

# The Credit Crisis and Cycle-Proof Regulation

Raghuram G. Rajan

This article was originally presented as the Homer Jones Memorial Lecture, organized by the Federal Reserve Bank of St. Louis, St. Louis, Missouri, April 15, 2009.

Federal Reserve Bank of St. Louis *Review*, September/October 2009, 91(5, Part 1), pp. 397-402.

**F**irst, I would like to thank the St. Louis Fed, especially Kevin Kliesen, and the National Association for Business Economics for inviting me to give this talk. I share with Homer Jones an affiliation with the University of Chicago. He was an important influence on Milton Friedman, and if that were all he did, he would deserve a place in history. But in addition, he was a very inquisitive economist with a reputation for thinking outside the box. He made major contributions to monetary economics. It is an honor to be asked to deliver a lecture in his name, especially at this critical time in the nation's regulatory history.

## WHAT CAUSED THE CRISIS?

The current financial crisis can be blamed on many factors and even some particular players in financial markets and regulatory institutions. But in pinning the disaster on specific agents, we could miss the cause that links them all. I argue that this common cause is cyclical euphoria; and, unless we recognize this, our regulatory efforts are likely to fall far short of preventing the next crisis.

Let me start at the beginning. There is some consensus that the proximate causes of the crisis are as follows: (i) The U.S. financial sector misallocated resources to real estate, financed through

the issuance of exotic new financial instruments. (ii) A significant portion of these instruments found their way, directly or indirectly, onto commercial and investment bank balance sheets. (iii) These investments were financed largely with short-term debt. (iv) The mix was potent and caused large-scale disruption in 2007. On these matters, there is broad agreement. But let us dig a little deeper.

This is a crisis born in some ways from previous financial crises. A wave of crises swept through the emerging markets in the late 1990s: East Asian economies collapsed, Russia defaulted, and Argentina, Brazil, and Turkey faced severe stress. In response to these problems, emerging markets became far more circumspect about borrowing from abroad to finance domestic demand. Instead, their corporations, governments, and households cut back on investment and reduced consumption. Formerly net absorbers of financial capital from the rest of the world, a number of these countries became net exporters of financial capital. Combined with the savings of habitual exporters such as Germany and Japan, these circumstances created what Chairman Bernanke referred to as a "global saving glut" (Bernanke, 2005).

Clearly, the net financial savings generated in one part of the world must be absorbed by

---

Raghuram G. Rajan is the Eric Gleacher Distinguished Service Professor of Finance at the Booth School of Business, University of Chicago.

© 2009, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

deficits elsewhere. Corporations in industrialized countries initially absorbed these savings by expanding investment, especially in information technology, but this proved unsustainable and investment was cut back sharply after the collapse of the information technology bubble.

Extremely accommodative monetary policy by the world's central banks, led by the Federal Reserve, ensured the world did not suffer a deep recession. Instead, the low interest rates in a number of countries ignited demand in interest-sensitive sectors such as automobiles and housing. House prices started rising, as did housing investment.

U.S. price growth was by no means the highest. Housing prices reached higher values relative to rent or incomes in Ireland, Spain, the Netherlands, the United Kingdom, and New Zealand, for example. Then why did the crisis first manifest itself in the United States? Probably because the United States went further with financial innovation, thus drawing more buyers with marginal credit quality into the market.

Holding a home mortgage loan directly is very hard for an international investor because it requires servicing, is of uncertain credit quality, and has a high propensity for default. Securitization dealt with some of these concerns. If the mortgage was packaged together with mortgages from other areas, diversification would reduce the risk. Furthermore, the riskiest claims against the package could be sold to those with the capacity to evaluate them and an appetite for bearing the risk, while the safest AAA-rated portions could be held by international investors.

Indeed, because of the demand from international investors for AAA paper, securitization focused on squeezing out the most AAA paper from an underlying package of mortgages: The lower-quality securities issued against the initial package of mortgages were repackaged once again with similar securities from other packages, and a new range of securities, including a large quantity rated AAA, was issued by this "collateralized debt obligation."

The "originate-to-securitize" process had the unintended consequence of reducing the due diligence undertaken by originators. Of course,

originators could not completely ignore the true quality of borrowers because they were held responsible for initial defaults, but because house prices were rising steadily over this period, even this source of discipline weakened.

If the buyer could not make even the nominal payments involved on the initial low mortgage teaser rates, the lender could repossess the house, sell it quickly in the hot market, and recoup any losses through the price appreciation. In the liquid housing market, as long as the buyer could scrawl an "X" on the dotted line, he or she could own a home.

The slicing and dicing through repeated securitization of the original package of mortgages created very complicated securities. The problems in valuing these securities were not obvious when house prices were rising and defaults were few. But as house prices stopped rising and defaults started increasing, the valuation of these securities became very complicated.

## **MALEVOLENT BANKERS OR FOOLISH NAÏFS?**

It was not entirely surprising that bad investments would be made in the housing boom. What was surprising was that the originators of these complex securities—the financial institutions that should have understood the deterioration of the underlying quality of mortgages—held on to so many of the mortgage-backed securities (MBS) in their own portfolios. Simply: Why did the sausage-makers, who knew what was in the sausage, keep so many sausages for personal consumption?

The explanation has to be that at least one arm of the bank thought these securities were worthwhile investments, despite their risk. Investment in MBS seemed to be part of a culture of excessive risk-taking that had overtaken banks. A key factor contributing to this culture is that, over short periods of time, it is very hard, especially in the case of new products, to tell whether a financial manager is generating true excess returns adjusting for risk or whether the current returns are simply compensation for a risk that

has not yet shown itself but will eventually materialize. Such difficulty could engender excess risk-taking both at the top of and within the firm.

For instance, the performance of CEOs is evaluated in part on the basis of the earnings they generate relative to their peers. To the extent that some leading banks can generate legitimately high returns, this puts pressure on other banks to keep up. CEOs of “follower” banks may take excessive risks to boost various observable measures of performance.

Indeed, even if managers recognize that this type of strategy is not truly value creating, a desire to pump up their bank’s stock prices and their own reputations may nevertheless make it their most attractive option. There is anecdotal evidence of such pressure on top management—perhaps most famously from Citigroup chairman, Chuck Prince, in describing why his bank continued financing buyouts despite mounting risks: “When the music stops, in terms of liquidity, things will be complicated. But, as long as the music is playing, you’ve got to get up and dance. We’re still dancing” (Wighton, 2007).

Even if top management wants to maximize long-term bank value, it may be difficult to create incentives and control systems that steer subordinates in this direction. Given the competition for talent, traders have to be paid generously based on performance, but many of the compensation schemes paid for short-term, risk-adjusted performance. This setting gave traders an incentive to take risks that were not recognized by the system, so they could generate income that appeared to stem from their superior abilities, even though it was in fact only a market-risk premium.

The classic case of such behavior is to write insurance on infrequent events such as defaults, assuming what is termed “tail” risk. If traders are allowed to boost bonuses by treating the entire insurance premium as income, instead of setting aside a significant fraction as a reserve for an eventual payout, they have an excessive incentive to engage in this sort of trade.

Indeed, traders who bought AAA-rated MBS were essentially getting the additional spread on

these instruments relative to corporate AAA securities (the spread being the insurance premium) while ignoring the additional default risk entailed in these untested securities. The traders in AIG’s financial products division took all this to an extreme by writing credit default swaps, pocketing the premiums as bonuses, and not bothering to set aside reserves in case the bonds covered by the swaps actually defaulted.

This is not to say that risk managers in banks were unaware of such incentives. However, they may have been unable to fully control them, because tail risks are by their nature rare and therefore hard to quantify with precision before they occur. Although the managers could try to impose crude limits on the activities of the traders taking maximum risk, these types of trades were likely to have been very profitable (before the risk actually was realized) and any limitations on such profits are unlikely to sit well with a top management that is being pressured for profits.

Finally, all these shaky assets were financed with short-term debt. Why? Because in good times, short-term debt seems relatively cheap compared with long-term capital, and the market is willing to supply it because the costs of illiquidity appear remote. Markets seem to favor a bank capital structure that is heavy on short-term leverage. In bad times, though, the costs of illiquidity seem to be more salient, while risk-averse (and burnt) bankers are unlikely to take on excessive risk. The markets then encourage a capital structure that is heavy on capital. Given the conditions that led banks to hold large quantities of MBS and other risky loans (such as those to private equity financed with a capital structure heavy on short-term debt), the crisis had a certain degree of inevitability.

As house prices stopped rising, and indeed started falling, mortgage defaults started increasing. MBS fell in value and became more difficult to price, and their prices became more volatile. They became hard to borrow against, even over the short term. Banks became illiquid and eventually insolvent. Only heavy intervention has kept the financial system afloat, and though the market seems to believe that the worst is over, its relief may be premature.

## **The Blame Game**

Who is to blame for the financial crisis? As my discussion suggests, there are many possible suspects—the exporting countries that still do not understand that their thrift is a burden and not a blessing to the rest of the world; the U.S. households that have spent way beyond their means in recent years; the monetary and fiscal authorities who were excessively ready to intervene to prevent short-term pain, even though they only postponed problems into the future; the bankers who took the upside and left the downside to the taxpayer; the politicians who tried to expand their vote banks by extending homeownership to even those who could not afford it; the markets that tolerated high leverage in the boom only to become risk averse in the bust...The list goes on.

There are plenty of suspects and enough blame to spread. But if all are to blame, should we also not admit they all had a willing accomplice—the euphoria generated by the boom? After all, who is there to stand for stability and against the prosperity and growth in a boom?

Internal risk managers, who repeatedly pointed to risks that never materialized during an upswing, have little credibility and influence—that is, if they still have jobs. It is also very hard for contrarian investors to bet against the boom: As Keynes said, the market can stay irrational longer than investors can stay solvent. Politicians have an incentive to ride the boom, indeed to abet it, through the deregulation sought by bankers. After all, bankers have not only the money to influence legislation but also the moral authority conferred by prosperity.

And what of regulators? When everyone is “for” the boom, how can regulators stand against it? They are reduced to rationalizing why it would be technically impossible for them to stop it.

Everyone is therefore complicit in the crisis because, ultimately, they are aided and abetted by cyclical euphoria. And unless we recognize this, the next crisis will be hard to prevent. For we typically regulate in the midst of a bust when righteous politicians feel the need to do something, when bankers’ frail balance sheets and

vivid memories make them eschew any risk, and when regulators’ backbones are stiffened by public disapproval of past laxity.

## **THE ROLE OF REGULATION**

We reform under the delusion that the regulated—and the markets they operate in—are static and passive and that the regulatory environment will not vary with the cycle. Ironically, faith in draconian regulation is strongest at the bottom of the cycle—when there is little need for participants to be regulated. By contrast, the misconception that markets will take care of themselves is most widespread at the top of the cycle—the point of maximum danger to the system. We need to acknowledge these differences and enact cycle-proof regulation, for a regulation set against the cycle will not stand.

Consider the dangers of ignoring this point. Recent studies such as the Geneva Report (Brunnermeier et al., 2009) have argued for “countercyclical” capital requirements—raising bank capital requirements significantly in good times, while allowing them to fall somewhat in bad times. Although this approach is sensible *prima facie*, these proposals may be far less effective than intended.

To see why this is so, we need to recognize that in boom times, the market demands very low levels of capital from financial intermediaries, in part because euphoria makes losses seem remote. So when regulated financial intermediaries are forced to hold more costly capital than the market requires, they have an incentive to shift activity to unregulated intermediaries, as did banks in setting up structured investment vehicles and conduits during the current crisis.

### **Changes in Regulation**

Even if regulations are strengthened to detect and prevent this shift in activity, banks can subvert capital requirements by assuming risk the regulators do not see or do not penalize adequately with capital requirements. Attempts to reduce capital requirements in busts are equally fraught. The risk-averse market wants banks to hold much

more capital than regulators require, and its will naturally prevails. Even the requirements themselves may not be immune to the cycle. Once memories of the current crisis fade and the ideological cycle turns, the political pressure to soften capital requirements or their enforcement will be enormous.

To have a better chance of creating stability through the cycle—of being cycle-proof—new regulations should be comprehensive, contingent, and cost effective. Regulations that apply comprehensively to all levered financial institutions are less likely to encourage the drift of activities from heavily regulated to lightly regulated institutions over the boom, a source of instability because the damaging consequences of such drift come back to hit the heavily regulated institutions during the bust through channels no one foresees.

Regulations should also be contingent so they have maximum force when the private sector is most likely to do itself harm but bind less the rest of the time. This will make regulations more cost-effective, which also makes them less prone to arbitrage or dilution.

Consider some examples of such regulations. First, instead of asking institutions to raise *permanent* capital, ask them to arrange for capital to be infused when the institution or the system is in trouble. Because these “contingent capital” arrangements will be contracted in good times (when the chances of a downturn seem remote), they will be relatively cheap (compared with raising new capital in the midst of a recession) and thus easier to enforce. Also, because the infusion is seen as an unlikely possibility, firms cannot go out and increase their risks by using the future capital as backing. Finally, because the infusions occur in bad times when capital is really needed, they protect the system and the taxpayer in the right contingencies.

One version of contingent capital is requiring banks to issue debt that would automatically convert to equity when two conditions are met: first, when the system is in crisis, either based on an assessment by regulators or based on objective indicators; and second, when the bank’s capital ratio falls below a certain value (Squam Lake Working Group on Financial Regulation, 2009).

The first condition ensures that banks that do badly because of their own idiosyncratic errors, and not when the system is in trouble, do not avoid the disciplinary effects of debt. The second condition rewards well-capitalized banks by allowing them to avoid the forced conversion (the number of shares to which the debt converts will be set at a level to substantially dilute the value of old equity), while also giving banks that anticipate losses an incentive to raise new equity well in advance.

Another version of contingent capital is requiring systemically important levered financial institutions to buy fully collateralized insurance policies (from unlevered institutions, foreigners, or the government) that will infuse capital into these institutions when the system is in trouble (Kashyap, Rajan, and Stein, 2009).

Here is one way this type of system could operate. Megabank would issue capital insurance bonds—say, to sovereign wealth funds—and invest the proceeds in Treasury bonds, which would then be placed in a custodial account in State Street Bank. Every quarter, Megabank would pay a pre-agreed insurance premium (contracted at the time the capital insurance bond is issued) which, together with the interest accumulated on the Treasury bonds held in the custodial account, would be paid to the sovereign fund.

If the aggregate losses of the banking system exceed a certain prespecified amount, Megabank would start receiving a payout from the custodial account to bolster its capital. The sovereign wealth fund would then face losses on the principal it has invested, but on average, it would be compensated by the insurance premium.

Consider regulations aimed at “too big to fail” institutions. Regulations to limit their size and activities will become very onerous when growth is high, thus increasing the incentive to dilute these regulations. Perhaps, instead, a more cyclically sustainable regulation would be to make these institutions easier to close down. What if systemically important financial institutions were required to develop a plan that would enable them to be resolved over a weekend?

Such a “shelf bankruptcy” plan would require banks to track, and document, their exposures

## Rajan

much more carefully and in a timely manner, probably through much better use of technology. The plan would require periodic stress testing by regulators and the support of enabling legislation—such as facilitating an orderly transfer of a troubled institution’s swap books to precommitted partners. Not only would the requirement to develop resolution plans give these institutions the incentive to reduce unnecessary complexity and improve management, it also would not be much more onerous in the boom cycle and might indeed force management to think the unthinkable at such times.

## CONCLUSION

A crisis offers us a rare window of opportunity to implement reforms—it is a terrible thing to waste. The temptation will be to overregulate, as we have done in the past. This creates its own perverse dynamic. For as we start eliminating senseless regulations once the recovery takes hold, we will find deregulation adds so much economic value that it further empowers the deregulatory camp. Eventually, though, the deregulatory momentum will cause us to eliminate regulatory muscle rather than fat. Perhaps rather than swinging maniacally between too much and too little regulation, it would be better to think of cycle-proof regulation.

## REFERENCES

- Bernanke, Ben S. “The Global Saving Glut and the U.S. Current Account Deficit.” Remarks by Governor Ben S. Bernanke at the Homer Jones Memorial Lecture, St. Louis, Missouri, April 14, 2005; [www.federalreserve.gov/boarddocs/speeches/2005/20050414/default.htm](http://www.federalreserve.gov/boarddocs/speeches/2005/20050414/default.htm).
- Brunnermeier, Markus K.; Crockett, Andrew; Goodhart, Charles A.; Persaud, Avinash D. and Shin, Hyun Song. *The Fundamental Principles of Financial Regulation: Geneva Reports on the World Economy 11*. London: Centre for Economic Policy Research, 2009.
- Kashyap, Anik K.; Rajan, Raghuram G. and Stein, Jeremy C. “Rethinking Capital Regulation” in Federal Reserve Bank of Kansas City Symposium, *Maintaining Stability in a Changing Financial System*, February 2009, pp. 431-71; [www.kc.frb.org/publicat/sympos/2008/KashyapRajanStein.03.12.09.pdf](http://www.kc.frb.org/publicat/sympos/2008/KashyapRajanStein.03.12.09.pdf).
- Squam Lake Working Group on Financial Regulation. “An Expedited Resolution Mechanism for Distressed Financial Firms: Regulatory Hybrid Securities.” Working paper, Council on Foreign Relations, Center for Geoeconomic Studies; April 2009; [www.cfr.org/content/publications/attachments/Squam\\_Lake\\_Working\\_Paper3.pdf](http://www.cfr.org/content/publications/attachments/Squam_Lake_Working_Paper3.pdf).
- Wighton, David. “Citigroup Chief Stays Bullish on Buy-Outs.” *Financial Times*, July 9, 2007; [www.ft.com/cms/s/0/80e2987a-2e50-11dc-821c-0000779fd2ac.html?ncllick\\_check=1](http://www.ft.com/cms/s/0/80e2987a-2e50-11dc-821c-0000779fd2ac.html?ncllick_check=1).



# Systemic Risk and the Financial Crisis: A Primer

James Bullard, Christopher J. Neely, and David C. Wheelock

How did problems in a relatively small portion of the home mortgage market trigger the most severe financial crisis in the United States since the Great Depression? Several developments played a role, including the proliferation of complex mortgage-backed securities and derivatives with highly opaque structures, high leverage, and inadequate risk management. These, in turn, created systemic risk—that is, the risk that a triggering event, such as the failure of a large financial firm, will seriously impair financial markets and harm the broader economy. This article examines the role of systemic risk in the recent financial crisis. Systemic concerns prompted the Federal Reserve and U.S. Department of the Treasury to act to prevent the bankruptcy of several large financial firms in 2008. The authors explain why the failures of financial firms are more likely to pose systemic risks than the failures of nonfinancial firms and discuss possible remedies for such risks. They conclude that the economy could benefit from reforms that reduce systemic risks, such as the creation of an improved regime for resolving failures of large financial firms. (JEL E44, E58, G01, G21, G28)

Federal Reserve Bank of St. Louis *Review*, September/October 2009, 91(5, Part 1), pp. 403-17.

**T**he financial crisis of 2008-09—the most severe since the 1930s—had its origins in the housing market. After several years of rapid growth and profitability, banks and other financial firms began to realize significant losses on their investments in home mortgages and related securities in the second half of 2007. Those losses triggered a full-blown financial crisis when banks and other lenders suddenly demanded much higher interest rates on loans to risky borrowers, including other banks, and trading in many financial instruments declined sharply. A string of failures and near-failures of major financial institutions—including Bear Stearns, IndyMac Federal Bank, the Federal National Mortgage Association (Fannie Mae), the Federal Home Loan Mortgage Corporation (Freddie Mac), Lehman Brothers, American

International Group (AIG), and Citigroup—kept financial markets on edge throughout much of 2008 and into 2009. The financial turmoil is widely considered the primary cause of the economic recession that began in late 2007.

As individual firms lurched toward collapse, market speculation focused on which firms the government would consider “too big” or “too connected” to allow to fail. Why should any firm, large or small, be protected from failure? For financial firms, the answer centers on *systemic risk*. Systemic risk refers to the possibility that a triggering event, such as the failure of an individual firm, will seriously impair other firms or markets and harm the broader economy.

Systemic risk concerns were at the heart of the Federal Reserve’s decision to facilitate the

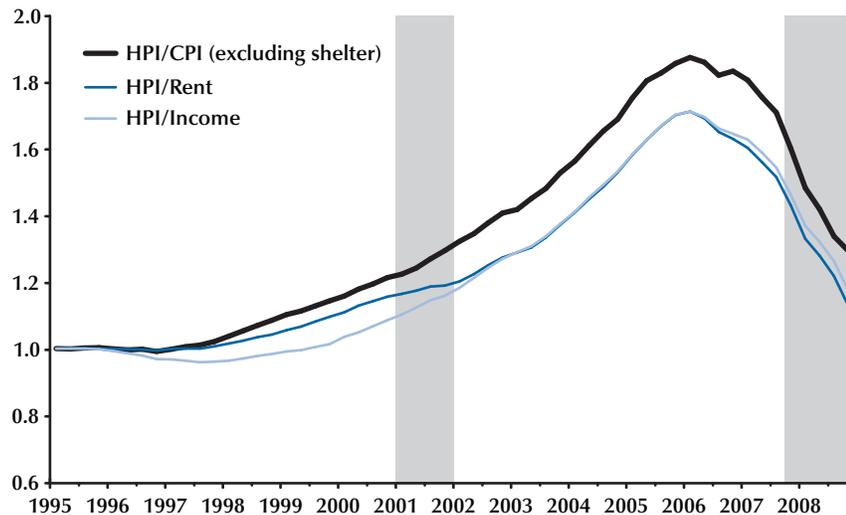
---

James Bullard is president and chief executive officer of the Federal Reserve Bank of St. Louis. Christopher J. Neely is an assistant vice president and economist and David C. Wheelock is a vice president and economist at the Federal Reserve Bank of St. Louis. The authors thank Richard Anderson, Rajdeep Sengupta, and Yi Wen for comments on a previous draft of this article. Craig P. Aubuchon provided research assistance.

© 2009, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

## Figure 1

### U.S. House Prices Relative to the CPI, Rents, and Median Family Income (1995:Q1–2008:Q4)



NOTE: The house price index (HPI) shown in the figure is the S&P/Case-Shiller National Home Price Index; the consumer price index (CPI) data exclude the shelter component of the index; the rent index is a separate component of the CPI; median family income is an aggregated monthly series from the National Association of Realtors; and recession dates (vertical gray bars) are from the National Bureau of Economic Research.

acquisition of Bear Stearns by JPMorgan Chase in March 2008 and the U.S. Department of the Treasury's decisions to place Fannie Mae and Freddie Mac into conservatorship<sup>1</sup> and to assume control of AIG in September 2008. Federal Reserve Chairman Bernanke (2008b) explained the Fed's decision to facilitate the acquisition of Bear Stearns as follows:

Our analyses persuaded us...that allowing Bear Stearns to fail so abruptly at a time when the financial markets were already under considerable stress would likely have had extremely adverse implications for the financial system and for the broader economy. In particular, Bear Stearns' failure under those circumstances would have seriously disrupted certain key secured funding markets and derivatives mar-

kets and possibly would have led to runs on other financial firms.

This article describes how the failure of a single financial firm or market could endanger the entire U.S. financial system and economy and how this possibility influenced the response of policymakers to the recent crisis. Further, we explain why failures of financial institutions are more likely to pose systemic risks than failures of nonfinancial firms and discuss possible remedies for the systemic risks exposed by this particular financial crisis.<sup>2</sup>

## A BRIEF GUIDE TO THE FINANCIAL CRISIS

We begin with a brief review of the evolution of the financial crisis and its origins in the hous-

<sup>1</sup> A conservatorship is a legal arrangement in which one party is given control of another party's legal or financial affairs. In this case, the Federal Housing Finance Agency was appointed conservator of Fannie Mae and Freddie Mac by the U.S. Treasury Department in accordance with the Federal Housing Finance Regulatory Reform Act of 2008.

<sup>2</sup> This article is based on and extends "Systemic Risk and the Macroeconomy" (see Bullard, 2008).

ing market to understand systemic risk in the context of this crisis.

U.S. house prices began to rise far above historical values in the late 1990s. Figure 1 shows the growth in an index of house prices relative to the consumer price index (CPI), an index of residential rents, and median family income, all normalized to equal 1 in the first quarter of 1995. House prices rose rapidly relative to consumer price inflation, rents, and median family income between 1998 and 2006. Analysts attribute the rapid growth in the demand for homes and the associated rise in house prices to unusually low interest rates, large capital inflows, rapid income growth, and innovations in the mortgage market.<sup>3</sup>

A rapid rise in the share of nonprime loans, especially nonprime loans with unconventional terms, was a key feature of the mortgage market during the housing boom. Nonprime loans increased from 9 percent of new mortgage originations in 2001 to 40 percent in 2006 (DiMartino and Duca, 2007). Most nonprime mortgage loans were made to homebuyers with weak credit histories, minimal down payments, low income-to-loan ratios, or other deficiencies that prevented them from qualifying for a prime loan.<sup>4</sup> Many nonprime loans also had adjustable interest rates or other features that kept the initial payments low but subjected borrowers to risk if interest rates rose or house prices declined.

The rise in nonprime loans was accompanied by a sharp increase in the percentage of nonprime loans that originating lenders sold to banks and

other financial institutions. The practice of selling conventional prime mortgages has been common since the 1930s, when the federal government established Fannie Mae to promote the flow of capital to the mortgage market.<sup>5</sup> The federal government chartered Freddie Mac in 1970 to compete with Fannie Mae, which had been sold to private investors in 1968. Both firms purchase large amounts of prime mortgage loans, which they finance by selling bonds in the capital markets. Before the 1990s, Fannie Mae, Freddie Mac, and other firms rarely purchased nonprime loans. Instead, the originating lenders held most nonprime loans, which comprised a relatively small portion of the mortgage market, until they matured.<sup>6</sup>

When a lender sells a loan rather than holding it until maturity, the lender has less incentive to ensure that the borrower is creditworthy. Many analysts contend that lax underwriting standards contributed to the high rate of nonprime loan delinquencies.<sup>7</sup> Although purchasers of loans do have an incentive to verify the creditworthiness of borrowers, many evidently failed to appreciate or manage the level of risk in their portfolios during the recent housing boom (Bernanke, 2008a). In some instances, investors may have relied too heavily on the judgments of credit rating agencies.<sup>8</sup>

The banks and other financial institutions that purchased nonprime mortgage loans typically created residential mortgage-backed securities (RMBSs) based on pools of mortgage loans. An

<sup>3</sup> Bernanke (2005) describes the “global saving glut” and changing pattern of international capital flows during the 1990s and early 2000s, and Caballero, Farhi, and Gourinchas (2008) discuss the role of capital inflows in fueling the housing boom. Taylor (2009), by contrast, blames the housing boom primarily on loose monetary policy during 2002-05.

<sup>4</sup> Mortgage loans are typically classified as prime or nonprime, depending on the risk that a borrower will default on the loan. Nonprime loans are further distinguished between “subprime” and “alternative-A” (Alt-A), again depending on credit risk. Generally, borrowers qualify for prime mortgages if their credit scores are 660 or higher and the loan-to-value ratio is below 80 percent. Borrowers with lower credit scores or other financial deficiencies, such as a previous record of delinquency, foreclosure or bankruptcy, or higher loan-to-value ratios, are more likely to qualify only for a nonprime loan. See Sengupta and Emmons (2007) for more information about nonprime mortgage lending.

<sup>5</sup> Wheelock (2008) discusses the establishment of Fannie Mae and other agencies and programs to alleviate home mortgage distress during the Great Depression.

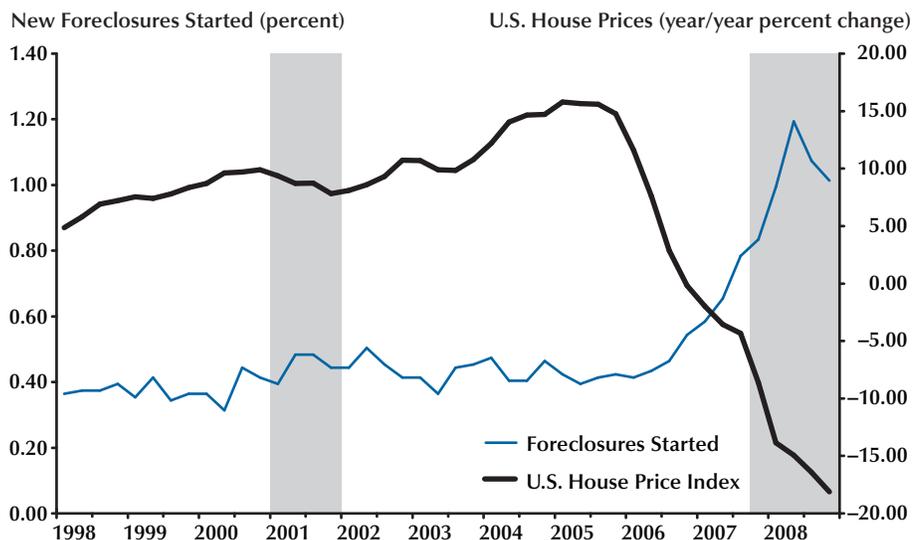
<sup>6</sup> Fannie Mae and Freddie Mac are not permitted to purchase loans that exceed a specific limit (currently \$417,000) except in designated high-cost areas. Further, Fannie Mae and Freddie Mac require minimum documentation and other standards on the loans they purchase, and hence they purchase relatively few nonprime loans.

<sup>7</sup> Demyanyk and Van Hemert (2008) and Bhardwaj and Sengupta (2008) provide alternative perspectives on the role of lax underwriting of nonprime loans.

<sup>8</sup> Critics charge that the rating agencies had a conflict of interest because bond issuers paid for the ratings (*New York Times*, 2007; Fons, 2008a,b). In addition, the rating agencies used inadequate risk models that did not account for a possibility of a serious drop in housing prices. See Fons (2008a,b).

**Figure 2**

**U.S. House Prices and Foreclosures**



NOTE: Foreclosures data are from the Mortgage Bankers Association; the house price index (HPI) is the S&P/Case-Shiller National Home Price Index. Vertical gray bars indicate recessions.

RMBS redistributes the income stream from the underlying mortgage pool among bonds that differ by the seniority of their claim. Sometimes additional securities, known as collateralized mortgage obligations (CMOs) or collateralized debt obligations, are created by combining multiple RMBSs (or parts of RMBSs) and then selling portions of the income streams derived from the mortgage pool or RMBSs to investors with different appetites for risk.

The securities rating agencies assigned high ratings to many of the mortgage-related securities created to finance purchases of nonprime loans. As long as house prices were rising, most nonprime loans performed well because borrowers were usually able to refinance or sell their house—at a higher price—if they were unable to make their loan payments.<sup>9</sup> When house prices began

to fall, many borrowers found that they owed more on their house than it was worth. This situation made it impossible for some borrowers to repay their loan by selling their house or refinancing their mortgage, and it also created an incentive simply to default. Consequently, loan defaults and foreclosures rose sharply, as shown in Figure 2, which plots data on the percentage of home mortgages entering foreclosure in a given quarter and the year-over-year percentage change in the S&P/Case-Shiller National Home Price Index.

Rising loan delinquencies caused many RMBSs and CMOs backed by home mortgage loans to default, and investment banks and other investors that held large portfolios of RMBSs and CMOs experienced substantial losses. Ultimately, the decline in house prices and the increase in mortgage loan defaults that began in 2006 were the root cause of the financial crisis. The following sections explore how systemic risks caused losses on nonprime mortgages and mortgage-related securities to disrupt the entire financial system.

<sup>9</sup> Most nonprime loan originations were refinances of existing mortgages in which borrowers withdrew accumulated equity from their homes (a phenomenon known as a “cash-out” refinance). See Bhardwaj and Sengupta (2009).

## SYSTEMIC RISK

### *Systemic Risk, Counterparty Risk, and Asymmetric Information*

In the recent financial crisis, the most important type of risk to the financial system has been “counterparty risk,” which is also known as “default risk.”<sup>10</sup> Counterparty risk is the danger that a party to a financial contract will fail to live up to its obligations.

Counterparty risk exists in large part because of *asymmetric information*. Individuals and firms typically know more about their own financial condition and prospects than do other individuals and firms. Much of the recent concern about systemic risk has focused on investment banks that deal in complex financial contracts. Consider the following example: Suppose Bank A purchases an option from Bank B to hedge the risk of a change in the term structure of interest rates. If Bank B later fails, perhaps because of bad investments in home mortgages, then the option sold by Bank B may lose value or even become worthless. Thus, Bank A—which thought it was carefully hedging its risk—is adversely affected by Bank B’s problems in housing markets.

Of course, financial firms can protect themselves to some degree in such simple situations. The logic of self-interested behavior combined with market clearing would lead to an appropriate pricing of risk; Bank A would have considered the possibility of the failure of Bank B and taken this into account in its contingency plan. For example, Bank A might require Bank B to post collateral to protect the value of the option in case Bank B failed. But in actual financial markets, arrangements are so complex that the nature of risk that firms face might not be obvious. In addition, the value of collateral fluctuates and thus even carefully collateralized deals are subject to some risk.<sup>11</sup>

<sup>10</sup> Taylor (2009) argues that the financial crisis was associated mainly with an increase in counterparty risk and not a shortage of liquidity.

<sup>11</sup> Kiyotaki and Moore (1997) and Pintos and Wen (2008) discuss how procyclical fluctuations in the value of collateral can exacerbate financial booms and busts and contribute to macroeconomic fluctuations.

### *Systemic Risk and Information Cascades*

Sophisticated investors and counterparties will cease to do business with a firm once the firm’s weak condition becomes known, as they did with Bear Stearns and Lehman Brothers. However, the inability to sort perfectly among good and bad risks can lead banks and other investors to pull away from nearly all lending during a crisis. The tendency of lenders to seek safe investments during a crisis explains why trading in risky assets declined sharply and their market yields rose relative to yields on federal government debt during 2007-08.

Sometimes all firms in an industry are “tarred by the same brush” and one firm’s failure leads investors to shun an entire industry. For example, before the introduction of federal deposit insurance in 1933, the failure of individual banks sometimes caused the public to shift a large portion of its funds from bank deposits into cash. Why should the failure of a single firm cause the public to suspect an entire industry? Again, the answer is related to the fact that people have imperfect information. Because depositors lack complete information about the condition of their bank, the failure of one bank can trigger mass withdrawals by depositors of other banks to avoid losses in the event their own bank fails. Indeed, even if a particular depositor believed that his bank was fundamentally sound, it would still make sense for him to withdraw his money if he thought that withdrawals by other depositors might cause the bank to fail. Banking panics are especially dangerous because large-scale deposit withdrawals can make bank failures more likely, as well as cause banks to reduce their lending in an effort to boost liquidity. Several severe banking panics during the nineteenth and early twentieth centuries resulted in widespread bank failures, financial distress, and economic contractions.<sup>12</sup>

Federal deposit insurance has largely ended the problem of banking panics. When IndyMac Bank was rumored to be near failure in 2008,

<sup>12</sup> Calomiris and Gorton (2000), Dwyer and Gilbert (1989), and Wicker (2000) are among the numerous studies of the causes and effects of U.S. banking panics. Diamond and Dybvig (1983) provide an important theoretical analysis of banking panics.

many depositors withdrew their funds from the bank. However, rather than holding their funds as cash, IndyMac's depositors merely moved their deposits to other banks. Similarly, the run on IndyMac did not trigger mass deposit withdrawals at other banks.

Panic-like phenomena have occurred during the recent financial crisis, however. For example, when the Reserve Primary Fund, a large money market mutual fund, halted investor redemptions after the net asset value of its shares fell below \$1 in September 2008, share redemptions rose sharply at other money market mutual funds. Although most money market mutual funds had ample reserves and good assets, investors interpreted the troubles of the Reserve Primary Fund (which held a large amount of Lehman Brothers debt) as a possible indicator of problems at other mutual funds. The federal government quickly guaranteed the value of existing accounts in money market mutual funds to discourage panic withdrawals from such funds.

The dramatic declines in trading volume and liquidity in the markets for mortgage-related securities during the recent financial crisis also reflected investor panic. Trading in all RMBSs declined sharply when defaults and ratings downgrades made investors wary of RMBSs in general. Heavy reliance by investors on the evaluation of mortgage instruments by the rating agencies may have exacerbated swings in market liquidity. For example, a ratings downgrade, especially of a previously highly rated security, could induce panic selling by signaling possible downgrades or losses on similar securities. Ratings downgrades and declining asset values can also force borrowers to post additional collateral to maintain a given level of borrowing. AIG collapsed in September 2008 when it was unable to raise additional collateral in the wake of a downgrade of its debt rating (Son, 2008). In general, deterioration in the collateral value of borrower assets was an important amplification mechanism during the recent financial crisis. Falling asset prices caused lenders to demand more collateral, which caused borrowers to dump risky assets, thereby exacerbating declines in their market values and leading to further demands for more collateral (Brunnermeier, 2008).

## ***Why the Financial System Is Special***

Many aspects of systemic risk are not unique to financial institutions or markets. The failure of a nonfinancial firm, such as an automobile manufacturer, will affect the firm's suppliers and dealerships, as well as the local economies where manufacturing plants and other operations are located. By the same token, a default by an airline company on its debt obligations might cause investors to shun the debt of other airline companies if investors believe that the default reflected an industry-wide problem, such as rising fuel prices. Still, over the past decade, some very large firms have failed, including Enron, WorldCom, and several major airlines, yet none caused significant problems beyond its immediate shareholders, employees, suppliers, and customers. The failure of a nonfinancial firm would rarely threaten the solvency of a competitor, let alone significantly affect the economy more broadly. Instead, the failure of a large firm could increase the market shares and profitability of the remaining firms in an industry, as well as provide opportunities for smaller firms to enter previously inaccessible markets.

Why do we think the failure of a large financial firm presents systemic risks that the failure of a nonfinancial firm does not? There are at least three reasons.

The first is interconnectedness. In the normal course of business, large commercial and investment banks lend and trade with each other through interbank lending and deposit markets, transactions in over-the-counter (OTC) derivatives, and wholesale payment and settlement systems. Settlement risk—the risk that one party to a financial transaction will default after the other party has delivered—is a major concern for large financial institutions whose daily exposures routinely run into many billions of dollars. The lightning speed of financial transactions and the complex structures of many banks and securities firms make it especially difficult for a firm to fully monitor the counterparties with which it deals, let alone the counterparties of counterparties. The rapid failure of a seemingly strong bank could potentially expose other firms to large losses.

Even firms that do not transact directly with the affected bank can be exposed through their dealings with affected third parties.<sup>13</sup>

A second reason why the financial sector is especially vulnerable to systemic risk is leverage. Compared with most nonfinancial firms, banks and other financial institutions are highly leveraged—that is, they fund a substantial portion of their assets by issuing debt rather than selling equity. During the housing boom, many banks, hedge funds, and other firms that invested heavily in mortgage-related securities financed their holdings by borrowing heavily in debt markets. Investment banks were especially highly leveraged before the crisis, with debt-to-equity ratios of approximately 25 to 1. That is, for every dollar of equity, investment banks issued an average of \$25 of debt. By comparison, commercial banks, which are subject to minimum capital requirements, had leverage ratios of approximately 12 to 1.<sup>14</sup> High leverage meant that financial firms enjoyed high rates of return on equity when times were good but also a high risk of failing when markets turned against them.

Because investment banks held a mere \$4 of equity for every \$100 of assets on their balance sheets, a relatively modest (4 percent) decline in the value of an investment bank's assets would wipe out the bank's equity, forcing it to raise additional capital and/or sell some of its assets. Many investment banks and other financial institutions sustained large losses on their portfolios of RMBSs and were forced to raise additional capital to remain solvent. Similarly, Fannie Mae and Freddie Mac ran into financial difficulties in part because of their extreme leverage. The federal government placed both Fannie Mae and Freddie Mac into conservatorship in July 2008 because losses on their portfolios of mortgages and RMBSs drove the firms to the brink of insolvency. Had those firms held more capital, they could have withstood larger losses without becoming insolvent.

<sup>13</sup> Lagunoff and Schreft (2001) present a model in which a financial crisis can arise as losses spread among firms whose portfolios are linked to those of other firms.

<sup>14</sup> See *Economic Report of the President* (2009, p. 71).

A third reason why the financial sector is especially vulnerable to systemic risk is the tendency of financial firms to finance their holdings of relatively illiquid long-term assets with short-term debt. Not only are financial institutions typically highly leveraged, but the nature of their business entails an inherent mismatch in the maturities of their assets and liabilities that can make them vulnerable to interest rate or liquidity shocks. Most financial intermediaries borrow short and lend long—that is, they fund long-term, relatively illiquid investments with short-term debt. For example, commercial banks traditionally have used demand deposits, which depositors can withdraw at any time, to fund loans and other long-term investments. Many investment banks and securities firms rely heavily on commercial paper, repurchase agreements (repos),<sup>15</sup> and other short-term funding sources to finance long-term investments. If depositors suddenly pull their funds from a commercial bank or lenders refuse to purchase a securities firm's commercial paper or repos, the bank or securities firm could be forced into bankruptcy. Bear Stearns collapsed when investors refused to purchase the firm's short-term debt. Other firms faced sharply higher funding costs in 2007-08 as markets reevaluated the creditworthiness of borrowers. The speed with which the markets can "turn off the tap" makes financial institutions especially vulnerable to temporary disruptions of liquidity in financial markets.<sup>16</sup>

## MITIGATING SYSTEMIC RISK

Recognizing the problem of systemic risk, financial firms have long cooperated to limit risks associated with the failures of other financial firms. For example, before the creation of the Federal Reserve System in 1913, commercial banks devised clearinghouse arrangements in an

<sup>15</sup> A repo is a trade in which one party agrees to sell securities to a second party and to buy those securities back at a prespecified price and date. It amounts to collateralized borrowing.

<sup>16</sup> Acharya, Gale, and Yorulmazer (2009) present a model that can explain a sudden collapse of liquidity in a financial market associated with a change in the information structure of the assets traded in the market.

attempt to protect themselves from banking panics. The primary purpose of a clearinghouse is to clear checks and other forms of payment among member banks. In the nineteenth century, clearinghouses developed mechanisms to protect their members from banking panics and to provide additional liquidity for banks facing deposit runs. For example, clearinghouse members could borrow certificates to settle their balances with other member banks in lieu of cash or other reserves. Further, clearinghouse members collectively guarantee the payment obligations of members threatened by deposit withdrawals.<sup>17</sup>

Financial market exchanges, such as the Chicago Mercantile Exchange, are also private arrangements that limit systemic risks. Securities and commodities exchanges arose centuries ago to settle trades efficiently under clear, fixed rules. Exchanges are the central counterparty to every transaction. Like bank clearinghouses, exchanges reduce default risk by requiring their members to meet minimum capital and disclosure requirements. If a member of the exchange does default, the other members bear that firm's obligations according to the exchange's loss-sharing rules. Thus, membership requirements and loss-sharing arrangements lessen the risk that default by one firm will adversely affect other members of the exchange.

Many derivatives trade in OTC markets, which consist of financial institutions doing business directly with each other rather than through an exchange. Many analysts have identified weaknesses in OTC derivatives markets, especially in the market for credit default swaps, as important contributors to the recent financial crisis.<sup>18</sup>

The use of credit default swaps and other financial derivatives has grown enormously in recent years. Although useful for hedging risks, the proliferation of OTC derivatives is widely believed to have increased systemic risks in the financial system by increasing the extent to which

large financial firms are interconnected and by reducing transparency. Many analysts believe that these risks could be substantially reduced by establishing a central exchange or clearinghouse for derivatives trading.<sup>19</sup> Because exchange-traded derivatives are standardized contracts that are traded among many parties every day, they could be valued more precisely than the custom products traded among individual firms on OTC markets. In addition, the requirement for exchange participants to post margins against potential losses and mark positions to market daily would help reduce counterparty risks. Exchange participants that cannot cover their losses will have their positions closed out before the losses become too large.

Cooperative arrangements, such as clearinghouses and exchanges, are one way of reducing systemic risks. However, in many circumstances private measures might be insufficient to ameliorate systemic risk. For example, individual firms could be reluctant to reveal private business information to competitors, which might impair a loss-sharing agreement. Further, firms often have little incentive to mitigate costs borne by others. Thus, a firm whose failure poses systemic risk will tend to behave less cautiously than society would desire and, hence, government involvement might be necessary to limit systemic risks.<sup>20</sup>

### ***Proposals for Government Policies to Control Systemic Risks***

The recent financial crisis has prompted numerous proposals for enhanced government regulation and supervision of large financial firms and markets to address systemic risks. Many proposals call for increased supervision of systemically important financial institutions, as well as new rules for resolving insolvent firms.

---

<sup>17</sup> See Gorton (1985), Timberlake (1984), and White (1983, pp. 74-83) for more information about the role of clearinghouses in mitigating banking panics.

<sup>18</sup> Wallison (2008) discusses the credit default insurance (or swap) market, and Schinasi et al. (2001) describe the OTC derivatives market.

---

<sup>19</sup> For example, see Bernanke (2008c) and Counterparty Risk Management Policy Group III (2008).

<sup>20</sup> Systemic risk constitutes a "negative externality" in the sense that the actions of one firm harm others. The situation is analogous to a firm that pollutes the environment. Because others bear at least some of the costs of the pollution, the firm will tend to pollute more than it would if it had to compensate others for these costs. Negative externalities are an example of a market failure that may require government intervention to ameliorate.

Other proposals recommend regulation to limit risk-taking and to ensure ample liquidity in financial markets. This section reviews some of the regulatory and legal proposals suggested in response to the recent crisis.

Many reform proposals call for the creation of a systemic financial regulator with responsibilities for “macroprudential” oversight of the financial system. A macroprudential regulator would consider broad economic trends and the impact of a firm’s actions on the entire financial system, not just the firm’s own risk of default (Bernanke, 2008c). To some extent, regulators already consider broad economic trends and effects, but several proposals argue for bringing all large or systemically important financial institutions under the umbrella of a systemic regulator.<sup>21</sup>

One justification for the regulation and supervision of systemically important firms is that governments are unlikely to permit such firms to fail, or if they do fail, the government will substantially protect many, if not all, of the firm’s creditors from loss. Such a government guarantee—either explicit or implicit—can encourage firms to take greater risks than they otherwise would, which increases the likelihood of their failure.<sup>22</sup> Consequently, regulation and supervision is required to offset the incentive to take excessive risk.

Federal deposit insurance is one example of a government guarantee that can encourage excessive risk-taking. Without deposit insurance, rational, fully informed depositors would require banks with risky assets to hold more capital or pay higher deposit rates than banks with less-risky

assets. However, with insurance, depositors have little incentive to monitor the risks their banks take; hence, deposit insurance gives banks an incentive to assume greater risks than they otherwise would.<sup>23</sup>

Whereas deposit insurance is an explicit guarantee, the public’s expectation that the federal government would stand behind the liabilities of Fannie Mae and Freddie Mac is an example of an implicit guarantee. The perception that the government would guarantee the liabilities of Fannie Mae and Freddie Mac enabled those firms to borrow at relatively low interest rates to fund their purchases of mortgages and RMBSs, including securities backed by nonprime mortgages. Fannie Mae and Freddie Mac grew rapidly and operated with much lower capital ratios than other financial firms. Ultimately, financial losses eroded their thin capital cushions and pushed both Fannie Mae and Freddie Mac to the brink of failure before they were placed into government conservatorship.<sup>24</sup>

Many policymakers and analysts have called for new rules for shutting down large financial firms that become insolvent. The current bankruptcy regime is widely criticized as inadequate for dealing with failures of systemically important financial institutions.<sup>25</sup> Delays and uncertainties inherent in the bankruptcy process of a systemically important firm could precipitate or exacerbate a financial crisis.

Several reform proposals advocate subjecting nonbank financial firms to “prompt corrective action” in the event their capital ratios fall below prescribed levels. The Federal Deposit Insurance Corporation Improvement Act of 1991 already

<sup>21</sup> For example, the Group of Thirty (2009, p. 17) argues that at the start of 2008, there were five U.S. investment banks (Bear Stearns, Goldman Sachs, Lehman Brothers, Merrill Lynch, and Morgan Stanley), one insurance company (AIG), and two government-sponsored enterprises (Fannie Mae and Freddie Mac) that were systemically significant and therefore should have been subject to stringent regulation and supervision. During 2008, all but two of those firms (Goldman Sachs and Morgan Stanley) failed or suffered large losses that required government intervention, and both Goldman Sachs and Morgan Stanley became bank holding companies.

<sup>22</sup> “Moral hazard” describes the idea that individuals and firms engage in riskier behavior when they are protected from the danger that such behaviors create. For example, a person who purchases fire insurance might be less concerned with fire hazards than one who would personally bear the full cost of a fire.

<sup>23</sup> Merton (1977) shows that banks maximize the value of deposit insurance to themselves by maximizing their risk. Capital requirements and other measures can limit the excessive risk-taking encouraged by deposit insurance. Many analysts blame lax regulation and supervision, coupled with an increase in deposit insurance coverage, for increased risk-taking and the high failure rates of banks and, especially, savings and loan associations during the 1980s. For example, see Kane (1989) and White (1991).

<sup>24</sup> Poole (2002, 2003, and 2007) was among those warning of the risks inherent with the implicit government guarantee of Fannie Mae and Freddie Mac debt. Stern and Feldman (2004) discuss the effects of “too big to fail” policies in general.

<sup>25</sup> For example, see Bernanke (2008c) and Congressional Oversight Panel (2009, p. 24).

mandates prompt corrective action for commercial banks. For example, bank supervisors can limit the growth, executive compensation, and payment of dividends by undercapitalized banks. Supervisors can also place critically undercapitalized banks into conservatorship or receivership.<sup>26</sup> Federal Reserve Chairman Bernanke (2008c) and others argue that prompt corrective action could reduce systemic risks and discourage large financial holding companies and nonbank financial firms from taking excessive risks. Further, the authority to place a critically undercapitalized firm into conservatorship or receivership would enable the government to resolve failures in an orderly way that imposes the failing firm's losses on the firm's creditors and equity holders rather than on taxpayers.

Prompt corrective action is one potential component of a general strengthening of the oversight of large financial firms. Another potential component is a more comprehensive approach to the supervision of complex and systemically important financial firms. Proponents argue that broader supervision of systemically significant firms might have prevented the failure of AIG, which required a government rescue to avoid bankruptcy in September 2008.

AIG is a large financial conglomerate with global operations. The traditional business of AIG is insurance—automobile, life, and so on. In the United States, state government authorities regulate insurance firms—New York State in the case of AIG. State insurance regulations and supervision are designed to ensure the solvency of insurance companies so that they are fairly certain to meet their contingent claims. But insurance regulators have little or no oversight of the other subsidiaries and operations of conglomerates such as AIG. Besides owning an insurance company, AIG also owns a federally chartered savings bank (AIG Bank, FSB), which places AIG under the supervision of the Office of Thrift Supervision. Bank and thrift regulators, however, traditionally

have focused on the condition of the depository institution rather than on the systemic risks posed by its parent holding company. The Office of Thrift Supervision has neither the resources to supervise the activities of the entire conglomerate nor the mandate to regulate the extent to which AIG poses systemic risk to the financial system.

AIG's unregulated activities, notably the underwriting of credit default insurance, created substantial losses as the housing market slumped badly in 2006-08. These unregulated operations had grown so large that government officials feared that AIG's sudden collapse could impose severe losses on other firms and seriously impair the functioning of the entire financial system. To avoid this outcome the U.S. Treasury and Federal Reserve provided AIG with loans and a capital injection in September 2008 when it appeared that the firm would default on its outstanding debts.

Many proposals for reforming financial regulation call for the supervision of large, complex financial institutions such as AIG by strong regulators with sweeping oversight and enforcement powers that can focus on the systemic risks posed by such organizations.<sup>27</sup> Brunnermeier et al. (2009) argue that an effective macroprudential regulator must have the political independence to impose unpopular measures. To limit discretion, the study argues, regulation should follow preset rules as much as possible. Writing rules to cover every possible contingency is difficult if not impossible, however, and before assigning sweeping oversight and enforcement authority to a systemic regulator, the scope of the regulator's authority would have to be carefully delineated.

In addition to enhanced macroprudential oversight, proposals for mitigating systemic risks in the financial system include the imposition of minimum capital requirements on large financial firms, regulations on the use of short-term debt to finance holdings of long-term assets, and changes to market value accounting rules.

Many analysts contend that extreme leverage contributed to the recent financial crisis by making large financial firms especially vulnerable to

<sup>26</sup> Aggarwal and Jacques (1998) and Spong (2000, pp. 84-98) provide additional information about commercial bank capital requirements and prompt corrective action. Evanoff and Wall (2003) argue that the use of subordinated debt spreads might be useful to trigger prompt corrective action.

<sup>27</sup> For example, see Brunnermeier et al. (2009), Congressional Oversight Panel (2009), Group of Thirty (2009), and Paulsen et al. (2008).

losses. This view has prompted proposals to strengthen capital requirements for commercial banks and to extend those requirements to previously unregulated financial firms, such as investment banks. Some analysts argue that large, systemically significant firms should be required to hold more capital as a percentage of their assets than smaller firms (e.g., Congressional Oversight Panel, 2009, p. 26).

Some proposals call for discouraging the funding of long-term, illiquid assets with short-term debt. A firm that cannot roll over its short-term debt could be forced to sell assets, and if many firms are in the same predicament, then asset prices could decline sharply. Such price declines would impose further losses on firms, forcing a spiral of still more sales and further price declines. As the recent financial crisis intensified, especially in September 2008, firms that relied heavily on short-term debt faced sharply higher interest rates as banks suddenly became less willing to lend and investors fled to the safety of U.S. Treasury securities.<sup>28</sup> Future systemic risks could be reduced by discouraging excessive leveraging and the use of short-term debt to fund long-term asset holdings, for example, by requiring firms to hold more capital against long-term, relatively illiquid assets funded with short-term debt than against more-liquid assets or assets funded with long-term debt (e.g., Brunnermeier et al., 2009, pp. 38-39). Kotlikoff and Leamer (2009) offer a more radical solution to the problem of short-term debt financing illiquid assets: “limited-purpose banking.” This scheme would convert all financial firms to mutual funds so that individual depositors, not the financial firms, would bear the risk of the asset holdings.

Noting the tendency of financial firms to increase their use of leverage when asset prices are rising and to reduce leverage when prices are falling, some analysts argue that capital requirements should become increasingly stringent when asset prices are rising. Some proposals call for tying capital requirements explicitly to the growth in the value of a bank’s assets (e.g., Congressional

Oversight Panel, 2009, pp. 27-28); others call on bank supervisors to encourage banks to build capital and liquidity when times are good and allow banks to draw down their buffers during difficult times. For example, the Group of Thirty (2009, p. 43) recommends capital requirements “expressed as a broad range within which capital ratios should be managed, with the expectation that, as part of supervisory guidance, firms will operate at the upper end of such a range in periods when markets are exuberant and tendencies for underestimating and underpricing risk are great.”

One of the more hotly debated issues surrounding the recent financial crisis is the extent to which fair value accounting rules contributed to the crisis. In textbook financial markets, valuations are the considered outcomes of the views of rational, relatively risk-tolerant speculators with deep pockets. In the real world, however, imperfect information and limited risk tolerance are facts of life that can inhibit the rational speculation necessary to drive prices back to long-term fundamental values. Trading in certain assets might cease during a crisis or trades might occur at widely disparate prices, making the determination of their market value problematic just when the value of transparency is greatest. In addition, by forcing financial firms to realize declines in asset prices immediately, mark-to-market rules might exacerbate a crisis by encouraging asset sales when prices are already falling, leading to further write-downs and financial losses.<sup>29</sup>

The Group of Thirty (2009, p. 46) calls for applying “more realistic” accounting guidelines to less-liquid assets and distressed markets but is generally supportive of fair value accounting. Similarly, Brunnermeier et al. (2009, pp. 36-37) advocate a “mark-to-funding” approach to fair value accounting in which the value of an asset is tied to the funding of that asset. For example, if an asset that matures in 20 years is financed

<sup>28</sup> The danger of issuing short-term debt is not limited to firms. Neely (1996) describes the role of short-term debt in triggering Mexico’s December 1994 peso crisis.

<sup>29</sup> Blanchard (2008) compares the current financial crisis with nineteenth-century bank runs. He points out that the opacity of the mortgage-backed assets has served to amplify the financial crisis by making those assets particularly difficult to value, lowering their resale price, and increasing uncertainty about financial firms’ solvency. Similarly, the high degree of leverage of financial institutions increases the probability that any losses will lead to insolvency.

with debt that matures in 30 days, the asset should be valued at the expected price of the asset in 30 days. Of course, calculating expected future prices in any reasonable way is difficult and the authors acknowledge that their scheme would give firms some discretion over the valuation of their assets. However, they argue that it would more accurately relate the value of assets to funding risks.

## CONCLUSION

The recent financial crisis has claimed many victims. Several prominent firms, including Bear Stearns, Lehman Brothers, AIG, Fannie Mae, and Freddie Mac, have gone bankrupt or required government intervention to prevent their failure. When the U.S. Treasury Department and the Federal Reserve intervened to prevent a failure, their goal was to protect the financial system—and the economy—from systemic risk.

Financial firms are much more susceptible to systemic risk than nonfinancial firms because financial firms are typically highly interconnected with one another, highly leveraged, and tend to use short-term debt to finance their holdings of long-term, relatively illiquid assets. In the recent crisis, the possible failure of counterparties in complex transactions created systemic risk.

Financial firms are cognizant of systemic risk and traditionally have tried to reduce their vulnerability to it by participating in clearinghouses or trading through financial exchanges. Nevertheless, because firms do not bear all the costs of their own failure, government has a role to play in limiting systemic risk in the financial system to protect the broader economy. Analysts have proposed regulatory reforms to reduce the danger from systemic risk in the future. In particular, some advocate the creation of a powerful macroprudential regulator that considers a firm's impact on the stability of the entire financial system. Other ideas for reducing systemic risk include limiting the use of leverage and short-term debt and revising market value accounting rules.

It is too soon to fully determine the causes of the recent financial crisis. Asset price booms and busts that impair the financial system and

the entire economy have occurred before. However, the complex nature of recently developed financial instruments has transmitted the consequences of the housing bust to the entire financial system and, ultimately, to the overall economy. Accordingly, many analysts favor measures to increase the use of organized exchanges for trading derivatives.

An improved regime for resolving large insolvent financial firms would limit systemic risk and excessive risk-taking. When the government has intervened to protect the economy from the failure of a large systemically important financial firm, the shareholders of these firms usually received little or no value for their equity and their senior managers were dismissed or subject to compensation limits. However, bondholders doubtless received more compensation than they would have in the absence of government intervention. A legal reform that permits rapid resolution of failing financial firms, including appropriate reductions in payments to bondholders, would help to create incentives for bondholders to be mindful of the risk of their investments. This, in turn, would discourage excessive risk-taking by increasing the borrowing costs for risky firms. The economy could benefit from reforms that reduce the risks to the financial system imposed by firms that are “too big to fail.”

## REFERENCES

- Acharya, Viral; Gale, Douglas and Yorulmazer, Tanju. “Rollover Risk and Market Freezes.” New York University and Federal Reserve Bank of New York Working Paper, October 2008, updated February 2009; [www.newyorkfed.org/research/conference/2009/cblt/Acharya-Gale-Yorulmazer.pdf](http://www.newyorkfed.org/research/conference/2009/cblt/Acharya-Gale-Yorulmazer.pdf).
- Aggarwal, Raj and Jacques, Kevin T. “Assessing the Impact of Prompt Corrective Action on Bank Capital and Risk.” Federal Reserve Bank of New York *Economic Policy Review*, October 1998, pp. 23-32; [www.newyorkfed.org/research/epr/98v04n3/9810agga.pdf](http://www.newyorkfed.org/research/epr/98v04n3/9810agga.pdf).
- Bernanke, Ben S. “The Global Saving Glut and the U.S. Current Account Deficit.” Remarks at the Homer Jones Memorial Lecture, St. Louis, April 14,

- 2005; [www.federalreserve.gov/boarddocs/speeches/2005/20050414/default.htm](http://www.federalreserve.gov/boarddocs/speeches/2005/20050414/default.htm).
- Bernanke, Ben S. "Risk Management in Financial Institutions." Speech at the Federal Reserve Bank of Chicago's Annual Conference on Bank Structure and Competition, Chicago, May 15, 2008a; [www.federalreserve.gov/newsevents/speech/bernanke20080515a.htm](http://www.federalreserve.gov/newsevents/speech/bernanke20080515a.htm).
- Bernanke, Ben S. "Financial Regulation and Financial Stability." Speech at the Federal Deposit Insurance Corporation's Forum on Mortgage Lending for Low and Moderate Income Households, Arlington, Virginia, July 8, 2008b; [www.federalreserve.gov/newsevents/speech/bernanke20080708a.htm](http://www.federalreserve.gov/newsevents/speech/bernanke20080708a.htm).
- Bernanke, Ben S. "Reducing Systemic Risk." Speech at the Federal Reserve Bank of Kansas City's Annual Economic Symposium in Jackson Hole, Wyoming, August 22, 2008c; [www.federalreserve.gov/newsevents/speech/bernanke20080822a.htm](http://www.federalreserve.gov/newsevents/speech/bernanke20080822a.htm).
- Bhardwaj, Geetesh and Sengupta, Rajdeep. "Where's the Smoking Gun? A Study of Underwriting Standards for U.S. Subprime Mortgages." Working Paper No. 2008-036B, Federal Reserve Bank of St. Louis, October 27, 2008; revised May 2009; <http://research.stlouisfed.org/wp/2008/2008-036.pdf>.
- Bhardwaj, Geetesh and Sengupta, Rajdeep. "Did Prepayments Sustain the Subprime Market?" Working Paper No. 2008-039B, Federal Reserve Bank of St. Louis, April 2009; <http://research.stlouisfed.org/wp/2008/2008-039.pdf>.
- Blanchard, Olivier J. "The Crisis: Basic Mechanisms, and Appropriate Policies." Working Paper No. 09-01, MIT Department of Economics, December 29, 2008; <http://ssrn.com/abstract=1324280>.
- Brunnermeier, Markus. "Deciphering the 2007-08 Liquidity and Credit Crunch." Working paper, Princeton University, May 2008.
- Brunnermeier, Markus K.; Crockett, Andrew; Goodhart, Charles A.; Persaud, Avinash D. and Shin, Hyun. *The Fundamental Principles of Financial Regulation: Geneva Reports on the World Economy 11* (Preliminary Conference Draft). London: Centre for Economic Policy Research, 2009; [www.voxeu.org/reports/Geneva11.pdf](http://www.voxeu.org/reports/Geneva11.pdf).
- Bullard, James. "Systemic Risk and the Macroeconomy: An Attempt at Perspective." Speech at Indiana University on October 2, 2008; [www.stlouisfed.org/newsroom/speeches/2008\\_10\\_02.cfm](http://www.stlouisfed.org/newsroom/speeches/2008_10_02.cfm).
- Caballero, Ricardo J.; Farhi, Emmanuel and Gourinchas, Pierre-Olivier. "Financial Crash, Commodity Prices and Global Imbalances." NBER Working Paper 14521, National Bureau of Economic Research, November 17, 2008; [www.nber.org/papers/w14521.pdf?new\\_window=1](http://www.nber.org/papers/w14521.pdf?new_window=1).
- Calomiris, Charles W. and Gorton, Gary. "The Origins of Banking Panics: Models, Facts and Bank Regulation," in Charles W. Calomiris, ed., *U.S. Bank Deregulation in Historical Perspective*. New York: Cambridge University Press, 2000, pp. 93-163.
- Congressional Oversight Panel. "Special Report on Regulatory Reform." January 2009; <http://cop.senate.gov/documents/cop-012909-report-regulatoryreform.pdf>.
- Counterparty Risk Management Policy Group III. "Containing Systemic Risk: The Road to Reform." August 6, 2008; [www.crmpolicygroup.org/docs/CRMPG-III.pdf](http://www.crmpolicygroup.org/docs/CRMPG-III.pdf).
- Demyanyk, Yuliya S. and Van Hemert, Otto. "Understanding the Subprime Mortgage Crisis." Advance Access published online on May 4, 2009, in *Review of Financial Studies*; doi:10.1093/rfs/hhp033.
- Diamond, Douglas and Dybvig, Philip. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, June 1983, 91(3), pp. 401-19.
- DiMartino, Danielle and Duca, John V. "The Rise and Fall of Subprime Mortgages." Federal Reserve Bank of Dallas *Economic Letter*, November 2007, 2(11); [www.dallasfed.org/research/eclett/2007/el0711.html](http://www.dallasfed.org/research/eclett/2007/el0711.html).
- Dwyer, Gerald P. Jr. and Gilbert, R. Alton. "Bank Runs and Private Remedies." Federal Reserve Bank of St. Louis *Review*, May/June 1989, pp. 43-61; [http://research.stlouisfed.org/publications/review/89/05/Remedies\\_May\\_Jun1989.pdf](http://research.stlouisfed.org/publications/review/89/05/Remedies_May_Jun1989.pdf).

**Bullard, Neely, Wheelock**

- Economic Report of the President*, 2009. Washington, DC: U.S. Government Printing Office, January 2009; [www.gpoaccess.gov/eop/2009/2009\\_erp.pdf](http://www.gpoaccess.gov/eop/2009/2009_erp.pdf).
- Evanoff, Douglas D. and Wall, Larry D. "Subordinated Debt and Prompt Corrective Regulatory Action." Working Paper Series No. WP 2003-03, Federal Reserve Bank of Chicago, 2003; [www.chicagofed.org/publications/workingpapers/papers/wp2003-03.pdf](http://www.chicagofed.org/publications/workingpapers/papers/wp2003-03.pdf).
- Fons, Jerome S. "Rating Competition and Structured Finance." *Journal of Structured Finance*, Fall 2008a, 14(3).
- Fons, Jerome S. Testimony of Jerome S. Fons Before the Committee on Oversight and Government Reform United States House of Representatives, October 22, 2008b; <http://oversight.house.gov/documents/20081022102726.pdf>.
- Gorton, Gary. "Clearinghouses and the Origin of Central Banking in the United States." *Journal of Economic History*, June 1985, 45(2), pp. 277-83.
- Group of Thirty. *Financial Reform: A Framework for Financial Stability*. Washington, DC: The Group of Thirty, 2009; [www.group30.org/pubs/recommendations.pdf](http://www.group30.org/pubs/recommendations.pdf).
- Kane, Edward J. *The S&L Insurance Mess: How Did It Happen?* Washington, DC: Urban Institute Press, 1989.
- Kiyotaki, Nobuhiro and Moore, John. "Credit Cycles." *Journal of Political Economy*, April 1997, 105(2), pp. 211-48.
- Kotlikoff, Laurence J. and Leamer, Edward. "A Banking System We Can Trust." *Forbes*, April 23, 2009; [www.forbes.com/2009/04/22/loan-mortgage-mutual-fund-wall-street-opinions-contributors-bank.html](http://www.forbes.com/2009/04/22/loan-mortgage-mutual-fund-wall-street-opinions-contributors-bank.html).
- Lagunoff, Roger and Schreft, Stacey L. "A Model of Financial Fragility." *Journal of Economic Theory*, July/August 2001, 99(1-2), pp. 220-64.
- Merton, Robert C. "An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees: An Application of Modern Option Pricing Theory." *Journal of Banking and Finance*, June 1977, 1(1), pp. 3-11.
- Neely, Christopher J. "The Giant Sucking Sound: Did NAFTA Swallow the Peso?" *Federal Reserve Bank of St. Louis Review*, July/August 1996, 78(4), pp. 33-47; <http://research.stlouisfed.org/publications/review/96/07/9607cn.pdf>.
- New York Times*. "Senators Accuse Rating Agencies of Conflicts of Interest in Market Turmoil," September 26, 2007; [www.nytimes.com/2007/09/26/business/worldbusiness/26iht-credit.4.7646763.html](http://www.nytimes.com/2007/09/26/business/worldbusiness/26iht-credit.4.7646763.html).
- Paulsen, Henry M. Jr.; Steel, Robert K.; Nason, David G. et al: *The Department of the Treasury Blueprint for a Modernized Financial Regulatory Structure*, March 2008; [www.treas.gov/press/releases/reports/Blueprint.pdf](http://www.treas.gov/press/releases/reports/Blueprint.pdf).
- Pintus, Patrick A. and Wen, Yi. "Excessive Demand and Boom-Bust Cycles." Working Paper 2008-014B, Federal Reserve Bank of St. Louis, June 2008; <http://research.stlouisfed.org/wp/2008/2008-014.pdf>.
- Poole, William. "Financial Stability." Remarks at the Council of State Governments Southern Legislative Conference Annual Meeting, New Orleans, Louisiana, August 4, 2002; <http://fraser.stlouisfed.org/historicaldocs/wp2002/download/41143/20020804.pdf>.
- Poole, William. "Housing in the Macroeconomy." Remarks at the Office of Federal Housing Enterprise Oversight Symposium, Washington, DC, March 10, 2003; <http://fraser.stlouisfed.org/historicaldocs/wp2003/download/41135/20030310.pdf>.
- Poole, William. "Reputation and the Non-Prime Mortgage Market." Remarks at the St. Louis Association of Real Estate Professionals. July 20, 2007; <http://fraser.stlouisfed.org/historicaldocs/wp2007/download/40917/20070720.pdf>.
- Schinasi, Garry J.; Craig, R. Sean; Drees, Burkhard and Kramer, Charles. "Modern Banking and OTC Derivatives Markets: The Transformation of Global Finance and Its Implications for Systemic Risk." Occasional Paper No. 203, International Monetary

- Fund, January 9, 2001;  
[www.imf.org/external/pubs/nft/op/203/](http://www.imf.org/external/pubs/nft/op/203/).
- Sengupta, Rajdeep and Emmons, William R. "What Is Subprime Lending?" Federal Reserve Bank of St. Louis, *Economic Synopses* No. 13, 2007;  
<http://research.stlouisfed.org/publications/es/07/ES0713.pdf>.
- Son, Hugh. "AIG Plunges as Downgrades Threaten Quest for Capital." Bloomberg.com, September 16, 2008; [www.bloomberg.com/apps/news?pid=20601087&sid=aP5rm0.62wqo](http://www.bloomberg.com/apps/news?pid=20601087&sid=aP5rm0.62wqo).
- Spong, Kenneth. *Banking Regulation: Its Purposes, Implementation, and Effects*. Fifth edition. Kansas City, MO: Federal Reserve Bank of Kansas City, 2000; [www.kansascityfed.org/banking/bankingpublications/RegsBook2000.pdf](http://www.kansascityfed.org/banking/bankingpublications/RegsBook2000.pdf).
- Stern, Gary H. and Feldman, Ron J. *Too Big to Fail: The Hazards of Bank Bailouts*. Washington, DC: Brookings Institution Press, 2004.
- Taylor, John B. *Getting Off Track: How Government Actions and Intervention Caused, Prolonged, and Worsened the Financial Crisis*. Stanford, CA: Hoover Institution Press, 2009.
- Timberlake, Richard H. Jr. "The Central Banking Role of Clearinghouse Associations." *Journal of Money, Credit, and Banking*, February 1984, 16(1), pp. 1-15.
- Wallison, Peter J. "Everything You Wanted to Know about Credit Default Swaps—But Were Never Told" (American Enterprise Institute for Public Policy Research Outlook Series). December 2008;  
[www.aei.org/docLib/20090107\\_12DecFSOg.pdf](http://www.aei.org/docLib/20090107_12DecFSOg.pdf).
- Wheelock, David C. "The Federal Response to Home Mortgage Distress: Lessons from the Great Depression." Federal Reserve Bank of St. Louis *Review*, May/June 2008, 90(3 Part 1), pp. 133-48;  
<http://research.stlouisfed.org/publications/review/08/05/Wheelock.pdf>.
- White, Eugene N. *The Regulation and Reform of the American Banking System, 1900-1929*. Princeton, NJ: Princeton University Press, 1983.
- White, Lawrence J. *The S&L Debacle: Public Policy Lessons for Bank and Thrift Regulation*. New York: Oxford University Press, 1991.
- Wicker, Elmus. *Banking Panics of the Gilded Age*. Cambridge: Cambridge University Press, 2000.





# Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature

David C. Wheelock and Mark E. Wohar

This article surveys recent research on the usefulness of the term spread (i.e., the difference between the yields on long-term and short-term Treasury securities) for predicting changes in economic activity. Most studies use linear regression techniques to forecast changes in output or dichotomous choice models to forecast recessions. Others use time-varying parameter models, such as Markov-switching models and smooth transition models, to account for structural changes or other nonlinearities. Many studies find that the term spread predicts output growth and recessions up to one year in advance, but several also find its usefulness varies across countries and over time. In particular, many studies find that the ability of the term spread to forecast output growth has diminished in recent years, although it remains a reliable predictor of recessions. (JEL C53, E37, E43)

Federal Reserve Bank of St. Louis *Review*, September/October 2009, 91(5, Part 1), pp. 419-40.

Information about a country's future economic activity is important to consumers, investors, and policymakers. Since Kessel (1965) first discussed how the term structure of interest rates varies with the business cycle, many studies have examined whether the term structure is useful for predicting various measures of economic activity. The term spread (the difference between the yields on long-term and short-term Treasury securities) has been found useful for forecasting such variables as output growth, inflation, industrial production, consumption, and recessions, and the ability of the spread to predict economic activity has become something of a "stylized fact" among macroeconomists.

This article surveys recent research investigating the ability of the term spread to forecast output growth and recessions.<sup>1</sup> The article briefly discusses theoretical explanations for why the

spread might predict future economic activity and then surveys empirical studies that investigate how well the spread predicts output growth and recessions. The survey describes the data and methods used in various studies to investigate the predictive power of the term spread, as well as key findings. In general, the literature has not reached a consensus about how well the term spread predicts output growth. Although many studies do find that the spread predicts output growth at one-year horizons, studies also find considerable variation across countries and over time. In particular, many studies find that the ability of the spread to forecast output growth has declined since the mid-1980s. The empirical literature provides more consistent evidence that

---

<sup>1</sup> Surveys of the older literature include Berk (1998), Dotsey (1998), Estrella and Hardouvelis (1991), Plosser and Rouwenhorst (1994), and Stock and Watson (2003). Stock and Watson (2003) also survey research on the usefulness of asset prices for forecasting inflation.

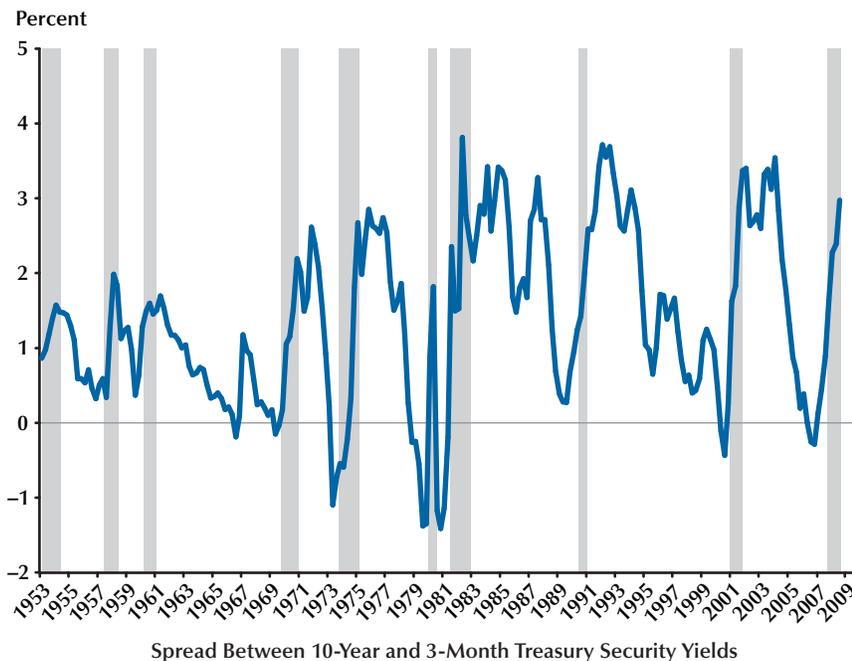
---

David C. Wheelock is a vice president and economist at the Federal Reserve Bank of St. Louis. Mark E. Wohar is a professor of economics at the University of Nebraska at Omaha. The authors thank Michael Dueker, Massimo Guidolin, and Dan Thornton for comments on a previous draft of this article. Craig P. Aubuchon provided research assistance.

© 2009, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

**Figure 1**

**U.S. Term Spread and Recessions**



NOTE: The term spread is calculated as the difference between the yields on 10-year and 3-month Treasury securities. The shaded areas denote recessions as determined by the National Bureau of Economic Research.

the term spread is useful for predicting recessions. Furthermore, the relationship appears robust to the inclusion of other variables and nonlinearities in the forecasting model.

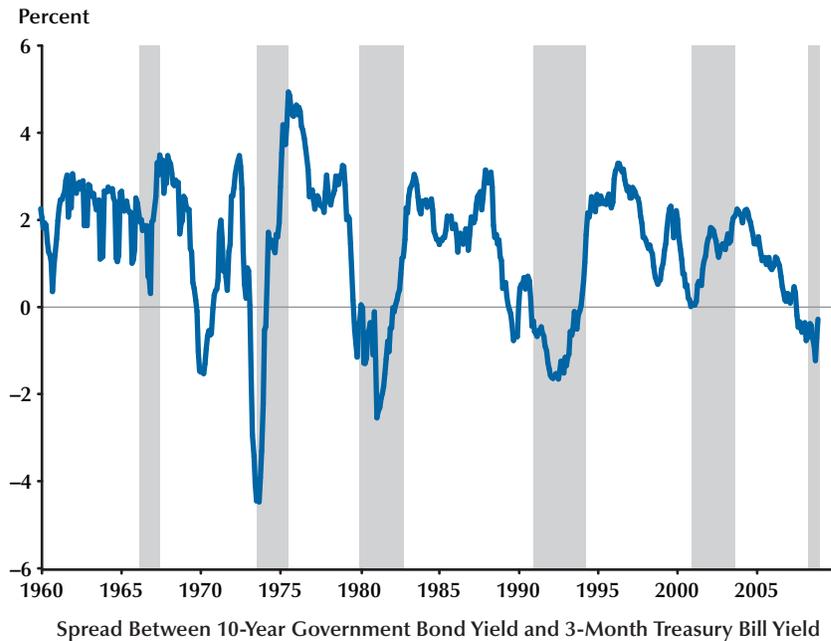
**A LOOK AT THE DATA**

Yields on long-term securities typically exceed those on otherwise comparable short-term securities, reflecting the preference of most investors to hold instruments with shorter maturities. Hence, the yield curve, which is a plot of the yields on otherwise comparable securities of different maturities, is typically upward sloping. Analysts have long noted, however, that most recessions are preceded by a sharp decline in the slope of the yield curve and frequently by an inversion of the yield curve (i.e., by short-term yields rising above those on long-term securities).

Figure 1 shows the difference between the yields on 10-year and 3-month U.S. Treasury securities for 1953-2008. The shaded regions indicate recession periods as defined by the National Bureau of Economic Research.<sup>2</sup> As Figure 1 shows, every U.S. recession since 1953 was preceded by a large decline in the yield on 10-year Treasury securities relative to the yield on 3-month Treasury securities, and several recessions were preceded by an inversion of the yield curve. Moreover, the only occasion when the 3-month Treasury security yield exceeded the (constant-maturity) 10-year Treasury yield without a subsequent recession was in December 1966.

Similar data for Germany and the United Kingdom are shown in Figures 2 and 3, respec-

<sup>2</sup> National Bureau of Economic Research, "Information on Recessions and Recoveries, the NBER Business Cycle Dating Committee, and Related Topics"; [www.nber.org/cycles/main.html](http://www.nber.org/cycles/main.html).

**Figure 2****German Term Spread and Recessions**

NOTE: The term spread is calculated as the difference between the yields on 10-year and 3-month Treasury securities. The shaded areas denote recessions as determined by the Economic Cycle Research Institute.

tively. Germany experienced recessions beginning in 1966, 1974, 1980, 1991, 2000, and 2008. All but the 1966 recession were preceded by a sharp decline in long-term Treasury security yields relative to short-term yields that resulted in a flat or inverted yield curve. The only inversion that was not followed by a recession occurred in 1970.

The United Kingdom experienced recessions beginning in 1974, 1979, 1990, and 2008. All were preceded by or coincided with a yield curve inversion. However, large inversions in 1985 and 1997-98 were not followed by recessions.<sup>3</sup>

Table 1 summarizes additional information about the association between the term spread and economic activity. The table presents correlations between the term spread (measured as a

quarterly average of monthly observations) and the year-over-year percentage change in real gross domestic product (GDP) for the United States, Germany, and the United Kingdom. The table presents the contemporaneous correlation between the two variables, as well as correlations at various leads and lags of the term spread relative to GDP growth. The top panel of the table reports correlations between GDP growth in one quarter and the term spread in the same quarter ( $t$ ) and in six preceding quarters ( $t-1$  and so on). The bottom panel reports the correlations between GDP growth in one quarter and the term spread in the same quarter and in the six subsequent quarters ( $t+1$  and so on).

The contemporaneous correlation between GDP growth and the term spread is not statistically different from zero for any of the three countries (column 1 in Table 1). By contrast, the correlations between GDP growth and the term spread lagged

<sup>3</sup> Recession dates for Germany and the United Kingdom are from the Economic Cycle Research Institute, as reported by Haver Analytics. Interest rate data for Germany and the United Kingdom are from Global Insight.

**Figure 3**

**U.K. Term Spread and Recessions**



NOTE: The term spread is calculated as the difference between the yields on 10-year and 3-month Treasury securities. The shaded areas denote recessions as determined by the Economic Cycle Research Institute.

from one to six quarters are uniformly positive and statistically significant (indicated by *p*-values of 0.10 or less) for all three countries, except for the correlation between U.S. GDP growth and the term spread lagged by one quarter. Thus, the correlations indicate that, in general, the higher the yield on 10-year Treasury securities relative to the yield on 3-month Treasury securities—that is, the more steeply sloped the yield curve—the higher the rate of future GDP growth. Similarly, the less steeply sloped the yield curve, the lower the subsequent rate of GDP growth.

The correlations between current GDP growth and *future* term spreads shown in the lower panel are negative and for the most part statistically significant for all three countries. Thus, a higher GDP growth rate in one quarter is associated with a less steeply sloped yield curve in subsequent quarters.

As discussed in more detail in the following section, the pattern of positive correlation between current GDP growth and lagged term spreads and

negative correlation between current GDP growth and future term spreads is consistent with more than one explanation of the relationship between the yield curve and output growth. Further, although the unconditional correlation between output growth and the term spread is high, the correlation might reflect the influence of some other variable, in which case the term spread would not forecast output growth if that other influence is included in the forecasting model. After discussing why the term spread might forecast economic activity in the next section, we review empirical research on the usefulness of the term spread for forecasting output growth and recessions in subsequent sections.

**WHY MIGHT THE TERM SPREAD FORECAST ECONOMIC ACTIVITY?**

Although many empirical studies find that the term spread predicts future economic activity,

**Table 1**  
**Correlation of GDP Growth and Lagged and Future Term Spreads by Country**

	Lagged term spread						
	Term						
	$t$	$(t - 1)$	$(t - 2)$	$(t - 3)$	$(t - 4)$	$(t - 5)$	$(t - 6)$
United States	-0.0449 (0.5047)	0.0999 (0.1379)	0.2557 (0.0001)	0.3605 (0.0001)	0.4141 (0.0001)	0.3957 (0.0001)	0.3196 (0.0001)
Germany	-0.0003 (0.9970)	0.1641 (0.0455)	0.2991 (0.0002)	0.3689 (0.0001)	0.3845 (0.0001)	0.3649 (0.0001)	0.3421 (0.0001)
United Kingdom	0.0723 (0.3319)	0.1816 (0.0144)	0.2486 (0.0008)	0.3025 (0.0001)	0.3379 (0.0001)	0.3166 (0.0001)	0.2607 (0.0005)

	Future term spread						
	Term						
	$t$	$(t + 1)$	$(t + 2)$	$(t + 3)$	$(t + 4)$	$(t + 5)$	$(t + 6)$
United States	-0.0449 (0.5047)	-0.1428 (0.0335)	-0.2374 (0.0004)	-0.2994 (0.0001)	-0.3372 (0.0001)	-0.3538 (0.0001)	-0.3421 (0.0001)
Germany	-0.0003 (0.9970)	-0.1722 (0.0357)	-0.3414 (0.0001)	-0.4424 (0.0001)	-0.4548 (0.0001)	-0.4545 (0.0001)	-0.4110 (0.0001)
United Kingdom	0.0723 (0.3319)	-0.0364 (0.6244)	-0.1366 (0.0652)	-0.2116 (0.0040)	-0.2306 (0.0017)	-0.2204 (0.0001)	-0.2261 (0.0021)

NOTE: U.S. data are for 1953:Q1–2008:Q4; German data are for 1973:Q1–2008:Q2 (West Germany, 1973–1991); U.K. data are for 1958:Q1–2008:Q2. Numbers in parentheses represent  $p$ -values.

there is no universally agreed-upon theory as to why a relationship between the term spread and economic activity should exist. To a large extent, the usefulness of the spread for forecasting economic activity remains a “stylized fact in search of a theory” (Benati and Goodhart, 2008, p. 1237).

The expectations hypothesis of the term structure is the foundation of many explanations of the term spread’s usefulness in forecasting output growth and recessions. The expectations hypothesis holds that long-term interest rates equal the sum of current and expected future short-term interest rates plus a term premium. The term premium explains why the yield curve usually slopes upward—that is, why the yields on long-term securities usually exceed those on short-term securities. However, the yield curve flattens or inverts—slopes downward—if the public expects short-term interest rates to fall. In that case, investors bid up the prices of longer-term securi-

ties, which causes their yields to fall relative to current yields on short-term securities.

Many studies attribute the apparent ability of the term spread to forecast economic activity to actions by monetary authorities to stabilize output growth. For example, monetary policy tightening causes both short- and long-term interest rates to rise. Short-term rates are likely to rise more than long-term rates, however, if policy is expected to ease once economic activity slows or inflation declines. Hence, a policy tightening is likely to cause the yield curve to flatten or possibly invert. Monetary policy explanations usually have been stated with little underlying theory.<sup>4</sup> However, as noted by Feroli (2004), Estrella (2005), and Estrella and Trubin (2006), the extent to which

<sup>4</sup> For example, Estrella and Hardouvelis (1991) and Berk (1998) refer to simple dynamic IS-LM models but do not explicitly derive testable hypotheses from those models (see also Bernanke and Blinder, 1992; Dueker, 1997; and Dotsey, 1998).

the term spread is a good predictor of output growth depends on the monetary authority's policy objectives and reaction function. For example, the term spread forecasts output growth better the more responsive the monetary authority is to deviations of output growth from potential. The spread forecasts less accurately if monetary authorities concentrate exclusively on controlling inflation. Further, changes in the relative responsiveness of the monetary authority to either output growth or inflation could cause changes in the ability of the term spread to forecast output growth.

In contrast to explanations that focus on monetary policy, theories of intertemporal consumption derive a relationship between the slope of the yield curve and future economic activity explicitly from the structure of the economy (e.g., Harvey, 1988; Hu, 1993). The central assumption of Harvey (1988), for example, is that individuals prefer stable consumption rather than high consumption during periods of rising income and low consumption when income is falling. Thus, when consumers expect a recession one year in the future, they will sell short-term financial instruments and purchase one-year discount bonds to obtain income during the recession year. As a result the term structure flattens or inverts.<sup>5</sup>

The theoretical implications of consumption-smoothing models apply to the *real* term structure, that is, the term structure adjusted for expected inflation. However, much of the empirical evidence on the information content of the term structure pertains to the *nominal* term structure. The consistency of the empirical evidence linking the nominal yield curve to changes in output with the theoretical relationship depends on the persistence of inflation. If inflation were a random walk, implying that shocks to inflation are permanent, then inflation shocks would have no impact on the slope of the nominal yield curve because expected inflation would change by an identical amount at all horizons. However, if infla-

tion has little persistence, an inflation shock will affect near-term expected inflation more than long-term expected inflation, causing the slope of the nominal yield curve to change. Hence, the extent to which changes in the slope of the nominal yield curve reflect changes in the real yield curve depends on the persistence of inflation which, in turn, reflects the underlying monetary regime.<sup>6</sup>

Much of the empirical literature has focused on estimating the precision with which the term spread forecasts economic activity, rather than on attempting to discriminate between the monetary policy and consumption-smoothing explanations. Laurent (1988, 1989) argues that the yield curve reflects the stance of monetary policy and finds that the term spread predicts changes in the growth rate of real GDP. On the other hand, several studies find that the term spread has significant predictive power for economic growth independent of the information contained in measures of current and future monetary policy, suggesting that monetary policy alone cannot explain all of the observed relationship (see, e.g., Estrella and Hardouvelis, 1991; Plosser and Rouwenhorst, 1994; Estrella and Mishkin, 1997; Benati and Goodhart, 2008).

Harvey (1988) and Rendu de Lint and Stolin (2003) offer support for the consumption-smoothing explanation by showing that the slope of the yield curve is useful for forecasting both consumption and output growth. Benati and Goodhart (2008), however, find that changes over time in the marginal predictive content of the nominal term spread for output growth do not match changes in inflation persistence, which they argue is evidence against the consumption-smoothing explanation.

Several studies find that the spread has forecast output growth less accurately since the mid-1980s, which some attribute to greater stability of output growth and other key macroeconomic data (e.g., D'Agostino, Giannone, and Surico, 2006). It remains to be seen how incorporating data for

<sup>5</sup> Rendu de Lint and Stolin (2003) study the relationship between the term structure and output growth in a dynamic equilibrium asset pricing model. They find that the term spread predicts future consumption and output growth at long horizons in a stochastic endowment economy model augmented with endogenous production.

<sup>6</sup> Under fiat monetary regimes, inflation has tended to be highly persistent. However, inflation tends to exhibit little persistence under metallic and inflation-targeting regimes (see, e.g., Shiller and Siegel, 1977; Barsky, 1987; Bordo and Schwartz, 1999; and Benati, 2006, 2008).

the recession that began in 2007 affects the performance of forecasting models that use the term spread to predict economic activity and whether the additional information sheds light on alternative explanations for the forecasting relationship.

## DOES THE TERM SPREAD FORECAST OUTPUT GROWTH?

Numerous studies using a wide variety of data and methods investigate how well the term spread forecasts output growth. Although many studies use post-World War II U.S. data, several recent studies investigate how well the term spread predicts future economic activity using data from other countries or time periods. Such efforts can indicate whether the association between the term spread and output growth is an artifact of the postwar U.S. experience and shed light on the validity of alternative explanations for why the spread might forecast economic activity. Our survey focuses primarily on the literature published or written since the mid-1990s. However, we briefly discuss some earlier studies to set the stage for a more detailed discussion of recent work.

Much of the evidence on the accuracy of the term spread in forecasting output growth comes from the estimation of linear models, such as the following linear regression, or some variant of it:

$$(1) \quad \Delta Y_t = \alpha + \beta \text{Spread} + \gamma(L)\Delta Y_{t-1} + \varepsilon_t,$$

where  $\Delta Y_t$  is the growth rate of output (e.g., real GDP); *Spread* is the difference between the yields on long-term and short-term Treasury securities;  $\gamma(L)$  is a lagged polynomial, typically of length four (current and three lags, assuming quarterly data);<sup>7</sup> and  $\varepsilon_t$  is an error term.

Laurent (1988), Harvey (1988, 1989), and Estrella and Hardouvelis (1991) were among the first to present empirical evidence on the strength of the relationship between the term spread and output growth using U.S. data. Harvey (1989), for example, finds that the spread between the yields on 5-year and 3-month U.S. Treasury securities predicts real gross national product growth from

1 to 5 quarters ahead. Similarly, Estrella and Hardouvelis (1991) find that the spread between yields on 10-year and 3-month Treasury securities is useful for forecasting U.S. output growth and recessions, as well as consumption and investment, especially at 4- to 6-quarter horizons.

## Evidence from Outside the United States

Although the earliest studies were based on U.S. data, several others have explored the usefulness of the spread for forecasting output growth using data from other countries. Often these studies show considerable variation across countries in how well the spread forecasts output growth. For example, Plosser and Rouwenhorst (1994) find that term spreads are useful for predicting GDP growth in Canada and Germany, as well as the United States, but not in France or the United Kingdom. Plosser and Rouwenhorst (1994) also find that foreign term spreads help predict future changes in output in individual countries.

Davis and Fagan (1997) find that the term spread has statistically significant within-sample explanatory power for output growth in six of nine European Union countries, but that the spread improves out-of-sample forecasts and satisfies conditions for statistical significance and stability in only three countries (Belgium, Denmark, and the United Kingdom). A related study by Berk and van Bergeijk (2001) examines 12 euro-area countries over the period 1970-98 and finds that the term spread contains only limited information about future output growth.

Several studies examine whether the term spread contains information about future output growth in Japan. Harvey (1991) finds that the spread contains no information about future economic activity in Japan for the period 1970-89. By contrast, Hu (1993) finds a positive correlation between the term spread and future economic activity in Japan for the period from January 1957 to April 1991, but that lagged changes in stock prices and output growth have more explanatory power than the term spread. Kim and Limpaphayom (1997) argue that heavy regulation prevented interest rates from reflecting market expectations before 1984. Their study finds that the spread is useful for predicting output growth up to five

<sup>7</sup> For example,  $\gamma(L) = \gamma_1 L^1 + \gamma_2 L^2 + \gamma_3 L^3 + \gamma_4 L^4$ , where  $L^i \Delta Y_t = \Delta Y_{t-i}$ .

quarters ahead during 1984-91 (see also Nakaota, 2005).

### **Evidence from Multivariate Models**

Several studies examine the marginal predictive content of the term spread in models that also include other explanatory variables. Estrella and Hardouvelis (1991), Plosser and Rouwenhorst (1994), Estrella and Mishkin (1997), Hamilton and Kim (2002), and Feroli (2004) are among several studies that find the term spread has significant predictive power for economic growth even when a short-term interest rate or other measure of the stance of monetary policy is included as an additional explanatory variable. These results suggest that monetary policy alone does not explain why the term spread predicts output growth. However, Stock and Watson (2003) show that including other explanatory variables does not improve forecasts obtained from a bivariate model of the term spread and output growth.<sup>8</sup>

Aretz and Peel (2008) include both the term spread and professional forecasts in a model of output growth and find that both variables individually forecast real GDP growth and that the term spread contains information not captured by professional forecasts. However, Aretz and Peel (2008) find that the term spread contributes no information beyond that in the professional forecasts in models that assume that forecasters' loss functions become more skewed as the forecast horizon lengthens.

Hamilton and Kim (2002) note that (i) the term spread consists of an expected interest rate component and a term premium component and (ii) determining the relative usefulness of one or the other component for forecasting output growth could help distinguish among alternative hypotheses for why the term spread predicts output growth. Hamilton and Kim (2002) find that the expected change in the short-term interest rate and the time-varying term premium both contribute to forecasts of real GDP growth up to eight

quarters ahead. However, expected changes in short-term rates explain significantly more of the output growth than does the term premium. Hence, the most important reason that an inverted yield curve predicts slower output growth in the future is that a low term spread implies falling future short-term interest rates, rather than, say, an increase in the term premium associated with higher interest rate volatility near the end of economic expansions.

### **Recent Research on the Stability of the Forecasting Relationship**

Table 2 summarizes the methods and principal findings of several recent studies of the ability of the term spread to forecast output growth. Much of the research during the past decade focuses on the stability of the forecasting relationship over time. Several studies find that the spread has been less useful for forecasting output growth since the mid-1980s, at least for the United States.<sup>9</sup> For example, Dotsey (1998) finds that the spread forecasts cumulative output growth up to two years in the future, but does so less accurately for 1985-97 than for earlier years. Further, Dotsey (1998) finds that the spread forecasts less accurately when past values of output growth and short-term interest rates are included in the forecasting model and contributes no information to forecasts for the 1985-97 period.

Estrella, Rodrigues, and Schich (2003) test for unknown breakpoints in the in-sample forecasting relationship between the term spread and output growth using data for the United States and Germany. Although the study detects a generally strong relationship between the term spread and output growth one year in the future for both countries, it identifies a break in September 1983 for the United States using models with one-year forecast-horizons. Estrella, Rodrigues, and Schich (2003), however, detect no breaks in longer-horizon forecasting models for the United States or in short- or long-horizon models estimated using data for Germany.

<sup>8</sup> Similarly, Cozier and Tkacz (1994) and Hamilton and Kim (2002) find that the spread predicts future changes in output growth in forecasting models that include the output gap and changes in the price of oil, respectively, as an explanatory variable.

<sup>9</sup> In addition to the studies summarized in Table 2, other studies that find a break in the forecasting relationship in the mid-1980s include Haubrich and Dombrosky (1996), Estrella and Mishkin (1997), and Smets and Tsatsaronis (1997).

**Table 2****Selective Summary of Studies of the Usefulness of the Term Spread for Predicting Output Growth**

Study	Methodology	Data* (years)	Principal finding(s)	Notes
Dotsey (1998)	Single-equation linear and nonlinear regression	U.S., quarterly (1955-97)	Spread is useful for predicting cumulative GDP growth up to 2 years ahead, but less accurate during 1985-97 than previously.	Spread has marginal predictive power only up to 6 quarters. Adding the spread to a VAR containing lagged output growth increases forecast errors.
Galbraith and Tkacz (2000)	Single-equation linear regression and smooth transition nonlinear asymmetric threshold model	G-7 developed countries, quarterly (1960s–late 1990s; varies by country)	Spread predicts changes in output. Evidence for the U.S. and Canada of asymmetric nonlinear behavior, where the impact of the spread is greater on one side of a threshold than on the other.	Across a variety of specifications, the spread has its most significant predictive power when it is negative.
Shaaf (2000)	Single-equation linear models and neural networks	U.S., quarterly (1959-97)	Spread forecasts output growth: A 5 percent increase in the yield spread results in a 9.33 percent increase in output growth.	Out-of-sample simulations indicate that the forecast of the artificial neural networks is more accurate and has less error and lower variation than forecasts from linear models.
Berk and van Bargeijk (2001)	Single-equation linear models	Twelve developed countries and the euro area, quarterly (1970-98)	Term spread has little information about future output growth beyond that contained in lagged output growth for most countries. The U.S. is an exception.	Evidence of parameter instability for the U.S. in the latter part of the sample but not for other countries or the euro area.
Tkacz (2001)	Neural networks	Canada, quarterly (1968-99)	Four-quarter forecasts of output growth outperform 1-quarter forecasts.	Neural network models outperform linear models at a 4-quarter horizon but not at a 1-quarter horizon.
Hamilton and Kim (2002)	Linear regression and GARCH models	U.S., quarterly (1953-98)	Cyclical behavior of interest rate volatility is an important determinant of the spread and the term premium and a useful predictor of future interest rates.	Cyclical movements in volatility are unable to account for the spread and the term premium in forecasting output growth.
Estrella, Rodrigues, and Schich (2003)	Single-equation linear models	U.S. and Germany, monthly industrial production (1955-98 for U.S.; 1967-98 for Germany)	Spread forecasts output growth well at 1-year horizons in both countries but less accurately at 2- and 3-year horizons.	Results are robust across several maturity combinations for the spread. Little evidence of instability for Germany, but a break in 1983 for the U.S. at a 1-year horizon.

NOTE: \*Unless otherwise noted, the dependent variable in each study is the growth rate of real GDP. GARCH, generalized autoregressive conditional heteroskedasticity; GNP, gross national product; VAR, vector autoregression; VAR-VECM, VAR–vector error correction model.

**Table 2, cont'd****Selective Summary of Studies of the Usefulness of the Term Spread for Predicting Output Growth**

Study	Methodology	Data* (years)	Principal finding(s)	Notes
Stock and Watson (2003)	Linear regression and combination forecasts	Canada, France, Germany, Italy, Japan, U.K., and U.S., quarterly (1959-99)	Some asset prices have predictive content for output growth, but results vary across time and by country. Forecasts based on individual indicators are unstable.	Simple combination forecasts, such as computing the median or trimmed mean of a panel of forecasts, seem to circumvent issues of instability in that they yield smaller errors than the autoregressive benchmark model. Combination forecasts are stable even though the individual predictive relations are unstable.
Venetis, Paya, and Peel (2003)	Smooth nonlinear transition models, regime-switching models, and time-varying models	U.S., U.K., and Canada, quarterly (Early 1960s to 2000; varies by country)	Threshold effects for the U.S., the U.K., and Canada. The term spread–output growth relationship is stronger when past values of the term spread do not exceed a positive threshold value.	Spread is less useful for predicting output growth in recent years.
Jardet (2004)	Single-equation linear model; VAR-VECM to identify sources of structural breaks	U.S., monthly industrial production and employment (1957-2001)	Spread forecasts output growth well, especially at 1-year horizons. Structural break occurs in 1984 with diminished forecasting strength thereafter.	VAR estimates suggest that a structural break is due to a drop in the contributions of monetary policy and supply shocks to the covariance between the spread and output growth.
Duarte, Venetis, and Paya (2005)	Linear and nonlinear threshold models	Euro area and U.S., quarterly (1970-2000)	Significant nonlinearity exists in the term yield spread–output growth relation with respect to time and past output growth. Nonlinear model outperforms linear model in 1-year out-of-sample forecasts.	With linear models, the term spread is a useful indicator of future output growth for the euro area. Linear models show signs of instability. Spreads are successful in predicting output growth when output growth has slowed.
Nakaoto (2005)	Single-equation linear model	Japan, monthly industrial production (1985-2001)	Spread forecasts output at 1- to 24-month horizons in models that account for a structural break in July 1991.	Usefulness of the spread is robust to inclusion of other variables. Expected future changes in short-term rates appear to contribute useful information both before and after 1991, but the term premium is useful only after 1991.

NOTE: \*Unless otherwise noted, the dependent variable in each study is the growth rate of real GDP. GARCH, generalized autoregressive conditional heteroskedasticity; GNP, gross national product; VAR, vector autoregression; VAR-VECM, VAR–vector error correction model.

**Table 2, cont'd****Selective Summary of Studies of the Usefulness of the Term Spread for Predicting Output Growth**

Study	Methodology	Data* (years)	Principal finding(s)	Notes
Ang, Piazzesi, and Wei (2006)	Linear models and VARs	U.S., quarterly (1952-2001)	Recommends using the longest yield spread to predict output growth regardless of forecast horizon. Results indicate that the level of the short-term rate contains more information about output growth than any yield spread.	VAR model forecasts are superior to linear model forecasts both in and out of sample. The factor structure appears largely responsible for most of the efficiency gains. The lagged spread does not predict output growth in the 1990s, but high short-term rates forecast negative output growth.
D'Agostino, Giannone, and Surico (2006)	Single-equation linear model and bivariate VAR	U.S., monthly personal income, industrial production, unemployment rate, and employment (1959-2003)	Spread dominates other variables in forecasting output and employment at 12-month horizons during 1959-84 but not during 1985-2003.	A general decline occurs in forecast accuracy for the spread, other variables, and professional forecasts after 1984 relative to a random walk.
Giacomini and Rossi (2006)	Structural break tests; both single break and multiple breaks	U.S., monthly industrial production (1965-2001)	Evidence of forecast breakdown in the relation between yield spread and output growth, especially during the Burns-Miller and Volker monetary policy regimes.	Results parallel the empirical evidence on structural breaks of the relation between spread and output growth documented in the literature.
Aretz and Peel (2008)	Single-equation linear model	U.S., quarterly GDP/GNP (1981-2006)	Spread forecasts output growth at various horizons and includes information beyond that in the Survey of Professional Forecasters.	Results are robust to the use of real-time or vintage data. The spread contributes no information in models that assume forecasters have asymmetric loss functions.
Benati and Goodhart (2008)	Bayesian VARs with time-varying parameters	U.S. and U.K., quarterly (1875-2005); euro area, quarterly (1970-2003); Australia, quarterly (1957-2005); Canada, quarterly (1975-2005)	Spread has considerable marginal predictive content for the U.S. before World War I and in the 1980s, but little during the interwar period or before or after the 1980s.	Similar parameter instability is found in forecasts for other countries and in models that also include inflation and a short-term interest rate. Results fail to distinguish clearly between leading explanations for why the spread may be useful for predicting output growth.
Bordo and Haubrich (2008)	Single-equation linear model	U.S., quarterly GNP, spread between corporate bonds and 6-month commercial paper (1875-1997)	Spread improves forecasting model in only three of nine subperiods: 1875-1913, 1971-84, and, to a lesser extent, 1985-97.	Spread performs somewhat better in forecasts based on rolling regressions.

NOTE: \*Unless otherwise noted, the dependent variable in each study is the growth rate of real GDP. GARCH, generalized autoregressive conditional heteroskedasticity; GNP, gross national product; VAR, vector autoregression; VAR-VECM, VAR-vector error correction model.

Stock and Watson (2003) examine the stability of the forecasting relationship between the term spread and output growth for the United States and other countries and consider both in-sample and out-of-sample forecasts. Like prior studies, Stock and Watson (2003) find that the term spread forecasts U.S. output growth less accurately after 1985. The study also finds that the spread forecasts output less accurately during 1985-99 than a simple autoregressive model.

A recent study by Giacomini and Rossi (2006) reexamines the forecasting performance of the yield curve for output growth using forecast breakdown tests developed by Giacomini and Rossi (2009). Giacomini and Rossi (2006) show that output growth models are characterized by a breakdown of predictability. In particular, they find strong evidence of forecast breakdowns at the one-year horizon during 1974-76 and 1979-87.

Several studies that find diminished performance of the term spread forecasts of output growth in recent years point to the increased stability of output growth and other macroeconomic variables since the mid-1980s (at least until 2007) as a possible reason for the apparent change. As noted previously, a change in the relative responsiveness of monetary policy to output growth and inflation could affect how well the term spread predicts output growth. Bordo and Haubrich (2004, 2008) investigate the ability of the term spread to forecast U.S. output growth across different monetary regimes from 1875 to 1997. The authors examine periods distinguished by major changes in the monetary and interest rate environment, including the founding of the Federal Reserve System in 1914, World War II, the Treasury-Fed Accord of 1951, and the closing of the U.S. gold window and collapse of the Bretton Woods system in 1971. Bordo and Haubrich (2004, 2008) find that the term spread improves the forecast of output growth, as indicated by the mean squared forecast error, in three of the nine subperiods they consider: (i) the period preceding the establishment of the Federal Reserve System (1875-1913), (ii) the first 13 years after the collapse of the Bretton Woods system (1971-84), and, to a lesser extent, (iii) the 1985-97 period.<sup>10</sup> The term spread does not improve forecasts of output

growth during the interwar period or the Bretton Woods era that followed World War II.

Bordo and Haubrich (2004, 2008) find that the term spread tends to forecast output growth better during periods when the persistence of inflation was relatively high, such as the first 13 years after the collapse of the Bretton Woods system. In such periods, inflation shocks increase both short- and long-term interest rates and thus do not affect the slope of the yield curve. Real shocks that are expected to be temporary, however, increase short-term rates by more than long-term rates and signal a future downturn in economic activity. Bordo and Haubrich (2004, 2008) find that the term spread forecasts output growth less accurately when inflation persistence is relatively low, as it was during the interwar period and the Bretton Woods era. In such periods, both inflation and real shocks increase short-term interest rates more than long-term rates. Bordo and Haubrich argue, however, that only real shocks are likely to affect future output growth and, hence, the lower the persistence of inflation, the noisier the signal produced by the term spread about future output growth.

Benati and Goodhart (2008) extend the work of Bordo and Haubrich (2004, 2008) by (i) considering the marginal predictive content of the term spread for forecasting output growth in a multivariate model and (ii) attempting to date more precisely changes in the marginal predictive content of the spread over time. Whereas Bordo and Haubrich (2004, 2008) estimate bivariate regression models similar to equation (1), Benati and Goodhart (2008) estimate Bayesian time-varying parameter vector autoregressions (VARs).

Benati and Goodhart (2008) find that the term spread forecasts U.S. output growth better during the 1880s and 1890s than during the first two decades of the twentieth century. Further, like Bordo and Haubrich (2004, 2008), Benati and Goodhart (2008) find that the spread has almost no predictive content for the interwar years or the

---

<sup>10</sup> Bordo and Haubrich (2004, 2008) also estimate rolling regressions with 24-quarter windows and find that the term spread predicts output less accurately during the pre-Fed period than suggested by their original estimates. However, their results for the post-Bretton Woods era are robust to the use of rolling regressions.

Bretton Woods era. In addition, the study finds that the term spread contains significant predictive information about output growth during 1979-87 but none for other postwar years. Benati and Goodhart (2008) also find that estimates of the marginal predictive content of the spread are sensitive to whether a short-term interest rate and inflation are included in the forecasting model, and they find considerable variation in the marginal predictive content of the term spread over time for other countries and for different forecast horizons. Thus, like Bordo and Haubrich (2004, 2008), Benati and Goodhart (2008) find numerous breaks in the relationship between the term spread and future changes in output over time. However, unlike Bordo and Haubrich (2004, 2008), the breaks identified by Benati and Goodhart (2008) are not clearly associated with changes in the monetary regime or inflation persistence.

### **Evidence from Nonlinear Models**

Much of the literature investigating the performance of the term spread in forecasting output growth relies on linear models. However, variation over time in the ability of the term spread to forecast output growth suggests possible nonlinearities in the forecasting relationship and some recent studies using data for the United States and Canada find this to be the case. Further, researchers are beginning to use models that capture such nonlinearities. For example, Galbraith and Tkacz (2000) find evidence of a threshold effect in the relationship between the term spread and conditional expectations of output growth for the United States and Canada but not for other major developed countries. Specifically, the authors find a large and statistically significant impact of the term spread on conditional expectations of output growth. However, the marginal effect that an increase in the spread has on predicted output growth is lower when the level of the term spread rises above a certain point.

Shaaf (2000) and Tkacz (2001) use neural network models to account for nonlinearity in the relationship between the term spread and output growth. Both studies find that this class of models produces smaller forecast errors than linear models. Venetis, Paya, and Peel (2003) use non-

linear smooth transition models that can accommodate regime-type nonlinear behavior and time-varying parameters to examine the predictive power and stability of the term spread–output growth relationship. Using data for the United States, United Kingdom, and Canada, Venetis, Paya, and Peel (2003) find that the term spread–output growth relationship is stronger when past values of the term spread do not exceed a positive threshold value.<sup>11</sup>

Duarte, Venetis, and Paya (2005) use both linear regression and nonlinear models to examine the predictive accuracy of the term spread–output growth relationship among euro-area countries. The authors find that linear indicator and nonlinear threshold indicator models predict output growth well at four-quarter horizons and that the term spread is a useful indicator of future output growth and recessions in the euro area. The linear models show signs of instability, however, and the authors find evidence of significant nonlinearities with respect to time and lagged output growth. Further, the authors' nonlinear model outperforms their linear model in out-of-sample forecasts of one-year-ahead output growth.

Ang, Piazzesi, and Wei (2006) point out that the regressions typically used to investigate the predictive content of the term spread are unconstrained, and the authors argue for a model that treats both the term spread and output growth as endogenous variables. Ang, Piazzesi, and Wei (2006) build a dynamic model of GDP growth and bond yields that completely characterizes expectations of GDP growth. Using quarterly U.S. data for 1952-2001, the authors find that, contrary to previous research, the short-term interest rate outperforms the term spread in forecasting real GDP growth both in and out of sample and that including the term spread does not significantly improve forecasts of output growth.

In summary, the recent empirical literature on the usefulness of the term spread for forecasting output growth finds that the spread predicts output growth less accurately in some countries and some periods than in others. Notably, several

<sup>11</sup> For a discussion of smooth transition regression, see Granger and Teräsvirta (1993) or Teräsvirta (1998).

studies find that the term spread's power to forecast output has diminished since the mid-1980s. Several recent studies find evidence of significant nonlinearities, such as threshold effects, in the empirical relationship between the term spread and output growth.

## DOES THE TERM SPREAD FORECAST RECESSIONS?

As an alternative to using the term spread to forecast output growth, many studies examine the extent to which the term spread is useful for forecasting the onset of recessions. Several of those studies are summarized in Table 3.

Most recession-forecasting studies estimate a probit model of the following type, in which the dependent variable is a categorical variable set equal to 1 for recession periods and to 0 otherwise:

$$(2) \quad P(\text{recession}_t) = F(\alpha_0 + \alpha_1 S_{t-k}),$$

where  $F$  indicates the cumulative normal distribution function. If the coefficient  $\alpha_1$  is statistically significant, then the term spread,  $S_{t-k}$ , is deemed useful for forecasting a recession  $k$  periods ahead.

Models of the following form are often used to test how well the spread predicts recessions when additional explanatory variables are included in the model:

$$(3) \quad P(\text{recession}_t) = F(\alpha_0 + \alpha_1 S_{t-k} + \alpha_2 X_{t-k}),$$

where  $X_{t-k}$  is a vector of additional explanatory variables. If  $\alpha_1$  is significant in equation (2) but not in equation (3), then the ability of the spread to predict recessions is not robust to the inclusion of other variables.

Using probit estimation, Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998) find that the term spread significantly outperforms other financial and macroeconomic variables in forecasting U.S. recessions. Estrella and Hardouvelis (1991) show that the spread between the yields on 10-year and 3-month Treasury securities is a useful predictor of recessions, as well as of future growth of output, consumption, and investment. Estrella and Mishkin

(1998) compare the ability of several financial variables, including interest rates, interest rate spreads, stock prices, and monetary aggregates, to predict U.S. recessions out of sample. They find that stock prices are useful for predicting recessions at one- to three-quarter horizons but that the term spread outperforms all other variables beyond a one-quarter forecast horizon. Moreover, based on U.S. data for 1955-98 and German data for 1967-98, Estrella, Rodrigues, and Schich (2003) find that models that use the term spread to predict recessions are more stable than forecasting models for continuous variables, such as GDP growth and industrial production.

The term spread appears useful for predicting recessions in many countries. Using probit estimation, Bernard and Gerlach (1998) find that the term spread forecasts recessions up to two years ahead in eight countries (Belgium, Canada, France, Germany, Japan, the Netherlands, United Kingdom, and United States) over the 1972-93 period. Similarly, Moneta (2005) finds that the spread is useful for predicting recession probabilities for the euro area as a whole, as well as in individual countries.<sup>12</sup>

Several studies test whether the term spread remains useful for predicting recessions in multivariate forecasting models. For example, Dueker's (1997) probit model includes the change in an index of leading economic indicators, real money stock growth, the spread between the 6-month commercial paper and Treasury bill rates, and the percentage change in a stock price index, as well as the difference in yields on 30-year Treasury bonds and 3-month Treasury bills as a measure of the term spread. Dueker (1997) finds that among the variables, the term spread is the dominant predictor of recessions at horizons beyond three months.

Bernard and Gerlach (1998) include both an index of leading indicators and foreign interest rate term spreads in a recession-forecasting model. The index of leading indicators contains information beyond that in the term spreads, but the

<sup>12</sup> Moneta (2005) examines the predictive power of 10 yield spreads, representing different segments of the yield curve, and finds that the spread between the yield on 10-year government bonds and the 3-month interbank rate outperforms all other spreads in predicting recessions in the euro area.

information is useful only for forecasting recessions in the immediate future. Bernard and Gerlach (1998) find that in addition to the domestic term spread, the term spreads of Germany and the United States are particularly useful for forecasting recessions in Japan and the United Kingdom, respectively.

Sensier et al. (2004) use logit models to predict recessions in four European countries. The authors find that international data (in particular, the U.S. index of leading indicators and short-term interest rate) are useful for predicting business cycles in the four countries. The domestic term spread helps forecast recessions in Germany when international variables are included in the model, and short- and long-term interest rates entered separately help forecast recessions in France and the United Kingdom.

Wright (2006) confirms previous studies in finding that the term spread is highly statistically significant in a bivariate probit recession model estimated on U.S. data for 1964-2005. However, Wright (2006) also finds that a model that includes both the federal funds rate and term spread fits the data much better than the bivariate model and provides superior out-of-sample recession forecasts. Similarly, King, Levin, and Perli (2007) find that a model that includes a corporate credit spread produces superior in- and out-of-sample recession forecasts compared with a model that includes only the term spread. In addition, they find that the multivariate model produces a much lower incidence of false-positive recession predictions.

Rosenberg and Maurer (2008) investigate whether recession forecasts can be improved by distinguishing between the interest rate expectations and term premium components of the term spread. Their approach is similar to that of Hamilton and Kim (2002) discussed previously. If changes in the term premium distort the empirical relationship between the spread and recessions, a model that isolates interest rate expectations might yield superior recession forecasts. Rosenberg and Maurer (2008) find that the expectations component is more useful for forecasting recessions than the term premium and that only the coefficient on the expectations

component is statistically significant in the probit model. Their study finds, however, that the term spread and expectations component generally produce similar recession probability forecasts. Moreover, between August 2006 and May 2007, the term spread model predicted a significantly higher recession probability than did the expectations component model.

Several recent studies investigate nonlinearities in recession-forecasting models. For example, Dueker (1997) estimates a probit model with Markov-switching coefficient variation and a lagged dependent variable. He finds that allowing for Markov-switching coefficient variation on the term spread improves forecast accuracy, especially at longer horizons, while including the lagged value of the recession indicator improves the model's fit and forecast accuracy, especially at 3- to 12-month horizons. Further, Dueker (1997) finds that the nonlinear model produces fewer false warnings of recessions than a linear model.

Ahrens (2002) estimates a probit forecasting model in which the term spread is assumed to follow a two-state Markov process. Using data for 1970-96 for eight countries among the Organisation for Economic Co-operation and Development (Canada, France, Germany, Italy, Japan, the Netherlands, the United Kingdom, and the United States), Ahrens (2002) finds that the term spread is a reliable predictor of business cycle peaks and troughs. Like Dueker (1997), Ahrens (2002) finds that the regime-switching framework produces more-accurate estimates of recession probabilities.

Other studies that estimate augmented probit (or logit) models, or compare results from probit estimation with those obtained using other methods, include Chauvet and Potter (2005), Galvao (2006), and Dueker (2005).

Chauvet and Potter (2005) compare recession forecasts obtained using four different probit model specifications: (i) a time-invariant conditionally independent version, (ii) a business cycle-specific conditionally independent model, (iii) a time-invariant probit model with autocorrelated errors, and (iv) a business cycle-specific probit model with autocorrelated errors. Chauvet and Potter (2005) find evidence in favor of the

**Table 3****Selective Summary of Studies of the Usefulness of the Term Spread for Predicting Recessions**

Study	Methodology	Data (years)	Principal finding(s)	Notes
Estrella and Hardouvelis (1991)	Probit model	U.S. (1955-88)	Spread is useful for forecasting recessions 4 quarters ahead.	Results are robust to including short-term interest rate and other variables in model.
Dueker (1997)	Dynamic probit with Markov switching	U.S. (1959-95)	Spread is useful for prediction up to 12 months ahead.	Results are robust to including other variables, including lagged recession indicator and regime switching.
Dotsey (1998)	Probit model	U.S. (1955-97)	Spread is useful for prediction; outperforms naive model.	Spread failed to accurately forecast 1990-91 recession.
Estrella and Mishkin (1998)	Probit model	U.S. (1959-95)	Spread is useful for prediction, especially at 2- to 6-quarter horizons.	Spread dominates other financial variables for out-of-sample prediction.
Bernard and Gerlach (1998)	Probit model	Eight industrialized countries (1972-93)	Spread is useful for prediction at 4- to 8-quarter horizons.	Foreign spreads add little information, except for Japan (German spread) and the U.K. (U.S. spread).
Ahrens (2002)	Probit model with Markov switching	Eight industrialized countries (1971-96)	Spread is useful for prediction, especially cycle peaks.	Regime-switching framework allows onset and ending of recessions to be determined endogenously.
Estrella, Rodrigues, and Schich (2003)	Probit model	U.S. (1955-98) and Germany (1967-98)	Spread is useful for prediction at 12-month horizons, less so at 24- and 36-month horizons.	Results are generally robust to alternative term spreads, with little evidence of instability over time.
Sensier et al. (2004)	Logistic regression model	Germany, France, Italy, and U.K. (1970-2001)	Interest rates generally predict recessions at 3-month horizon.	Short- and long-term rates entered separately; U.S. and German interest rates were useful for predicting recessions in other countries.
Chauvet and Potter (2005)	Variants of probit model allowing for multiple structural breaks and autoregression	U.S. (1954-2001)	Spread is useful for prediction at 12-month horizon.	Model with breakpoints and autocorrelated errors fits better in sample than basic probit model.

NOTE: EMU, European Monetary Union.

**Table 3, cont'd****Selective Summary of Studies of the Usefulness of the Term Spread for Predicting Recessions**

<b>Study</b>	<b>Methodology</b>	<b>Data (years)</b>	<b>Principal finding(s)</b>	<b>Notes</b>
Duarte, Venetis, and Paya (2005)	Dynamic probit model	Euro area (1970-2004)	Spread is useful for prediction at 3-quarter horizon.	Both EMU and U.S. spreads useful, but EMU spread dominates.
Moneta (2005)	Standard and dynamic probit model	Euro area, Germany, France, and Italy (1970-2002)	Spread is useful for predicting at 1-year horizon; dynamic model outperforms standard probit model.	Spread between 10-year and 3-month Treasury securities dominates other spreads in forecasts.
Galvao (2006)	Structural break threshold VAR model	U.S. (1953-2003)	Spread is useful for predicting output at 2-quarter horizons.	Model allowing for structural breaks and nonlinearities outperforms standard VAR both in and out of sample.
Wright (2006)	Probit model	U.S. (1964-2005)	Spread is useful for predicting recessions.	Models that include the level of the federal funds rate produced superior in- and out-of-sample forecasts.
Rosenberg and Maurer (2008)	Probit model	U.S. (1961-2006)	The expectations component of the spread is more accurate than the term premium component at forecasting recessions.	The spread remains useful when the federal funds rate is included in the model.

NOTE: EMU, European Monetary Union.

business cycle–specific probit model with autocorrelated errors, which allows for multiple structural breaks across business cycles and autocorrelation.

Galvao (2006) estimates a recession-forecasting model that accounts for time-varying nonlinearity and structural breaks in the relationship between the term spread and recessions. The author finds that a model with time-varying thresholds predicts the timing of recessions better than models with a constant threshold or that allow only a structural break.

Finally, Dueker (2005) proposes a VAR (“Qual-VAR”) model to forecast recessions using data on the term spread, GDP growth, inflation, and the federal funds rate. He finds that the model fits well in sample and accurately forecasts the 2001 recession out of sample.

In summary, most empirical research to date finds that the term spread is useful for forecasting recessions—both for the United States and other countries—and that the spread predicts recessions more reliably than it does output growth. However, a few studies find that multivariate models that include other financial indicators besides the term spread improve recession-forecasting performance, as do models that account for threshold effects or other nonlinearities in the empirical relationship between the term spread and recessions.

## **CONCLUSION**

The literature on the relationship between the yield curve and economic activity is large and expanding rapidly. Much of the literature examines empirically how well the term spread forecasts output growth or recessions, with less emphasis on *why* the yield curve predicts economic activity. To a great extent, the observation that changes in the slope of the yield curve appear to forecast changes in economic activity remains, as Benati and Goodhart (2008, p. 1237) contend, “a stylized fact in search of a theory.”

Does the yield spread forecast output growth? Does it forecast recessions? The answer to both questions is a *qualified* “yes.” Early studies based on estimation of linear forecasting models using

postwar U.S. data, as well as several recent studies, find that the term spread forecasts output growth well. Much research finds that the term spread is useful for forecasting output growth, especially at horizons of 6 to 12 months, and that the term spread remains useful even if other variables, including measures of monetary policy, are added to the forecasting model. However, several recent studies also find considerable variation in the ability of the spread to forecast output growth across countries and time periods. In particular, several studies find that the spread’s ability to predict output growth has diminished since the mid-1980s. The literature also provides considerable evidence of nonlinearities and structural breaks in the relationship between the term spread and output growth.

In general, studies show that the term spread is a more reliable predictor of recessions than of output growth and that the spread provides good recession forecasts, especially up to one year ahead. Researchers generally obtain superior forecasting performance from (i) probit models that include a lagged recession indicator and Markov-switching coefficients or other nonlinearities and (ii) other nonlinear approaches, such as smooth transition regression and multivariate adaptive regression splines estimation.

The literature has not reached a consensus regarding the reasons for structural breaks or nonlinearities in the empirical relationship between the term spread and future economic activity. Several studies note that the relationship between the nominal yield curve and future economic activity is likely to depend on the nature of the monetary regime, including the relative responsiveness of the monetary authority to output and inflation. For example, the term spread is likely to forecast output growth better when the monetary authority is more responsive to output than inflation and when inflation is relatively persistent. Further estimation refinements, as well as additional research based on dynamic structural models (Ang, Piazzesi, and Wei, 2006), might provide insights into the interactions among the policy regime, financial variables, and output growth that help explain the questions posed by the empirical literature.

## REFERENCES

- Ahrens, Ralf. "Predicting Recessions with Interest Rate Spreads: A Multicountry Regime-Switching Analysis." *Journal of International Money and Finance*, August 2002, 21(4), pp. 519-37.
- Ang, Andrew; Piazzesi, Monika and Wei, Min. "What Does the Yield Curve Tell Us About GDP Growth?" *Journal of Econometrics*, March/April 2006, 131(1/2), pp. 359-403.
- Aretz, Kevin and Peel, David A. "Spreads versus Professional Forecasters as Predictors of Future Output Change." Working paper, Lancaster University, 2008; <http://ssrn.com/abstract=1123949>.
- Barsky, Robert B. "The Fisher Hypothesis and the Forecastability and Persistence of Inflation." *Journal of Monetary Economics*, January 1987, 19(1), pp. 3-24.
- Benati, Luca. "UK Monetary Regimes and Macroeconomic Stylized Facts." Bank of England Working Paper 290, Bank of England, March 2006; [www.bankofengland.co.uk/publications/workingpapers/wp290.pdf](http://www.bankofengland.co.uk/publications/workingpapers/wp290.pdf).
- Benati, Luca. "Investigating Inflation Persistence Across Monetary Regimes," *Quarterly Journal of Economics*, August 2008, 123(3), pp. 1005-60.
- Benati, Luca and Goodhart, Charles. "Investigating Time-Variation in the Marginal Predictive Power of the Yield Spread." *Journal of Economic Dynamics and Control*, April 2008, 32(4), pp. 1236-72.
- Berk, Jan M. "The Information Content of the Yield Curve for Monetary Policy: A Survey." *De Economist*, July 1998, 146(2), pp. 303-20.
- Berk, Jan M. and van Bergeijk, Peter A.G. "On the Information Content of the Yield Curve: Lessons for the Eurosystem?" *Kredit und Kapital*, 2001, 1, pp. 28-47.
- Bernanke, Ben S. and Blinder, Alan S. "The Federal Funds Rate and the Channels of Monetary Transmission." *American Economic Review*, September 1992, 82(4), pp. 901-21.
- Bernard, Henri and Gerlach, Stefan. "Does the Term Structure Predict Recessions? The International Evidence." *International Journal of Finance and Economics*, July 1998, 3(3), pp. 195-215.
- Bordo, Michael D. and Haubrich, Joseph G. "The Yield Curve, Recessions, and the Credibility of the Monetary Regime: Long-Run Evidence, 1875-1997." NBER Working Paper No. 10431, National Bureau of Economic Research, April 2004; [www.nber.org/papers/w10431.pdf?new\\_window=1](http://www.nber.org/papers/w10431.pdf?new_window=1).
- Bordo, Michael D. and Haubrich, Joseph G. "The Yield Curve as a Predictor of Growth: Long-run Evidence, 1875-1997." *Review of Economics and Statistics*, February 2008, 90(1), pp. 182-85.
- Bordo, Michael D. and Schwartz, Anna J. "Monetary Policy Regimes and Economic Performance: The Historical Record," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*. Chap. 2. Amsterdam: North Holland, 1999.
- Chauvet, Marcelle and Potter, Simon. "Forecasting Recessions Using the Yield Curve." *Journal of Forecasting*, March 2005, 24(2), pp. 77-103.
- Cozier, Barry and Tkacz, Greg. "The Term Structure and Real Activity in Canada." Bank of Canada Working Paper No. 94-3, Bank of Canada, March 1994; [www.bankofcanada.ca/en/res/wp/1994/wp94-3.pdf](http://www.bankofcanada.ca/en/res/wp/1994/wp94-3.pdf).
- D'Agostino, Antonello; Domenico, Giannone and Surico, Paolo. "(Un)predictability and Macroeconomic Stability." European Central Bank Working Paper No. 605, European Central Bank, April 2006; [www.ecb.int/pub/pdf/scpwps/ecbwp605.pdf](http://www.ecb.int/pub/pdf/scpwps/ecbwp605.pdf).
- Davis, E. Phillip and Fagan, Gabriel. "Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in EU Countries?" *Journal of Applied Econometrics*, November/December 1997, 12(6), pp. 701-14.
- Dotsey, Michael. "The Predictive Content of the Interest Rate Yield Spread for Future Economic Growth." *Federal Reserve Bank of Richmond Economic Quarterly*, Summer 1998, 84(3), pp. 31-51; [www.richmondfed.org/publications/research/economic\\_quarterly/1998/summer/pdf/dotsey.pdf](http://www.richmondfed.org/publications/research/economic_quarterly/1998/summer/pdf/dotsey.pdf).

## Wheelock and Wohar

- Duarte, Augustin; Venetis, Ioannis A. and Paya, Ivan. "Predicting Real Growth and the Probability of Recession in the Euro Area Using the Yield Spread." *International Journal of Forecasting*, April/June 2005, 21(2), pp. 262-77.
- Dueker, Michael J. "Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions." Federal Reserve Bank of St. Louis *Review*, March/April 1997, 79(2), pp. 41-51; <http://research.stlouisfed.org/publications/review/97/03/9703md.pdf>.
- Dueker, Michael J. "Dynamic Forecasts of Qualitative Variables: A Qual VAR Model of U.S. Recessions." *Journal of Business and Economics Statistics*, January 2005, 23(1), pp. 96-104.
- Estrella, Arturo. "Why Does the Yield Curve Predict Output and Inflation?" *Economic Journal*, July 2005, 115(505), pp. 722-44.
- Estrella, Arturo and Hardouvelis, Gikas A. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance*, June 1991, 46(2), pp. 555-76.
- Estrella, Arturo and Mishkin, Frederic S. "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank." *European Economic Review*, July 1997, 41(7), pp. 1375-401.
- Estrella, Arturo and Mishkin, Frederic S. "Predicting U.S. Recessions: Financial Variables as Leading Indicators." *Review of Economics and Statistics*, February 1998, 80(1), pp. 45-61.
- Estrella, Arturo; Rodrigues, Anthony P. and Schich, Sebastian. "How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States." *Review of Economics and Statistics*, August 2003, 85(3), pp. 629-44.
- Estrella, Arturo and Trubin, Mary R. "The Yield Curve as a Leading Indicator: Some Practical Issues." Federal Reserve Bank of New York *Current Issues in Economics and Finance*, July/August 2006, 12(5), pp. 1-7; [www.newyorkfed.org/research/current\\_issues/ci12-5.pdf](http://www.newyorkfed.org/research/current_issues/ci12-5.pdf).
- Feroli, Michael. "Monetary Policy and the Information Content of the Yield Spread." *Topics in Macroeconomics*, September 2004, 4(1), Article 13.
- Galbraith, John W. and Tkacz, Greg. "Testing for Asymmetry in the Link Between the Yield Spread and Output in the G-7 Countries." *Journal of International Money and Finance*, October 2000, 19(5), pp. 657-72.
- Galvão, Ana Beatriz C. "Structural Break Threshold VARs for Predicting U.S. Recessions Using the Spread." *Journal of Applied Econometrics*, May/June 2006, 21(4), pp. 463-87.
- Giacomini, Raffaella and Rossi, Barbara. "How Stable Is the Forecasting Performance of the Yield Curve for Output Growth?" *Oxford Bulletin of Economics and Statistics*, December 2006, 68(Suppl. 1), pp. 783-95.
- Giacomini, Raffaella and Rossi, Barbara. "Detecting and Predicting Forecast Breakdown." *Review of Economic Studies*, April 2009, 76(2), pp. 669-705.
- Granger, Clyde W. J. and Teräsvirta, Timo. *Modeling Nonlinear Economic Relationships*. New York: Oxford University Press, 1993.
- Hamilton, James D. and Kim, Dong H. "A Re-Examination of the Predictability of the Yield Spread for Real Economic Activity." *Journal of Money, Credit, and Banking*, May 2002, 34(2), pp. 340-60.
- Harvey, Campbell R. "The Real Term Structure and Consumption Growth." *Journal of Financial Economics*, December 1988, 22(2), pp. 305-33.
- Harvey, Campbell R. "Forecasts of Economic Growth From the Bond and Stock Markets." *Financial Analysts Journal*, September/October 1989, 45(5), pp. 38-45.
- Harvey, Campbell R. "The Term Structure and World Economic Growth." *Journal of Fixed Income*, June 1991, 1(1), pp. 7-19.
- Haubrich, Joseph G. and Dombrosky, Ann M. "Predicting Real Growth Using the Yield Curve."

- Federal Reserve Bank of Cleveland *Economic Review*, First Quarter 1996, 32(1), pp. 26-35; [www.clevelandfed.org/Research/Review/1996/96-q1-haubrich.pdf](http://www.clevelandfed.org/Research/Review/1996/96-q1-haubrich.pdf).
- Hu, Zuli. "The Yield Curve and Real Economic Activity." *IMF Staff Papers*, December 1993, 40(4), pp. 781-806.
- Jardet, Caroline. "Why Did the Term Structure of Interest Rates Lose Its Predictive Power?" *Economic Modelling*, May 2004, 21(3), pp. 509-24.
- Kessel, Reuben, A. "The Cyclical Behavior of the Term Structure of Interest Rates." NBER Occasional Paper 91, National Bureau of Economic Research, 1965.
- Kim, Kenneth A. and Limpaphayom, Piman. "The Effect of Economic Regimes on the Relation Between Term Structure and Real Activity in Japan." *Journal of Economics and Business*, July/August 1997, 49(4), pp. 379-92.
- King, Thomas B.; Levin, Andrew T. and Perli, Roberto. "Financial Market Perceptions of Recession Risk." Finance and Economics Discussion Series No. 2007-57, Board of Governors of the Federal Reserve System; [www.federalreserve.gov/pubs/feds/2007/200757/index.html](http://www.federalreserve.gov/pubs/feds/2007/200757/index.html).
- Laurent, Robert D. "An Interest Rate-Based Indicator of Monetary Policy." Federal Reserve Bank of Chicago *Economic Perspectives*, 1988, 12(1), pp. 3-14; [www.chicagofed.org/publications/economicperspectives/1988/ep\\_jan\\_feb1988\\_part1\\_laurent.pdf](http://www.chicagofed.org/publications/economicperspectives/1988/ep_jan_feb1988_part1_laurent.pdf).
- Laurent, Robert D. "Testing the 'Spread.'" Federal Reserve Bank of Chicago *Economic Perspectives*, 1989, 13(4), pp. 22-34; [www.chicagofed.org/publications/economicperspectives/1989/ep\\_jul\\_aug1989\\_part3\\_laurent.pdf](http://www.chicagofed.org/publications/economicperspectives/1989/ep_jul_aug1989_part3_laurent.pdf).
- Moneta, F. "Does the Yield Spread Predict Recession in the Euro Area?" *International Finance*, Summer 2005, 8(2), pp. 263-301.
- Nakaota, Hiroshi. "The Term Structure of Interest Rates in Japan: The Predictability of Economic Activity." *Japan and the World Economy*, August 2005, 17(3), pp. 311-26.
- Plosser, Charles I. and Rouwenhorst, K. Geert. "International Term Structures and Real Economic Growth." *Journal of Monetary Economics*, February 1994, 33(1), pp. 133-55.
- Rendu de Lint, Christel and Stolin, David. "The Predictive Power of the Yield Curve: A Theoretical Assessment." *Journal of Monetary Economics*, October 2003, 50(7), pp. 1603-22.
- Rosenberg, Joshua V. and Maurer, Samuel. "Signal or Noise? Implications of the Term Premium for Recession Forecasting." Federal Reserve Bank of New York *Economic Policy Review*, July 2008, 14(1), pp. 1-11; [www.newyorkfed.org/research/epr/08v14n1/0807rose.pdf](http://www.newyorkfed.org/research/epr/08v14n1/0807rose.pdf).
- Sensier, Marianne; Artis, Michael J.; Osborn, Denise R. and Birchenhall, Chris. "Domestic and International Influences on Business Cycle Regimes in Europe." *International Journal of Forecasting*, April/June 2004, 20(2), pp. 343-57.
- Shaaf, Mohamad. "Predicting Recessions Using the Yield Curve: An Artificial Intelligence and Econometric Comparison." *Eastern Economic Journal*, Spring 2000, 26(2), pp. 171-90.
- Shiller, Robert J. and Siegel, Jeremy J. "The Gibson Paradox and Historical Movements in Real Long-term Interest Rates." *Journal of Political Economy*, October 1977, 85(1), pp. 11-30.
- Smets, Frank and Tsatsaronis, Kostas. "Why Does the Yield Curve Predict Economic Activity? Dissecting the Evidence for Germany and the United States." CEPR Discussion Paper No. 1758, Centre for Economic Policy Research, December 1997.
- Stock, James H. and Watson, Mark W. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature*, September 2003, 41(3), pp. 788-829.
- Teräsvirta, Timo. "Modeling Economic Relationships with Smooth Transition Regressions," in Aman Ullah and David E.A. Giles, eds., *Handbook of*

**Wheelock and Wohar**

*Applied Economic Statistics*. Chap. 15. New York: Marcel Dekker, 1998, pp. 507-32.

Tkacz, Greg. "Neural Network Forecasting of Canadian GDP Growth." *International Journal of Forecasting*, January/March 2001, 17(1), pp. 57-69.

Venetis, Ioannis A.; Paya, Ivan and Peel, David A. "Re-Examination of the Predictability of Economic Activity Using the Yield Spread: A Nonlinear Approach." *International Review of Economics and Finance*, 2003, 12(2), pp. 187-207.

Wright, Jonathan, H. "The Yield Curve and Predicting Recessions." Finance and Economics Discussion Series No. 2006-07, Federal Reserve Board of Governors, February 2006; [www.federalreserve.gov/pubs/feds/2006/200607/200607pap.pdf](http://www.federalreserve.gov/pubs/feds/2006/200607/200607pap.pdf).



# Mexico's Integration into NAFTA Markets: A View from Sectoral Real Exchange Rates

Rodolphe Blavy and Luciana Juvenal

The authors use a threshold autoregressive model to confirm the presence of nonlinearities in sectoral real exchange rate dynamics across Mexico, Canada, and the United States for the periods before and after the North American Free Trade Agreement (NAFTA). Although trade liberalization is associated with reduced transaction costs and lower relative price differentials among countries, the authors find, by using estimated threshold bands, that Mexico still faces higher transaction costs than its developed counterparts. Other determinants of transaction costs are distance and nominal exchange rate volatility. The authors' results show that the half-lives of sectoral real exchange rate shocks, calculated by Monte Carlo integration, imply much faster adjustment in the post-NAFTA period. (JEL F31, F36, F41)

Federal Reserve Bank of St. Louis *Review*, September/October 2009, 91(5, Part 1), pp. 441-64.

**T**he analysis of relative price differentials across countries and sectors offers a way to evaluate the degree of market integration. The law of one price (LOOP) states that identical goods should sell for the same price across countries when prices are expressed in a common currency. Evidence has shown, however, that prices of goods fail to fully equalize between countries, indicating that markets are not perfectly integrated.

Prices of homogeneous goods tend to differ across countries because the presence of transaction costs—such as transport costs and (explicit or implicit) trade barriers—limits price arbitrage. The study of the LOOP among members of the North American Free Trade Agreement (NAFTA) is of particular interest because it allows an assessment of whether regional trade liberalization results in faster price convergence and smaller

price differentials across countries and in greater market integration.

This paper focuses on three issues. First, we assess the degree of market integration between the United States, Mexico, and Canada by analyzing the validity of the LOOP between the countries. Second, we determine whether markets became more integrated, with reduced transaction costs, after the introduction of NAFTA. Finally, we analyze whether transaction costs are related to economic determinants.

Our study focuses on the role of transaction costs in modeling deviations from the LOOP. Several theoretical studies (see Dumas, 1992; Sercu and Raman, 1995; and O'Connell, 1998) show that because of transaction costs, it may not be profitable to arbitrage away relative price differences across countries when the marginal costs of arbitrage exceed the marginal benefits. This

---

Rodolphe Blavy is an economist at the International Monetary Fund and Luciana Juvenal is an economist at the Federal Reserve Bank of St. Louis. The authors thank the staff of the Banco de Mexico for their helpful comments, Steven Phillips for his contributions at various stages of preparation of this paper, and Roberto Benelli, Roberto Garcia-Saltos, David J. Robinson, Lucio Sarno, and seminar participants at the International Monetary Fund and at the Latin American and Caribbean Economic Association 2007 conference for comments. Volodymyr Tulin and Douglas Smith provided research assistance.

© 2009, The Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

situation generates a band of no trade where prices in two locations fail to equalize. Outside this threshold band, arbitrage is profitable and the sectoral real exchange rate (SRER) can become mean-reverting. This dynamic implies nonlinearities in SRERs and is well captured by using a threshold autoregressive (TAR) model for each sectoral relative price (see Tong, 1990; and Hansen, 1996 and 1997). The TAR model allows for deviations from the LOOP to exhibit unit root behavior inside the threshold band and to become mean-reverting outside the band. If there is no mean reversion in the outer regime, relative prices fail to equalize between countries—a sign of weak market integration. In this way, the estimated threshold bands provide a measure of transaction costs.

The empirical methodology analyzes dynamics in relative price adjustment and innovates by taking the perspective of an emerging market—Mexico.<sup>1</sup> Motivated by the previous literature, we investigate the presence of threshold-type nonlinearities in deviations from the LOOP by comparing the monthly real U.S. dollar/Mexican peso exchange rate, U.S. dollar/Canadian dollar exchange rate, and monthly real Mexican peso/Canadian dollar exchange rate over 1980-2006.

Nonlinearities are captured using a self-exciting threshold autoregressive (SETAR) model. More precisely, we estimate SETAR models for each SRER for the pre- and post-NAFTA periods. This estimation gives a measure of transaction costs (threshold band) and the autoregressive parameter outside the band. We determine whether deviations from the LOOP show mean-reverting properties by testing whether the nonlinear specification is superior to a nonstationary model for each subsample. This requires testing whether the autoregressive process outside the band is significantly different from the random walk observed inside the band. We also test whether the threshold bands are significantly wider for each SRER in the pre- and post-NAFTA periods,

thus allowing assessment of whether NAFTA led to higher market integration.

The results show that transaction costs are larger for the Mexico-U.S. and Mexico-Canada country pairs than for the Canada-U.S. pair, thus suggesting a higher degree of market integration between the United States and Canada. We also find that NAFTA significantly reduced transaction costs and price differentials between the United States and Mexico, although this was not uniform across sectors. Finally, our estimated transaction costs are negatively related to trade liberalization, commonly shared geographic borders, and lower exchange rate volatility.

To measure the speed of mean reversion, we use generalized impulse response functions to compute the half-life of exchange rates, which is the time it takes for 50 percent of the effect of a shock to dissipate (see Koop, Pesaran, and Potter, 1996). We find that half-lives are substantially reduced after the introduction of NAFTA, especially for the Mexico-U.S. country pair. This implies that reduced arbitrage costs were accompanied by faster adjustments in price differentials.

The remainder of the paper is organized as follows. The next section reviews theoretical considerations on nonlinear dynamics in SRERs and presents the corresponding econometric methodology. The following sections first discuss the results and then provide a battery of robustness tests. The last section concludes.

## **NONLINEARITIES: MOTIVATION AND EMPIRICAL FRAMEWORK**

According to the LOOP, similar goods should be priced the same across countries when prices are expressed in a common currency. At the aggregate level, the LOOP translates into purchasing power parity. The LOOP is based on the assumption of frictionless goods arbitrage—an environment in which there are no impediments to trade or transaction costs that would prevent perfect arbitrage.

Ample empirical evidence (Isard, 1977; Richardson, 1978; and Giovannini, 1988) suggests

---

<sup>1</sup> There is now an established literature on the nonlinear behavior of SERSs for developed markets (see Obstfeld and Taylor, 1997; Imbs et al., 2003; Sarno, Taylor, and Chowdhury, 2004; and Juvenal and Taylor, 2008).

that relative prices do not converge, or do so only with a very long-term horizon, and that price differentials are persistent. These studies also find that relative price differentials are significant and highly correlated with exchange rate movements.

One reason that prices of homogeneous commodities may not be the same across different countries is the existence of transaction costs arising from transport costs, tariffs, and nontariff barriers.<sup>2</sup> A number of theoretical papers suggest the importance of transport and trade barriers in creating price differences between countries (e.g., Dumas, 1992; Sercu and Raman, 1995; and O'Connell, 1998). The models described in such studies have incorporated different assumptions regarding the nature of trade costs. Overall, price differences driven by transaction costs can be expressed as  $S^i P_j^i = P_j^R + A_j$ , where  $S^i$  is the nominal exchange rate between country  $i$ 's currency and the reference country,  $P_j^i$  is the price of good  $j$  in country  $i$ ,  $P_j^R$  is the price of good  $j$  in the reference country, and  $A_j$  is the marginal transaction cost. In particular,  $A_j$  shows the minimum price difference that makes arbitrage profitable between country  $i$  and the reference country. In the presence of perfectly competitive markets and constant returns to scale technology and in the absence of sellers' pricing power, price differences that are higher than the transaction costs will be arbitrated. Thus,

$$(1) \quad -A_j \leq S^i P_j^i - P_j^R \leq A_j.$$

In this framework, transaction costs generate two regimes: (i) when price differentials are smaller than transaction costs, there is a regime of no arbitrage described by equation (1) and (ii) when price differences exceed transaction costs, arbitrage is profitable and equation (1) does not hold. This implies that price differentials behave in a nonlinear fashion. Price differentials follow a nonstationary process within the transaction costs band (or threshold band), and outside the

band they are mean reverting toward the band because of arbitrage effects.

The condition expressed in equation (1) can be written in terms of each SRER as

$$(2) \quad 1 - \frac{A_j}{P_j^R} \leq \frac{S^i P_j^i}{P_j^R} \leq 1 + \frac{A_j}{P_j^R},$$

where

$$\frac{S^i P_j^i}{P_j^R}$$

is the SRER between country  $i$ 's currency and the reference country for good  $i$ . Condition (2) implies that transaction cost bands and nonlinearities are both good-specific and country pair-specific.

Based on the previous theoretical framework, a number of empirical studies analyze the nonlinear nature of deviations from the LOOP in terms of a TAR model (e.g., Tong, 1990). The TAR model allows for the presence of a threshold band within which arbitrage is not profitable. Consequently, deviations from the LOOP follow a unit root process. Outside the band the process can become mean-reverting.

Recent contributions that use this model to analyze SRER dynamics of developed markets include Obstfeld and Taylor (1997), Sarno, Taylor, and Chowdhury (2004), Imbs et al. (2003), and Juvenal and Taylor (2008). In particular, Obstfeld and Taylor (1997), who used disaggregated data on clothing, food, and fuel, find evidence of nonlinearities in a sample of 32 locations. Sarno et al. (2004) provide support for nonlinear mean reversion with considerable cross-country and sectoral heterogeneity. They use annual price data interpolated into quarterly data for nine sectors and quarterly data on five exchange rates vis-à-vis the U.S. dollar. Juvenal and Taylor (2008) study the presence of nonlinearities in deviations from the LOOP for 19 sectors in 10 European countries and find significant evidence of threshold adjustment with transaction costs varying considerably across sectors and countries.

## Empirical Framework

**Data.** We use disaggregated monthly data on consumer price indices (CPIs) for 18 sectors from

<sup>2</sup> Heckscher (1916) first pointed out the possibility of nonlinearities in relative prices in the presence of trade frictions. In the case of Mexico, González and Rivadeneyra (2004) investigate the LOOP between Mexican cities and provide empirical evidence that transactions costs (including tariff and nontariff barriers) explain departures from the LOOP.

January 1980 to December 2006 for Mexico, the United States, and Canada. Data on CPIs were obtained from the Bank of Mexico, the U.S. Bureau of Labor Statistics, and Statistics Canada. The sectors analyzed are *bread, meat, fish, dairy, fruits, veg* (vegetables), *nonalco* (nonalcoholic beverages), *alco* (alcoholic beverages), *tobac* (tobacco), *clothw* (women’s clothing), *clothm* (men’s clothing), *foot* (footwear), *fuel, furniture, medic* (medication), *vehicles, gasoline, and photo* (photographic equipment). Table 1 lists the sectors analyzed in this study and the description of the category for each country. Monthly nominal exchange rates are period averages from International Financial Statistics of the International Monetary Fund.

**Model.** We model deviations from the LOOP using a SETAR model for each sectoral exchange rate to analyze the patterns in relative price convergence. More precisely, we investigate the presence of nonlinearities in deviations from the LOOP using a threshold-type model with two regimes.

Our model process involves four steps. First, we estimate TAR models for each SRER. Second, we explore the validity of the nonlinear threshold model with respect to a null hypothesis of unit root process. This allows us to test for the existence of some degree of price convergence as opposed to no price convergence at all.<sup>3</sup> Third, when we find evidence that a nonlinear specification is superior to a nonstationary model, we determine whether price convergence is characterized by an asymmetric threshold adjustment consistent with arbitrage arguments. That is, we test whether a nonlinear model fits the data better than a stationary linear one. Finally, when we find evidence of nonlinear price convergence in

the pre- and post-NAFTA periods, we determine whether the size of the threshold band is equal in both periods.

The existence of transaction costs, in the form of transport costs or trade barriers, is one explanation for the lack of price convergence. As described previously, frictions to trade imply the presence of significant nonlinearities in SRER dynamics. That is, transaction costs generate a band in which the marginal costs of arbitrage exceed the marginal benefit. Within this band, there is a zone of no trade and consequently prices in two locations fail to equalize. Outside this band, arbitrage is profitable and the SRER can become mean-reverting. Empirically, this pattern is described by a TAR model, which was originally popularized by Balke and Fomby (1997) in the context of testing for purchasing power parity and the LOOP.

Let  $x_{jt}^i$  be the deviation from the LOOP for a sector  $j$  in country  $i$  at time  $t$ , defined as follows:

$$(3) \quad x_{jt}^i = s_t^i + p_{jt}^i - p_{jt}^R,$$

where  $s_t^i$  is the logarithm of the nominal exchange rate between country  $i$ 's currency and the reference country,  $p_{jt}^i$  is the logarithm of the price of good  $j$  in country  $i$  at time  $t$ , and  $p_{jt}^R$  is the logarithm of the price of good  $j$  in the reference country at time  $t$ .

A simple three-regime TAR model may be written as

$$(4) \quad q_{jt}^i = \alpha q_{jt-1}^i + \varepsilon_{jt}^i \text{ if } |q_{jt-d}^i| \leq \kappa$$

$$(5) \quad q_{jt}^i = \kappa(1 - \rho) + \rho q_{jt-1}^i + \varepsilon_{jt}^i \text{ if } q_{jt-d}^i > \kappa$$

$$(6) \quad q_{jt}^i = -\kappa(1 - \rho) + \rho q_{jt-1}^i + \varepsilon_{jt}^i \text{ if } q_{jt-d}^i < -\kappa$$

$$(7) \quad \varepsilon_{jt}^i \sim N(0, \sigma^2),$$

where  $q_{jt}^i$  is the demeaned component of the relative price difference,  $x_{jt}^i$ , given by  $x_{jt}^i = c_j^i + q_{jt}^i$  ( $q_{jt}^i$  is estimated as an ordinary least squares [OLS] residual),  $\kappa$  is the threshold parameter,<sup>4</sup> and  $q_{jt-d}^i$  is the threshold variable for sector  $j$  and country  $i$ . The parameter  $d$  accounts for the delay with

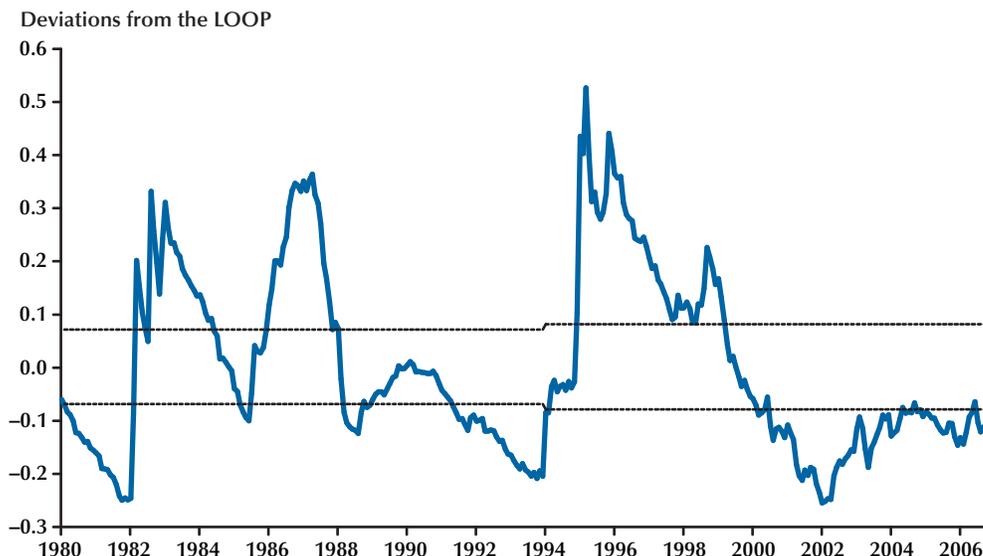
<sup>3</sup> A failure to reject the unit root hypothesis implies that deviations from the LOOP are a uniform unit root process and, thus, prices in two locations are disconnected. This test allows identification of any difference in the autoregressive parameters between the inner band and the outer band regimes. This test is an important addition to the methodology generally used in the literature. Earlier studies directly test for nonlinearity with respect to a linear model but do not determine whether the outer regime is nonstationary. An exception is found in Peel and Taylor (2002), who present a procedure to test for unit root to study covered interest parity. We use the procedure developed by Enders and Granger (1998) to test for the null hypothesis of nonstationarity against an alternative of stationarity with threshold adjustment.

<sup>4</sup> Note that  $\kappa$  is country and sector specific.

**Table 1**  
**Categories of Goods in the CPIs**

Sector	Mexico	United States	Canada
<i>Bread</i>	Bread, tortillas, and cereals	Cereals and bakery products	Bakery and other cereal products
<i>Meat</i>	Meat	Meat	Meat
<i>Fish</i>	Fish and seafood	Fish and seafood	Fish and other seafood
<i>Dairy</i>	Milk, dairy products, and eggs	Dairy and related products	Dairy products and eggs
<i>Fruits</i>	Fresh fruits	Fresh fruits	Fruit, fruit preparation, and nuts
<i>Veg</i>	Fresh vegetables	Fresh vegetables	Fresh vegetables
<i>Nonalco</i>	Sugar, coffee, and packaged refreshments	Nonalcoholic beverages	—
<i>Alco</i>	Alcoholic beverages	Alcoholic beverages	Alcoholic beverages
<i>Tobac</i>	Tobacco	Tobacco	Tobacco products and smokers' supplies
<i>Clothw</i>	Women's clothing	Women's apparel	Women's wear
<i>Clothm</i>	Men's clothing	Men's apparel	Men's wear
<i>Foot</i>	Footwear	Footwear	Footwear
<i>Fuel</i>	Electricity and fuel	Fuel and utilities	Water, fuel, and electricity
<i>Furniture</i>	Furniture	Furniture and bedding	Furniture
<i>Medic</i>	Medications and equipment	Medical care commodities	—
<i>Vehicles</i>	Acquisition of vehicles	New vehicles	Purchase of automotive vehicles
<i>Gasoline</i>	Gasoline and lubricants/oil	Gasoline (all types)	Gasoline
<i>Photo</i>	Photographic equipment and material	Photographic equipment and supplies	—

**Figure 1**  
**Footwear Real Exchange Rate and Threshold Bands**



which economic agents react to real exchange rate deviations.

Hereafter, we restrict the value of  $\alpha$  to unity, so that, inside the band, deviations from the LOOP are persistent and follow a random walk.<sup>5</sup> Outside the band, when  $|q_{jt-d}^i| > \kappa$ , the process becomes mean-reverting as long as  $\rho < 1$ . The model described is a TAR (1, 2,  $d$ ), where 1 is the autoregressive order, 2 is the number of thresholds, and  $d$  is the delay parameter. Further, because the threshold variable is assumed to be the lagged dependent variable, the model is called SETAR (1, 2,  $d$ ) with the given parameters.

Figure 1 shows an example of the estimated model. The graph contains the time series for  $q_{jt}^i$  (solid line), which represents the demeaned real exchange rate between Mexico and the United States for the footwear sector and the estimated  $\kappa$  (dashed lines).

**Estimation.** Using indicator functions  $1(q_{jt-d}^i > \kappa)$  and  $1(q_{jt-d}^i < -\kappa)$ , which take the

value of 1 when the inequality is satisfied, the model in equations (4) through (7) can be simplified to equation (8):

$$(8) \quad \Delta q_{jt}^i = [(\rho - 1)(q_{jt-1}^i - \kappa)]1(q_{jt-d}^i > \kappa) + [(\rho - 1)(q_{jt-1}^i + \kappa)]1(q_{jt-d}^i < -\kappa) + \varepsilon_{jt}^i.$$

Note that the model in equation (8) is assumed to be symmetric. Thus, deviations from the LOOP outside the threshold band are the same regardless of whether prices are higher in the United States or in another country. This specification assumes that reversion is toward the edge of the band.

Let us rewrite equation (8) as

$$(9) \quad \Delta q_{jt}^i = B_{jt}^i(\kappa, d)' \Gamma + \varepsilon_{jt}^i,$$

where  $B_{jt}^i(\kappa, d)'$  is a  $(1 \times 2)$  row vector that describes the behavior of  $\Delta q_{jt}^i$  in the outer regime and  $\Gamma$  is a  $(2 \times 1)$  vector containing the autoregressive parameters to be estimated. More precisely,

$$(10) \quad B_{jt}^i(\kappa, d)' = [X'1(q_{jt-d}^i > \kappa) \quad Y'1(q_{jt-d}^i < -\kappa)],$$

where

<sup>5</sup> This restriction is widely used in the literature; see Obstfeld and Taylor (1997), Imbs et al. (2003), Sarno, Taylor, and Chowdhury (2004), and Juvenal and Taylor (2008).

$$\begin{aligned} X' &= [q_{t-1} - \kappa] \\ Y' &= [q_{t-1} + \kappa] \end{aligned}$$

and

$$(11) \quad \Gamma' = [\rho - 1 \quad \rho - 1].$$

The parameters of interest are  $\Gamma$ ,  $\kappa$ , and  $d$ . Equation (8) is a regression equation nonlinear in parameters that can be estimated using least squares. For given values of  $\kappa$  and  $d$ , the least-squares estimate of  $\Gamma$  is

$$(12) \quad \hat{\Gamma}(\kappa, d) = \left( \sum_{t=1}^T B_{jt}^i(\kappa, d) B_{jt}^i(\kappa, d)' \right)^{-1} \left( \sum_{t=1}^T B_{jt}^i(\kappa, d) \Delta q_{jt}^i \right),$$

with residuals

$$\hat{\varepsilon}_{jt}^i(\kappa, d) = \Delta q_{jt}^i - B_{jt}^i(\kappa, d)' \hat{\Gamma}(\kappa, d),$$

and residual variance

$$(13) \quad \hat{\sigma}^2(\kappa, d) = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_{jt}^i(\kappa, d)^2.$$

Because the values of  $\kappa$  and  $d$  are not given, they should be estimated together with the autoregressive parameter,  $\rho$ . Hansen (1997) suggests a methodology to identify the model in equation (9) that consists of the simultaneous estimation of  $\kappa$ ,  $d$ , and  $\rho$  via a grid search over  $\kappa$  and  $d$ . The model is estimated by sequential least squares for values of  $d$  from 1 to 6. The values of  $\kappa$  and  $d$  that minimize the sum of squared residuals are chosen. The range for the grid search is selected to contain the 15th and 85th percentiles of the threshold variable. This can be written as

$$(14) \quad (\hat{\kappa}, \hat{d}) = \arg \min_{\kappa \in \Theta, d \in \Psi} \hat{\sigma}^2(\kappa, d),$$

where  $\Theta = [\underline{\kappa}, \bar{\kappa}]$ .

The least-squares estimator of  $\Gamma$  is  $\hat{\Gamma} = \hat{\Gamma}(\hat{\kappa}, \hat{d})$  with residuals

$$\hat{\varepsilon}_{jt}^i(\hat{\kappa}, \hat{d}) = \Delta q_{jt}^i - B_{jt}^i(\hat{\kappa}, \hat{d})' \hat{\Gamma}(\hat{\kappa}, \hat{d})$$

and residual variance

$$\hat{\sigma}^2(\hat{\kappa}, \hat{d}) = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_{jt}^i(\hat{\kappa}, \hat{d})^2.$$

**Testing Procedures.** Before explaining the results, it is important to determine whether the TAR-type nonlinear model is superior when tested against a unit root process and against a linear AR(1) process. These tests require pre-estimation of both the linear model under the null hypothesis and the TAR model under the alternative.

First, we determine whether the SETAR specification is superior to a unit root process for each SRER using the Enders and Granger (1998) threshold unit root test.<sup>6</sup> The method is a generalization of the Dickey-Fuller test. The null hypothesis is

$$H_0^A : \rho = 1$$

against an alternative of stationarity with threshold adjustment. This test allows identification of any difference in the autoregressive parameters between the inner and outer regimes. Its main advantage is that it is generally more powerful than the Dickey-Fuller test. A failure to reject the unit root null hypothesis implies that the LOOP does not hold and prices in two locations are disconnected. We interpret this as conveying that transaction costs are so high that the entire series are included within the threshold bands. Thus, the inner and outer regimes cannot be distinguished.

When the unit root null hypothesis is rejected, we continue with our analysis. Our second step is to test a linear AR(1) specification against a nonlinear stationary SETAR. Let  $\beta$  be the autoregressive parameter implied by the linear AR(1). The linear null hypothesis is

$$H_0^B : \beta = \rho.$$

<sup>6</sup> Other tests for the null hypothesis of the unit root against a nonlinear model have been proposed in the literature. Recent contributions include Kapetanios and Shin (2006) and Bec, Guay, and Guerre (2008). In particular, Kapetanios and Shin (2006) propose a Wald statistic to test a unit root null hypothesis against a three-regime SETAR process. Bec, Guay, and Guerre (2008) develop a more general procedure that consists of an adaptive threshold SupWald unit root test. We emphasize that the decision to use the Enders and Granger (1998) test does not represent a criticism of other methods. Overall, simulations have not provided evidence in favor of one test or another and this analysis is beyond the scope of our paper.

When we find evidence of nonlinearities in the pre- and post-NAFTA periods, we determine whether the size of the threshold band is equal in both periods. Let  $\tau_j^i$  be the threshold variable in the post-NAFTA period and  $\theta_j^i$  be the threshold variable in the pre-NAFTA period. The null hypothesis is

$$H_0^C : \tau_j^i = \theta_j^i.$$

As noted in Hansen (1997), testing hypotheses  $H_0^B$  and  $H_0^C$  is not straightforward. A statistical problem is present because conventional tests have asymptotic nonstandard distributions. To overcome inference problems, the asymptotic distribution of the conventional  $F$ -statistic must be calculated using a Monte Carlo simulation. Following Hansen (1997) and Peel and Taylor (2002), if the errors are i.i.d., the null hypothesis  $H_0^B$  and  $H_0^C$  can be tested using the statistic

$$(15) \quad F_T(\kappa, d) = T \left( \frac{\tilde{\sigma}^2 - \hat{\sigma}^2(\kappa, d)}{\hat{\sigma}^2(\kappa, d)} \right),$$

where  $F_T$  is the  $F$ -statistic when  $\kappa$  and  $d$  are known,  $T$  is the sample size, and  $\hat{\sigma}^2(\kappa, d)$  and  $\tilde{\sigma}^2$  are the unrestricted and restricted estimates of the residual variance, respectively. Hence,  $\hat{\sigma}^2(\kappa, d)$  is obtained from the unconstrained nonlinear least-squares estimation of equation (8) and  $\tilde{\sigma}^2$  results from the estimation of equation (8) with the restriction to be tested imposed.

Because  $\kappa$  and  $d$  are not identified under the null hypothesis, the distribution of  $F_T(\kappa, d)$  is not a standard chi-square distribution. Hansen (1997) shows that the asymptotic distribution of  $F_T(\kappa, d)$  may be approximated using the following bootstrap procedure: (i) generate  $y_{jt}^{i*}, t = 1, \dots, T$  from i.i.d.  $N(0, 1)$  random draws; (ii) set  $q_{jt}^{i*} = y_{jt}^{i*}$ ; (iii) using  $q_{jt-1}^{i*}$  for  $t = 1, \dots, T$ , regress  $y_{jt}^{i*}$  on  $q_{jt-1}^{i*}$  and estimate the restricted and unrestricted models and obtain the residual variances  $\tilde{\sigma}^{*2}$  and  $\hat{\sigma}^{*2}(\kappa, d)$ , respectively; (iv) with these residual variances, it is possible to calculate the following  $F$ -statistic:

$$(16) \quad F_T^*(\kappa, d) = T \left( \frac{\tilde{\sigma}^{*2} - \hat{\sigma}^{*2}(\kappa, d)}{\hat{\sigma}^{*2}(\kappa, d)} \right).$$

The bootstrap approximation to the asymptotic  $p$ -value of the test is calculated by counting the

number of bootstrap samples for which  $F_T^*(\kappa, d)$  exceeds the observed  $F_T(\kappa, d)$ .

## ESTIMATION RESULTS

### Testing for Nonlinearity

Tables 2A, 2B, and 2C show the results of the estimation of the SETAR model for the Mexico-U.S., Canada-U.S., and Mexico-Canada country pairs, respectively. The first step consists of testing the null hypothesis of a unit root using the Enders and Granger (1998) threshold unit root test. Essentially, this allows us to determine whether the autoregressive process is the same outside and inside the threshold band. A failure to reject the null hypothesis implies that the SRER is nonstationary and consequently prices in two locations are disconnected. Thus, the LOOP does not hold. Our interpretation of such a case is that transaction costs are so large that arbitrage is not profitable and the threshold band is wide enough to contain the entire time series of the SRER.

For the Mexico-U.S. country pair, the test rejects the unit root null hypothesis in half of the series for the pre-NAFTA period. By contrast, in the post-NAFTA period nonstationarity is found in four of the sectors. We interpret these results as evidence that NAFTA has been associated with greater integration between the United States and Mexico.

The behavior of relative prices between Mexico and Canada shows a similar pattern even though the degree of market integration has not improved as much in the post-NAFTA period as in the case of the United States and Mexico.

The deviations from the LOOP in the Canada-U.S. country pair show a different behavior. The unit root null hypothesis is rejected in 73 percent of the series in the pre-NAFTA period and in all the series except one in the post-NAFTA period. These results suggest that the Canadian and American markets have been more closely integrated, with a slight improvement with NAFTA.

To further test for the validity of the SETAR model, the second step consists of testing whether the nonlinear model is superior to a linear AR(1) process applying the Hansen test described pre-

**Table 2A****SETAR Estimation Results: Mexico–United States**

Sector	Pre-NAFTA				Post-NAFTA				
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	$p$ -Value $H_0^C$
<i>Bread</i>	—	—	0.52	—	—	—	0.24	—	—
<i>Meat</i>	0.27	0.92	—	0.00	0.09	0.96	—	0.00	0.00
<i>Fish</i>	—	—	0.15	—	0.02	0.96	—	0.00	—
<i>Dairy</i>	0.28	0.85	—	—	0.10	0.75	—	0.00	0.00
<i>Fruits</i>	—	—	0.25	—	0.05	0.84	—	0.00	—
<i>Veg</i>	0.09	0.78	—	0.00	0.15	0.70	—	0.00	0.05
<i>Nonalco</i>	—	—	0.35	—	0.15	0.81	—	0.00	—
<i>Alco</i>	0.10	0.92	—	0.00	—	—	0.11	—	—
<i>Tobac</i>	0.32	0.73	—	0.00	0.14	0.86	—	0.00	0.00
<i>Clothw</i>	0.18	0.86	—	0.00	0.09	0.83	—	0.00	0.01
<i>Clothm</i>	—	—	0.13	—	0.16	0.87	—	0.00	—
<i>Foot</i>	0.07	0.95	—	0.02	0.08	0.87	—	0.00	0.64
<i>Fuel</i>	—	—	0.34	—	—	—	0.59	—	—
<i>Furniture</i>	—	—	0.28	—	0.18	0.86	—	0.01	—
<i>Medic</i>	—	—	0.14	—	0.20	0.85	—	0.00	—
<i>Vehicles</i>	0.14	0.75	—	0.00	0.12	0.64	—	0.00	0.39
<i>Gasoline</i>	—	—	0.23	—	—	—	0.11	—	—
<i>Photo</i>	0.19	0.97	—	0.03	0.19	0.85	—	0.00	0.00

NOTE: This table shows the results from the estimation of the SETAR (1, 2,  $d$ ) model in equation (8).  $\kappa$  is the value of the threshold and  $\rho$  is the outer root of the TAR process. The estimation of  $\kappa$ ,  $\rho$ , and  $d$  is done simultaneously via a grid search over  $\kappa$  and  $d$  as described in the text. The  $p$ -values  $H_0^A$ ,  $H_0^B$ , and  $H_0^C$  represent, respectively, the marginal significance levels of the null hypothesis of unit root in the outer regime, null hypothesis of linearity, and null hypothesis of equality of thresholds during pre- and post-NAFTA periods.

**Table 2B****SETAR Estimation Results: Canada–United States**

Sector	Pre-NAFTA				Post-NAFTA				
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	$p$ -Value $H_0^C$
<i>Bread</i>	—	—	0.36	—	0.09	0.93	—	0.00	—
<i>Meat</i>	0.06	0.91	—	0.00	0.04	0.94	—	0.00	0.39
<i>Fish</i>	0.08	0.85	—	0.00	0.04	0.90	—	0.00	0.08
<i>Dairy</i>	0.07	0.91	—	0.00	0.07	0.95	—	0.00	—
<i>Fruits</i>	0.16	0.95	—	0.02	0.09	0.79	—	0.00	—
<i>Veg</i>	0.14	0.80	—	0.00	0.05	0.79	—	0.00	0.01
<i>Alco</i>	0.15	0.89	—	0.00	0.14	0.93	—	0.00	0.47
<i>Tobac</i>	—	—	0.14	—	—	—	0.41	—	—
<i>Clothw</i>	0.05	0.94	—	0.00	0.13	0.81	—	0.00	0.07
<i>Clothm</i>	—	—	0.23	—	0.14	0.93	—	0.00	—
<i>Foot</i>	—	—	0.18	—	0.08	0.96	—	0.00	—
<i>Fuel</i>	0.08	0.95	—	0.00	0.04	0.94	—	0.00	0.07
<i>Furniture</i>	0.16	0.91	—	0.00	0.10	0.95	—	0.01	0.02
<i>Vehicles</i>	0.08	0.92	—	0.00	0.07	0.94	—	0.00	0.54
<i>Gasoline</i>	0.27	0.79	—	0.00	0.28	0.72	—	0.00	0.46

NOTE: See Table 2A. In some cases, fewer sectors are shown because data were not available.

**Table 2C****SETAR Estimation Results: Mexico-Canada**

Sector	Pre-NAFTA				Post-NAFTA				
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	$p$ -Value $H_0^C$
<i>Bread</i>	—	—	0.34	—	—	—	0.53	—	—
<i>Meat</i>	0.24	0.90	—	0.00	0.76	—	—	0.00	0.03
<i>Fish</i>	0.14	0.87	—	0.00	0.14	—	—	0.01	—
<i>Dairy</i>	0.30	0.80	—	0.00	0.19	—	—	0.00	0.00
<i>Fruits</i>	—	—	0.17	—	0.15	—	—	0.00	—
<i>Veg</i>	0.15	0.71	—	0.00	0.21	—	—	0.00	0.07
<i>Alco</i>	0.23	0.92	—	0.00	0.27	—	—	0.00	0.58
<i>Tobac</i>	—	—	0.14	—	—	—	0.25	—	—
<i>Clothw</i>	0.15	0.80	—	0.00	0.21	—	—	0.00	0.14
<i>Clothm</i>	0.17	0.90	—	0.00	0.20	—	—	0.00	0.19
<i>Foot</i>	0.10	0.90	—	0.00	0.20	—	—	0.00	0.03
<i>Fuel</i>	—	—	0.27	—	—	—	0.61	—	—
<i>Furniture</i>	—	—	0.16	—	0.22	—	—	0.00	0.01
<i>Vehicles</i>	—	—	0.18	—	—	—	0.66	—	—
<i>Gasoline</i>	—	—	0.13	—	—	—	0.24	—	—

NOTE: See Table 2A.

viously. We conduct this test only for cases in which the Enders and Granger (1998) test rejects the unit root null hypothesis.<sup>7</sup> Our results show that the outcomes of the Hansen test are in line with those of the Enders and Granger (1998) test. When the Enders and Granger test finds evidence of threshold behavior, the Hansen test rejects the linear null hypothesis.

A few sectoral-level points should be highlighted. For the Mexico-U.S. country pair, the following sectors show evidence of unit root behavior: bread, a low-cost subsidized food sector; sectors subject to intervention through taxation, such as alcoholic and nonalcoholic beverages; and a sector with a high degree of differentiation, such as furniture. Interestingly, nonstationary behavior is found in sectors such as gasoline and fuel, which are characterized by a high degree of monopolistic power. Similarly, for the Mexico-Canada country pair there is evidence of unit root in gasoline and bread, further suggesting the potential role of specific regulations in price differences.

In the Canada-U.S. country pair, nonstationary behavior is present in sectors subject to government intervention, such as tobacco, clothing, and footwear. By contrast, threshold adjustment is significant in food products sectors except for bread.

### **Estimated Transaction Costs**

Tables 2A, 2B, and 2C show the estimated threshold bands for each SRER for the three country pairs. These bands are interpreted as a measure of transaction costs and thus reflect the degree of market integration.

Evidence of a strong NAFTA effect is found for the Mexico-U.S. SRERs. Transaction costs bands and the heterogeneity of the threshold values are significantly reduced after the introduction of NAFTA. In the pre-NAFTA period, they range from 7 percent (footwear) to 32 percent (tobacco). By contrast, in the post-NAFTA period, threshold values range from 2 percent (fish products) to 20 percent (medical commodities). At an indi-

vidual level, in sectors such as nonalcoholic beverages, clothing, furniture, and medication, transaction costs decrease from “very large” (unit root process) in the pre-NAFTA period to “measurable” with a threshold model in the post-NAFTA period. In sectors that exhibit significant nonlinear behavior in both periods, threshold bands are significantly smaller in the post-NAFTA period for meat, dairy, vegetables, tobacco, women’s clothing, and photo equipment. The reduction in the transaction costs bands suggests a greater market integration.

Considering those sectors in which nonlinearities are detected, average transaction costs in the Mexico-U.S. pair are smaller than those for the Mexico-Canada pair. Moreover, the latter pair shows evidence of unit root behavior in a greater number of sectors. This means that transaction costs are so high arbitrage is not worthwhile.

Transaction costs between the United States and Canada are the lowest among the three country pairs examined. Overall, average transaction costs are 34 percent higher between the United States and Mexico than between the United States and Canada. This result confirms previous evidence that the United States and Canada are the most integrated among NAFTA members.<sup>8</sup> We also find less dispersion in the threshold bands in the pre- and post-NAFTA periods. The fact that the integration between Canada and the United States started before the introduction of NAFTA could explain this result.

A further look at sectoral characteristics confirms that highly homogeneous sectors such as fish and fruits show relatively low threshold bands. This is a standard result in the literature, reported in studies for other country pairs (see Juvenal and Taylor, 2008). Compared with the work of Juvenal

<sup>7</sup> The Hansen test requires that the series are stationary; this is why we apply this test only for the series in which the unit root null hypothesis is rejected.

<sup>8</sup> One possible alternative explanation for the lower thresholds between the United States and Canada than between Mexico and the United States may be that goods are more homogeneous between the first two countries. More generally, the comparability of the sectors may vary across country pairs. First, wealth effects may be at play. The relatively large income differences between Mexico and the United States and Canada affect the specific goods sampled in each CPI category. This disparity may complicate the analysis with the varying composition among luxury, middle, and ordinary products across countries. Second, statistical differences exist in the compilation of price-level data, notably in adjustments for quality changes. A solution to this problem is to look at more disaggregated price indices and SRERs.

**Table 3A****Half-Lives: Mexico–United States**

Sector	Pre-NAFTA					Post-NAFTA				
	Shock (%)					Shock (%)				
	10	20	30	40	50	10	20	30	40	50
<i>Bread</i>	—	—	—	—	—	—	—	—	—	—
<i>Meat</i>	36	26	20	17	15	29	25	23	22	21
<i>Fish</i>	—	—	—	—	—	19	18	18	18	18
<i>Dairy</i>	20	15	11	9	8	7	5	5	5	5
<i>Fruits</i>	—	—	—	—	—	6	5	5	5	5
<i>Veg</i>	4	4	4	4	4	5	5	5	5	5
<i>Nonalco</i>	—	—	—	—	—	7	7	6	6	6
<i>Alco</i>	13	12	12	11	11	—	—	—	—	—
<i>Tobac</i>	18	12	8	7	6	8	7	7	7	7
<i>Clothw</i>	10	10	10	9	9	5	5	5	5	5
<i>Clothm</i>	—	—	—	—	—	10	8	8	7	7
<i>Foot</i>	18	17	16	16	16	6	6	6	6	6
<i>Fuel</i>	—	—	—	—	—	—	—	—	—	—
<i>Furniture</i>	—	—	—	—	—	14	10	8	8	8
<i>Medic</i>	—	—	—	—	—	8	8	8	8	7
<i>Vehicles</i>	6	5	5	4	3	6	4	4	4	4
<i>Gasoline</i>	—	—	—	—	—	—	—	—	—	—
<i>Photo</i>	55	49	44	40	37	24	14	10	9	8
Average	20	17	14	13	12	11	9	8	8	8

NOTE: This table shows the estimated half-lives of deviations from the LOOP for five shocks of various percentages: 10, 20, 30, 40, and 50. The half-lives were calculated conditional on average initial history using the generalized impulse response functions procedure developed by Koop et al. (1996).

and Taylor (2008), threshold bands among NAFTA members are on average slightly lower than those between the United States and European countries.

### **Half-Lives of Relative Price Adjustment**

A usual measure of the speed of mean reversion is the half-life, which is the time required for the effect of 50 percent of a shock to die out. Tables 3A, 3B, and 3C report the estimated half-lives (in terms of months) of price deviations from the LOOP for the Mexico-U.S., Canada-U.S. and Mexico-Canada SRERs.<sup>9</sup>

The speed of mean reversion is generally computed by taking into account the adjustment in the outer regime, which depends on the value

of  $\rho$ . In this case, the half-life is calculated as if it were a linear model, that is,  $\ln(0.5)/\ln(\rho)$ . Lo and Zivot (2001) emphasize the uncertainty of whether the computation of half-lives for linear models is applicable for nonlinear models. However, studies based on a SETAR model generally use this measure (see, for example, Taylor, 2001). As highlighted in Juvenal and Taylor (2008), although the estimated half-lives of the outer regime yield some insights on the speed of mean reversion, this measure is limited because it does not consider the regime switching within the SETAR model.

<sup>9</sup> We compute the half-lives only for cases in which we find evidence of threshold behavior.

**Table 3B****Half-Lives: Canada–United States**

Sector	Pre-NAFTA					Post-NAFTA				
	Shock (%)					Shock (%)				
	10	20	30	40	50	10	20	30	40	50
<i>Bread</i>	—	—	—	—	—	14	12	12	11	11
<i>Meat</i>	11	10	10	10	9	13	12	12	12	12
<i>Fish</i>	6	5	4	4	4	9	8	8	8	8
<i>Dairy</i>	12	10	10	10	10	16	15	15	14	14
<i>Fruits</i>	27	24	21	20	19	5	5	5	5	5
<i>Veg</i>	7	6	6	6	6	5	5	5	5	5
<i>Alco</i>	13	10	9	9	9	17	16	15	14	13
<i>Tobac</i>	—	—	—	—	—	—	—	—	—	—
<i>Clothw</i>	14	13	12	12	11	7	7	6	6	6
<i>Clothm</i>	—	—	—	—	—	18	15	14	13	13
<i>Foot</i>	—	—	—	—	—	25	22	20	20	19
<i>Fuel</i>	17	15	15	15	15	12	12	12	12	11
<i>Furniture</i>	21	15	13	12	12	29	24	21	19	18
<i>Vehicles</i>	13	12	11	11	11	14	13	13	13	12
<i>Gasoline</i>	8	7	6	6	6	7	5	5	5	5
Average	14	12	11	10	10	12	11	11	10	10

NOTE: See Table 3A.

Thus, we compute the half-life using generalized impulse response functions proposed by Koop, Pesaran, and Potter (1996). This method considers the nonlinear nature of the SETAR model and the different adjustment speeds in the inner and outer regimes. The SETAR model exhibits an infinite half-life within the threshold band and depends on  $\rho$  outside the band. A shock may cause the model to switch regimes, and this adjustment is not captured by the first methodology.

Following Taylor, Peel, and Sarno (2001), we compute the impulse response functions conditional on average initial history using Monte Carlo integration for shocks of 10, 20, 30, 40, and 50 percent. For the Mexico-U.S. pair, the average relative price adjustment is significantly faster in the post-NAFTA period. For example, for a 10 percent shock, the average pre-NAFTA half-life is 20 months, whereas the average is reduced to 11

months in the post-NAFTA period (see Table 3A). Our results also yield additional observations. In the post-NAFTA period, the speed of mean reversion varies less across different shock sizes than in the pre-NAFTA period. This suggests that relative prices adjust more quickly, independent of the size of the price shock. Half-lives vary substantially across sectors. Relative prices adjust fairly quickly for homogeneous goods, such as food products. The relative price of more high-end products (e.g., furniture and photographic equipment) takes longer to adjust.

The speed of relative price adjustment in the post-NAFTA period is comparable for the Mexico-U.S. and the Canada-U.S. pairs. For a 10 percent shock, the average half-lives are 11 months and 12 months, respectively. This contrasts with significant differences in the pre-NAFTA period when Mexico-U.S. relative prices were much slower to adjust than Canada-U.S. prices (see

**Table 3C****Half-Lives: Mexico-Canada**

Sector	Pre-NAFTA					Post-NAFTA				
	Shock (%)					Shock (%)				
	10	20	30	40	50	10	20	30	40	50
<i>Bread</i>	—	—	—	—	—	—	—	—	—	—
<i>Meat</i>	24	17	13	12	11	7	6	6	6	6
<i>Fish</i>	10	8	7	7	6	16	14	12	12	12
<i>Dairy</i>	9	7	6	5	5	11	9	9	8	8
<i>Fruits</i>	—	—	—	—	—	5	4	4	4	4
<i>Veg</i>	4	4	4	4	4	5	4	4	4	4
<i>Alco</i>	16	14	13	12	11	16	15	14	14	14
<i>Tobac</i>	—	—	—	—	—	—	—	—	—	—
<i>Clothw</i>	10	10	9	8	8	11	10	9	8	8
<i>Clothm</i>	12	11	11	10	9	14	13	12	12	11
<i>Foot</i>	9	8	8	8	7	15	13	12	12	11
<i>Fuel</i>	—	—	—	—	—	—	—	—	—	—
<i>Furniture</i>	—	—	—	—	—	8	6	6	5	5
<i>Vehicles</i>	—	—	—	—	—	—	—	—	—	—
<i>Gasoline</i>	—	—	—	—	—	—	—	—	—	—
Average	12	10	9	8	8	11	10	9	9	9

NOTE: See Table 3A.

Tables 3A and 3B). The half-lives of the Mexico-Canada country pairs are also less persistent in the post-NAFTA period (see Table 3C).

### Determinants of Thresholds

Based on the estimates of the SETAR models, we assess whether transaction costs are related to economic variables. To do this, we estimate a regression explaining the threshold parameter obtained from the section on estimated transaction costs:

$$(17) \quad \kappa = \lambda_j^i + \sum_{c=1}^C \Phi_j^i(c) z_j^i(c) + \varepsilon_j^i,$$

where  $\kappa$  is the threshold parameter and  $z_j^i$  is a vector of explanatory variables. In equation (17) we assess whether transaction costs, measured by the estimated thresholds, are explained by selected explanatory variables.

The explanatory variables are intended to capture the size and nature of transaction costs. The first variable (distance) is a proxy for shipping costs. Given the small number of country pairs and their relative proximity, distance appears to be a poor measure. Instead, we include a dummy variable that takes value 1 when countries share a common border. The second variable is the volatility of the nominal exchange rate, which intends to capture the uncertainty about the macroeconomic environment. It is measured as the standard deviation of monthly exchange rate observations. Third, we include a measure of “tradability,” defined as the sum of imports and exports relative to the total output in a sector for a given country sourced from the United Nations Industrial Development Organization (UNIDO) database. Fourth, we use the number of establishments in each sector as a proxy for competition, or concentration, obtained from the UNIDO data-

**Table 4**  
**Threshold Regressions**

Variables	(1)	(2)
Distance	-0.042 (0.054)*	-0.036 (0.058)*
Dummy post-NAFTA	-0.105 (0.002)**	-0.111 (0.001)**
Exchange rate volatility	4.468 (0.000)***	4.266 (0.000)***
Firms	-0.002 (0.477)	—
Tradability	-0.045 (0.259)	—
R <sup>2</sup>	0.34	0.33
N	89	89

NOTE: This table shows the results from the estimation of equation (17); *p*-values are shown in parentheses. \*, \*\*, and \*\*\* denote significance at the 10, 5, and 1 percent levels, respectively.

base. Finally, a dummy for the post-NAFTA period is included.

We examine the determinants of thresholds for the entire sample, including all three country pairs.<sup>10,11</sup> The results, shown in Table 4, indicate that three variables are significant: the post-NAFTA dummy, the shared border, and nominal exchange rate volatility. These variables are significant in all specifications. We find that the thresholds are lower when countries share a border. Nominal exchange rate volatility is also significant. This indicates that uncertainty about the macroeconomic environment limits arbitrage. The post-NAFTA dummy is also highly significant: The negative coefficient indicates that the introduction of NAFTA is associated with lower transaction costs. Neither the number of firms in a sector nor the degree of “tradability” in a sector is statistically significant (column 1 in Table 4).<sup>12</sup>

<sup>10</sup> Because we cannot obtain data on firms and tradability disaggregated for clothing (women) and clothing (men) but for only a generic clothing sector, we consider the average threshold value of clothing (women) and clothing (men) as the  $\hat{\kappa}$  value for clothing.

<sup>11</sup> When we find evidence of unit root behavior in deviations from the LOOP, we consider  $\kappa$  to be the highest value of the threshold variable in the grid search. This implies that transaction costs are so high that the entire SRER series is within the threshold band.

In column 2, these two variables are excluded with little change in the results.

Overall, thresholds appear to be determined by distance (border) and exchange rate volatility. These results are consistent with findings in the literature. For example, Imbs et al. (2003) find that distance and exchange rate volatility explain the threshold values.

Another strand of the literature analyzed the determinants of relative price differentials between the United States and Canada using different types of models. Our results are consistent with the findings of these studies. As an example, Engel and Rogers (1996) study the nature of deviations from the LOOP using CPI data for 14 goods sectors for different U.S. and Canadian cities. This study shows that the Canadian and U.S. markets are not perfectly integrated and that distance and border are major determinants of price differences. In a related study, Engel et al. (2005) investigate the LOOP between U.S. and Canadian cities using actual prices (instead of price indices). They find that absolute price differences between U.S. and Canadian prices are higher than 7 percent. In addition, their results show border plays a significant role in explaining price differentials between cities.

## ROBUSTNESS OF RESULTS

We conduct three robustness checks to gauge the sensitivity of empirical results to underlying assumptions and variable definitions. First, we consider the possibility of long-run trends in the measured price differentials arising from aggregation issues in price indices or the presence of nontradable components or quality differences. We define  $q_{jt}^i$  as the detrended and demeaned component of the price difference,  $x_{jt}^i$ , given by  $x_{jt}^i = +c_j^i + \theta t + q_{jt}^i$ . As described previously, it is estimated as an OLS residual.

Overall, our baseline findings prove robust to using detrended SRERs instead of the demeaned series. Tables 5A, 5B, and 5C show the results of the estimation of the SETAR model with detrended

<sup>12</sup> Poor data quality is a probable explanation for the lack of significance.

**Table 5A****SETAR Estimation Results (Detrended Data): Mexico–United States**

Sector	Pre-NAFTA				Post-NAFTA			
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$
<i>Bread</i>	—	—	0.31	—	—	—	0.14	—
<i>Meat</i>	0.26	0.92	—	0.00	0.03	0.94	—	0.00
<i>Fish</i>	—	—	0.18	—	0.03	0.95	—	0.00
<i>Dairy</i>	0.29	0.84	—	—	0.09	0.83	—	0.00
<i>Fruits</i>	—	—	0.13	—	0.02	0.82	—	0.00
<i>Veg</i>	0.06	0.77	—	0.00	0.15	0.78	—	0.00
<i>Nonalco</i>	—	—	0.16	—	0.10	0.76	—	0.00
<i>Alco</i>	0.22	0.79	—	0.00	—	—	0.17	—
<i>Tobac</i>	—	—	0.15	0.00	0.16	0.90	—	0.00
<i>Clothw</i>	0.17	0.88	—	0.00	0.18	0.80	—	0.00
<i>Clothm</i>	—	—	0.33	—	0.15	0.77	—	0.00
<i>Foot</i>	0.11	0.93	—	0.02	0.09	0.88	—	0.00
<i>Fuel</i>	—	—	0.22	—	—	—	0.70	—
<i>Furniture</i>	—	—	0.46	—	0.16	0.81	—	0.01
<i>Medic</i>	—	—	0.27	—	0.15	0.88	—	0.00
<i>Vehicles</i>	0.16	0.79	—	0.00	0.09	0.70	—	0.00
<i>Gasoline</i>	—	—	0.19	—	—	—	0.17	—
<i>Photo</i>	0.16	0.96	—	0.02	0.17	0.90	—	0.00

NOTE: See Table 2A.

**Table 5B****SETAR Estimation Results (Detrended Data): Canada–United States**

Sector	Pre-NAFTA				Post-NAFTA			
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$
<i>Bread</i>	—	—	0.40	—	0.15	0.83	—	0.00
<i>Meat</i>	—	—	0.23	—	0.03	0.95	—	0.00
<i>Fish</i>	0.11	0.85	—	0.00	0.02	0.94	—	0.00
<i>Dairy</i>	0.05	0.94	—	0.00	0.07	0.92	—	0.00
<i>Fruits</i>	0.11	0.88	—	0.02	0.09	0.83	—	0.00
<i>Veg</i>	0.04	0.72	—	0.00	0.03	0.85	—	0.00
<i>Alco</i>	0.08	0.91	—	0.00	0.10	0.82	—	0.00
<i>Tobac</i>	—	—	0.22	—	—	—	0.22	—
<i>Clothw</i>	0.04	0.90	—	0.00	0.09	0.80	—	0.00
<i>Clothm</i>	0.06	0.88	—	0.00	0.11	0.94	—	0.00
<i>Foot</i>	—	—	0.12	—	0.05	0.90	—	0.00
<i>Fuel</i>	0.05	0.90	—	0.00	0.09	0.86	—	0.00
<i>Furniture</i>	0.08	0.87	—	0.00	0.16	0.91	—	0.00
<i>Vehicles</i>	0.09	0.80	—	0.00	0.10	0.95	—	0.00
<i>Gasoline</i>	0.16	0.97	—	0.00	0.05	0.80	—	0.00

NOTE: See Table 2A.

**Table 5C****SETAR Estimation Results (Detrended Data): Mexico-Canada**

Sector	Pre-NAFTA				Post-NAFTA			
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$
<i>Bread</i>	0.28	0.82	—	0.00	0.21	0.72	—	0.00
<i>Meat</i>	0.22	0.92	—	0.00	0.11	0.88	—	0.00
<i>Fish</i>	—	—	0.13	—	0.12	0.92	—	0.00
<i>Dairy</i>	0.31	0.91	—	0.00	0.20	0.87	—	0.00
<i>Fruits</i>	—	—	0.11	—	0.08	0.78	—	0.00
<i>Veg</i>	0.08	0.75	—	0.00	0.12	0.70	—	0.00
<i>Alco</i>	0.22	0.83	—	0.00	0.25	0.93	—	0.01
<i>Tobac</i>	—	—	0.19	—	—	—	0.55	—
<i>Clothw</i>	0.24	0.94	—	0.02	0.24	0.72	—	0.00
<i>Clothm</i>	0.23	0.93	—	0.01	0.24	0.82	—	0.00
<i>Foot</i>	0.15	0.85	—	0.00	0.20	0.92	—	0.00
<i>Fuel</i>	—	—	0.35	—	—	—	0.31	—
<i>Furniture</i>	—	—	0.19	—	0.18	0.86	—	0.00
<i>Vehicles</i>	—	—	0.17	—	—	—	0.15	—
<i>Gasoline</i>	—	—	0.18	—	—	—	0.39	—

NOTE: See Table 2A.

**Table 6A****SETAR Estimation Results (Different Mean during Tequila Crisis): Mexico–United States**

Sector	Post-NAFTA			
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$
<i>Bread</i>	—	—	0.54	—
<i>Meat</i>	0.14	0.82	—	0.00
<i>Fish</i>	0.13	0.91	—	0.00
<i>Dairy</i>	0.07	0.71	—	0.00
<i>Fruits</i>	0.05	0.77	—	0.00
<i>Veg</i>	0.04	0.83	—	0.00
<i>Nonalco</i>	0.14	0.78	—	0.00
<i>Alco</i>	0.11	0.93	—	0.00
<i>Tobac</i>	0.08	0.89	—	0.00
<i>Clothw</i>	0.09	0.83	—	0.00
<i>Clothm</i>	0.10	0.79	—	0.00
<i>Foot</i>	0.08	0.94	—	0.00
<i>Fuel</i>	0.14	0.75	—	0.00
<i>Furniture</i>	0.11	0.90	—	0.00
<i>Medic</i>	0.17	0.77	—	0.00
<i>Vehicles</i>	0.12	0.83	—	0.00
<i>Gasoline</i>	—	—	0.25	—
<i>Photo</i>	0.12	0.91	—	0.00

NOTE: See Table 2A.

SRERs. The conceptual problem with including a trend in the real exchange rate is that it implies that the real exchange rate converges to a different mean across time. This implication is somewhat contradictory to the LOOP. Hence, our preferred measure is the demeaned series. The stability of our results with the different measures indicates that the trend component may not be of the utmost importance.

Second, we test the sensitivity of the results to a structural break in the Mexican series over the study period (1980-2006) during the Tequila Crisis. The results reported herein assume a constant mean over the period, consistent with the LOOP hypothesis. However, as a robustness check, we also test the sensitivity of the results to two conditions: (i) allowing for a different mean over the Tequila Crisis (1994:12–1995:12) and (ii)

restricting the estimation period to 1996-2006. This was intended to assess whether the Tequila Crisis would significantly affect our findings. Our baseline findings are again robust to these checks. Tables 6A, 6B, and 6C report the estimated thresholds for each SRER, allowing for a different mean for the real exchange rate during the Tequila Crisis. Across sectors, homogeneous goods have lower transaction costs than other goods in the sample. Across country pairs, average transaction costs among NAFTA members are 27 percent higher between the United States and Mexico than between the United States and Canada, slightly less than the results when the Tequila Crisis is ignored. The results of the latter robustness analysis (not reported here but available upon request) are broadly consistent with the ones discussed here; thus, the Tequila Crisis does not significantly affect our findings.

**Table 6B****SETAR Estimation Results (Different Mean during Tequila Crisis): Canada–United States**

Sector	Post-NAFTA			
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$
<i>Bread</i>	0.09	0.93	—	0.00
<i>Meat</i>	0.04	0.94	—	0.00
<i>Fish</i>	0.04	0.90	—	0.00
<i>Dairy</i>	0.07	0.95	—	0.00
<i>Fruits</i>	0.09	0.79	—	0.00
<i>Veg</i>	0.05	0.79	—	0.00
<i>Alco</i>	0.14	0.93	—	0.00
<i>Tobac</i>	0.05	0.95	—	0.03
<i>Clothw</i>	0.13	0.81	—	0.00
<i>Clothm</i>	0.14	0.93	—	0.00
<i>Foot</i>	0.08	0.96	—	0.00
<i>Fuel</i>	0.04	0.94	—	0.00
<i>Furniture</i>	0.10	0.95	—	0.00
<i>Vehicles</i>	0.07	0.94	—	0.00
<i>Gasoline</i>	0.26	0.72	—	0.00

NOTE: See Table 2A.

**CONCLUSION**

Using a SETAR model, we find strong evidence of nonlinearities in SREER dynamics across Mexico, Canada, and the United States in the pre- and post-NAFTA periods. This result is consistent with the predictions of theoretical models that incorporate some form of market segmentation. Overall, mean reversion occurs when deviations from the LOOP are significant and the benefits of arbitrage are higher than transaction costs.

We obtain two key parameters from the estimation of SETAR models. The first parameter is the threshold, taken as a measure of transaction costs. The second parameter is the autoregressive parameter in the outer regime, which determines the speed of mean reversion. We obtain these parameters for each SREER corresponding to the three country pairs for both periods.

Our findings indicate that the value of transaction costs is highly heterogeneous for different

sectors and countries. The estimated price thresholds range from 2 percent to 32 percent for the Mexico-U.S. and Canada-U.S. country pairs. The results generally confirm that highly homogeneous sectors, such as fish and fruits, show low threshold bands. Overall, average transaction costs among NAFTA members are 34 percent higher between the United States and Mexico than between the United States and Canada. This indicates that Mexico and the United States are relatively less integrated than Canada and the United States. In turn, threshold bands are higher for the Mexico-Canada pair.

We relate the value of the threshold band to plausible economic determinants. Our results show that the border effect and exchange rate volatility are significant determinants of transaction costs. The dummy post-NAFTA is also strongly significant and negative, confirming that the introduction of NAFTA is associated with lower transaction costs.

**Table 6C****SETAR Estimation Results (Different Mean during Tequila Crisis): Mexico-Canada**

Sector	Post-NAFTA			
	Threshold $\kappa$	Outer regime $\rho$	Unit root test $p$ -value $H_0^A$	Hansen test $p$ -value $H_0^B$
<i>Bread</i>	—	—	0.74	—
<i>Meat</i>	0.20	0.92	—	0.00
<i>Fish</i>	0.13	0.91	—	0.00
<i>Dairy</i>	0.08	0.97	—	0.05
<i>Fruits</i>	0.08	0.83	—	0.00
<i>Veg</i>	0.04	0.80	—	0.00
<i>Alco</i>	0.06	0.95	—	0.02
<i>Tobac</i>	—	—	0.25	—
<i>Clothw</i>	0.10	0.90	—	0.00
<i>Clothm</i>	0.11	0.89	—	0.00
<i>Foot</i>	0.06	0.95	—	0.02
<i>Fuel</i>	0.14	0.77	—	0.01
<i>Furniture</i>	—	—	0.16	—
<i>Vehicles</i>	—	—	0.13	—
<i>Gasoline</i>	—	—	0.07	—

NOTE: See Table 2A.

To shed some light on the mean-reverting properties of the SREs, we consider the regime switching that occurs inside and outside the band in the SETAR model and compute the half-lives using generalized impulse response functions. Overall, the speed of mean reversion depends on the size of the shock. Larger shocks mean-revert much faster than smaller ones. On average, the half-lives are substantially reduced after the introduction of NAFTA. For the Mexico-U.S. country pair, the average half-life is reduced from 20 months in the pre-NAFTA period to 11 months in the post-NAFTA period. The post-NAFTA period shows less variation in the speed of mean reversion across different shock sizes than in the pre-NAFTA period.

Our analysis therefore supports the arguments that (i) emerging markets—in this case, Mexico—still face higher transaction costs than their developed counterparts and (ii) trade liberalization

may help in lower relative price differentials between countries. We suspect that lack of competition may be a major determinant of high price thresholds but cannot prove this matter empirically.

The main conclusion of our analysis is that Mexico has made progress but still has considerable room for improvement in reducing barriers to goods market integration and achieving the full benefits of globalization. Future research should focus on why transactions costs between Mexico and the United States continue to exceed those between Canada and the United States for many types of goods and whether these costs can be reduced through policy actions. Examples of such actions include developing logistics, transportation, and internal distribution mechanisms or enhancing the state of competition among domestic firms and reducing remaining barriers to external trade.

## REFERENCES

- Bec, Frédérique; Guay, Alain and Guerre, Emmanuel. "Adaptive Consistent Unit-Root Tests Based on Autoregressive Threshold Model." *Journal of Econometrics*, January 2008, 142(1), pp. 94-133.
- Balke, Nathan S. and Fomby, Thomas B. "Threshold Cointegration." *International Economic Review*, August 1997, 38, pp. 627-45.
- Dumas, Bernard. "Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World." *Review of Financial Studies*, June 1992, 5(2), pp. 153-80.
- Enders, Walter and Granger, C.W.J. "Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates." *Journal of Business and Economic Statistics*, July 1998, 16, pp. 304-12.
- Engel, Charles and Rogers, John H. "How Wide Is the Border?" *American Economic Review*, December 1996, 86(5), pp. 1112-25.
- Engel, Charles; Rogers, John H. and Wang, Shing-Yi. "Revisiting the Border: An Assessment of the Law of One Price Using Very Disaggregated Consumer Price Data," in Rebecca Driver; Peter Sinclair and Christoph Thoenissen, eds., *Exchange Rates, Capital Flows and Policy*. London: Routledge, 2005, pp. 187-203.
- Giovannini, Alberto. "Exchange Rates and Traded Goods Prices." *Journal of International Economics*, February 1988, 24(1-2), pp. 45-68.
- González, Marco and Rivadeneyra, Francisco. "La Ley de un Solo Precio en México: Un Análisis Empírico." *Gaceta de Economía*, 2004, 19, pp. 91-115.
- Hansen, Bruce E. "Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis." *Econometrica*, March 1996, 64, pp. 413-30.
- Hansen, Bruce E. "Inference in TAR Models." *Studies in Nonlinear Dynamics and Econometrics*, April 1997, 2(1), pp. 1-14.
- Heckscher, Eli F. "Växelkursens Grundval vid Pappersmyrfot." *Economisk Tidskrift*, 1916, 18(10), pp. 309-12.
- Imbs, Jean; Mumtaz, Haroon; Ravn, Morten O. and Rey, Hélène. "Nonlinearities and Real Exchange Rate Dynamics." *Journal of the European Economic Association*, April 2003, 1(2-3), pp. 639-49.
- Isard, Peter. "How Far Can We Push the Law of One Price?" *American Economic Review*, December 1977, 67(5), pp. 942-48.
- Juvenal, Luciana and Taylor, Mark P. "Threshold Adjustment of Deviations from the Law of One Price." *Studies in Nonlinear Dynamics and Econometrics*, September 2008, 12(3), Article 8.
- Kapetanios, George and Shin, Yongcheol. "Unit Root Tests in Three-Regime SETAR Models." *Econometrics Journal*, June 2006, 9(2), pp. 252-78.
- Koop, Gary; Pesaran, M. Hashem and Potter, Simon M. "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics*, September 1996, 74(1), pp. 119-47.
- Lo, Ming Chien and Zivot, Eric. "Threshold Cointegration and Nonlinear Adjustment to the Law of One Price." *Macroeconomic Dynamics*, September 2001, 5(4), pp. 533-76.
- Obstfeld, Maurice and Taylor, Alan M. "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited." *Journal of Japanese and International Economics*, December 1997, 11(4), pp. 441-79.
- O'Connell, Paul G.J. "The Overvaluation of the Purchasing Power Parity." *Journal of International Economics*, February 1998, 44(1), pp. 1-19.
- Peel, David A. and Taylor, Mark P. "Covered Interest Rate Arbitrage in the Interwar Period and the Keynes-Einzig Conjecture." *Journal of Money, Credit, and Banking*, February 2002, 34(1), pp. 51-75.
- Richardson, J. David. "Some Empirical Evidence on Commodity Arbitrage and the Law of One Price." *Journal of International Economics*, May 1978, 8(2), pp. 341-51.

## **Blavy and Juvenal**

Sarno, Lucio; Taylor, Mark P. and Chowdhury, Ibrahim. "Nonlinear Dynamics in Deviations from the Law of One Price: A Broad-Based Empirical Study." *Journal of International Money and Finance*, February 2004, 23(1), pp. 1-25.

Sercu, Piet and Uppal Raman. "The Exchange Rate in the Presence of Transaction Costs: Implications for Tests of Purchasing Power Parity." *The Journal of Finance*, September 1995, 50(4), pp. 1309-319.

Taylor, Alan M. "Potential Pitfalls for the Purchasing-Power-Parity Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the Law of One Price." *Econometrica*, March 2001, 69(2), pp. 473-98.

Taylor, Mark P.; Peel, David A. and Sarno, Lucio. "Nonlinear Mean-Reversion in Real Exchange Rates: Towards a Solution to the Purchasing Power Parity Puzzles." *International Economic Review*, November 2001, 42(4), 1015-42.

Tong, Howell. *Nonlinear Time Series: A Dynamic System Approach*. Oxford, UK: Clarendon Press, July 1993.