatively few loans, and those loans total relatively few dollars. Moreover, within the MSA groups, there is no clear pattern of lower in-market shares for smaller areas. The relatively high in-market shares for MSAs, by themselves, though not conclusive, are consistent with the findings of Brevoort and Hannan (2006) that proximity is conducive to small business lending. The in-market shares also are consistent with the view that a geographically based

Table 1
Small Business Loans, 2004: Sample Means

	In-market			
	$ volume (millions)	# loans	# lenders	Avg. size (thousands)
States (51)	3,103	45,542	128	85***
Large (18)	5,635	65,495	216	92***
Medium (17)	2,847	60,863	125	79***
Small (16)	525	6,817	32	83***
MSAs (362)	338	5,337	26	84***
Large (124)	747	12,616	46	87***
Medium (119)	170	2,074	18	85***
Small (119)	81	1,015	12	81
	Out-of-market			
States (51)	144	2,543	100	70
Large (18)	290	5,249	165	79
Medium (17)	98	1,628	92	71
Small (16)	27	472	36	58
MSAs (362)	28	438	35	73
Large (124)	61	996	62	68
Medium (119)	14	194	25	74
Small (119)	8	101	17	77

Notes: The null hypothesis is that the in-market sample is from a population with the same distribution as the out-of-market sample; *** indicates rejection at a 1-percent level, based on the Wilcoxon rank-sum test. State and MSA size categories are determined by the number of small businesses in each.

Table 2
Average Percent Share of Small Business Lending Provided by In-Market Lenders, 2004

	Average percent share according to		
	$ volume	# loans	# lenders
States	94.8	92.7	51.7
Large	94.3	92.1	55.0
Medium	96.0	94.7	53.7
Small	94.0	91.4	45.8
MSAs	90.4	89.2	41.0
Large	90.7	88.3	39.2
Medium	90.7	89.9	41.4
Small	89.7	89.4	42.4

Note: State and MSA size categories are determined by the number of small businesses in each.

7. Banks report business credit card lines of credit, whether drawn on or not, on the CRA reporting form. In contrast, personal credit card lines of credit, even if used for business purposes (for example, lines through a small business owner's personal credit card), are not reported on the CRA reporting form.

8. The nonparametric statistical tests I conduct in this paper are the Wilcoxon rank-sum test, which tests the hypothesis that two samples are drawn from populations with identical distributions, and the median test, which tests whether two samples are drawn from populations with the same median.

market definition is warranted and that MSAs are an appropriate upper bound on the geographic size of small business lending markets.

Moreover, it appears that roughly half of the dollar volume of out-of-MSA lending may come from within the same state as the MSA. (About 10 percent of lending is from outside the MSA, and about 5 percent is from outside the state, leaving about 5 percent from inside the state.) If lending from within the same state is an important component of out-of-MSA lending, this also would be consistent with the role of proximity in lending. At the same time, the sizable contribution of in-state banks to out-of-market lending suggests that, although MSAs likely are an appropriate, workable upper bound on the geographic size of small business lending markets, the geographic borders of these markets actually are not very finely drawn. On this basis alone, it makes sense to consider out-of-market lenders when measuring small business lending competition in local markets.

5.3. Market Concentration

Below, I present further evidence regarding whether out-of-market small business loans are similar to in-market loans. In this section, I simply treat them as the same for the purpose of measuring their effect on competition.

I measure competition with the Hirschman-Herfindahl Index (HHI) of market concentration. In the classic structure-conduct-performance paradigm of industrial organization theory, when market shares are more concentrated at the top, competition is weaker, and the HHI is a convenient and widely used measure of concentration. The HHI, which is the sum of the squares of the percent market shares of all the firms in a market, increases with the variance of market shares, holding constant the number of firms.[9]

However, while an increase in the number of firms, holding the variance constant, often decreases the HHI, this is not always the case. In fact, whether the inclusion of out-of-market lenders in the calculation of the HHI increases or decreases it depends on several factors, including not only the number of additional lenders but also the change in the variance of market shares due to the inclusion of out-of-market lenders, the change in the number of lenders *times* the variance, the variance of market shares including only in-market lenders, and the number of in-market lenders.[10]

Table 3
MSA Market Concentrations Measured by Hirschman-Herfindahl Index, 2004

	Mean	Median	Standard deviation
Excluding out-of-market loans	2,282***	2,023***	1,163
Including out-of-market loans	1,924	1,695	927

Note: The null hypothesis is that the in-market-only sample is from a population with the same distribution or median as the sample including out-of-market loans; *** indicates rejection at a 1-percent level, based on the Wilcoxon rank-sum or median test.

To investigate the effect of the current method of including out-of-market loans on concentration, I compare the HHI for MSA-based markets without out-of-market loans to that for MSA-based markets with out-of-market loans in Table 3. Even though out-of-market lending constitutes a relatively small proportion of total lending, it does have a statistically significant and meaningful effect on the small business lending HHI for MSAs, decreasing it from 2,282 to 1,924 at the mean. At the median, the HHI including only in-market loans is in the highly concentrated range, whereas the HHI including both in- and out-of-market loans indicates moderate concentration. The apparent effect of out-of-market lending on competition in MSA-based small business lending markets supports the current practice of including out-of-market loans in the calculation of market shares, as opposed to excluding them.

5.4. Out-of-Market Lenders and Out-of-Market Loans

I argue above that the importance of in-state lenders for out-of-market lending supports giving some consideration to out-of-market lending when measuring market concentration. In this section, I present further evidence on the characteristics of out-of-market lenders and loans. This evidence also is relevant to the issue of whether out-of-market loans are good substitutes for in-market loans.

Table 4 shows a considerable amount of concentration in out-of-market small business lending. With nearly 1,600 out-of-market lenders, the top 50 account for almost 60 percent

9. Both the Federal Reserve and the Department of Justice (DOJ) use the HHI as a measure of market concentration when assessing the potential effects of a proposed bank merger on competition, and both use the DOJ's market concentration level definitions and its Horizontal Merger Guidelines. The DOJ defines a market with an HHI below 1,000 as "unconcentrated," one with an HHI between 1,000 and 1,800 as "moderately concentrated," and one with an HHI of at least 1,800 as "highly concentrated." A merger that would increase the HHI by more than 200 to a highly concentrated level would violate the Merger Guidelines. Typically, the Federal Reserve evaluates the potential effect of a proposed transaction on competition in small business lending whenever the transaction violates the Merger Guidelines as calculated on the basis of deposits. The DOJ more routinely performs both types of evaluations.

10. See Laderman (1995) for a more detailed discussion of the breakdown of the HHI.

Table 4
Percent Shares of National Small Business Loan Volume Held by Top Out-of-Market Lenders, 2004

	by $ volume	by number
Top 5	38.1	75.1
Top 10	42.8	78.1
Top 20	49.2	80.9
Top 50	59.4	84.3

Note: The total number of out-of-market lenders is 1,578.

of out-of-market loans by dollar volume and almost 85 percent of out-of-market loans by number.

Table 5 and Table 6 list the top ten out-of-market lenders by dollar volume and by number of loans, respectively. The top three alone account for more than one-third of the dollar volume and more than two-thirds of the total number of loans. Indeed, roughly 60 percent of lenders outside of the top ten by number of loans made ten or fewer out-of-market loans in 2004.

Several of the names in Table 5 and Table 6 are those of well-known large banks.[11] Indeed, Table 7 confirms that banks that do any out-of-market lending are, as a group, considerably larger than banks that do any in-market lending.[12] This is true both at the mean and at the median. The distinction is even stronger between banks that do any out-of-market lending and banks that do strictly in-market lending.

The small business loan-to-asset ratios in the third column of Table 7 suggest that out-of-market lenders also tend to be somewhat less intensely engaged in small business lending than in-market lenders and in-market only lenders. However, although all the differences are statistically significant, they are relatively small.[13] Moreover, the ratios for out-of-market

Table 5
Top Out-of-Market Lenders by Dollar Volume, 2004

	$ millions	Cum. share of total (%)	# loans	Avg. loan ($ thousands)
Wells Fargo Bank Northwest, N.A.	1,719.7	16.9	56,466	30
Wells Fargo Bank, N.A.	1,272.3	29.4	48,315	26
JPMorgan Chase Bank, N.A.	599.8	35.3	6,243	96
Bank of the West	155.5	36.8	1,078	144
Fleet National Bank	127.8	38.1	6,296	20
Amsouth Bank	114.3	39.2	1,039	110
Umpqua Bank	104.5	40.2	459	228
Comerica Bank	94.3	41.1	278	339
Wachovia Bank, N.A.	89.2	42.0	227	393
Branch Banking & Trust Co.	81.9	42.8	416	197
Remainder (1,568)	5,830.3	57.2	37,664	155

Table 6
Top Out-of-Market Lenders by Number of Loans, 2004

	# loans	Cum. share of total (%)	$ millions	Avg. loan ($ thousands)
Wells Fargo Bank Northwest, N.A.	56,466	35.6	1,719.7	30
Wells Fargo Bank, N.A.	48,315	66.1	1,272.3	26
Fleet National Bank	6,296	70.1	127.8	20
JPMorgan Chase Bank, N.A.	6,243	74.0	599.8	96
First Tennessee Bank N.A., Memphis	1,704	75.1	77.2	45
Washington Mutual Bk.	1,177	75.8	18.7	16
Bank of the West	1,078	76.5	155.6	144
Amsouth Bank	1,039	77.2	114.3	110
Netbank	833	77.7	65.1	78
Farm Bureau Bk., F.S.B.	626	78.1	11.9	19
Remainder (1,568)	34,704	21.9	6,027.2	174

lenders are larger than for big banks in general.[14] Relative to their peers, out-of-market small business lenders do emphasize small business lending.

Table 8 shows that out-of-market loans have a statistically significantly higher probability of being under $100,000 and a lower probability of being between $100,000 and $250,000 or between $250,000 and $1 million than in-market loans. (Means and medians are across MSAs.) The slightly greater

11. Some of the banks in Table 5 or Table 6, for example Branch Banking & Trust and First Tennessee Bank, were among the top credit card lenders in the country in 2004. However, based on available data, these banks' credit card lending did not constitute a large enough share of their total lending to justify a conclusion that the bulk of their small business lending was through credit cards.

12. Of course, many banks do some in-market lending and some out-of-market lending and are therefore included in both groups.

13. The third column of Table 7 shows the ratio of the total dollar volume outstanding of commercial and industrial and commercial real estate loans under $1 million to total assets as of June 2004. Although the in- and out-of-market designations used for all tables rely on CRA information, the actual data in Tables 7 and 10 are from commercial banks' Reports of Condition and Income (Call Reports) and savings banks' and savings and loans' Thrift Financial Reports. In addition, the loans under $1 million in Tables 7 and 10 may be to businesses of any size. In contrast, as stated above, the small business loan data presented up to this point, from the CRA reports, have been for loans under $1 million to businesses with revenues under $1 million.

14. As of June 2004, the mean small business loan-to-asset ratio for all banks with over $1 billion in assets was 0.094.

TABLE 7
LENDER SIZES AND LOANS UNDER $1 MILLION, 2004

	Assets ($ millions)	Asset share of loans <$1 mil.
Means		
Out-of-market lenders (1,576)	4,472.2	.162
In-market lenders (5,380)	1,516.7***	.185***
In-market-only lenders (3,986)	295.3***	.193***
Medians		
Out-of-market lenders (1,576)	515.9	.152
In-market lenders (5,380)	158.9***	.170***
In-market-only lenders (3,986)	111.6***	.182***

Note: The null hypothesis is that the in-market-only sample is from a population with the same distribution or median as the sample including out-of-market loans; *** indicates rejection at a 1-percent level, based on the Wilcoxon rank-sum or median test.

TABLE 8
SHARE OF NUMBER OF BUSINESS LOANS UNDER $1 MILLION, 2004

	< $100,000	$100,000 to $250,000	$250,000 to $1 million
Means			
Out-of-market loans	.845	.072	.083
In-market loans	.731***	.145***	.124***
Medians			
Out-of-market loans	.857	.063	.078
In-market loans	.740***	.141***	.119***

Note: The null hypothesis is that the in-market-only sample is from a population with the same distribution or median as the sample including out-of-market loans; *** indicates rejection at a 1-percent level, based on the Wilcoxon rank-sum or median test.

tendency of out-of-market loans to be under $100,000 than in-market loans to be under $100,000 is consistent with the evidence on differences in average loan sizes in Table 1.[15] To the degree that credit-scored loans are likely to be under $100,000, it also is consistent with the relationship seen in Table 7 between large banks and out-of-market lending and the literature's links between large banks, lending at a distance, and credit scoring.

However, as noted, the differences are slight, and both in-market and out-of-market loans fall heavily into the under $100,000 category. But, given the dominance of Wells Fargo Bank Northwest and Wells Fargo Bank in out-of-market lending, it may be important to investigate how these two banks influence this finding. Note that these two banks'

15. The data in Tables 8 and 9 are from the CRA report and therefore pertain to the flow of loans in 2004, but they include loans under $1 million to businesses of any size.

TABLE 9
LOAN SIZE DISTRIBUTION WITHOUT OUT-OF-MARKET LOANS FROM TOP TWO PROVIDERS, 2004

	Share of number of small business loans		
	< $100,000	$100,000 to $250,000	$250,000 to $1 million
Means			
Out-of-market loans	.736	.117	.147
In-market loans	.731	.145***	.124***
Medians			
Out-of-market loans	.743	.113	.143
In-market loans	.740	.141***	.119***

Note: The null hypothesis is that the in-market-only sample is from a population with the same distribution or median as the sample including out-of-market loans; *** indicates rejection at a 1-percent level, based on the Wilcoxon rank-sum or median test.

average loan sizes are among the smallest for the top out-of-market lenders (Table 5 and Table 6). Indeed, when I do exclude the Wells Fargo banks' out-of-market loans, the proportion of in-market loans that are under $100,000 is virtually identical to the proportion of out-of-market loans that are under $100,000 (Table 9).[16]

In-market lending also appears to be about as likely to be secured by commercial real estate as out-of-market lending (Table 10). Although, based on Call Report and Thrift Financial Report data, out-of-market lenders have a slightly higher commercial real estate loan-to-asset ratio than in-market lenders, out-of-market lenders have a slightly lower share of loans under $1 million that are backed by commercial real estate than do in-market lenders. But, the difference is quite small.

6. Conclusion

The quantity and character of out-of-market small business lending have important policy implications. Too much lending from far outside the market might call into question the

16. Regarding possible differences between the Wells Fargo banks' out-of-market lending and in-market lending, I note anecdotal evidence that Wells Fargo was one of the very first banks to use small business credit scoring and continues to use it extensively. However, credit scoring may mark a distinction without a difference. As argued in Berger and Udell (2006), credit scoring, asset-based lending, factoring, fixed-asset lending (such as lending secured by commercial real estate, discussed below), and leasing all are "transactions technologies" (lending based primarily on "hard" quantitative data) that enable banks to lend to businesses with little or no financial statement data ("opaque" businesses). Therefore, credit scoring may be an effective substitute for "relationship" lending, which is based primarily on "soft" qualitative information and usually directed toward opaque firms, as well as for the other transactions technologies named.

TABLE 10
LOANS SECURED BY COMMERCIAL REAL ESTATE (CRE), 2004

	All CRE loans, share of assets (by $ volume)	CRE loans <$1 mil., share of all loans <$1 mil. (by number)
Means		
Out-of-market lenders	.190	.331
In-market lenders	.165***	.356***
In-market-only lenders	.155***	.366***
Medians		
Out-of-market lenders	.173	.315
In-market lenders	.150***	.326*
In-market-only lenders	.133***	.332***

Note: The null hypothesis is that the in-market-only sample is from a population with the same distribution or median as the sample including out-of-market loans; *** (*) indicates rejection at a 1-percent (10-percent) level, based on the Wilcoxon rank-sum or median test.

geographically focused basis of the Federal Reserve's small business lending markets. Too much lending from nearby the market might argue for the expansion of the geographic boundaries of small business lending markets beyond the MSA. And no matter what the quantity of out-of-market lending, its distribution across banks will affect its contribution to competition, as measured by market concentration. In addition, the characteristics of out-of-market loans compared with in-market loans influence how well out-of-market loans might serve as substitutes for in-market loans.

I find that only about 10 percent of the dollar volume of small business loans is held by banks with no physical presence in the local, MSA-based banking market. This relatively small out-of-market share supports the use of a geographically based small business lending market, with the MSA as a reasonable upper bound on its size. However, given that about half of the dollar volume of out-of-market loans seems to come from banks with a physical presence in the same state as the MSA, that upper bound does appear slightly fuzzy. No matter what other data may say about the characteristics of out-of-market loans versus in-market loans, this point alone argues for some consideration being given to out-of-market loans. When these loans are included, as is current practice, market concentration tends to be appreciably lower than when these loans are excluded.

In any case, along most of the lines examined, out-of-market lenders and loans do appear to be quite similar to in-market lenders and loans, further suggesting that out-of-market lending is a good substitute for in-market lending. As a group, out-of-market lenders are considerably larger than in-market lenders, but, on every other count, differences are relatively modest. Out-of-market loans tend to be only a little smaller and a little more likely to be under $100,000 than in-market loans and only slightly less likely to be secured by commercial real estate, while out-of-market lenders are only a little less intensely engaged in small business lending than in-market lenders.

References

Akhavein, Jalal, W. Scott Frame, and Lawrence J. White. 2005. "The Diffusion of Financial Innovation: An Examination of the Adoption of Small Business Credit Scoring by Large Banking Organizations." *Journal of Business* 78(2) pp. 577–596.

Berger, Allen N., and Gregory F. Udell. 2006. "A More Complete Conceptual Framework for SME Finance." *Journal of Banking and Finance* 30(11) pp. 2,945–2,966.

Brevoort, Kenneth P., and Timothy H. Hannan. 2006. "Commercial Lending and Distance: Evidence from Community Reinvestment Act Data." *Journal of Money, Credit, and Banking* 38(8) pp. 1,991–2,012.

DeYoung, Robert, Dennis Glennon, and Peter Nigro. 2006. "Borrower-Lender Distance, Credit Scoring, and the Performance of Small Business Loans." Unpublished manuscript.

Frame, W. Scott, Michael Padhi, and Lynn Woosley. 2004. "Credit Scoring and the Availability of Small Business Credit in Low- and Moderate-Income Areas." *The Financial Review* 39(1) pp. 35–54.

Frame, W. Scott, Aruna Srinivasan, and Lynn Woosley. 2001. "The Effect of Credit Scoring on Small Business Lending." *Journal of Money, Credit, and Banking* 33(3) pp. 813–825.

Krainer, John, and John Beauchamp. 1999. "Small Business Lending Patterns in California." *FRBSF Economic Letter* 99-03 (January 22). http://www.frbsf.org/econrsrch/wklyltr/wklyltr99/el99-03.html

Laderman, Elizabeth S. 1995. "Changes in the Structure of Urban Banking Markets in the West." FRBSF *Economic Review* 95(1), pp. 21–34. http://www.frbsf.org/publications/economics/review/1995/95-1_22-34.pdf

Laderman, Elizabeth S. 2006. "The Geographic Scope of Small Business Lending: Evidence from the San Francisco Market." *FRBSF Economic Letter* 2006-36 (December 15). http://www.frbsf.org/publications/economics/letter/2006/el2006-36.html

Laderman, Elizabeth S., and Steven J. Pilloff. 2007. "Using County-Based Markets to Support and Federal Reserve Markets to Implement Bank Merger Policy." *Journal of Competition Law and Economics* 3(1) pp. 127–148.

Mach, Traci L., and John D. Wolken. 2006. "Financial Services Used by Small Businesses: Evidence from the 2003 Survey of Small Business Finances." *Federal Reserve Bulletin,* A167–A195. http://www.federalreserve.gov/pubs/bulletin/2006/smallbusiness/smallbusiness.pdf

Petersen, Mitchell A., and Raghuram G. Rajan. 2002. "The Information Revolution and Small Business Lending: Does Distance Still Matter?" *Journal of Finance* 57, pp. 2,533–2,570.

Wolken, John D., and Douglas Rohde. 2002. "Changes in the Location of Small Business's Financial Services Suppliers between 1993 and 1998." Federal Reserve Board Memorandum.

Working Papers Series Abstracts

Complete texts of papers in this series are available on the Bank's website at
http://www.frbsf.org/publications/economics/papers/index.php
Electronic subscriptions are available through this site; paper copies are no longer available.

WP 2007-01

Wealth Effects out of Financial and Housing Wealth: Cross Country and Age Group Comparisons

Eva Sierminska, *Luxembourg Income Study and DIW Berlin*
Yelena Takhtamanova, *FRB San Francisco*

To explore the link between household consumption and wealth, we use a new source of harmonized microdata (Luxembourg Wealth Study). We investigate whether there are differences in wealth effects from different types of wealth and across age groups. We consider three countries: Canada, Italy, and Finland. We find that the overall wealth effect from housing is stronger than the effect from financial wealth for the three countries in the sample. Additionally, in accordance with the life-cycle theory of consumption, we find the housing wealth effect to be significantly lower for younger households. We also find between-country differences in the wealth effect.

WP 2007-02

Currency Crises and Foreign Credit in Emerging Markets: Credit Crunch or Demand Effect?

Galina Hale, *FRB San Francisco*
Carlos Arteta, *Federal Reserve Board of Governors*

Currency crises of the past decade highlighted the importance of balance-sheet effects of currency crises. In credit-constrained markets such effects may lead to further declines in credit. Controlling for a host of fundamentals, we find a systematic decline in foreign credit to emerging market private firms of about 25 percent in the first year following currency crises, which we define as large changes in real value of the currency. This decline is especially large in the first five months, lessens in the second year, and disappears entirely by the third year. We identify the effects of currency crises on the demand and supply of credit and find that the decline in the supply of credit is persistent and contributes to about 8 percent decline in credit for the first two years, while the 35 percent decline in demand lasts only five months.

WP 2007-03

Marriage and Divorce: Changes and Their Driving Forces

Betsey Stevenson, *University of Pennsylvania*
Justin Wolfers, *University of Pennsylvania*

We document key facts about marriage and divorce, comparing trends through the past 150 years and outcomes across demographic groups and countries. While divorce rates have risen over the past 150 years, they have been falling for the past quarter century. Marriage rates have also been falling, but more strikingly, the importance of marriage at different points in the life cycle has changed, reflecting rising age at first marriage, rising divorce followed by high remarriage rates, and a combination of increased longevity with a declining age gap between husbands and wives. Cohabitation has also become increasingly important, emerging as a widely used step on the path to marriage. Out-of-wedlock fertility has also risen, consistent with declining "shotgun marriages." Compared with other countries, marriage maintains a central role in American life. We present evidence on some of the driving forces causing these changes in the marriage market: the rise of the birth control pill and women's control over their own fertility; sharp changes in wage structure, including a rise in inequality and partial closing of the gender wage gap; dramatic changes in home production technologies; and the emergence of the internet as a new matching technology. We note that recent changes in family forms demand a reassessment of theories of the family and argue that consumption complementarities may be an increasingly important component of marriage. Finally, we discuss how these facts should inform family policy debates.

WP 2007-04
Monetary Policy in a Small Open Economy with a Preference for Robustness

Richard Dennis, *FRB San Francisco*
Kai Leitemo, *Norwegian School of Management*
Ulf Söderström, *Bocconi University*

We use robust control techniques to study the effects of model uncertainty on monetary policy in an estimated, semistructural, small open-economy model of the U.K. Compared to the closed economy, the presence of an exchange rate channel for monetary policy not only produces new trade-offs for monetary policy, but it also introduces an additional source of specification errors. We find that exchange rate shocks are an important contributor to volatility in the model, and that the exchange rate equation is particularly vulnerable to model misspecification, along with the equation for domestic inflation. However, when policy is set with discretion, the cost of insuring against model misspecification appears reasonably small.

WP 2007-05
Innovations in Mortgage Markets and Increased Spending on Housing

Mark Doms, *FRB San Francisco*
John Krainer, *FRB San Francisco*

Innovations in the mortgage market since the mid-1990s have effectively reduced a number of financing constraints. Coinciding with these innovations, we document a significant change in the propensity for households to own their homes, as well as substantial increases in the share of household income devoted to housing. These changes in housing expenditures are especially large for those groups that faced the greatest financial constraints, and are robust across the changing composition of households and their geographic location. We present evidence that young, constrained households may have used newly designed mortgages to finance their increased expenditures on housing.

WP 2007-06
Productivity Shocks in a Model with Vintage Capital and Heterogeneous Labor

Milton Marquis, *Florida State University*
Bharat Trehan, *FRB San Francisco*

We construct a vintage capital model in which worker skills lie along a continuum and workers can be paired with different vintages (as technology evolves) under a matching rule of "best worker with the best machine." Labor reallocation in response to technology shocks has two key implications for the wage premium. First, it limits both the magnitude and duration of change in the wage premium following a (permanent) embodied technology shock, so empirically plausible shocks do not lead to the kind of increases in the wage premium observed in the United States during the 1980s and early 1990s (though an increase in labor force heterogeneity does). Second, positive disembodied technology shocks tend to push up the wage premium as well, and while this effect is small, it does mean that a higher premium does not provide unambiguous information about the underlying shock. Labor reallocation also means that if embodied technology comes to play a larger role in long-run growth, investment and savings tend to fall in steady state, with little effect on output and employment, enabling the household to increase consumption without sacrificing leisure. The short-run effects are more conventional: permanent shocks to disembodied technology induce a strong wealth effect that reduces savings and induces a consumption boom while permanent shocks to embodied technology induce dominant substitution effects and an expansion characterized by an investment boom.

WP 2007-07
Market Power and Relationships in Small Business Lending

Elizabeth Laderman, *FRB San Francisco*

The empirical research literature regarding the effects of market structure on small business lending has yielded ambiguous results. This paper empirically tests for the presence of countervailing effects of increases in market concentration on small business loan volume. Countervailing effects would be expected if both the traditional structure-conduct-performance (SCP) paradigm of industrial organization and a paradigm whereby market power benefits the formation of lending relationships (the relationship hypothesis) are at work. Using Community Reinvestment Act (CRA) data on

small loans to small businesses, this paper finds that, on average, across MSAs, SCP effects dominate. But, as predicted by the relationship hypothesis, the negative effects of increases in concentration on small business loan volume are weaker, the greater the presence of young firms and the higher the business failure rate. Relationship effects due to business failure appear to come from highly concentrated MSAs. Endogeneity concerns are further addressed with the estimation of a regression that separates out the effects of changes in the number of lenders from the effects of changes in the sum of squared deviations of market shares.

WP 2007-08
Robust Monetary Policy with Imperfect Knowledge

Athanasios Orphanides, *Central Bank of Cyprus*
John C. Williams, *FRB San Francisco*

Published in *Journal of Monetary Economics* 54(5) (July 2007) pp. 1,406–1,435.
See p. 59 for the abstract of this paper.

WP 2007-09
Model Uncertainty and Monetary Policy

Richard Dennis, *FRB San Francisco*

Model uncertainty has the potential to change importantly how monetary policy should be conducted, making it an issue that central banks cannot ignore. In this paper, I use a standard new Keynesian business cycle model to analyze the behavior of a central bank that conducts policy with discretion while fearing that its model is misspecified. I begin by showing how to solve linear-quadratic robust Markov-perfect Stackelberg problems where the leader fears that private agents form expectations using the misspecified model. Next, I exploit the connection between robust control and uncertainty aversion to present and interpret my results in terms of the distorted beliefs held by the central bank, households, and firms. My main results are as follows. First, the central bank's pessimism leads it to forecast future outcomes using an expectations operator that, relative to rational expectations, assigns greater probability to extreme inflation and consumption outcomes. Second, the central bank's skepticism about its model causes it to move forcefully to stabilize inflation following shocks. Finally, even in the absence of misspecification, policy loss can be improved if the central bank implements a robust policy.

WP 2007-10
Rational and Near-Rational Bubbles without Drift

Kevin J. Lansing, *FRB San Francisco*

This paper derives a general class of intrinsic rational bubble solutions in a standard Lucas-type asset pricing model. I show that the rational bubble component of the price-dividend ratio can evolve as a geometric random walk without drift. The volatility of bubble innovations depends exclusively on fundamentals. Starting from an arbitrarily small positive value, the rational bubble expands and contracts over time in an irregular, wholly endogenous fashion, always returning to the vicinity of the fundamental solution. I also examine a near-rational solution in which the representative agent does not construct separate forecasts for the fundamental and bubble components of the asset price. Rather, the agent constructs only a single forecast for the total asset price that is based on a geometric random walk without drift. The agent's forecast rule is parameterized to match the moments of observable data. In equilibrium, the actual law of motion for the price-dividend ratio is stationary, highly persistent, and nonlinear. The agent's forecast errors exhibit near-zero autocorrelation at all lags, making it difficult for the agent to detect a misspecification of the forecast rule. Unlike a rational bubble, the near-rational solution allows the asset price to occasionally dip below its fundamental value. Under mild risk aversion, the near-rational solution generates pronounced low-frequency swings in the price-dividend ratio, positive skewness, excess kurtosis, and time-varying volatility—all of which are present in long-run U.S. stock market data. An independent contribution of the paper is to demonstrate an approximate analytical solution for the fundamental asset price that employs a nonlinear change of variables.

WP 2007-11
Welfare-Maximizing Monetary Policy under Parameter Uncertainty

Rochelle Edge, *Federal Reserve Board of Governors*
Thomas Laubach, *Federal Reserve Board of Governors*
John C. Williams, *FRB San Francisco*

This paper examines welfare-maximizing monetary policy in an estimated micro-founded general equilibrium model of the U.S. economy where the policymaker faces uncertainty about model parameters. Uncertainty about parameters describing preferences and technology implies not only uncer-

tainty about the dynamics of the economy. It also implies uncertainty about the model's utility-based welfare criterion and about the economy's natural rate measures of interest and output. We analyze the characteristics and performance of alternative monetary policy rules given the estimated uncertainty regarding parameter estimates. We find that the natural rates of interest and output are imprecisely estimated. We then show that, relative to the case of known parameters, optimal policy under parameter uncertainty responds less to natural-rate terms and more to other variables, such as price and wage inflation and measures of tightness or slack that do not depend on natural rates.

WP 2007-12

Relative Status and Well-Being: Evidence from U.S. Suicide Deaths

Mary C. Daly, *FRB San Francisco*
Daniel J. Wilson, *FRB San Francisco*
Norman J. Johnson, *U.S. Census Bureau*

This paper empirically assesses the theory of interpersonal income comparison using individual-level data on suicide deaths in the United States. We model suicide as a choice variable, conditional on exogenous risk factors, reflecting an individual's assessment of current and expected future utility. Our empirical analysis considers whether suicide risk is systematically related to the income of others, holding own income and other individual factors fixed. We estimate proportional hazards and probit models of the suicide hazard using two separate and independent data sets: (1) the National Longitudinal Mortality Study and (2) the Detailed Mortality Files combined with the 5 percent Public Use Micro Sample of the 1990 decennial census. Results from both data sources show that, controlling for own income and individual characteristics, individual suicide risk rises with reference group income. This result holds for reference groups defined broadly, such as by county, and more narrowly by county and one demographic marker (e.g., age, sex, race). These findings are robust to alternative specifications and cannot be explained by geographic variation in cost of living, access to emergency medical care, mismeasurement of deaths by suicide, or by bias due to endogeneity of own income. Our results confirm findings using self-reported happiness data and are consistent with models of utility featuring "external habit" or "keeping up with the Joneses" preferences.

WP 2007-13

The Composition of Capital Inflows when Emerging Market Firms Face Financing Constraints

Katherine A. Smith, *U.S. Naval Academy*
Diego Valderrama, *FRB San Francisco*

The composition of capital inflows to emerging market economies tends to follow a predictable dynamic pattern across the business cycle. In most emerging market economies, total inflows are procyclical, with debt and portfolio equity flowing in first, followed later in the expansion by foreign direct investment. To understand the timing of these flows, we use a small open economy framework to model the composition of capital inflows as the equilibrium outcome of emerging market firms' financing decisions. We show how costly external financing and foreign direct investment search costs generate a state-contingent cost of financing, so that the "cheapest" source of financing depends on the phase of the business cycle. In this manner, the financial frictions are able to explain the interaction between the types of flows and deliver a time-varying composition of flows, as well as other standard features of emerging market business cycles. If, as this work suggests, flows are an equilibrium outcome of firms' financing decisions, then volatility of capital inflows is not necessarily "bad" for an economy. Furthermore, using capital controls to shut down one type of flow and encourage another is certain to have both long- and short-run welfare implications.

WP 2007-14

Empirical Analysis of Corporate Credit Lines

Gabriel Jiménez, *Banco de España*
Jose A. Lopez, *FRB San Francisco*
Jesús Saurina, *Banco de España*

Since bank credit lines are a major source of corporate funding and liquidity, we examine the determinants of credit line usage with a database of Spanish corporate credit lines. A line's default status is the primary factor driving its usage, which increases as a firm approaches default. Several lender characteristics suggest an important role for bank monitoring in firms' usage decisions. Credit line usage is found to be inversely related to macroeconomic conditions. Overall, while several factors influence corporate credit line usage, our analysis suggests that default and supply-side variables are the most important.

WP 2007-15
Real Wage Cyclicality in the PSID

Eric T. Swanson, *FRB San Francisco*

Published as "Real Wage Cyclicality in the Panel Study of Income Dynamics" in *Scottish Journal of Political Economy* 54(5) (November 2007) pp. 617–647.
See p. 56 for the abstract of this paper.

WP 2007-16
Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve

Glenn D. Rudebusch, *FRB San Francisco*
John C. Williams, *FRB San Francisco*

We show that professional forecasters have essentially no ability to predict future recessions a few quarters ahead. This is particularly puzzling because, for at least the past two decades, researchers have provided much evidence that the yield curve, specifically the spread between long- and short-term interest rates, does contain useful information at that forecast horizon for predicting aggregate economic activity and, especially, for signaling future recessions. We document this puzzle and suggest that forecasters have generally placed too little weight on yield curve information when projecting declines in the aggregate economy.

WP 2007-17
Do Countries Default in "Bad Times"?

Michael Tomz, *Stanford University*
Mark L.J. Wright, *University of California, Los Angeles*

This paper uses a new data set to study the relationship between economic output and sovereign default for the period 1820–2004. We find a negative but surprisingly weak relationship between output and default. Throughout history, countries have indeed defaulted during bad times (when output was relatively low), but they have also maintained debt service in the face of severe adverse shocks, and they have defaulted when domestic economic conditions were favorable. We show that this constitutes a puzzle for standard theories, which predict a much tighter negative relationship as default provides partial insurance against declines in output.

WP 2007-18
Imperfect Information, Self-Selection, and the Market for Higher Education

Tali Regev, *FRB San Francisco*

This paper explores how the steady trends in increasing tuition costs, college enrollment, and the college wage gap might be related to the quality of college graduates. The model shows that the signaling role of education might be an important yet largely neglected ingredient in these recent changes. I develop a special signaling model in which workers of heterogeneous abilities face the same costs, yet a larger proportion of able individuals self-select to attend college since they are more likely to get higher returns. With imperfect information, the skill premium is an outcome which depends on the equilibrium quality of college attendees and nonattendees. Incorporating a production function of college education, I discuss the properties of the college market equilibrium. A skill-biased technical change directly decreases self-selection into college, but the general equilibrium effect may overturn the direct decline, since increased enrollment and rising tuition costs increase self-selection. Higher initial human capital has an external effect on subsequent investment in school: All agents increase their education, and the higher equilibrium tuition costs increase self-selection and the college premium.

WP 2007-19
Learning and Optimal Monetary Policy

Richard Dennis, *FRB San Francisco*
Federico Ravenna, *University of California, Santa Cruz*

Forthcoming in *Journal of Economic Dynamics and Control.*
See p. 50 for the abstract of this paper.

WP 2007-20
The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models

Jens H.E. Christensen, *FRB San Francisco*
Francis X. Diebold, *University of Pennsylvania*
Glenn D. Rudebusch, *FRB San Francisco*

We derive the class of arbitrage-free affine dynamic term structure models that approximate the widely used Nelson-

Siegel yield-curve specification. Our theoretical analysis relates this new class of models to the canonical representation of the three-factor arbitrage-free affine model. Our empirical analysis shows that imposing the Nelson-Siegel structure on this canonical representation greatly improves its empirical tractability; furthermore, we find that improvements in predictive performance are achieved from the imposition of absence of arbitrage.

WP 2007-21

Regional Economic Conditions and the Variability of Rates of Return in Commercial Banking

Frederick T. Furlong, *FRB San Francisco*
John Krainer, *FRB San Francisco*

We develop new techniques to assess the relationship between commercial bank performance and the economic conditions in the markets in which they operate. In the analysis, we allow for heterogeneity in the responses of banks to regional economic conditions. We find a statistically significant relationship between bank performance and shocks to the regional markets in which they operate. We find that region-specific shocks have a significant and persistent effect on the cross-sectional variance of bank performance in the market. That is, shocks affecting average performance of banks in a region also tend to increase the dispersion of their performance. We demonstrate that this effect is due to heterogeneity in the banks' exposures to their regional economies. Moreover, by allowing for this heterogeneity, we find that systematic responses to regional economic effects are notably more important in explaining the variation in bank performance than suggested by analysis in which responses are constrained to be the same for all banks.

WP 2007-22

Determinants of Access to External Finance: Evidence from Spanish Firms

Raquel Lago Gonzáles, *Banco de España*
Jose A. Lopez, *FRB San Francisco*
Jesús Saurina, *Banco de España*

Access to external finance is a key determinant of a firm's ability to develop, operate, and expand. To date, the literature has examined a variety of macroeconomic and microeconomic factors that influence firm financing. In this paper, we examine access by Spanish firms to external financing, both from bank and nonbank sources. We use dynamic panel data estimation techniques to estimate our models over a sample of 60,000 firms during the period from 1992 to 2002. We find that Spanish firms are quite dependent on short-term nonbank financing (such as trade credit), which makes up about 65 percent of total firm debt. Our results indicate that this type of financing is less sensitive to firm characteristics than short-term bank financing. However, we also find that short-term bank debt seems to be accessed more during economic expansions, which may suggest a substitution away from nonbank financing as firm conditions improve. Short-term bank debt also seems to be accessed more as funding rates rise, possibly again suggesting a substitution away from higher priced nonbank alternatives. Using data from the Spanish Credit Register maintained by the Banco de España, we find that the impact of funding costs on access to external financing, whether from banks or nonbanks, is affected by the nature of borrowing firms' bank relationships and collateral. In particular, we provide evidence of a potential hold-up problem in loan markets. Moreover, collateral plays a key role in making long-term finance available to firms.

WP 2007-23

How Does Competition Impact Bank Risk-Taking?

Gabriel Jiménez, *Banco de España*
Jose A. Lopez, *FRB San Francisco*
Jesús Saurina, *Banco de España*

A common assumption in the academic literature and in the actual supervision of banking systems worldwide is that franchise value plays a key role in limiting bank risk-taking. As the underlying source of franchise value is assumed to be market power, reduced competition has been considered to promote banking stability. Boyd and De Nicolo (2005) propose an alternative view where concentration in the loan market could lead to increased borrower debt loads and a corresponding increase in loan defaults that undermine bank stability. Martinez-Miera and Repullo (2007) encompass both approaches by proposing a nonlinear relationship between competition and bank risk-taking. Using unique data sets for the Spanish banking system, we examine the empirical nature of that relationship. After controlling for macroeconomic conditions and bank characteristics, we find that standard measures of market concentration do not affect the ratio of nonperforming commercial loans (NPL), our measure of bank risk. However, using Lerner indexes based on bank-specific interest rates, we find a negative relationship

between loan market power and bank risk. This result provides evidence in favor of the franchise value paradigm.

WP 2007-24
Convergence and Anchoring of Yield Curves in the Euro Area

Michael Ehrmann, *European Central Bank*
Marcel Fratzscher, *European Central Bank*
Refet S. Gürkaynak, *Bilkent University*
Eric T. Swanson, *FRB San Francisco*

We study the convergence of European bond markets and the anchoring of inflation expectations in euro-area countries using high-frequency bond yield data for France, Germany, Italy, and Spain. We find that the Economic and Monetary Union (EMU) has led to substantial convergence in euro-area sovereign bond markets in terms of interest rate levels, unconditional daily fluctuations, and conditional responses to major macroeconomic data announcements. Our findings also suggest a substantial increase in the anchoring of long-term inflation expectations since EMU, particularly for Italy and Spain, which since monetary union have seen their long-term interest rates become much lower, much less volatile, and much better anchored in response to news. Finally, the reaction of far-ahead forward interest rates to macroeconomic announcements has converged substantially across euro-area countries and even been eliminated over time, thus underlining not only market integration but also the credibility that financial markets attach to monetary policy in the euro area.

WP 2007-25
Examining the Bond Premium Puzzle with a DSGE Model

Glenn D. Rudebusch, *FRB San Francisco*
Eric T. Swanson, *FRB San Francisco*

The basic inability of standard theoretical models to generate a sufficiently large and variable nominal bond risk premium has been termed the "bond premium puzzle." We show that the term premium on long-term bonds in the canonical dynamic stochastic general equilibrium (DSGE) model used in macroeconomics is far too small and stable relative to the data. We find that introducing long-memory habits in consumption as well as labor market frictions can help fit the term premium, but only by seriously distorting the DSGE model's ability to fit other macroeconomic variables, such as the real wage; therefore, the bond premium puzzle remains.

WP 2007-26
Pricing-to-Market, Trade Costs, and International Relative Prices

Andrew Atkeson, *University of California, Los Angeles*
Ariel Burstein, *University of California, Los Angeles*

Data on international relative prices from industrialized countries show large and systematic deviations from relative purchasing power parity. We embed a model of imperfect competition and variable markups in some of the recently developed quantitative models of international trade to examine whether such models can reproduce the main features of the fluctuations in international relative prices. We find that when our model is parameterized to match salient features of the data on international trade and market structure in the United States, it reproduces deviations from relative purchasing power parity similar to those observed in the data because firms choose to price-to-market. We then examine how pricing-to-market depends on the presence of international trade costs and various features of market structure.

WP 2007-27
Productivity and the Dollar

Giancarlo Corsetti, *European University Institute*
Luca Dedola, *European Central Bank*
Sylvain Leduc, *Federal Reserve Board of Governors*

This paper investigates the role of shocks to U.S. productivity and demand in driving the real value of the dollar, and the dynamics of the U.S. trade balance. Using sign restrictions based on robust predictions by standard theory, we identify shocks that increase domestic labor productivity and output in manufacturing (our measure of U.S. tradables), relative to an aggregate of other industrial countries including the rest of the G-7, while driving down (up in the case of demand) the relative price of tradables (in accord with Harrod-Balassa-Samuelson effects). Consistent with previous results based on different methodologies, we find that positive productivity differentials raise U.S. consumption and investment relative to the rest of the world and deteriorate net exports; both the U.S. real exchange rate and the U.S. terms of trade appreciate in response to these shocks. Demand shocks also appreciate the dollar but have negligible effects on absorption

and the trade balance. These findings question a common view in the literature, that a country's terms of trade deteriorate when its tradables supply grows, providing a mechanism to contain differences in national wealth even if productivity levels do not converge. They also provide an empirical contribution to the current debate on the adjustment of the U.S. current account position. Contrary to widespread presumptions, productivity growth in the U.S. tradable sector does not necessarily improve the U.S. trade deficit, nor deteriorate the U.S. terms of trade, at least in the short and medium run.

WP 2007-28

The Determinants of Household Saving in China: A Dynamic Panel Analysis of Provincial Data

Charles Horioka, *Osaka University*
Junmin Wan, *Osaka University*

In this paper, we conduct a dynamic panel analysis of the determinants of the household saving rate in China using a life-cycle model and panel data on Chinese provinces for the 1995–2004 period from China's household survey. We find that China's household saving rate has been high and rising and that the main determinants of variations over time and over space therein are the lagged saving rate, the income growth rate, and (in some cases) the real interest rate and the inflation rate. However, we find that the variables relating to the age structure of the population usually do not have a significant impact on the household saving rate. These results provide mixed support for the life-cycle hypothesis as well as the permanent income hypothesis, are consistent with the existence of inertia or persistence, and imply that China's household saving rate will remain high for some time to come.

WP 2007-29

Optimal Reserve Management and Sovereign Debt

Laura Alfaro, *Harvard Business School*
Fabio Kanczuk, *Universidade de São Paulo*

Most models currently used to determine optimal foreign reserve holdings take the level of international debt as given. However, given the sovereign's willingness-to-pay incentive problems, reserve accumulation may reduce sustainable debt levels. In addition, assuming constant debt levels does not allow addressing one of the puzzles behind using reserves as a means to avoid the negative effects of crisis: why don't sovereign countries reduce their sovereign debt instead? To study the joint decision of holding sovereign debt and reserves, we construct a stochastic dynamic equilibrium model calibrated to a sample of emerging markets. We obtain that the optimal policy is not to hold reserves at all. This finding is robust to considering interest rate shocks, sudden stops, contingent reserves, and reserve-dependent output costs.

WP 2007-30

Financial Integration in East Asia

Hiroshi Fujiki, *Bank of Japan*
Akiko Terada-Hagiwara, *Bank of Japan*

This paper examines the degree of integration into world financial markets and the impacts on several key macroeconomic variables of selected East Asian economies and draws policy implications. According to our analysis, the degrees of integration into world financial markets in those economies are increasing. Regarding the impacts of increasing integration into world financial markets on several macroeconomic variables, we find three results. First, casual two-way plots among macroeconomic variables do not support the theoretical prediction of reduction in relative consumption volatility. Second, the saving–investment correlation is higher than those in euro-area economies. Third, the degrees of smoothing of idiosyncratic shock by cross-holding of financial assets are lower than in euro-area economies. Those results suggest two policy implications. First, there's some room for improvement in welfare gains in those economies by further risk sharing. Second, holding all other conditions given, the increasing integration into world financial markets alone is unlikely to provide a sound ground for a currency union in East Asia at this stage.

WP 2007-31

Capital Controls: Myth and Reality, A Portfolio Balance Approach to Capital Controls

Nicolas Magud, *University of Oregon*
Carmen Reinhart, *University of Maryland*
Kenneth Rogoff, *Harvard University*

The literature on capital controls has (at least) four very serious apples-to-oranges problems: (1) There is no unified theoretical framework to analyze the macroeconomic consequences

of controls; (2) there is significant heterogeneity across countries and time in the control measures implemented; (3) there are multiple definitions of what constitutes a "success"; and (4) the empirical studies lack a common methodology—furthermore, these are significantly "overweighted" by a couple of country cases (Chile and Malaysia). In this paper, we attempt to address some of these shortcomings by being very explicit about what measures are construed as capital controls. Also, given that success is measured so differently across studies, we sought to "standardize" the results of over 30 empirical studies we summarize in this paper. The standardization was done by constructing two indices of capital controls: Capital Controls Effectiveness Index (CCE Index), and Weighted Capital Control Effectiveness Index (WCCE Index). The difference between them lies in that the WCCE controls for the differentiated degree of methodological rigor applied to draw conclusions in each of the considered papers. As much as possible, we bring to bear the experiences of less well-known episodes than those of Chile and Malaysia. Then, using a portfolio balance approach, we model the effects of imposing short-term capital controls. We find that there should exist country-specific characteristics for capital controls to be effective. From this simple perspective, this rationalizes why some capital controls were effective and some were not. We also show that the equivalence in effects of price- vs. quantity-capital control are conditional on the level of short-term capital flows.

WP 2007-32
Capital Account Liberalization: Theory, Evidence, and Speculation

Peter Henry, *Stanford University*

Writings on the macroeconomic impact of capital account liberalization find few, if any, robust effects of liberalization on real variables. In contrast to the prevailing wisdom, I argue that the textbook theory of liberalization holds up quite well to a critical reading of this literature. The lion's share of papers that find no effect of liberalization on real variables tell us nothing about the empirical validity of the theory because they do not really test it. This paper explains why it is that most studies do not really address the theory they set out to test. It also discusses what is necessary to test the theory and examines papers that have done so. Studies that actually test the theory show that liberalization has significant effects on the cost of capital, investment, and economic growth.

WP 2007-33
Subprime Mortgage Delinquency Rates

Mark Doms, *FRB San Francisco*
Frederick T. Furlong, *FRB San Francisco*
John Krainer, *FRB San Francisco*

We evaluate the importance of three different channels for explaining the recent performance of subprime mortgages. First, the riskiness of the subprime borrowing pool may have increased. Second, pockets of regional economic weakness may have helped push a larger proportion of subprime borrowers into delinquency. Third, for a variety of reasons, the recent history of local house price appreciation and the degree of house price deceleration may have affected delinquency rates on subprime mortgages. While we find a role for all three candidate explanations, patterns in recent house price appreciation are far and away the best single predictor of delinquency levels and changes in delinquencies. Importantly, after controlling for the current level of house price appreciation, measures of house price deceleration remain significant predictors of changes in subprime delinquencies. The results point to a possible role for changes in house price expectations for explaining changes in delinquencies.

Abstracts of Articles Accepted in Journals, Books, and Conference Volumes*

Government Consumption Expenditures and the Current Account

Michele Cavallo

Published in *Public Finance & Management* 7(1) (2007).

This paper focuses on the effects on the current account of changes to two distinct components of government consumption expenditures, expenditure on goods and expenditure on hours worked. I find that changes to government expenditure on hours do not directly affect the current account and that their effect is considerably smaller—one order of magnitude—than the effect of changes to government expenditure on goods. These findings indicate that considering government consumption as entirely expenditure on goods can lead to overestimating its role in accounting for movements in the current account balance.

Cross-National Trends in Earnings Inequality and Instability

Mary C. Daly
Robert G. Valletta

Forthcoming in *Economics Letters*.

We compare trends in earnings inequality in the United States, Germany, and Great Britain. Estimation of a heterogeneous growth model of permanent and transitory earnings variation reveals substantial convergence in the permanent component of inequality in these countries during the 1990s.

Optimal Policy in Rational Expectations Models: New Solution Algorithms

Richard Dennis

Published in *Macroeconomic Dynamics* 11(1) (February 2007) pp. 31–55.

This paper develops methods to solve for optimal discretionary policies and optimal commitment policies in rational expectations models. These algorithms, which allow the optimization constraints to be conveniently expressed in second-order structural form, are more general than existing methods and are simple to apply. We use several New Keynesian business cycle models to illustrate their application. Simulations show that the procedures developed in this paper can quickly solve small-scale models and that they can be usefully and effectively applied to medium- and large-scale models.

*The abstracts are arranged alphabetically by FRB San Francisco authors, whose names are in boldface.

Robust Control with Commitment: A Modification to Hansen-Sargent

Richard Dennis

Forthcoming in *Journal of Economic Dynamics and Control.*

This paper examines the Hansen and Sargent (2003) formulation of the robust Stackelberg problem and shows that their method of constructing the approximating equilibrium, which is central to any robust control exercise, is generally invalid. The paper then turns to the Hansen and Sargent (2007) treatment, which, responding to the problems raised in this paper, changes subtly, but importantly, how the robust Stackelberg problem is formulated. This paper proves, first, that their method for obtaining the approximating equilibrium is now equivalent to the one developed in this paper, and, second, that the worst-case specification errors are not subject to a time-inconsistency problem. Analyzing robust monetary policy in two New Keynesian business cycle models, the paper demonstrates that a robust central bank should primarily fear that the supply side of its approximating model is misspecified and that attenuation characterizes robust policymaking. Depending on how the robust Stackelberg problem is formulated, this paper shows that the Hansen-Sargent approximating equilibrium can be dynamically unstable and that robustness can be irrelevant, i.e., that the robust policy can coincide with the rational expectations policy.

Learning and Optimal Monetary Policy

Richard Dennis, with Federico Ravenna, *University of California, Santa Cruz*

Forthcoming in *Journal of Economic Dynamics and Control.*

To conduct policy efficiently, central banks must use available data to infer, or learn, the relevant structural relationships in the economy. However, because a central bank's policy affects economic outcomes, the chosen policy may help or hinder its efforts to learn. This paper examines whether real-time learning allows a central bank to learn the economy's underlying structure and studies the impact that learning has on the performance of optimal policies under a variety of learning environments. Our main results are as follows. First, when monetary policy is formulated as an optimal discretionary targeting rule, we find that the rational expectations equilibrium and the optimal policy are real-time learnable. This result is robust to a range of assumptions concerning private sector learning behavior. Second, when policy is set with discretion, learning can lead to outcomes that are better than if the model parameters are known. Finally, if the private sector is learning, then unannounced changes to the policy regime, particularly changes to the inflation target, can raise policy loss considerably.

Trend Breaks, Long-Run Restrictions, and Contractionary Technology Improvements

John G. Fernald

Published in *Journal of Monetary Economics* 54(8) (November 2007) pp. 2,467–2,485.

Structural vector autoregressions with long-run restrictions are extraordinarily sensitive to low-frequency correlations. Recent literature finds that the estimated effects of technology shocks are sensitive to how one treats hours per capita. However, after allowing for (statistically and economically significant) trend breaks in productivity, results are much less sensitive: hours fall when technology improves. The issue is that the common high-low-high pattern of productivity growth and hours (i.e., the low-frequency correlation) inevitably leads to a positive estimated response. The trend breaks control for this correlation. This example suggests a practical need for care in using long-run restrictions.

Information and Communications Technology as a General Purpose Technology: Evidence from U.S. Industry Data

John G. Fernald, with
Susanto Basu, *Boston College*

Published in *German Economic Review* 8(2) (May 2007) pp. 146–173.

Many people point to information and communications technology (ICT) as the key for understanding the acceleration in productivity in the United States since the mid-1990s. Stories of ICT as a general purpose technology (GPT) suggest that measured total factor productivity (TFP) should rise in ICT-using sectors (reflecting either unobserved accumulation of intangible organizational capital, spillovers, or both), but with a long lag. Contemporaneously, however, investments in ICT may be associated with lower TFP as resources are diverted to reorganization and learning. We find that U.S. industry results are consistent with GPT stories: the acceleration after the mid-1990s was broad-based—located primarily in ICT-using industries rather than ICT-producing industries. Furthermore, industry TFP accelerations in the 2000s are positively correlated with (appropriately weighted) industry ICT capital growth in the 1990s. Indeed, as GPT stories would suggest, after controlling for past ICT investment, industry TFP accelerations are negatively correlated with increases in ICT usage in the 2000s.

This article is published in this volume, pp. 1–15.

Global Price Dispersion: Are Prices Converging or Diverging?

Reuven Glick, with
Paul Bergin,
University of California, Davis

Published in *Journal of International Money and Finance* 26(5) (September 2007) pp. 703–729.

This paper documents significant time variation in the degree of global price convergence over the last two decades. In particular, there appears to be a general U-shaped pattern with price dispersion first falling and then rising in recent years, a pattern which is remarkably robust across country groupings and commodity groups. This time-variation is difficult to explain in terms of the standard gravity equation variables common in the literature, as these tend not to vary much over time or have not risen in recent years. However, regression analysis indicates that this time-varying pattern coincides well with oil price fluctuations, which are clearly time-varying and have risen substantially since the late 1990s. As a result, this paper offers new evidence on the role of transportation costs in driving international price dispersion.

A Model of Endogenous Nontradability and Its Implications for the Current Account

Reuven Glick, with
Paul Bergin,
University of California, Davis

Published in *Review of International Economics* 15(5) (November 2007) pp. 916–931.

This paper studies how nontraded goods limit the ability of a country to finance current account deficits. It uses an intertemporal model of the current account for a small open economy where goods are endogenously nontraded due to explicit trade costs. The economy has an endowment of two goods with differing trade costs, either of which can be traded or nontraded in equilibrium. The model implies that current account deficits impose a cost, in the form of raising the effective interest rate in the country. The findings differ from some recent studies: first, in that the interest rate rises even for countries with modest current account deficits; secondly, the interest rate cost eventually reaches an upper bound as current account deficits grow, and progressively more nontraded goods become traded to service the debt. Panel regression analysis of interest rate and current account data are consistent with our conclusions.

Tradability, Productivity, and International Economic Integration

Reuven Glick, with
Paul Bergin,
University of California, Davis

Published in *Journal of International Economics* 73(1) (September 2007) pp. 128–151.

This paper develops a model of endogenously tradable goods to study the implications of international integration for price dispersion and pricing to market. A distinctive feature of the model is heterogeneity in both trade costs and productivity. The model highlights the role of heterogeneity in shaping how new entrants at the extensive margin differ from incumbent traders, thereby giving extensive margin movements distinctive implications relative to the intensive margin. In particular, the model predicts that international integration mainly along the extensive margin should be associated with a more limited degree of price convergence. This prediction finds support in cross-sectional regressions on European data and offers insight into recent integration episodes.

Bonds or Loans? The Effect of Macroeconomic Fundamentals

Galina B. Hale

Published in *The Economic Journal* 117(516) (January 2007) pp. 196–215.

The costs of debt crises are not invariant to the foreign debt instrument composition: bank loans or bonds. The lending boom of the 1990s witnessed considerable variation over time and across countries in the debt instrument used by emerging market (EM) borrowers. This article tests how macroeconomic fundamentals affect the composition of international debt instruments used by EM borrowers. Analysis of micro-level data using an ordered probability model shows that macroeconomic fundamentals explain a significant share of variation in the ratio of bonds to loans for private borrowers but not for the sovereigns.

Sovereign Debt Crises and Credit to the Private Sector

Galina B. Hale, with
Carlos Arteta,
Federal Reserve Board of Governors

Published in *Journal of International Economics* 74(1) (January 2008) pp. 53–69.

We use micro-level data to analyze emerging markets' private sector access to international debt markets during sovereign debt crises. We find that these crises are systematically accompanied by a decline in foreign credit to domestic private firms, both during debt renegotiations and for over two years after restructuring agreements are reached. This decline is large, statistically significant, and robust. We find that this effect is concentrated in the nonfinancial sector and is different for firms in the exporting and in the non-exporting sectors. We also find that the magnitude of the effect depends on the type of debt restructuring agreement.

The Decision to First Enter the Public Bond Market: The Role of Firm Reputation, Funding Choices, and Bank Relationships

Galina B. Hale, with
João Santos, *FRB New York*

Forthcoming in
Journal of Banking and Finance.

This paper uses survival analysis to investigate the timing of a firm's decision to issue for the first time in the public bond market. We find that firms that are more creditworthy and have higher demand for external funds issue their first public bond earlier. We also find that issuing private bonds or taking out syndicated loans is associated with a faster entry to the public bond market. According to our results, the relationships that firms develop with investment banks in connection with their private bond issues and syndicated loans further speed up their entry to the public bond market. Finally, we find that a firm's reputation has a U-shaped effect on the timing of a firm's bond IPO. Consistent with Diamond's (1991) reputational theory, firms that establish a track record of high creditworthiness as well as those that establish a track record of low creditworthiness enter the public bond market earlier than firms with intermediate reputation.

Using Securities Market Information for Bank Supervisory Monitoring

John Krainer
Jose A. Lopez

Forthcoming in
International Journal of Central Banking.

Bank supervisors in the United States conduct comprehensive on-site inspections of bank holding companies (BHCs) and assign them a supervisory rating meant to summarize their overall condition. We develop an empirical forecasting model of these ratings that combines supervisory and securities market data. We find that securities market variables, such as BHC stock returns and bond yield spreads, improve the model's in-sample fit. We also find that debt market variables provide more information on supervisory ratings for BHCs closer to default, while equity market variables provide more information for those further from default. In out-of-sample forecasting, we find that the accuracy of the model with both equity and debt variables is little different from the accuracy of a model based on supervisory information alone. However, the model with securities market data identifies additional ratings downgrades, which supervisors would probably value enough to warrant the use of this extended model for off-site monitoring purposes.

Using County-Based Markets to Support and Federal Reserve Markets to Implement Bank Merger Policy

Elizabeth S. Laderman, with
Steven J. Pilloff, *Hood College*

Published in *Journal of Competition Law and Economics* 3(1) (March 2007) pp. 127–148.

In this paper, we consider three issues raised by the apparent inconsistency between the current research practice of using county-based markets (metropolitan statistical areas (MSAs) and non-MSA counties) to investigate the validity of the theoretical underpinnings of bank merger policy and the current regulatory practice of using Federal Reserve (FR) banking markets, which often do not follow county lines, to implement that policy. Using a national sample of bank and thrift branch deposit data, we find that county-based areas cannot simply substitute for FR markets in the implementation of bank merger policy. For example, numerous potential mergers would raise competitive issues in county-based areas, but not in FR markets, and vice versa. We also conclude that, because of the relative difficulty of assembling demographic data for non-county-based areas, it is impractical to consistently use FR markets in bank merger policy research. However, we do find that, despite the inconsistencies between the two types of markets, analysis that uses county-based areas is relevant for bank merger policy that is implemented with FR markets. For example, we find that profitability regression results using variables based on FR markets are similar to those found using variables based on MSAs and non-MSA counties.

Maintenance Expenditures and Indeterminacy under Increasing Returns to Scale

Kevin J. Lansing, with
Jang-Ting Guo,
University of California, Riverside

Published in *International Journal of Economic Theory* 3(2) (June 2007) pp. 147–158.

This paper develops a one-sector real business cycle model in which competitive firms allocate resources for the production of goods, investment in new capital, and maintenance of existing capital. Firms also choose the utilization rate of existing capital. A higher utilization rate leads to faster capital depreciation, and an increase in maintenance activity has the opposite effect. We show that as the equilibrium ratio of maintenance expenditures to GDP rises, the required degree of increasing returns for local indeterminacy declines over a wide range of parameter combinations. When the model is calibrated to match empirical evidence on the relative size of maintenance and repair activity, we find that local indeterminacy (and belief-driven fluctuations) can occur with a mild and empirically plausible degree of increasing returns: approximately 1.08.

Foreign Bank Lending and Bond Underwriting in Japan During the Lost Decade

Jose A. Lopez
Mark M. Spiegel

Forthcoming in *Proceedings of the 2006 Asian Pacific Economic Association Conference,* eds. Kar Yiu Wong and Yin Wong Cheung. London: Routledge.

We examine foreign intermediation activity in Japan during the so-called "lost decade" of the 1990s, contrasting the behavior of lending by foreign commercial banks and underwriting activity by foreign investment banks over that period. Foreign bank lending is shown to be sensitive to domestic Japanese conditions, particularly Japanese interest rates, more so than their domestic Japanese bank counterparts. During the 1990s, foreign bank lending in Japan fell, both in overall numbers and as a share of total lending. However, there was marked growth in foreign underwriting activity in the international yen-denominated bond sector. A key factor in the disparity between these activities is their different clientele: While foreign banks in Japan lent primarily to domestic borrowers, international yen-denominated bond issuers were primarily foreign entities with yen funding needs or opportunities for profitable swaps. Indeed, low interest rates that discouraged lending activity in Japan by foreign banks directly encouraged foreign underwriting activity tied to the so-called "carry trades." Regulatory reforms, particularly the "Big Bang" reforms of the 1990s, also play a large role in the growth of foreign underwriting activity over our sample period.

Macroeconomic Implications of Changes in the Term Premium

Glenn D. Rudebusch, and
Eric T. Swanson, with
Brian P. Sack, *Macroeconomic Advisers*

Published in the *FRB St. Louis Review* 89(4) (July/August 2007) pp. 241–269.

Linearized New Keynesian models and empirical no-arbitrage macro-finance models offer little insight regarding the implications of changes in bond term premiums for economic activity. This paper investigates these implications using both a structural model and a reduced-form framework. The authors show that there is no structural relationship running from the term premium to economic activity, but a reduced-form empirical analysis does suggest that a decline in the term premium has typically been associated with stimulus to real economic activity, which contradicts earlier results in the literature.

Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models

Glenn D. Rudebusch, with
Tao Wu, *FRB Dallas*

Published in *Journal of Money, Credit, and Banking* 39(2–3) (March–April 2007) pp. 395–422.

This paper examines a shift in the dynamics of the term structure of interest rates in the U.S. during the mid-1980s. We document this shift using standard interest rate regressions and using dynamic, affine, no-arbitrage models estimated for the pre- and post-shift subsamples. The term structure shift largely appears to be the result of changes in the pricing of risk associated with a "level" factor. Using a macro-finance model, we suggest a link between this shift in term structure behavior and changes in the dynamics and risk pricing of the Federal Reserve's inflation target as perceived by investors.

A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy

Glenn D. Rudebusch, with
Tao Wu, *FRB Dallas*

Forthcoming in *The Economic Journal*.

This paper develops and estimates a macro-finance model that combines a canonical affine no-arbitrage finance specification of the term structure with standard macroeconomic aggregate relationships for output and inflation. From this new empirical formulation, we obtain several important results: (1) the latent term structure factors from finance no-arbitrage models appear to have important macroeconomic and monetary policy underpinnings, (2) there is no evidence of monetary policy inertia or a slow partial adjustment of the policy interest rate by the Federal Reserve, and (3) both forward-looking and backward-looking elements play important roles in macroeconomic dynamics.

Offshore Financial Centres: Parasites or Symbionts?

Mark M. Spiegel, with
Andrew Rose,
University of California, Berkeley

Published in *The Economic Journal* 117(523) (October 2007) pp. 1,310–1,335.

This article analyses the causes and consequences of offshore financial centres (OFCs). While OFCs are likely to encourage bad behavior in source countries, they may also have unintended positive consequences, such as providing competition for the domestic banking sector. We derive and simulate a model of a home country monopoly bank facing a representative competitive OFC which offers tax advantages attained by moving assets offshore at a cost that is increasing in distance to the OFC. Our model predicts that proximity to an OFC is likely to be pro-competitive. We test and confirm the predictions empirically. OFC proximity is associated with a more competitive domestic banking system and greater overall financial depth.

Market Price Accounting and Depositor Discipline: The Case of Japanese Regional Banks

Mark M. Spiegel, with
Nobuyoshi Yamori, *Nagoya University*

Published in *Journal of Banking and Finance* 31(3) (March 2007) pp. 769–786.

We examine the determinants of Japanese regional bank pricing-to-market decisions and their impact on the intensity of depositor discipline, in the form of the sensitivity of deposit growth to bank financial conditions. To obtain consistent estimates, we first model and estimate the bank pricing-to-market decision and then estimate the intensity of depositor discipline after conditioning for that decision. We find that banks were less likely to adopt market price accounting the larger were their unrealized securities losses. We also find statistically significant evidence of depositor discipline among banks that elected to price to market. Our results indicate that depositor discipline was more intense for the subset of banks that adopted market price accounting.

Real Wage Cyclicality in the Panel Study of Income Dynamics

Eric T. Swanson

Published in *Scottish Journal of Political Economy* 54(5) (November 2007) pp. 617–647.

Previous studies of real wage cyclicality have made only sparing use of the micro-data detail that is available in the Panel Study of Income Dynamics (PSID). The present paper brings to bear this additional detail to investigate the robustness of the previous results and to examine whether there are important cross-sectional and demographic differences in wage cyclicality. Although real wages were procyclical across the entire distribution of workers from 1967 to 1991, the wages of lower-income, younger, and less-educated workers exhibited greater procyclicality. However, workers' straight-time hourly pay rates have been acyclical, suggesting that more variable pay margins such as bonuses, overtime, late shift premia, and commissions have played a substantial if not primary role in generating procyclicality.

Inflation Targeting and the Anchoring of Inflation Expectations in the Western Hemisphere

Eric T. Swanson, with
Refet S. Gürkaynak, *Bilkent University*
Andrew T. Levin, *Federal Reserve Board*
Andrew N. Marder, *Princeton University*

Published in *Series on Central Banking, Analysis, and Economic Policies X: Monetary Policy under Inflation Targeting,* eds. F. Mishkin and K. Schmidt-Hebbel. Santiago, Chile: Central Bank of Chile, 2007.

We investigate the extent to which long-run inflation expectations are well anchored in three western hemisphere countries—Canada, Chile, and the United States—using a high-frequency event-study analysis. Specifically, we use daily data on far-ahead forward inflation compensation—the difference between forward rates on nominal and inflation-indexed bonds—as an indicator of financial market perceptions of inflation risk and the expected level of inflation at long horizons. For the United States, we find that far-ahead forward inflation compensation reacts significantly to macroeconomic data releases, implying that long-run inflation expectations are not completely anchored. In contrast, the Canadian inflation compensation data do not exhibit significant sensitivity to either Canadian or U.S. macroeconomic news, confirming that inflation targeting in Canada has been successful in anchoring long-run inflation expectations. Finally, while the requisite data for Chile is only available for a limited sample period (2002–2005), our results are consistent with the hypothesis that inflation targeting in Chile has also succeeded in anchoring long-run inflation expectations.

Market-Based Measures of Monetary Policy Expectations

Eric T. Swanson, with
Refet S. Gürkaynak, *Bilkent University*
Brian P. Sack, *Macroeconomic Advisers*

Published in *Journal of Business and Economic Statistics* 25(2) (April 2007) pp. 201–212.

A number of recent articles have used different financial market instruments to measure near-term expectations of the federal funds rate and the high-frequency changes in these instruments around Federal Open Market Committee announcements to measure monetary policy shocks. This article evaluates the empirical success of a variety of financial market instruments in predicting the future path of monetary policy. All of the instruments we consider provide forecasts that are clearly superior to those of standard time series models at all of the horizons considered. Among financial market instruments, we find that federal funds futures dominate all the other securities in forecasting monetary policy at horizons out to six months. For longer horizons, the predictive power of many of the instruments we consider is very similar. In addition, we present evidence that monetary policy shocks computed using the current-month federal funds futures contract are influenced by changes in the timing of policy actions that do not influence the expected course of policy beyond a horizon of about six weeks. We propose an alternative shock measure that captures changes in market expectations of policy over slightly longer horizons.

Futures Prices as Risk-Adjusted Forecasts of Monetary Policy

Eric T. Swanson, with
Monika Piazzesi, *University of Chicago*

Forthcoming in
Journal of Monetary Economics.

Many researchers have used federal funds futures rates as measures of financial markets' expectations of future monetary policy. However, to the extent that federal funds futures reflect risk premia, these measures require some adjustment. In this paper, we document that excess returns on federal funds futures have been positive on average and strongly countercyclical. In particular, excess returns are surprisingly well predicted by macroeconomic indicators such as employment growth and financial business cycle indicators such as Treasury yield spreads and corporate bond spreads. Excess returns on eurodollar futures display similar patterns. We document that simply ignoring these risk premia significantly biases forecasts of the future path of monetary policy. We also show that risk premia matter for some futures-based measures of monetary policy shocks used in the literature.

On Using Relative Prices to Measure Capital Specific Technological Progress

Bharat Trehan, with
Milton Marquis, *Florida State University*

Forthcoming in
Journal of Macroeconomics.

Recently, Greenwood, Hercowitz, and Krusell (GHK) have identified the relative price of (new) capital with capital-specific technological progress. In a two-sector growth model, however, the relative price of capital equals the ratio of the productivity processes in the two sectors. Restrictions from this model are used with data on wages and prices to construct measures of productivity growth and test the GHK identification, which is easily rejected by the data. This raises questions about various measures of the contribution that capital-specific technological progress might make to the economy. This identification also induces a negative correlation between the resulting measures of capital-specific and economy-wide technological change, which potentially explains why papers employing this identification find that capital-specific technological change accelerated in the mid-1970s. We impose structure on the productivity measures based on their long-run behavior and find evidence of a slowdown in productivity in the 1970s that is common to both sectors and an acceleration in the mid-1990s that is exclusive to the capital sector.

Statistical Nonlinearities in the Business Cycle: A Challenge for the Canonical RBC Model

Diego Valderrama

Published in *Journal of Economic Dynamics and Control* 31(9) (September 2007) pp. 2,957–2,983.

The cyclical components of U.S. macroeconomic time series exhibit significant nonlinearities. Standard equilibrium models of business cycles cannot replicate nonlinear features of the data. Applying the efficient method of moments (Gallant and Tauchen, 1996, *Econometric Theory*) to build an algorithm that searches over the model's parameter space establishes the parameterization that best allows replication of all statistical properties of the data. The results show that under this parameterization, the model captures nonlinearities in investment but fails to account for observed properties of consumption.

Learning and Shifts in Long-Run Productivity Growth

John C. Williams, with Rochelle Edge, *Federal Reserve Board of Governors* Thomas Laubach, *Federal Reserve Board of Governors*

Published in *Journal of Monetary Economics* 54(8) (November 2007) pp. 2,421–2,438.

An extensive literature has analyzed the macroeconomic effects of shocks to the level of aggregate productivity; however, there has been little corresponding research on sustained shifts in the growth rate of productivity. In this paper, we examine the effects of shocks to productivity growth in a dynamic general equilibrium model where agents do not directly observe whether shocks are transitory or persistent. We show that an estimated Kalman filter model using real-time data describes economists' long-run productivity growth forecasts in the United States extremely well and that filtering has profound implications for the macroeconomic effects of shifts in productivity growth.

Inflation Targeting under Imperfect Knowledge

John C. Williams, with Athanasios Orphanides, *Central Bank of Cyprus*

Published in *Series on Central Banking, Analysis, and Economic Policies X: Monetary Policy under Inflation Targeting,* eds. F. Mishkin and K. Schmidt-Hebbel. Santiago, Chile: Central Bank of Chile, 2007, pp. 77–123.

A central tenet of inflation targeting is that establishing and maintaining well-anchored inflation expectations are essential. In this paper, we reexamine the role of key elements of the inflation targeting framework towards this end, in the context of an economy where economic agents have an imperfect understanding of the macroeconomic landscape within which the public forms expectations and policymakers must formulate and implement monetary policy. Using an estimated model of the U.S. economy, we show that monetary policy rules that would perform well under the assumption of rational expectations can perform very poorly when we introduce imperfect knowledge. We then examine the performance of an easily implemented policy rule that incorporates three key characteristics of inflation targeting: transparency, commitment to maintaining price stability, and close monitoring of inflation expectations, and find that all three play an important role in assuring its success. Our analysis suggests that simple difference rules in the spirit of Knut Wicksell excel at tethering inflation expectations to the central bank's goal and in so doing achieve superior stabilization of inflation and economic activity in an environment of imperfect knowledge.

Robust Monetary Policy with Imperfect Knowledge

John C. Williams, with Athanasios Orphanides, *Central Bank of Cyprus*

Published in *Journal of Monetary Economics* 54(5) (July 2007) pp. 1,406–1,435.

We examine the performance and robustness properties of monetary policy rules in an estimated macroeconomic model in which the economy undergoes structural change and where private agents and the central bank possess imperfect knowledge about the true structure of the economy. Policymakers follow an interest rate rule aiming to maintain price stability and to minimize fluctuations of unemployment around its natural rate but are uncertain about the economy's natural rates of interest and unemployment and how private agents form expectations. In particular, we consider two models of expectations formation: rational expectations and learning. We show that in this environment the ability to stabilize the real side of the economy is significantly reduced relative to an economy under rational expectations with perfect knowledge. Furthermore, policies that would be optimal under perfect knowledge can perform very poorly if knowledge is imperfect. Efficient policies that take account of private learning and misperceptions of natural rates call for greater policy inertia, a more aggressive response to inflation, and a smaller response to the perceived unemployment gap than would be optimal if everyone had perfect knowledge of the economy. We show that such policies are quite robust to potential misspecification of private sector learning and the magnitude of variation in natural rates.

Beggar Thy Neighbor? The In-State, Out-of-State, and Aggregate Effects of R&D Tax Credits

Daniel J. Wilson

Forthcoming in *Review of Economics and Statistics.*

The proliferation of research and development (R&D) tax incentives among U.S. states in recent decades raises two important questions: (1) Are these tax incentives effective in achieving their stated objective, to increase R&D spending within the state? (2) To the extent the incentives do increase R&D within the state, how much of this increase is due to drawing R&D away from other states? In short, this paper answers (1) "yes" and (2) "nearly all," with the implication that the net national effect of R&D tax incentives on R&D spending is near zero. The paper addresses these questions by exploiting the cross-sectional and time-series variation in R&D tax credits, and in turn the user cost of R&D, among U.S. states from 1981 to 2004 to estimate an augmented version of the standard R&D factor demand model. I estimate an in-state user cost elasticity around –2.5 (in the long-run), consistent with previous studies of the R&D cost elasticity. However, the R&D elasticity with respect to costs in neighboring states, which has not previously been investigated, is estimated to be around +2.5, suggesting a zero-sum game among states and raising concerns about the efficiency of state R&D credits from the standpoint of national social welfare.

IT and Beyond: The Contribution of Heterogeneous Capital to Productivity

Daniel J. Wilson

Forthcoming in *Journal of Business and Economic Statistics.*

This paper explores the relationship between capital composition and productivity using a unique and highly detailed data set on firm-level investment in the U.S. I develop a succinct methodology for modeling the separate effects of a large number of capital types in a production function framework. I then use this methodology, combined with recently developed techniques for accounting for unobserved productivity, to identify these effects and back out the implied marginal products of each capital type. The results indicate that information and communications technology (ICT) capital—specifically, computers, software, and communications equipment—are positively and statistically significantly associated with output, even after conditioning on total capital, labor, and various proxies for unobserved productivity. I compare the implied marginal products for different capital types to official data on rental prices and find that the marginal products of the ICT capital types are substantially above their measured rental prices. Last, I provide evidence of complementarities and substitutabilities, both among capital types—a rejection of the common assumption of perfect substitutability—and between certain capital types and labor.

State Investment Tax Incentives: What Are the Facts?

Daniel J. Wilson, with Robert S. Chirinko, *Emory University*

Forthcoming in *Proceedings of the 99th Annual Conference on Taxation,* National Tax Association. Also forthcoming as "State Investment Tax Incentives: A Few Facts" in *State Tax Notes.*

There is an ongoing debate in the United States among policymakers and the courts concerning the practical effects of state investment tax incentives. However, this debate often suffers from a lack of clear information on the extent of such incentives among states and how these incentives have evolved over time. This paper takes a first step toward addressing this shortcoming. Compiling information from all 50 states and the District of Columbia over the past 40 years, we are able to paint a picture of the variation in state investment tax incentives across states and over time. In particular, we document three stylized facts: (1) Over the last 40 years, state investment tax incentives have become increasingly large and increasingly common among states; (2) these incentives, as well as the level of the overall after-tax price of capital, are to a large extent clustered in certain regions of the country; and (3) states that enact investment tax credits tend to do so around the same time as their neighboring states.

Conferences

The San Francisco Fed's Research Department organized three conferences and two symposiums in 2007.

The Department's annual macroeconomic conference, "Monetary Policy, Transparency, and Credibility," focused on transparency and credibility and how central banks can better achieve their goals by effectively communicating their views on monetary policy as well as their views on the economy, which inherently involve some degree of uncertainty. Three of the papers focused on the benefits and limits of transparency, identifying circumstances where transparency may be helpful and those where it may be harmful. Another paper studied central bank communication in an environment where private agents have incomplete knowledge of the economy. A fifth paper analyzed policymaking in an economy whose parameters are uncertain. A final paper examined the role of the banking sector in the conduct of monetary policy.

The 2007 Annual Pacific Basin Conference sponsored by the Bank's Center for Pacific Basin Studies (CPBS) brought together papers on a variety of international topics, including international pricing behavior and exchange rates, foreign reserve management, the efficacy of capital controls, Asian financial market integration, and developments in China.

The Bank's Center for the Study of Innovation and Productivity (CSIP) sponsored one conference and two symposiums. The conference focused on "Recent Trends in Economic Volatility: Sources and Implications." Since the early 1980s, volatility of GDP in the United States and other advanced economies has declined considerably. The causes of this so-called "Great Moderation" are of significant interest in current economic research. A key focal point of this research is the relationship between these macroeconomic factors and their microeconomic underpinnings. The seven papers presented at this conference investigated this question from a number of perspectives, with an emphasis on the effects of technological change.

The CSIP symposium "The Costs and Value of New Medical Technologies" explored how new medical technologies contribute to the evolution of healthcare benefits and costs and how government policy may affect these trends. The second symposium, "The Economics of Private Equity Investments," examined the economic factors driving the heightened level of activity in private equity markets through the first half of 2007 and the slowdown subsequent to the credit market dislocations observed in July and August.

These conferences bring professional economists from the Federal Reserve System and from research institutions together with policymakers from the United States and abroad. Many of the papers presented are "works in progress" and therefore represent the latest research on policy-related issues.

Attendance at all of the conferences is by invitation only. In addition, the papers are chosen from submissions by a select group of noted researchers.

In this section are the conference agendas as well as summaries of the conferences that appeared in our *FRBSF Economic Letter*.

Monetary Policy, Transparency, and Credibility

Federal Reserve Bank of San Francisco
March 23–24, 2007

Sponsored by the Federal Reserve Bank of San Francisco

Papers presented at this conference can be found on the website
http://www.frbsf.org/economics/conferences/0703/

Keynote Speaker
Frederic Mishkin, *Federal Reserve Board of Governors*

The Limits of Transparency
Alex Cukierman, *Tel-Aviv University*

Discussants: Petra Geraats, *University of Cambridge*
Hyun Song Shin, *Princeton University*

Banking and Interest Rates in Monetary Policy Analysis: A Quantitative Exploration
Marvin Goodfriend, *Carnegie Mellon University*
Bennett McCallum, *Carnegie Mellon University*

Discussants: Stephen Cecchetti, *Brandeis University*
Huw Pill, *European Central Bank*

Monetary Policy and Its Informative Value
Romain Baeriswyl, *Ludwig-Maximilians Universität*
Camille Cornand, *BETA–Université Louis Pasteur*

Discussants: George-Marios Angeletos, *Massachusetts Institute of Technology*
Christian Hellwig, *University of California, Los Angeles*

Three Great American Disinflations
Michael Bordo, *Rutgers University*
Christopher Erceg, *Federal Reserve Board of Governors*
Andrew Levin, *Federal Reserve Board of Governors*
Ryan Michaels, *University of Michigan*

Discussants: Gauti Eggertsson, *FRB New York*
John Taylor, *Stanford University*

Central Bank Communication and Expectations Stabilization
Stefano Eusepi, *FRB New York*
Bruce Preston, *Columbia University*

Discussants: Michael Ehrmann, *European Central Bank*
Athanasios Orphanides, *Federal Reserve Board of Governors*

Welfare-Maximizing Monetary Policy under Parameter Uncertainty
Rochelle Edge, *Federal Reserve Board of Governors*
Thomas Laubach, *Federal Reserve Board of Governors*
John C. Williams, *FRB San Francisco*

Discussants: Christopher Sims, *Princeton University*
Ulf Söderström, *Università Bocconi*

2007 Annual Pacific Basin Conference

Federal Reserve Bank of San Francisco
June 8–9, 2007

Sponsored by the Center for Pacific Basin Studies (CPBS), Federal Reserve Bank of San Francisco

Papers presented at this conference can be found on the website http://www.frbsf.org/economics/conferences/0706/agenda.pdf

Keynote Speech: Rebalancing Economic Growth in China

Nicholas Lardy, *Peterson Institute for International Economics*

Pricing-to-Market, Trade Costs, and International Relative Prices

Andrew Atkeson, *University of California, Los Angeles*
Ariel Burstein, *University of California, Los Angeles*

Discussant: Paolo Pesenti, *FRB New York*

Productivity and the Dollar

Giancarlo Corsetti, *European University Institute*
Luca Dedola, *European Central Bank*
Sylvain Leduc, *Federal Reserve Board of Governors*

Discussants: Nelson Mark, *University of Notre Dame*
Doireann Fitzgerald, *Stanford University*

The Determinants of Household Saving in China: A Dynamic Panel Analysis of Provincial Data

Charles Horioka, *Osaka University*
Junmin Wan, *Osaka University*

Discussants: Aart Kraay, *World Bank*
Galina Hale, *FRB San Francisco*

Optimal Reserve Management and Sovereign Debt

Laura Alfaro, *Harvard Business School*
Fabio Kanczuk, *Universidade de São Paulo*

Discussants: Manuel Amador, *Stanford University*
Jaewoo Lee, *International Monetary Fund*

Financial Integration in East Asia

Hiroshi Fujiki, *Bank of Japan*
Akiko Terada-Hagiwara, *Bank of Japan*

Discussants: Sebnem Kalemli-Ozcan, *University of Houston*
Jing Zhang, *University of Michigan*

Capital Controls: Myth and Reality, A Portfolio Balance Approach to Capital Controls

Nicolas Magud, *University of Oregon*
Carmen Reinhart, *University of Maryland*
Kenneth Rogoff, *Harvard University*

Discussants: Frank Warnock, *University of Virginia*
Jon Wongswan, *Barclays Global Investors*

Capital Account Liberalization: Theory, Evidence, and Speculation

Peter Henry, *Stanford University*

Discussants: Jay Shambaugh, *Dartmouth College*
Michael Hutchison, *University of California, Santa Cruz*

Panel on Asian Capital Markets

Barry Eichengreen, *University of California, Berkeley*
Robert McCauley, *Bank for International Settlements*

Recent Trends in Economic Volatility: Sources and Implications

Federal Reserve Bank of San Francisco
November 2–3, 2007

Sponsored by the Center for the Study of Innovation and Productivity (CSIP), Federal Reserve Bank of San Francisco

Papers presented at this conference can be found on the website
http://www.frbsf.org/economics/conferences/0711/

Keynote Speaker

Robert Shiller, *Yale University*

On the Sources of the Great Moderation

Luca Gambetti, *Universitat Autònoma de Barcelona*
Jordi Galí, *CREI and Universitat Pompeu Fabra*

Discussants: Mark W. Watson, *Princeton University*
Todd Walker, *Indiana University*

Technological Diversification

Miklós Koren, *FRB New York*
Silvana Tenreyro, *London School of Economics*

Discussants: Jonathan Eaton, *New York University*
David K. Levine, *Washington University in St. Louis*

Macroeconomic Implications of Changes in Micro Volatility

Steven J. Davis, *University of Chicago*
James A. Kahn, *FRB New York*

Discussants: Nick Bloom, *Stanford University*
Douglas Elmendorf, *Brookings Institution*

A Theory of Growth and Volatility at the Aggregate and Firm Level

Diego Comin, *Harvard Business School*
Sunil Mulani, *Commonfund Capital, Inc.*

Discussants: Rasmus Lentz, *University of Wisconsin*
Chad Jones, *University of California, Berkeley*

Scale without Mass: Business Process Replication and Industry Dynamics

Erik Brynjolfsson, *MIT Sloan*
Andrew McAfee, *Harvard Business School*
Michael Sorell, *Harvard Business School*
Feng Zhu, *Harvard Business School*

Discussants: Zhu Wang, *FRB Kansas City*
Shane Greenstein, *Northwestern University*

Business Volatility, Job Destruction, and Unemployment

Steven J. Davis, *University of Chicago*
R. Jason Faberman, *U.S. Bureau of Labor Statistics*
John Haltiwanger, *University of Maryland*
Ron Jarmin, *Center for Economic Studies, U.S. Census Bureau*
Javier Miranda, *Center for Economic Studies, U.S. Census Bureau*

Discussants: John Abowd, *Cornell University*
Robert E. Hall, *Stanford University*

The Roles of Comovement and Inventory Investment in the Reduction of Output Volatility

Owen Irvine, *Michigan State University*
Scott Schuh, *FRB Boston*

Discussants: Robert Gordon, *Northwestern University*
Valerie Ramey, *University of California, San Diego*

Symposium: The Costs and Value of New Medical Technologies

Federal Reserve Bank of San Francisco
May 25, 2007

Sponsored by the Federal Reserve Bank of San Francisco and the Center for the Study of Innovation and Productivity (CSIP)

Cost-Conscious Coverage for Medical Innovations

Alan Garber, *VA Palo Alto and Stanford University*

Health Status, Health Spending, and Future Medical Technology Risk

Dana Goldman, *RAND Corporation*

The Impact of Government Programs on Pharmaceutical Prices and Innovation

Fiona Scott Morton, *Yale University*

Learning Effects and the Diffusion of Medical Technology in a Regulated Environment

Vivian Ho, *Rice University and Baylor College of Medicine*

Symposium: The Economics of Private Equity Investment

Federal Reserve Bank of San Francisco
October 19, 2007

Sponsored by the Center for the Study of Innovation and Productivity (CSIP), Federal Reserve Bank of San Francisco

Some papers presented at this symposium can be found on the website http://www.frbsf.org/csip/research/symposium200710.pdf

Overview of the Private Equity Sector	Colin C. Blaydon, *Tuck School of Business, Dartmouth College*
Financing Private Equity Acquisitions	Christopher M. James, *University of Florida*
The Economics of Private Equity Funds	Ayako Yasuda, *The Wharton School, University of Pennsylvania*
Sources of Value in Private Equity Acquisitions	Peter Y. Chung, *Summit Partners*
Debt Markets and LBO Financing	Jonathan Coslet, *TPG*
Collateralized Debt Markets	Peter Rappoport, *JPMorgan*

Monetary Policy, Transparency, and Credibility: Conference Summary

Reprinted from FRBSF Economic Letter *2007-12, May 25, 2007.*

This *Economic Letter* summarizes the papers presented at a conference on "Monetary Policy, Transparency, and Credibility" held at the Federal Reserve Bank of San Francisco on March 23 and 24, 2007.

At this year's conference, academic researchers and policymakers gathered to discuss six research papers that focused on transparency and credibility and how central banks can achieve their goals by effectively communicating their views on monetary policy as well as their views on the economy, which inherently involve some degree of uncertainty.

Three of the papers focus on the benefits and limits of transparency, identifying circumstances where transparency may be helpful and those where it may be harmful. Another paper studies central bank communication in an environment where private agents have incomplete knowledge of the economy. A fifth paper analyzes policymaking in an economy whose parameters are uncertain. A final paper examines the role of the banking sector in the conduct of monetary policy.

The limits of transparency

It is increasingly common for central banks to be transparent about their long-run inflation goals. In addition to democratic accountability, underlying this transparency is the hope that by publicly announcing a target for inflation the central bank will establish more quickly a reputation for price stability and that this reputation will provide a firmer anchor for inflation expectations. By being more open about its goals, procedures, and forecasts, the central bank hopes to convince households and firms that it is committed to price stability, making inflation stabilization less costly. However, even central banks admired for their transparency are not necessarily all that transparent, invariably withholding key information about their policy objectives and their assessment of the economy and its future prospects.

Although transparency is generally thought to be a good thing, Cukierman examines the limits of monetary policy transparency, focusing on two main dimensions: feasibility and desirability. With respect to feasibility, Cukierman argues that uncertainty about the economy, about the effects monetary policy has on the economy, and about the measurement of key variables like potential output, the output gap, and the natural rate of unemployment make it extremely difficult for even well-intentioned central banks to be fully transparent. In Cukierman's words, "the 'science of monetary policy' is not yet in a stage at which it can replace the 'art of monetary policy'" (p. 32). With respect to desirability, Cukierman argues that a compelling case for secrecy can be made when the central bank has private information about threats to financial stability, such as about the health of banks. There, too much disclosure may lead to contagion, jeopardizing the wider banking system.

Monetary policy and its informative value

It is sometimes argued that households and firms may place too much weight on the central bank's assessment of the economy, which can be problematic when the central bank's information about the economy is imprecise. If its views about the economy are overly influential, then it may be optimal for a central bank to not reveal its views, to not be transparent. However, because central banks base their policy decisions on their assessment of the economy, policy interventions intended to stabilize the economy cannot help but convey information about the economy, even if the interventions are not accompanied by formal policy statements. Of course, it is generally not possible for private agents to infer unambiguously the central bank's information about the economy simply by observing the policy interest rate, but the fact remains that the very act of conducting stabilization policy inevitably reveals information.

Recognizing that the policy interest rate has a stabilization role and an information role, Baeriswyl and Cornand analyze jointly the optimal monetary policy and the optimal level of transparency. In their framework, the central bank conducts monetary policy to stabilize prices and output, but an opaque central bank does not divulge its information about the economy while a fully transparent central bank does. Employing a small-scale model in which fluctuations are caused by demand and supply shocks, Baeriswyl and Cornand show that greater transparency is desirable when supply shocks are not too volatile, when the central bank is more focused on stabilizing prices than output, and when firms already have relatively precise information about the economy.

Three great American disinflations

Although there is little doubt that episodes of deflation or disinflation can be costly for the real economy, there is less agreement about the factors that contribute to the high real cost or about why the real cost varies across episodes. Naturally, disagreement about the factors that influence the real cost of deflation (or disinflation) stimulates debate about how inflation might best be lowered. For example, during the 1970s and 1980s some argued that inflation should be lowered gradually while others argued for an aggressive monetary tightening intended to lower inflation sharply.

To uncover the factors that govern the costs associated with deflation or disinflation, Bordo, Erceg, Levin, and Michaels analyze three episodes of deliberate monetary contraction: the 1870s post-Civil War deflation; the 1920–1921 post-WWI deflation; and the early 1980s disinflation under Federal Reserve Chairman Volcker. In the case of the 1870s deflation, the authors argue that the highly transparent policy objective coupled with a credible commitment allowed a decline in the price level to occur alongside robust real output growth. In contrast, the abrupt shift to a contractionary policy stance in 1920 produced a rapid decline in prices, but at the cost of a sharp fall in output. Here, the authors argue, deflation came at a higher cost because the Federal Reserve departed sharply from the expansionary policy that it had pursued previously. For the Volcker disinflation, the authors argue that a lack of policy credibility, brought about by the rise in inflation that occurred during the late 1960s and 1970s, contributed importantly to the large real cost associated with inflation's decline.

Central bank communication

Since the early 1990s, central banks have increasingly adopted inflation targeting as a framework for conducting monetary policy. A cornerstone of inflation targeting is a publicly announced numerical value, or range, for some measure of inflation. Some, but not all, inflation targeting central banks also make public the forecasts, or projections, upon which their policy decisions are based. The underlying rationale is that central banks can more firmly anchor inflation expectations if they provide private agents with guidance about monetary policy, and by anchoring inflation expectations firmly, the central bank can help prevent undesired fluctuations in the economy and mitigate the possibility of economic instability. But is announcing an inflation target sufficient to anchor inflation expectations, or does the central bank also need to articulate, in some form or other, how the inflation target is to be achieved? Does the central bank need to reveal any trade-offs it perceives in meeting the inflation objective against other policy objectives?

Eusepi and Preston study these issues using a model in which households and firms have incomplete knowledge of the economy and must learn about monetary policy before they can make decisions. In their framework, central bank communication involves revealing to private agents information that they can use to help learn and forecast the economy. They begin by showing that self-fulfilling expectations often arise if the central bank does not communicate with private agents. Alternatively, by communicating the entire policy decision process—which in this model is the coefficients and variables that enter the policy rule—the optimal policy is successfully implemented and instability is mitigated. For intermediate cases, the authors find that communicating to private agents the inflation target and the variables that enter the central bank's policy rule garners the same benefits as communicating the entire policy process. However, in a key result, the authors demonstrate that communicating the inflation target only is insufficient to anchor inflation expectations.

Monetary policy and uncertainty

An issue that central banks are increasingly grappling with is how to best formulate policy when there is uncertainty about the economy. One reason that uncertainty about the economy, especially uncertainty about the parameters that govern the policy transmission mechanism, is troublesome for central banks is that it raises doubts about the timing and magnitude with which policy actions affect the economy. Another subtle, and less widely recognized, reason that parameter uncertainty is troublesome is that it can render uncertain the very goals and objectives to which monetary policy should be directed. Taking the position that monetary policy should attempt to maximize the welfare of a stand-in representative household, Edge, Laubach, and Williams argue that uncertainty about the parameters that govern the household's preferences and the economy's production technology will affect the economy's dynamic behavior, key variables like the output gap and natural rate of interest, and the policy objective function.

To understand the impact of parameter uncertainty on policy design, Edge, Laubach, and Williams study a simulated economy in which parameter uncertainty has the three effects described above. They show that parameter uncertainty leads to the economy's potential output and natural rate of interest being imprecisely estimated. Imprecision about the natural rate of interest makes it difficult for the central bank to determine the appropriate level of interest rates, while imprecision about potential output makes it harder for the central bank to assess whether the economy's productive resources are under- or overutilized. In terms of optimal policymaking, they show that parameter uncertainty means that

policymakers should rely less on estimates of the output gap and more on variables like prices and wages that can be measured with greater precision.

Banking and interest rates in monetary policy analysis

Modern studies examining the design and conduct of monetary policy generally employ models, or frameworks, in which a significant role for monetary aggregates and financial intermediation is absent. Instead, monetary policy is invariably analyzed in terms of how to set a short-term nominal interest rate, with the central bank then supplying the quantity of money required to satisfy demand. Moreover, the banking section is invariably taken to be perfectly competitive or simply omitted, such that the economy effectively contains a single short-term nominal interest rate. Although this approach to modeling monetary policy is widely accepted among central banks and academia, it may prove misleading if factors such as collateral, financial intermediation, or a need by banks to monitor loans give rise to an array of interest rates with differing effects on the economy.

To assess whether such factors may be important, Goodfriend and McCallum develop a model suitable for policy analysis that contains a banking sector in addition to the usual goods-producing sector. In the banking sector, loan production requires both collateral (with capital less useful than bonds as collateral) and loan-monitoring inputs, giving rise to an endogenous external finance premium. Accordingly, a monetary policy that stimulates economic activity may either raise or lower the external finance premium, depending on model parameters. By raising the value of collateral, the stimulus may lower the external finance premium, generating a "banking accelerator" or, by raising the demand for bank deposits, the stimulus may raise the external finance premium, generating a "banking attenuator." With the rates of return on government bonds, deposits, collateralized loans, and uncollateralized loans varying from each other and from the return on physical capital, the key result in the paper is to show that in response to a shock to goods-sector productivity a monetary policy that ignores the distinction between these various rates of return could go terribly awry.

Richard Dennis
Senior Economist

John C. Williams
Senior Vice President and Advisor

Conference papers

Papers are available in pdf format at
http://www.frbsf.org/economics/conferences/0703/index.html

Baeriswyl, Romain, and Camille Cornand. 2007. "Monetary Policy and Its Informative Value." London School of Economics Financial Markets Group Discussion Paper #569.

Bordo, Michael, Christopher Erceg, Andrew Levin, and Ryan Michaels 2007. "Three Great American Disinflations." National Bureau of Economic Research Working Paper #12982.

Cukierman, Alex. 2007. "The Limits of Transparency." Manuscript.

Edge, Rochelle, Thomas Laubach, and John Williams. 2007. "Welfare-Maximizing Monetary Policy under Parameter Uncertainty." FRB San Francisco, manuscript.

Eusepi, Stefano, and Bruce Preston. 2007. "Central Bank Communication and Expectations Stabilization." FRB York Staff Report #199.

Goodfriend, Marvin, and Bennett McCallum. 2007. "Banking and Interest Rates in Monetary Policy Analysis: A Quantitative Exploration." Manuscript.

2007 Annual Pacific Basin Conference: Summary

Reprinted from FRBSF Economic Letter *2008-03, February 1, 2008.*

This *Economic Letter* summarizes the papers presented at the 2007 Annual Pacific Basin conference held June 8–9, 2007, at the Federal Reserve Bank of San Francisco under the sponsorship of the Bank's Center for Pacific Basin Studies (CPBS).

This year's Pacific Basin conference brought together papers on a variety of international topics, including international pricing behavior and exchange rates, foreign reserve management, the efficacy of capital controls, Asian financial market integration, and developments in China.

International prices and exchange rates

Atkeson and Burstein address a puzzle in international macroeconomics: why are changes in a country's terms of trade—that is, the relative price of its exports to its imports—generally much smaller than changes in the relative prices of goods produced in the United States and goods produced abroad? For example, between 1985 and 1988, the price of manufactured goods produced abroad rose by roughly 40 percent relative to the average price of manufactured goods produced in the United States; movements of similar magnitudes occurred again in the late 1990s and more recently after 2002. In contrast, the terms of trade between exports and imports of U.S.-manufactured goods have been much less volatile. This is consistent with the stylized fact that, even though manufactured goods are heavily traded across countries, the prices that U.S. consumers pay for imports of manufactures move much less than one-for-one with the prices that foreign producers charge in their own markets.

The authors formulate a quantitative model of international trade to explain this behavior. In their model, producers engage in "pricing-to-market" behavior; that is, they raise the price at which they sell in foreign markets much less than any increase in their costs. The reason firms do not fully pass through changes in their marginal costs to their prices is because their desired markup depends on their share of the foreign sales market. By lowering markups, firms can avoid losing too much market share as they adjust prices in response to greater costs.

Corsetti, Dedola, and Leduc examine whether U.S. economic growth causes the dollar to depreciate or appreciate. Several traditional international macro models predict that the domestic currency should depreciate as a country grows. For example, models in which domestic demand is stimulated by monetary policy often predict higher inflation and depreciation of the currency. Other models in which productivity shocks expand domestic output also imply an increase in the supply of exportable goods, causing the relative price of exports to decline in world markets. This decline in the country's terms of trade also depreciates the currency. However, these predictions are at odds with the recent experience of the U.S., in which the high productivity and domestic output boom of the second half of the 1990s was accompanied by a strong real appreciation of the dollar.

The authors seek to reconcile the empirical evidence with theory by investigating the role of U.S. productivity in a real business cycle model. They find that U.S. productivity shocks do, indeed, increase investment and output. At the same time, however, these shocks also generate strong domestic wealth effects that boost demand for domestic goods as well that (under certain conditions) also leads to an appreciation of the dollar.

Foreign reserves

There is a renewed interest in policy and academic circles about whether some countries are now holding too much foreign reserves. International reserves holdings by developing countries, for example, have risen rapidly in recent years, amounting to 20 percent of GDP in 2005, quadruple the level in high-income countries. A common explanation advanced for foreign reserve accumulation is that it provides an insurance mechanism against the risk of shocks, such as sudden spikes in foreign interest rates or "stops" of capital inflows.

Alfaro and Kanczuk argue that foreign borrowing also provides some of the same functions as holding reserves. They study optimal reserve policy in a stochastic dynamic general equilibrium model that recognizes the potential benefits of holding reserves or borrowing abroad as needed. Calibrating the parameters of the model to accord with a typical developing economy, they conclude that the optimal policy is not to hold reserves at all; contingent borrowing provides sufficient insurance. This contrasts with actual behavior, of course. The authors suggest explanations for the contrast between this theoretical prediction and actual behavior: devel-

oping countries often face limits to the extent of their foreign borrowing, and they may be motivated to hold reserves, not just for insurance reasons, but also for political economy considerations, such as desired spending on public works.

Capital controls

The effectiveness of capital controls is the subject of ongoing debate. Magud, Reinhart, and Rogoff argue that it is hard to compare results from existing studies on capital controls, since there are significant differences across countries and time in the control measures implemented, there is no common empirical methodology, nor is there any clear definition of what constitutes "success." The authors seek to fill this void and measure the effectiveness of controls on short-term capital flows by whether they reduce the volume of capital flows, decrease the proportion of short-term capital flows, reduce real exchange rate pressures, and/or allow for more monetary policy independence. They find that capital controls on inflows seem to reduce the share of short-term capital flows, reduce real exchange rate pressures, and make monetary policy more independent, but capital controls on outflows do not have any systematic effect, with the exception of Malaysia. Hence imposing capital controls on outflows need not always be effective.

Henry surveys the literature on capital flows and finds little evidence that capital account openness is associated with higher economic growth. However, he argues that traditional theory implies that capital account liberalization should have only a temporary, not a permanent, effect on growth. He shows that opening the capital account leads countries to temporarily invest more and grow faster than they did when their capital accounts were closed. Allowing foreign investors into emerging market equity markets lowers the cost of capital, raises the optimal level of the capital stock, and increases steady-state per capita income. During the transition to the new steady-state (higher) level of capital stock, growth rates will increase above normal before eventually returning to trend.

Asian financial markets

Fujiki and Terada-Hagiwara examine the degree of integration of East Asian economies with world financial markets. They find that, while East Asia has become more integrated with world financial markets since the Asia crisis of 1997–1998, domestic saving and investment still are much more highly correlated within East Asia compared to the euro area. They also find that the cross-holding of financial assets is lower within Asia than it is within the euro area. In addition, they find no evidence of any decline in consumption volatility in Asia as one might expect if greater financial integration were enabling greater risk-sharing with the rest of the world. These results suggest that there is room for welfare gains in Asia via further asset flows and risk-sharing within the region as well as with the rest of the world. The results also imply that increased integration into world financial markets alone is unlikely to provide a firm basis for a currency union in East Asia at this stage.

Two panelists offered presentations on Asian capital markets. Robert McCauley of the Bank for International Settlements observed that Asian financial markets have become more integrated with the rest of the world. Equity investors from the U.S. and Europe invest heavily in the region. In addition, dollar-denominated bonds and syndicated loan markets show significant regional integration, with 40 percent of dollar bonds sold by Asian issuers bought by Asian residents, and similar fractions of syndicated loans for Asian borrowers taken up by banks from Asia. However, the role of foreign investors in local currency bond markets is very limited, because of either explicit inflow restrictions or withholding tax requirements. McCauley argued that foreign investors in domestic bond markets could provide a more diverse investor base to support domestic growth (though at the risk of greater exposure of local markets to global bond market strains and possibly large inflows and outflows).

Barry Eichengreen of the University of California at Berkeley also emphasized the importance of developing national bond markets in local currency. He noted that Asia has progressed slowly in developing local markets for corporate debt, particularly in terms of market liquidity. He provided evidence that bond market growth in developing countries depends on the extent of banking sector development, macroeconomic stability, creditor rights, and corporate governance. He attributed the limited development of corporate bond markets in Asia to slow progress in several of these areas.

China saving

China's overall saving rate is now nearly 50 percent, by far the highest in the world. China's domestic investment rate has also been high, but not as high as saving, resulting in net current account surpluses which rose from 4 percent of GDP in 2004 to 7 percent in 2007. The corresponding trade deficits with its trading partners, particularly the United States, imply that China's high saving rate has important ramifications for its economic relations with other countries.

Horioka and Wan analyze the determinants of the household saving rate in China using panel data on Chinese provinces for the period 1995–2004. They find China's saving rate is very persistent and strongly related to income growth and the interest rate (in the case of rural, but not urban, households). However, they do not find that the variables re-

lating to the age structure of the population have any significant impact on the household saving rate (in part because the shortness of their sample limits the time series variation in demographic variables). Thus it is not clear how much saving will fall once the aging of the population is completed. Further research is warranted on the effects of China's aging population, its one-child policy, and its currently high corporate saving levels.

Nicholas Lardy of the Peterson Institute for International Economics delivered the keynote address and discussed China's recent efforts to alter fundamentally the country's growth strategy by expanding domestic consumption in place of investment and exports. In his view the success of this new policy requires policies to reduce China's currently high saving rate. He argued, however, that initiatives to implement these policies have thus far been modest in scope.

Reuven Glick
Group Vice President

Conference papers

Papers are available in pdf format at
http://www.frbsf.org/economics/conferences/0706/agenda.pdf

Alfaro, Laura, and Fabio Kanczuk. "Optimal Reserve Management and Sovereign Debt."

Atkeson, Andrew, and Ariel Burstein. "Pricing-to-Market, Trade Costs, and International Relative Prices."

Corsetti, Giancarlo, Luca Dedola, and Sylvain Leduc. "Productivity and the Dollar."

Fujiki, Hiroshi, and Akiko Terada-Hagiwara. "Financial Integration in East Asia."

Henry, Peter Blair. "Capital Account Liberalization: Theory, Evidence, and Speculation."

Horioka, Charles Yuji, and Junmin Wan. "The Determinants of Household Saving in China: A Dynamic Panel Analysis of Provincial Data."

Lardy, Nicholas. "Rebalancing Economic Growth in China."

Magud, Nicolas, Carmen Reinhart, and Kenneth Rogoff. "Capital Controls: Myth and Reality, A Portfolio Balance Approach to Capital Controls."

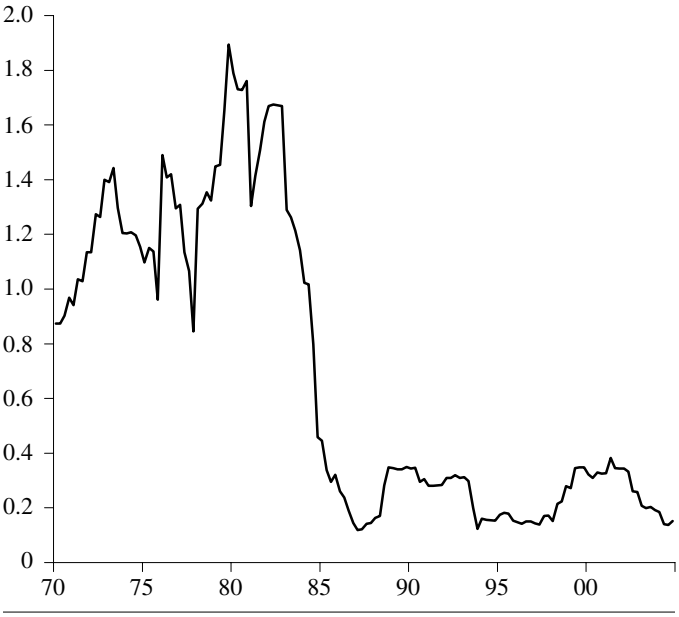
Recent Trends in Economic Volatility: Conference Summary

Reprinted from FRBSF Economic Letter *2008-06, February 15, 2008.*

This *Economic Letter* summarizes the papers presented at a conference on "Recent Trends in Economic Volatility" held at the Federal Reserve Bank of San Francisco by the Bank's Center for the Study of Innovation and Productivity (CSIP) on November 2–3, 2007.

Over the last 25 years, the U.S. economy has become much less volatile; that is, the swings from boom to bust have been greatly reduced, as has the pain typically associated with such cycles. As Figure 1 illustrates, the volatility of GDP growth has fallen by more than half since 1985. Many observers refer to this phenomenon as the "Great Moderation." To what can we credit this improved environment? Researchers have uncovered several potential drivers, including improved technology (especially related to inventory and supply chain management), better monetary policy, and simple good luck, but to date they have found little consensus on which factor is most important. Also in dispute is the extent to which the decline in aggregate volatility has been mirrored in the microeconomic data on income and employment. In other words, have households and businesses also experienced a decline in volatility? The seven papers presented at the Center for the Study of Innovation and Productivity's conference on "Recent Trends in Economic Volatility" investigate these questions. Although the debate is not over, the papers have moved the research forward and highlighted key questions for future work.

FIGURE 1
VARIANCE OF QUARTERLY REAL GDP GROWTH
(FIVE-YEAR MOVING VARIANCE)



Source: Bureau of Economic Analysis.

Structural change vs. good luck in explaining the Great Moderation

The first paper of the conference, by Galí and Gambetti, begins with a useful summary of the various explanations for the Great Moderation, placing them into two broad categories: structural changes and "good luck." Structural changes include changes in the way monetary policy is conducted and technology-driven changes that affect the way firms operate. "Good luck" essentially means smaller and fewer economic shocks. Galí and Gambetti go on to use a standard empirical model known as a structural vector autoregression in order to characterize the correlations in post-World War II data among key U.S. macroeconomic variables. They posit that if declining volatility is merely the result of "good luck," then the data should show no change in the correlations between them. Their model, however, finds that this is not the case, as correlations between output, labor hours, and productivity have indeed changed since the early 1980s. Having eliminated good luck as an explanation, they attribute most of the decline in volatility to a decline in nontechnology shocks, which have come about due to a change in the Federal Reserve's monetary policy "rules" (specifically, an increased emphasis on fostering low and stable inflation in addition to strong economic growth) as well as a reduction in labor adjustment costs. It is worth noting that the authors' finding that reduced labor adjustment costs may have played an important role in the Great Moderation is consistent with evidence, discussed below, provided by the paper of Davis et al., which explores the secular decline in labor market volatility.

The role of technological change

Several papers ascribe a key role to technological progress in explaining declining volatility. The first of these papers, by Koren and Tenreyro, looks at how development of new technologies affects both the rate of growth and the volatility of growth in an economy. Their model posits that, just as households benefit from investing in a diversified portfolio of stocks (smoothing their returns and minimizing losses stemming from shocks to specific assets), having a larger and more diverse "menu" of technologies available to firms in a country means that each specific technology plays less of a role in production. The diversification of technologies in an economy makes it easier for firms to offset price or supply shocks to specific inputs (oil, for example) by substituting with other technologies that rely less on those inputs. In this way, technological advances reduce firm-level volatility, which consequently reduces overall volatility. Technological change also boosts the level of growth, since it allows firms to move to a new technology before reaching the point of diminishing returns in their old technology. While sensible and consistent with data that Koren and Tenreyro bring to bear, this finding contrasts sharply with the conclusions of previous research, which point to an explicit tradeoff between risk (volatility) and return (fast growth).

Comin and Mulani also examine the effects of technological change on economic growth and volatility and, similarly, find that technological change leads to both faster growth and lower volatility. But in contrast to the previous paper, Comin and Mulani argue that this good result holds only for the national, or macro, measures. Indeed, predictions from their model suggest that firm-level, or micro, volatility should increase as the pace of technological innovation increases. To get this result, they consider an economy with two types of technologies: general innovations (GIs), which are not patentable and are used by all firms in the economy, and research and development innovations (RDI), which are patentable and used by a limited number of firms. They then assume that GIs are produced by large, stable firms and RDIs are produced by smaller, more volatile firms. Under these conditions, they show that increases in RDIs (for example, due to government research and development (R&D) subsidies) lead to market "shake-up," whereby smaller firms gain market share and perhaps even leapfrog ahead of the previous market leaders. Since GI activity relies on the presence of stable market leaders, this shake-up creates both firm-level volatility and lower GI activity. The decline in GIs, which by definition help all firms, reduces the comovement between firms in the economy, ultimately reducing the volatility of aggregate outcomes. Said more simply, if the increase in the innovative activity comes from small firms jockeying for position in the industry, aggregate volatility will go down, as winners and losers will offset each other, but microvolatility will rise, as losing firms compete to get back on top. Comin and Mulani provide empirical evidence showing that increased R&D activity in the U.S. has coincided with increased volatility in sales and market shares for publicly traded firms, reduced comovement across industries, and reduced volatility in aggregate economic growth.

Turning to the purely micro data, Brynjolfsson et al. analyze the impact of information technology (IT) on industry volatility or turbulence. Use of IT allows an innovation to diffuse rapidly throughout a firm, increasing productivity and market share faster than was previously possible. Although first movers on an innovation are able to gain market share quickly creating the opportunity for concentration, the speed of diffusion that IT affords also enables new entrants to leapfrog ahead of leaders in a given sector, thus increasing sectoral turnover rates (turbulence). Empirically, IT-intensive industries have indeed experienced both greater concentration and turbulence. This evidence is consistent with the findings of Comin and Mulani that firms in more R&D-intensive industries tend to have more volatile sales and market shares, since there is a strong correlation between an industry's R&D intensity and its IT intensity.

Supply chain management

The role of supply chain management in the Great Moderation is the subject of a paper by Davis and Kahn as well as one by Irvine and Schuh. Davis and Kahn argue that dramatic technology-driven improvements in supply chain management in the durable goods sector, combined with a secular shift away from domestic durable goods manufacturing and toward services, is the explanation for the decline in aggregate volatility. They suggest that changes in monetary policy, on the other hand, played a minimal role. Their model of the firm's inventory decision process mirrors observed declines in output and sales volatility, as well as the sales-to-output ratio, and the authors suggest that a shorter lead time for materials orders (more precise inventory control) is the key mechanism through which this change has occurred.

Irvine and Schuh also find that improvement in supply chain management likely played the predominant role in reduced aggregate volatility. Using a multi-sector, vector-autoregression empirical model, they find that a decline in the comovement of output among inventory-holding industries (for example, manufacturing and wholesale trade) can explain a substantial share of the decline in aggregate output volatility. Their model suggests that a change in structural relationships between inventory-holding industries seems to be the cause of this decline, and industries in which firms share supply and distribution chains exhibited the largest decline in covariance in volatility. As in Davis and Kahn, Irvine and

Schuh find little evidence that changes in monetary policy or "good luck" are major factors behind the Great Moderation.

Volatility in the labor market

In the final paper of the conference, Davis et al. establish and attempt to explain two interesting facts from the data. The first fact is that volatility of employment levels within firms, particularly those not publicly traded, has declined over the past 25 years. The second fact is that the flows of individuals into unemployment have fallen over time. In the early 1980s about 4 percent of employed persons fell into unemployment (either voluntarily or involuntarily) in the average month; by the early 1990s, this figure had dropped to just 2 percent. The focus of their paper, then, is to investigate whether the decline in volatility in employment demand by businesses is responsible for the decline in unemployment inflows. Using industry-level data, they find a strong statistical association between an industry's volatility in employment demand, as measured by the variance in its job destruction rate, and the industry's unemployment inflow rate (the rate at which workers in the industry go into unemployment in a given period). They conclude that the decline in firm level employment volatility likely has reduced flows into unemployment.

Conclusion

While there is broad agreement that aggregate economic volatility has declined over the last 25 years, the relative roles of economywide factors, such as changes in monetary policy and technological change, remain topics of dispute. Also in dispute is the extent to which this decline in aggregate volatility is mirrored in microeconomic variables, such as income and employment. Reductions in aggregate and firm-level volatility do not necessarily translate into a reduction in volatility at the individual level. Rather, some studies argue that household consumption and individual earnings have become more volatile in recent decades, not less. The linkages between the disparate trends in volatility at the aggregate level and at the individual level remain important areas of economic research.

Charles Notzon
Research Associate

Dan Wilson
Senior Economist

Conference papers

Papers are available in pdf format at
http://www.frbsf.org/economics/conferences/0711/

Brynjolfsson, Erik, Andrew McAfee, Michael Sorell, and Feng Zhu. "Scale without Mass: Business Process Replication and Industry Dynamics."

Comin, Diego, and Sunil Mulani. "A Theory of Growth and Volatility at the Aggregate and Firm Level."

Davis, Steven J., and James A. Kahn. "Changes in the Volatility of Economic Activity at the Macro and Micro Levels."

Davis, Steven J., R. Jason Faberman, John Haltiwanger, Ron Jarmin, and Javier Miranda. "Business Volatility, Job Destruction, and Unemployment."

Galí, Jordi, and Luca Gambetti. "On the Sources of the Great Moderation."

Irvine, Owen, and Scott Schuh. "The Roles of Comovement and Inventory Investment in the Reduction of Output Volatility."

Koren, Miklós, and Silvana Tenreyro. "Technological Diversification."

The Costs and Value of New Medical Technologies: Symposium Summary

Reprinted from FRBSF Economic Letter *2007-18, July 6, 2007.*

This *Economic Letter* summarizes the presentations made at a symposium by the same title sponsored by the Center for the Study of Innovation and Productivity (CSIP) and held at the Federal Reserve Bank of San Francisco on May 25, 2007.

Health care is among the most technologically advanced sectors, and it also constitutes a large and growing share of the U.S. economy. Between 1960 and 2005, the share of health-care spending in U.S. gross domestic product more than tripled, growing from 5.2 percent to 16 percent; this growth is likely to continue, with health care conceivably expanding to encompass up to one-third of national output by the year 2050 (Jones 2005).

Much of this growth is demand driven, as purchasers of health care spend increasing amounts of money to pay for new, technologically advanced medical procedures and drugs that extend life and improve its quality. At the same time, however, rising costs mean lower affordability: coverage under private health plans, mostly through employers, has declined in recent years, putting added strain on already strapped public programs (Buchmueller and Valletta 2006). These trade-offs are likely to intensify over time, raising a host of issues for policymakers and the public alike.

To help improve our understanding of how new medical technologies contribute to the evolution of health-care benefits and costs and how government policy may affect these trends, the Center for the Study of Innovation and Productivity convened a conference that brought together four leading scholars to discuss various aspects of the development and use of new medical technologies.

Responses to rising costs

Alan Garber, from Stanford University and the Palo Alto VA hospital, presented his work on "Cost-Conscious Coverage for Medical Innovation." His presentation focused on the role that new medical technologies have played in the rapid rise in health-care costs and how to alter the incentives in U.S. health care so that costs associated with new technologies are controlled but the quality of services is not undermined. As U.S. health-care costs have risen in recent years, out-of-pocket costs for the insured have grown rapidly: for example, premium contributions for workers covered under plans provided by their employers grew about 50 percent between 2000 and 2003. Such cost sharing has the potential to curb utilization, which may help contain cost growth in an efficient manner. However, the overall containment potential of cost sharing is limited because the highest-cost claims account for a large share of total spending and are relatively insensitive to cost sharing. Moreover, increased cost sharing offsets the risk-protection and risk-pooling intent of insurance plans.

Garber's preferred strategy for cost control relies on modifying the process used to determine which medical procedures and therapies are covered under insurance plans. For U.S. private and public health plans, this determination currently is based on an assessment of whether the technology or procedure yields greater improvement in health outcomes than do established alternatives. This approach entails various problems, including the possibility of mistaken assessments due to limitations of accepted experimental designs and statistical evaluations. Most importantly, the existing framework for evaluating the effectiveness of health interventions does not take into account considerations of relative cost: procedures with similar impacts on health outcomes can be regarded as equally meritorious despite large differences in the costs of their use.

The exclusion of cost considerations likely has contributed to rapid increases in U.S. health-care costs. Garber therefore recommends the use of "cost-conscious coverage" policies, whereby health interventions are evaluated in terms of their relative cost effectiveness in addition to their impact on medical outcomes (for example, Garber 2004). Evidence on cost effectiveness of different health interventions currently is available and could be used to initiate a switch toward cost-based coverage, resulting in immediate cost savings. Moreover, these savings are likely to grow substantially over time, as health plan designers, consumers, and medical innovators respond to the newly available information and modified incentives.

Future medical technologies

Dana Goldman, a director at the RAND Corporation and adjunct professor at UCLA, discussed his work on "The Costs and Benefits of Future Medical Technologies." He first es-

tablished that expanding technology has been by far the largest contributor to the rapid increase in health-care costs since 1960, accounting for about one-half of the increase. Like Garber, however, Goldman also emphasized the large variation in benefit/cost ratios that is evident across medical procedures. In an ongoing RAND research study, his team models the effects of 34 key emerging medical technologies, including anti-aging compounds, stroke treatments, cancer therapies, and implantable heart defibrillators and pacemakers (see Goldman et al. 2005). For example, their model predicts that use of intra-ventricular cardio-defibrillators (ICDs) will expand dramatically in coming years, adding about $30 billion annually (3.7 percent) to U.S. medical spending through the year 2030. This makes ICDs an expensive technology relative to the value of resulting health improvements, but other advanced technologies, such as certain cancer treatments and pacemakers, are even more expensive.

In addition to their direct costs, medical innovations can have large indirect costs. For example, medical researchers currently are investigating the potential use of anti-aging compounds in humans, which could substantially extend life at relatively low cost. However, if such treatments prove successful, the size of the U.S. elderly population will swell, increasing the prevalence of old-age conditions (such as heart problems) and leading to large increases in overall health spending. Similar considerations apply to preventive health therapies such as smoking cessation and obesity control. Successful smoking cessation programs will save lives but be relatively expensive, since they entail limited savings in end-of-life treatments but increases in other forms of old-age care. By contrast, while successful obesity control may not greatly lengthen life spans, it is likely to produce substantial improvements in health and well-being that enable reductions in health-care costs more generally. The RAND model's predictions have important implications for government entitlement programs, suggesting that medical innovations are likely to increase Medicare spending but may not adversely affect the financing of the U.S. Social Security program.

Impacts of government programs

Fiona M. Scott Morton is a professor of economics at the Yale School of Management. Her talk, titled "The Impact of Government Programs on Pharmaceutical Prices and Innovation," addressed pharmaceutical markets and the role of the U.S. government's large Medicaid and Medicare programs. Within health care, spending growth has been especially rapid for pharmaceuticals, with innovation accounting for a large share of producer and consumer expenditures. Moreover, the government share of this market in the United States is large (about 50 percent) and likely to grow. Medicaid is the state-managed program to provide health care for low-income individuals. Drug prices in the program initially are set based on market prices, but with a 15-percent discount imposed on manufacturers. Subsequent price increases are limited to the prevailing inflation rate, unless a new form of the drug is introduced; the new form may consist only of minor modifications in dosage or packaging. Pharmaceutical companies specializing in expensive Medicaid drugs therefore face substantial incentives for frequent product modifications and high prices, which reduces their private sector sales but yields a higher price on Medicaid sales (with no quantity reduction because Medicaid recipients do not pay for their purchases). In recent research, Scott Morton finds direct empirical evidence of such shifts in the composition and pricing of prescription medications under the Medicaid program (Duggan and Scott Morton 2006).

Scott Morton also discussed pricing decisions for the new Medicare Part D prescription drug benefit, initiated in January 2006. Drug provision under Part D is similar to provision under private sector plans, with participants choosing among competing plans, drug makers competing for business, and participants paying a cost share (which is heavily subsidized for low-income enrollees); however, access to certain classes of drugs is guaranteed under Part D plans. Although direct empirical evidence is not yet available, it is likely that drug prices under Part D plans will be similar to those in the private sector, although deviations are likely among protected classes of drugs.

In the public as well as the private sector, development of cost-effective drug therapies faces substantial hurdles due to a lack of targeted coordination between insurers and health-care providers. Like the preceding speakers, Scott Morton therefore emphasized the importance of developing integrated frameworks for assessing the cost effectiveness of health interventions.

Learning effects

Vivian Ho, from Rice University and the Baylor College of Medicine, discussed her work on "Learning Effects and the Diffusion of Medical Technology in a Regulated Environment," which expanded on the earlier presentations by addressing the issue of how best to use new technologies. In particular, she focused on the well-known "volume-outcome" relationship for medical procedures, in which hospitals and surgeons that have greater experience with complex surgical procedures typically obtain better outcomes from those procedures (such as lower mortality rates). The two leading explanations for this relationship are: (i) "learning-by-doing" (LBD), which refers to the process by which repeated performance (by surgeons and hospitals) increases knowledge and skill, thereby directly improving quality; and (ii) "selective referral," whereby hospitals that provide the

highest quality service will attract more patients. Explanation (i) points toward beneficial effects of policies that encourage hospital specialization in specific procedures, whereas (ii) reverses the causation and undercuts arguments in favor of such policies.

These two explanations are difficult to distinguish empirically. Researchers have used volume changes over time for specific hospitals in an attempt to separate out reverse causation, but such studies are undermined by small changes in volume over time and confounding effects from changing technology. In recent work, however, Ho and colleagues (Gowrisankaran, Ho, and Town 2007) used an "instrumental variables" strategy, which relies on variations in procedure volumes across hospitals that are uniquely determined by the choices of individual patients. Their technique yields statistically precise estimates showing a substantial impact of volume on quality for several types of open heart and abdominal surgeries, providing strong evidence in favor of LBD.

Ho's findings suggest that medical policy guidelines that require or encourage hospitals to reach minimum volume thresholds for complex procedures may be advantageous to patients. On the other hand, regulations that attempt to capitalize on these gains may increase the market power of the high-volume providers, leading in turn to higher prices. In additional work, Ho and colleagues (Ho, Town, and Heslin 2007) found that increased market power partially offsets the value of health gains to patients, but substantial net benefits to volume remain. Overall, her findings suggest that learning is an important element for the successful use of new technologies, and that medical practitioners and policymakers should more systematically account for learning effects when developing health-care guidelines.

Discussion

Among the common themes identified by the presenters, it seems clear that advances in medical technology have generated large benefits relative to their costs in the United States in recent decades. However, incentive structures within the U.S. private and public systems for health-care delivery are not always ideal: market power among providers sometimes offsets consumer gains from new procedures, and cost control generally is not rewarded. Achieving greater cost control will be technically and politically challenging because it is likely to entail some degree of rationing in the supply of health-care services, but explicitly making such trade-offs may be necessary to ensure the spread of beneficial medical technologies to the widest possible population.

Rob Valletta
Research Advisor

References

URLs accessed March 2008.

Buchmueller, Tom, and Rob Valletta. 2006. "Health Insurance Costs and Declining Coverage." *FRBSF Economic Letter* 2006-25 (September 29). http://www.frbsf.org/publications/economics/letter/2006/el2006-25.html

Duggan, Mark, and Fiona Scott Morton. 2006. "The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing." *Quarterly Journal of Economics* 121(1) (February) pp. 1–30.

Garber, Alan M. 2004. "Cost-Effectiveness and Evidence Evaluation as Criteria for Coverage Policy." *Health Affairs* Web Exclusive, May 19. http://content.healthaffairs.org/cgi/reprint/hlthaff.w4.284v1

Goldman, Dana P., et al. 2005. "Consequences of Health Trends and Medical Innovation for the Future Elderly." *Health Affairs* Web Exclusive, September 26. http://content.healthaffairs.org/cgi/reprint/hlthaff.w5.r5v1

Gowrisankaran, Gautam, Vivian Ho, and Robert J. Town. 2007. "Causality, Learning, and Forgetting in Surgery." Working paper, University of Minnesota.

Ho, Vivian, Robert J. Town, and Martin J. Heslin. 2007. "Regionalization versus Competition in Complex Cancer Surgery." *Health Economics, Policy, and Law* 2, pp. 51–71.

Jones, Charles I. 2005. "More Life vs. More Goods: Explaining Rising Health Expenditures." *FRBSF Economic Letter* 2005-10 (May 27). http://www.frbsf.org/publications/economics/letter/2005/el2005-10.html

The Economics of Private Equity Investments: Symposium Summary

Reprinted from FRBSF Economic Letter *2008-08, February 29, 2008.*

This *Economic Letter* summarizes proceedings of a symposium held at the Federal Reserve Bank of San Francisco on October 19, 2007, sponsored by the Bank's Center for the Study of Innovation and Productivity (CSIP). The symposium brought together academic researchers and private equity practitioners, including representatives of private equity firms, investors in private equity, and lenders.

Private equity investment, particularly related to the purchase of private and public firms, has been a central component of so-called leveraged buyouts (LBOs) over the past several years. In the U.S., the dollar amounts of these LBOs increased markedly, from $24 billion in 2001 to $320 billion in 2006. By midyear 2007, they had reached almost $200 billion, but then activity slowed dramatically due to severe financial market conditions. While the private equity sector is relatively small compared to the entirety of the U.S. capital markets, it has been very prominent because of its rapid growth and the degree to which financial innovations have played a role.

In October 2007, the Economic Research Department's Center for the Study of Innovation and Productivity (CSIP) convened a symposium of academic researchers and industry experts to examine the economic factors driving the heightened level of activity through the first half of 2007 and the slowdown following the credit market dislocations observed starting in July and August. This *Economic Letter* summarizes the main themes discussed at the symposium.

Growth of private equity investment

Broadly defined, private equity investment refers to investments made by professional managers of investment funds in private companies. The two main categories of such investment are venture capital (VC), which concentrates on newer companies, and buyouts, which concentrate on more seasoned companies. In recent years, a large percentage of these buyouts have involved the purchase of publicly traded companies in their entirety and their conversion to private companies. The end-investors, or limited partners (LPs), in these funds are typically large institutional investors, such as pension funds, endowments, and foundations, as well as wealthy individuals. The LPs invest in funds managed by professionals from private equity firms, which typically manage several funds at once. These managers, or general partners (GPs), receive annually a percentage fee of the money under management in these funds as well as a portion of any realized gains, which are commonly realized through the sale of the firm to other investors or through a public stock offering. The term "leveraged buyouts" is more commonly used for these types of investments, since acquisitions of firms are financed using a combination of equity (from the LPs) and debt issued in the form of bonds and loans under the company's name.

Private equity investment funds have been around since the mid-1940s, but their growth in recent years has been remarkable. As noted at the conference by Peter Chung (Summit Partners), total private equity commitments globally totaled about $2.3 billion in 1969, increasing to nearly $335 billion in 2006. Focusing on buyouts, the annual amount invested in these funds was approximately $275 billion in 2006, according to Colin Blaydon (Tuck School of Business, Dartmouth).

Changes in the mix and composition of leverage

The mix of equity and debt used in LBOs has changed over time. Blaydon's presentation showed that the average share of equity used in LBOs rose through most the 1990s, reaching a peak of about 42 percent in 2001, and then declined to about 32 percent in 2006.

In addition, the composition of the debt used in LBOs has changed. Jonathan Coslet (TPG Capital), presented evidence on the availability of leverage for buyout transactions by highlighting the growth of loans, both by banks and especially by institutional investors. He reported that the overall size of the leveraged loan and high-yield bond sectors, which is where buyouts were mainly financed, had reached $324 billion in 1998, with respective shares of 68 percent and 32 percent. Furthermore, the leveraged loan market was dominated by bank lending. These markets began to shrink in size shortly after that, reaching a low of $201 billion in 2002. During this period, the bank lending share decreased to 40 percent, but the leveraged loan market share grew in level and percentage to $59 billion and just short of 30 percent, respectively.

After 2002, the supply of leverage increased dramatically up through 2006. Overall market size surged by a factor of more than three to $656 billion. While all three categories grew over this period, leveraged loans by institutional investors grew by a factor of nearly 5.5 from $59 billion to $321

billion. Correspondingly, this sector's share of overall financing availability grew from 30 percent to nearly 50 percent. Preliminary numbers for the first half of 2007 indicated a continuation of this trend.

Coslet noted that the steady growth of institutional lenders at the primary expense of bank lenders from the mid-1990s to the mid-2000s was due in part to the development of loan securitizations, commonly referred to as collateralized debt obligations (CDOs). These securitization vehicles were very common in the mortgage and consumer debt markets and migrated over into the commercial and buyout loan markets over the past few years. He found that the number of these institutional loan vehicles grew from 150 in 1999 to just over 800 by mid-2007. Peter Rappoport (JPMorgan) reported that, in dollar terms, CDO issuance that focused on corporate lending went from roughly $15 billion in 2003 to over $100 billion in 2006. The preliminary number for the first half of 2007 was nearly $60 billion.

These changes in financing sources were accompanied by financing terms that were more favorable for private equity firms. The analysis by Christopher James (University of Florida), for example, suggests that the general decline in interest rates and risk premiums made debt financing more attractive. Other terms of lending also eased over this period. One indication was the increased origination of so-called "covenant lite" loans in which various types of loan covenants (i.e., conditions placed on the borrower by the lender) were either scaled down or excluded from the loan contract entirely. Blaydon noted that while only 5 percent of U.S. corporate loans could be categorized as "covenant lite" at year-end 2006, that ratio had increased sharply to over 25 percent in the first-half of 2007.

Perhaps the most notable indicator of the degree of easing in financing conditions for LBOs is the increased value of debt firms raised relative to their operating cash flow, or earnings before interest, taxes, depreciation, and amortization (EBITDA). Blaydon noted that, for LBOs, the average ratio of debt to EBITDA increased from about 4 in 2001 to 7 in the first part of 2007.

As discussed by several symposium participants, the credit events of August and September 2007, which originated in the U.S. subprime mortgage markets, spread quickly to the markets for buyout financing (as well as other debt markets) through CDOs and related securitization vehicles. As noted by Rappoport, purchasers of the highest-rated CDO securities, particularly investors in short-term commercial paper backed by these CDO securities, began to require more compensation for taking on the additional funding risks that came to the fore during that period. Once these investors slowed down their securities purchases, the funding of buyout-related transactions, both those already initiated and new ones, slowed dramatically.

Sources of value added

A central issue for the symposium was how private equity investors are able to generate value. In particular, the question is: What elements of private equity acquisitions make a firm worth more than the pre-acquisition stock price of publicly traded firms or the value that private owners place on the firm?

One place to start is with the leverage itself. In LBOs, the financing structures of the acquired firms are changed dramatically. As noted by James, before the LBO, a typical firm's capital structure is roughly two-thirds equity and one-third debt, and after the LBO it is just the reverse, one-third equity and two-thirds debt. Everything else equal, with the tax advantages of debt financing (i.e., fully deductible interest expenses), the value of a firm should increase. Indeed, James finds the LBO acquisition price relative to operating cash flow is strongly positively related to the debt-to-EBITDA ratio. Using a different data source, Chung presented similar trends in the rise in buyout debt-to-earnings ratios and also showed that buyout purchase prices rose accordingly.

Since the tax advantages to debt financing are generally available, the question of why private equity firms are able and willing to use more leverage remains. Several related reasons were suggested during the symposium. One is that private equity firms are able to concentrate the management of the firm on improving performance. In particular, Chung argued that conversion of the firm's management into much more direct ownership within the form of a private firm, as opposed to a public firm, realigned management incentives to emphasize improved performance. For example, Blaydon stated that the absence of regular reporting of performance measures to public shareholders via public filings removes an "earnings myopia" and allows managers to focus more directly on firm profitability. More generally, private equity ownership can be seen as better aligning management incentives to maximize the value of the acquired firms.

Beyond these incentives and the increased latitude for firms to have longer-term decision horizons, the discussion at the symposium also suggested that private equity acquisition involves identifying firms in which operational efficiency can be improved.

Compensation

Another important economic component of the private equity investment business is how the investment managers, the GPs, are compensated. As one might expect, the contract terms are affected by and influence their behavior and performance. Ayako Yasuda (Wharton School of Business at the University of Pennsylvania) presented her research on this topic based on a database of 238 buyout and VC firms

from 1992 to 2006. The standard setup of a private equity investment is that the investor (or LP) commits a fixed level of capital at the inception of a fund, but not all of the commitment is drawn at once. Over the contractual life of the fund, capital is drawn down for making specific investments or for paying the annual management fees, which tend to be 2 percent of the total commitment per year. Typically, funds make investments during the first five years, hold a given investment for three to seven years, and exit them before the fund expires. The GPs receive variable compensation, known as "carry," that is typically 20 percent of the investment return above the original committed capital amount. For example, on a $100 million investment that grows to $150 million over a ten-year investment horizon, the LP pays the GP $20 million (= $100 million x 2 percent x 10 years) in management fees and $10 million (= ($150–$100) million x 20 percent) in carry fees.

Yasuda's research showed that private equity funds expect to receive about 60 percent of their revenue from management fees (and other fixed revenue components) and the remaining 40 percent from carried interest (and other variable revenue components). There are key differences between buyout funds and VC funds. Buyout fund managers are found to earn lower revenue per dollar managed than do VC funds, but they earn substantially higher revenue per partner and per professional than do VC funds. The reason for this result is that buyout funds are more scalable and can grow to a larger size without compromising the abilities and success rates of the GPs. In other words, successful VCs can increase the size of funds, but not the size of individual investments; in contrast, successful buyout funds can increase both the size of the funds and the size of individual investments to generate larger revenues per partner.

This degree of scalability of private equity buyout firms is consistent with the increase in both the size of funds and the size of individual acquisitions discussed by Blaydon. However, it also was noted that the very largest buyout deals in recent years have involved partnerships among several private equity firms. This suggests some limits on scalability, perhaps owing to a goal to limit the degree of concentration of risk exposure for a given fund.

Conclusion

Private equity investment provides an alternative mechanism for corporate governance and financing. While the events of the latter half of 2007 clearly indicate that the degree to which this mechanism is used can be affected by general conditions in the overall capital markets, the economics underlying the approach suggest that buyout funds managed by private equity firms will remain an integral part of the global capital markets.

Jose A. Lopez
Research Advisor

Conference presentations

Some of the presentations listed below are available at http://www.frbsf.org/csip/research/symposium200710.pdf

Blaydon, Colin C. "Overview of the Private Equity Sector."

Chung, Peter Y. "Sources of Value in Private Equity Acquisitions."

Coslet, Jonathan. "Debt Markets and LBO Financing."

James, Christopher M. "Financing Private Equity Acquisitions."

Rappoport, Peter. "Collateralized Debt Markets."

Yasuda, Ayako. "The Economics of Private Equity Funds."

2007 *FRBSF Economic Letters*

Complete texts of *FRBSF Economic Letters* are available on the Bank's website at http://www.frbsf.org/publications/economics/letter/index.html

Subscriptions to the *FRBSF Economic Letter* are available via e-mail or postal delivery.

To subscribe, register at http://www.frbsf.org/tools/allsubscriptions/login.html and complete the form under Economic Research

or send requests to

Public Information Department
Federal Reserve Bank of San Francisco
PO Box 7702
San Francisco, CA 94120
phone (415) 974-2163
fax (415) 974-3341

Federal Reserve Bank of San Francisco
P.O. Box 7702
San Francisco, California 94120

Address Service Requested