Volume 8, Number 2 June 2012

IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING
Volume 8, Number 2
June 2012

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

IJCB International Journal of Central Banking

Volume 8, Number 2 June 2012

The contents of this journal, together with additional materials provided by article authors, are available without charge at ***www.ijcb.org****.*

The *International Journal of Central Banking* is published quarterly (ISSN: 1815-4654). Online access to the publication is available free of charge at **www.ijcb.org**. Individual print subscriptions are available. Orders may be placed by phone (001 415 974 2035), via fax (001 415 974 2168), or by e-mail (editor@ijcb.org).

Renewals, claims, address changes, and requests for permission to reprint material from this journal should be addressed to:

International Journal of Central Banking
Economic Research Department
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
USA

Phone: 001 415 974 2035
Fax: 001 415 974 2168
E-mail: editor@ijcb.org

ISSN: 1815-4654

Estimated Impact of the Federal Reserve's Mortgage-Backed Securities Purchase Program*

Johannes Stroebel and John B. Taylor
Stanford University

The largest credit or liquidity program created by the Federal Reserve during the financial crisis was the mortgage-backed securities (MBS) purchase program. In this paper, we examine the quantitative impact of this program on mortgage interest rate spreads. This is more difficult than frequently perceived because of simultaneous changes in prepayment risk and default risk. Our empirical results attribute a sizable portion of the decline in mortgage rates to such risks and a relatively small and uncertain portion to the program. For specifications where the existence or announcement of the program appears to have lowered spreads, we find no separate effect of the stock of MBS purchased by the Federal Reserve.

JEL Codes: E52, E58, G01.

1. Introduction

As part of its response to the financial crisis, the Federal Reserve introduced a host of new credit and liquidity programs in 2008 and 2009. The largest of the new programs was the mortgage-backed securities (MBS) purchase program. This program was part of a quantitative easing or credit easing policy which replaced the usual tool of monetary policy—the federal funds rate—when it hit the lower bound of zero. The mortgage-backed securities that the Federal Reserve purchased were guaranteed by Fannie Mae and Freddie

*We would like to thank Jim Dignan, Darrell Duffie, Peter Frederico, Frank Nothaft, Eric Pellicciaro, Josie Smith, and two anonymous referees for helpful comments. Author e-mails: JohnBTaylor@stanford.edu and stroebel@stanford.edu.

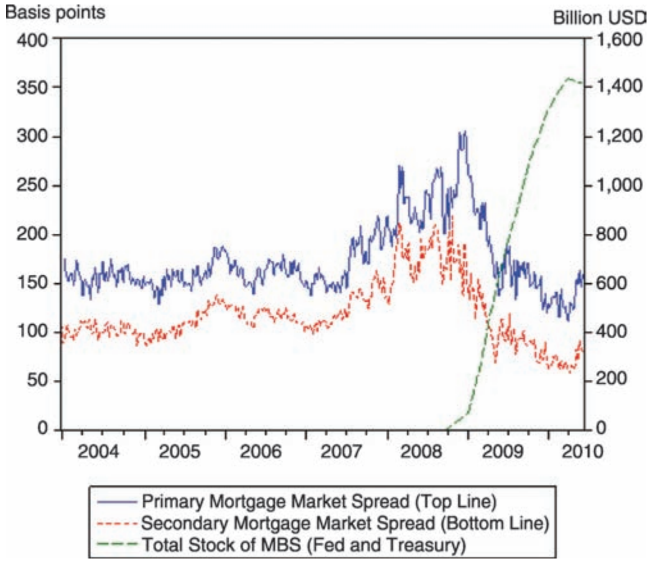
Mac, the two government-sponsored enterprises (GSEs) with this role, as well as by Ginnie Mae, the U.S. government-owned corporation within the Department of Housing and Urban Development. The program was set up with an initial limit of $500 billion but was later expanded to $1.25 trillion. It expired on March 31, 2010. The Federal Reserve also created a program to buy GSE debt—initially up to $100 billion and later expanded to $200 billion—and a program to purchase $300 billion of medium-term Treasury securities. The Federal Reserve's MBS purchases came on top of an earlier-announced MBS purchase program by the Treasury.

These programs were introduced with the explicit aim of reducing mortgage interest rates.[1] Figure 1 shows both primary and secondary mortgage interest rate spreads over Treasury yields during the financial crisis. Primary mortgage rates are the rates that are paid by the individual borrower. They are based on the secondary market rate but also include a fee for the GSE insurance, a servicing spread to cover the cost of the mortgage servicer, and an originator spread. Observe that mortgage spreads over U.S. Treasuries started rising in 2007 and continued rising until late 2008, when they reached a peak and started to decline. By July 2009 they had returned to their long-run average, or to slightly below that average.

In this paper we consider to what degree the decline in spreads in 2009 can be attributed to the purchases of mortgage-backed securities by the Federal Reserve and the Treasury. This question is very important for deciding whether or not to use such programs in the future as a tool of monetary policy. Determining whether central banks have the ability to affect the pricing of mortgage securities for extended periods is also an important input into the debate about the role, responsibilities, and powers of central banks (see, for example, the collection of essays on this subject in Ciorciari and Taylor 2009), and we see this paper as part of a larger empirical analysis of quantitative easing, or credit easing, at central banks during the crisis.

[1]The press release on November 25, 2008 announcing the MBS purchase program stated that "this action is being taken to reduce the cost and increase the availability of credit for the purchase of houses, which in turn should support housing markets and foster improved conditions in financial markets more generally."

Figure 1. Mortgage Spreads and Stock of MBS Purchases



Notes: This figure shows the primary market mortgage spread, the secondary market mortgage spread, and the total stock of MBS purchases by the Federal Reserve and Treasury. The primary market mortgage rate series comes from Freddie Mac's Primary Mortgage Market Survey, which surveys lenders each week on the rates and points for their most popular thirty-year fixed-rate, fifteen-year fixed-rate, 5/1 hybrid amortizing adjustable-rate, and one-year amortizing adjustable-rate mortgage products. The secondary market mortgage series is the Fannie Mae thirty-year current-coupon MBS (Bloomberg ticker: MTGEFNCL.IND). The spreads are created by subtracting the yield on ten-year U.S. Treasuries from both series. The maturity difference between these series captures the fact that most thirty-year mortgages are paid off or refinanced before their maturity. We add MBS to the total stock when they are contracted and reported by the Federal Reserve Bank of New York, not when they appear on the Federal Reserve's balance sheet.

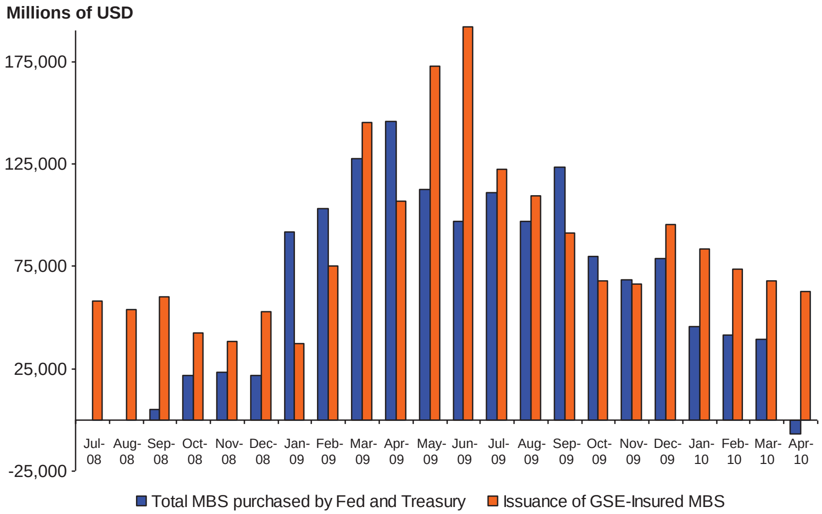
A common perception is that the MBS purchase program led to a significant reduction in mortgage rates. For example, early in the program, in January 2009, Ben Bernanke (2009) noted that "mortgage rates dropped significantly on the announcement of this program and have fallen further since it went into operation." Later, in December 2009, Brian Sack (2009) of the Federal Reserve Bank of New York reiterated that view. Figure 1 shows that the decline in the mortgage interest rate spread was contemporaneous to the expansion of the MBS purchase program.[2] Some also cite the large fraction of new agency-insured MBS issuance that the Federal Reserve has purchased each month since the start of the purchase program.[3] Figure 2 shows that MBS purchases in 2009 were up to 200 percent of new issuance of GSE-insured debt, and a significantly larger fraction of net issuance.

In our view, however, an evaluation of the program's impact requires an econometric analysis that controls for influences other than the MBS purchase program on mortgage spreads. In particular, any coherent story that links the decline in mortgage interest rates to the purchases of MBS by the Federal Reserve also needs to explain why mortgage spreads increased so dramatically between 2007 and late 2008, and consider whether those same factors may be responsible for at least part of the subsequent decline in 2009. It is conceivable that precisely those indicators of risk in mortgage lending that drove up mortgage spreads through 2007 and 2008 relaxed

[2]MBS purchases are primarily made in the "to be announced" (TBA) market in which the pool identity is unknown at the time of the purchase. The TBA contract defines the MBS that will be delivered only by the average maturity and coupon of the underlying mortgage pool, and by the GSE backing the MBS. For example, an investor might purchase $1 million worth of 8 percent, thirty-year Fannie Maes for delivery next month. Precise pool information is then "to be announced" forty-eight hours prior to the established trade settlement. This allows a lender to lock in the rate they can offer the mortgage borrowers by preselling their loans to investors, and thus to fund their origination pipeline. For more details on this market, see Boudoukh et al. (1999). The Federal Reserve Bank of New York announced MBS purchases when they contracted to buy; the Federal Reserve placed the MBS on its balance sheet (reported in the H.41 release) when the contract settled. This explains why at the end of the MBS purchase program, on March 31, 2010, the Federal Reserve had just over $1 trillion of MBS on its balance sheet, rather than $1.25 trillion, which is the overall size of the program. In this paper we record the volume of purchases when they are contracted and reported by the Federal Reserve Bank of New York, not when they appear on the Federal Reserve's balance sheet. A robustness check has shown that this does not affect our conclusions.

[3]This point was also made by Sack (2009): "How has the Federal Reserve been able to generate these substantial effects on longer-term interest rates? One word: size. The total amount of securities to be purchased under the LSAPs is quite large relative to the size of the relevant markets. That is particularly the case for mortgage-backed securities. Federal Reserve purchases to date have run at more than two times the net issuance of securities in this market."

Figure 2. Monthly Flows of GSE-Insured MBS Issues and Shares Bought by the Federal Reserve and Treasury



Note: This figure shows the monthly purchases of MBS by the Federal Reserve and Treasury, as well as the total monthly issuance of GSE-insured MBS.

throughout the first half of 2009, providing a coherent theory for both the rise and the subsequent fall of mortgage spreads, without a large role for the Federal Reserve's purchases. While identifying the effects of the Federal Reserve's MBS purchases is complicated by the many unusual developments in financial markets between 2007 and 2009, we attempt to address the issue empirically using statistical methods and available data.

A number of other recent papers have considered the effect of large-scale asset purchase (LSAP) programs since the initial publication of our results.[4] Gagnon et al. (2011) examine the cumulative effect of eight different announcements related to long-term asset

[4]Our original estimates of the impact of the MBS purchases on mortgage spreads were performed in "real time" while the Federal Reserve was still making purchases under the MBS program and were presented briefly in preliminary form in the NBER Feldstein lecture in July 2009 and circulated in December 2009 as NBER Working Paper No. 15626.

programs, including the MBS purchases. They find the current-coupon thirty-year agency MBS yield to decline by a total of 113 basis points (recall that we are considering spreads, not yields). This approach assumes that the markets correctly and completely price in the information contained in the announcement within the one-day window of the baseline analysis. We consider our analysis to be complimentary to approaches looking at announcement effects. Hancock and Passmore (2011) examine whether the MBS purchase program lowered mortgage rates, and conclude that the program's announcement reduced mortgage rates by about 85 basis points in the month following the announcement, and that it contributed an additional 50 basis points towards lowering risk premiums once the program had started. Fuster and Willen (2010) consider the movement of prices as well as quantities around the announcement of the MBS purchase program. They argue that the number of mortgage applications for refinancing increased around the announcement of the program. They find no effect of the announcement on the search and application for purchase mortgages. In addition, they use a high-frequency data set of loan offers to show that there was a wide variation in the effect of the announcement on mortgage rates. In particular, they detect a range from a fall of 40 basis points to an increase of 10 basis points. Aït-Sahalia et al. (2010) look at announcement effects of programs at a number of central banks. Duygan-Bump et al. (2010) examine short-term liquidity facilities.

In the next section we discuss the theory of the valuation of MBS, and we explain how the option-adjusted spread (OAS) can be used to control for the prepayment risk inherent in MBS valuation. We also discuss our approach to controlling for the default risk of the underlying mortgages. We then report our empirical results. We show that a sizable portion of the decline in mortgage spreads can be attributed to declines in default risk and prepayment risk, a result which is robust to alternative measures of mortgage spreads, including other OAS series and simple spreads between mortgage rates and Treasuries or interest rate swaps. We then show that the estimated size of the impact of the MBS program on mortgage spreads is sensitive to which interest rate the spread is measured relative to. We explore the reason for this difference and find that it can be traced to a shift in the spread between Treasuries and swaps which occurred around the time of the panic in October 2008.[5]

2. Valuing Mortgage-Backed Securities

Mortgage-backed securities are structured financial products that are secured by a collection of mortgages, most commonly on residential property. Mortgage loans made by individual lenders are assembled into pools by Fannie Mae, Freddie Mac, Ginnie Mae, or private entities. Mortgage-backed securities then represent a claim on the principal and interest of the mortgage loans in the pool. The Federal Reserve's MBS purchases concentrated on the market for mortgage-backed securities assembled and insured by Fannie Mae, Freddie Mac, and Ginnie Mae. These institutions guarantee the timely payment of principal and interest of those mortgage-backed securities, even if the underlying mortgages default (Passmore 2005).

Basic finance theory suggests that for given intertemporal preferences, there are two key determinants of the spread of mortgage-backed securities over the risk-free rate. These determinants are the prepayment risk and the default risk of the MBS. Most of the mortgages that collateralize a mortgage-backed security entail a prepayment option for the individual borrower, which gives the borrower the right to prepay the mortgage at any time prior to the maturity of the loan, and thereby to refinance at a favorable rate. This prepayment option gives mortgage-backed securities characteristics similar to those of a callable bond in which the issuer has the right to redeem prior to its maturity date (Windas 1996). In the case of MBS, when interest rates decline, mortgage holders might choose to prepay their mortgage and refinance at lower rates. This terminates the investors' source of above-market returns and requires them to reinvest at the lower prevailing rates. To compensate an investor for the presence of this prepayment risk, coupon payments on MBS must be adjusted upwards. To determine how much of the observed

[5] If the MBS program also lowered Treasuries or swaps, then the effects of the purchase program on mortgage rates could be larger than what we detect. Sack (2009) stresses that "a primary channel through which this effect takes place is by narrowing the risk premiums on the assets being purchased." But he also states that the effect "would be expected to spill over into other assets that are similar in nature."

fall in mortgage spreads can possibly be attributed to a decline in prepayment risk, we use several different option-adjusted spreads, which we explain below in section 2.1.

A second key determinant of MBS pricing is the default risk of those securities. Falling house prices, rising foreclosures, and increasing inventory in the housing market all contribute to a higher default probability for the underlying mortgages. As we described above, the Federal Reserve's purchases were limited to the market of mortgage-backed securities guaranteed by Fannie Mae, Freddie Mac, and Ginnie Mae. The default risk of these agency-insured MBS is thus not only affected by the default risk of the underlying mortgages but also by the perceived probability that the insuring entity will not be able to fulfill its insurance pledge. While Ginnie Mae securities are the only MBS that carry the full faith and credit guarantee of the U.S. government, many market participants believed that there was also an "implicit guarantee" for the MBS guaranteed by Fannie Mae and Freddie Mac. In section 2.2 we discuss a number of approaches we take to control for the default risk of agency-insured MBS.

2.1 Controlling for Prepayment Risk

As discussed above, the individual mortgage borrower usually has the option to prepay the mortgage in part or in full at any time prior to its maturity. Provisions allowing for borrower prepayment prior to the maturity of a loan are referred to as embedded options. To compensate the investor for the presence of this prepayment risk, coupon payments on MBS must be adjusted upwards. Pricing of a mortgage-backed security thus proceeds by modeling it as a combination of (i) a long position in a non-callable bond and (ii) a short position in a prepayment option. The combined valuation of those two parts determines the secondary market yield of MBS (for a discussion of the extent to which the option approach can explain default and prepayment behavior, see Deng, Quigley, and Van Order 2000).

The option-adjusted spread (OAS) is a natural way to control for these prepayment risks. It is calculated by considering the average discounted cash flow from the MBS along a number of possible interest rate scenarios (below we discuss how these scenarios are generated). To define the OAS, let r_{it} represent the short-term interest rate at time t under scenario i and let C_{it} represent the cash flow from the mortgage-backed security at time t under scenario i. (Note that the cash flow path depends on the interest rate path, as discussed below). The present value of the cash flows for each scenario i is given by

$$PV_i = \sum_{k=1}^{K} \frac{C_{ik}}{\prod_{j=1}^{k}(1 + r_{ij})}. \tag{1}$$

Hence, the theoretical value P_E of the MBS is equal to the probability-weighted average of the PVs of each scenario. Let $w(i)$ be the probability of each interest rate and cash flow scenario.[6] Then

$$P_E = \sum_{i=1}^{N} w(i) PV_i \quad \text{s.t.} \sum_{i=1}^{N} w(i) = 1. \tag{2}$$

If we denote the market value of an MBS by P_M, then the OAS is defined as the θ such that

$$P_M = \sum_{i=1}^{N} w(i) \left[\sum_{k=1}^{K} \frac{C_{ik}}{\prod_{j=1}^{k}(1 + r_{ij} + \theta)} \right]. \tag{3}$$

Thus, the OAS is the spread—over a term structure of interest rates—that equates the market price of the MBS to the probability-weighted average discounted present value of expected cash flows along a number of possible simulated future interest rates paths. In other words, the OAS is the number of basis points θ that the discount curve needs to be adjusted upwards until the theoretical price calculated using the "adjusted term structure" matches the market price of the security.

It is common to use an interest rate model based on the LIBOR swap curve[7] for the projection of r_{it}, in which case the OAS is

[6]If these interest rate scenarios are drawn using Monte Carlo methods, then each scenario would have an equal likelihood, and $w(i) = 1/N$ for all i.

[7]By the LIBOR swap curve, we mean the swap rate as a function of the maturity of the interest rate swap. The swap rate is the rate paid by a fixed-rate payer in return for receiving floating-rate three-month LIBOR rolled over during the maturity period of the swap. To emphasize that LIBOR is the floating-rate side of the interest rates swaps in this paper, we sometimes use the term LIBOR swap.

referred to as the *swap-OAS*. LIBOR is the most appropriate discount rate for most financial market participants who balance mortgage investments with other non-government investments. LIBOR thus provides a measure of the opportunity cost of most investors. Fabozzi and Mann (2001) argue that "funded investors use LIBOR as their benchmark interest rate. Most funded investors borrow at a spread over LIBOR. Consequently, if the LIBOR swap curve is used as the benchmark interest rate, the OAS reflects a spread relative to their funding costs."[8]

To make the OAS operational, multiple paths of future interest rates must be generated using a model of interest rates. The cash flows from the underlying mortgages can then be calculated using the generated interest rates. Three swap-OAS series are used in this paper. We first focus on a Bloomberg series which is widely used by market participants. The interest rate path and cash flow path for this series are calculated using the Bloomberg "two-factor interest rate" and "prepayment" models. We show that the results are robust to using two other swap-OAS series which are based on different models (one by Barclays Capital, the other by Deutsche Bank). The results from using these series are very similar to those obtained using the Bloomberg series. The OAS can also be calculated using an interest rate model based on Treasury securities rather than LIBOR, and we also consider this alternative measure, which we call *Treasury-OAS* in our analysis. However, Treasury rates and interest rate swap rates behaved quite differently during the financial crisis, and some of the results are different for this alternative, as we discuss below.

The interest rate model used to compute the Bloomberg OAS series, which is described in Belikoff et al. (2010), is a time-series model which assumes no-arbitrage conditions on the term structure of interest rates. For the swap-OAS, the no-arbitrage conditions are imposed using the LIBOR swap curve, and for the Treasury-OAS, the no-arbitrage conditions are imposed using the constant-maturity

[8] Belikoff et al. (2010) also address this issue. They argue that using the LIBOR swap curve has the additional advantage that "the swap market is quoted more uniformly and more densely [than the Treasury market]," which helps with calibrating the interest rate model used to determine the OAS. Consequently "the mortgage market has evolved to value securities relative to the swap market."

Treasury (CMT) curve. Brigo and Mercurio (2006) discuss the value of using more than one factor in such time-series models for the interest rates as well as the rationale for imposing no-arbitrage conditions. Rudebusch (2010) discusses the benefit of adding macro variables to these interest rate models.

The prepayment model, also described in Belikoff et al. (2010), takes into account refinancing as well as housing turnover, curtailment (when the debtor elects to pay more than the required mortgage payment), and default. Refinancing is the major interest-rate-dependent component. Prepayment increases when interest rates are low relative to the MBS's coupon, but it is also affected by credit quality (borrowers with poor credit are less able to refinance), a "media effect" (prepayment jumps when rates hit historic lows), and a "burnout effect" (pools that have experienced substantial prepayment are less likely to prepay in the future, since those members who are most likely to prepay have been removed). Housing turnover is modeled with a seasonally adjusted turnover rate modified by a lock-in effect in which housing turnover is reduced when it is more expensive to close out an existing mortgage. Further adjustments are made to account for the fact that prepayments first tend to increase and then level off over time.[9]

Figure 3 shows the Bloomberg swap-OAS, the Treasury-OAS, and the primary mortgage spread over Treasuries.[10] The gap between the OAS series and the primary mortgage spread partially captures changes in prepayment risk. The gap between the swap-OAS and the Treasury-OAS is driven by movements of the Treasury term structure relative to the swap-curve term structure, as we discuss further below.

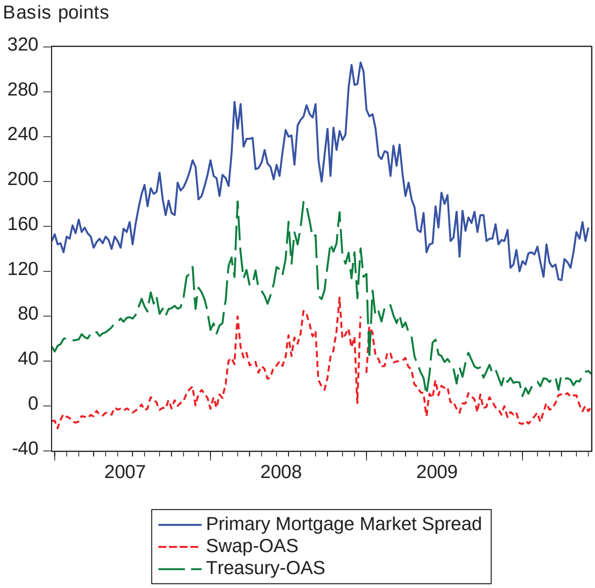
2.2 Controlling for Default Risk

While the prepayment models underlying the OAS endeavor to control for prepayment risk, they do not control for default risk.

[9] More details on the computation of option-adjusted spreads can be found in Kupiec and Kah (1999) and in Windas (1996).

[10] In particular, the swap-OAS is the NOASFNCL.IND series and the Treasury-OAS is the MOASFNCL.IND series in Bloomberg. Both series capture the OAS of Fannie Mae thirty-year current-coupon MBS, and are used widely by market participants. The swap-OAS uses the S23 swap curve and the Treasury-OAS uses the constant-maturity Treasury curve.

Figure 3. Option-Adjusted Spread and Primary Mortgage Market Spread



Notes: This figure shows the primary market mortgage spread over ten-year U.S. Treasuries. The primary market mortgage rate series comes from Freddie Mac's Primary Mortgage Market Survey, which surveys lenders each week on the rates and points for their most popular thirty-year fixed-rate, fifteen-year fixed-rate, 5/1 hybrid amortizing adjustable-rate, and one-year amortizing adjustable-rate mortgage products. The figure also shows the swap-OAS for Fannie Mae securities (Bloomberg ticker: NOASFNCL.IND) and the Treasury-OAS for Fannie Mae securities (Bloomberg ticker: MOASFNCL.IND).

Controlling for the default risk of agency-insured MBS is necessary to ensure that the decline in spreads in the OAS in 2009 was not driven by a decline in the default risk of the underlying securities. Finding a good, uncontaminated measure for default risk, however, is not easy. In the case of agency-insured MBS, the default risk is not only related to the default risk of the underlying mortgages but also to the potential of the insuring agency being unable to meet its guarantee obligations. The ability to fulfill such an insurance pledge is a function of the health of the housing market and of a number of political factors that determine whether the government would eventually act as a backstop to the guarantees. This uncertainty is more relevant for the GSEs Fannie Mae and Freddie Mac than for Ginnie Mae, which has the full faith and credit backing of the U.S. government.

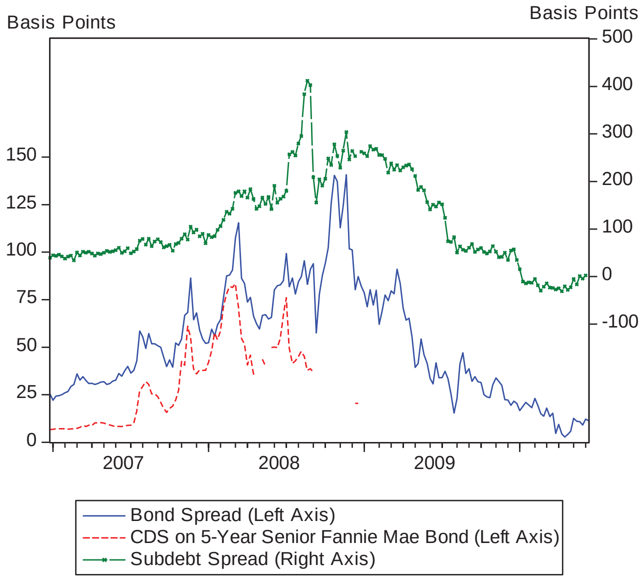
A good measure of the default risk of GSE-insured MBS is the credit default swap (CDS) series on the debt issued by the insuring institutions. When there is an increased risk of default of GSE debt, as measured by higher costs for CDS on that debt, the risk that the GSEs will not be able to fulfill their insurance pledge increases. Consequently, secondary market spreads on agency-insured MBS will increase. Unfortunately, placing Fannie Mae and Freddie Mac into conservatorship on September 7, 2008 was a trigger event for outstanding CDS, so the data series stops at that time. To our knowledge, no new CDS series have emerged since then that would allow us to directly measure GSE default risk.

An alternative proxy for the default risk of Fannie and Freddie is the spread between GSE debt and U.S. Treasury securities. One such series is the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities.[11] There are some concerns that such bond spread series might be picking up liquidity effects as well as changes in default risk (Krishnamurthy 2010). Figure 4 shows this bond spread series together with the associated CDS series prior to its discontinuation. For the time period that the two series coexist, they are highly correlated, which suggests that the bond spread series does pick up changes in the default risk of GSE-insured MBS, and does not just capture liquidity or other effects. Another complicating factor, however, is that in late 2008 the Federal Reserve also embarked on a program to purchase agency debt. While these interventions capture a much smaller fraction of the market than the purchases of agency-insured MBS, they may contaminate the usefulness of bond spreads as a pure measure of agency default risk during this period. To deal with this problem we take two approaches.

First, we instrument for the bond spread series with three instrumental variables: the level of the Case-Shiller house-price index, the

[11]This series is available with the Bloomberg ticker FNMGVN5.IND.

Figure 4. GSE Bond Spreads, Subdebt Spreads, and GSE CDS



Notes: The solid line in this figure shows the spread of five-year Fannie Mae bonds to U.S. Treasury active (on-the-run) securities, given by Bloomberg ticker FNMGVN5.IND. The dashed line shows the CDS series on senior five-year Fannie Mae bonds, given by Bloomberg ticker FNMA 5YR CDS SR Index. The solid line and the dashed line are plotted on the left axis. The solid crossed line shows the development of the spread between a Fannie Mae subordinated debt series (FNMA 4.625 05/01/13) and five-year Treasuries, and is plotted on the right axis.

change in this index, and Moody's AAA bond index.[12] We interpolate the monthly Case-Shiller index data to get weekly observations. A lower level of the house-price index and a large decline in the index should indicate a higher degree of mortgage default risk. Falling house prices will push borrowers into negative home equity,

[12]Moody's Long-Term Corporate Bond Yield Averages are derived from pricing data on a regularly replenished population of corporate bonds in the U.S. market, each with current outstandings over $100 million. The bonds have maturities as close as possible to thirty years; they are dropped from the list if their remaining life falls below twenty years, if they are susceptible to redemption, or if their ratings change.

increasing their incentives for strategic default, and thus increasing the risk of mortgage default. The Moody's AAA bond index captures the general degree of riskiness in the credit markets. Because these instruments are unlikely to be affected by Federal Reserve purchases of GSE debt or MBS and are highly correlated with the bond spread (the first-stage regression has an F-statistic of 99.92), they are good instruments in our view.[13] In addition, beyond its effect through capturing increased risk in the housing credit market, neither of the instruments should have a significant effect on the default probability of GSE debt. Thus the exclusion restrictions are likely to be met. We also ran robustness checks which use the CDS spreads for Bank of America, Wells Fargo, Citigroup, and JP Morgan—four large mortgage lenders in the United States—to instrument for the GSE debt spread, in place of the Moody's AAA index. The results are very similar to those reported with the Moody's AAA index as the instrument.[14]

In a second approach, we use the spreads of Fannie Mae's Subordinated Benchmark Note series to proxy for credit risk. Since the Federal Reserve's GSE debt purchases were focused on the senior debt market, they are less likely to have contaminated this subordinated debt as a proxy of risk. Fannie Mae started issuing subordinated debt in 2001, with the expressed goal of "enhancing market discipline, transparency and capital adequacy." The subordinated debt series is unsecured and ranks junior in priority of payment to all senior creditors, so "the price is sensitive to how the market views our [Fannie's] financial situation" (Fannie Mae 2001). Since MBS guarantees rank pari passu to senior bonds, the subordinated debt will only be repaid if the MBS insurances issued are fulfilled. This means that an increase in the subordinated debt spreads should

[13]One may be concerned that since the Moody's AAA index contains corporate debt, which did not suffer as much during the crisis, it will not pick up the adequate default risk. In addition, there might also be concerns that it could be affected through a portfolio-balance channel. While we do not believe that this is very likely, a robustness check of our results, in which we drop the index from our list of instruments, shows that the inclusion of the index does not affect our conclusions materially.

[14]The results are available on request. The series are CDS series on five-year senior bonds. They have the following Bloomberg tickers: BOFA CDS USD SR 5Y Corp, WELLFARGO CDS USD SR 5Y Corp, CINC CDS USD SR 5Y Corp, and JPMCC CDS USD SR 5Y Corp.

signal an increase in the probability of default for the GSE-insured MBS. The downside of looking at the subordinated debt series is its very small volume, which is usually around $1 billion per issuance and not comparable in liquidity to the senior GSE bonds. Therefore, the pricing of these securities may conflate liquidity elements with credit risk elements. Figure 4 also compares the development of the bond spread series with the subordinated debt spread series[15]—it appears as if the subordinated debt spread series moves more dramatically, especially in the period running up to the conservatorship of Fannie and Freddie, and may thus be more able to pick up changes in default risk.[16]

3. Empirical Analysis

In reporting our results, we first consider the swap-OAS and the simple spread of MBS yields over swap rates. Second, we consider the Treasury-OAS and the simple spread of MBS yields over Treasury yields. Third, we discuss shifts in the swap spread (the spread between swap rates and Treasury rates of the same maturity) that can help to understand the differences in the detected impact of the MBS program on the swap-OAS and Treasury-OAS.

3.1 Spreads over LIBOR Swaps

In the basic model, the option-adjusted spread is a function of the various measures of default risk discussed in section 2.2 (interest rate spreads on Fannie Mae senior debt as well as Fannie Mae subordinated debt, both with and without instruments) and the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market (about $5 trillion). Both the OAS and the other spreads are measured in basis points (1/100 of a percentage point). The observations are at a weekly frequency. For

[15]Spread of the Fannie Mae Subordinated Benchmark series, Maturity on 5.1.2013, Volume: $1 billion (Bloomberg ticker: FNMA 4.625 05/01/13 Corp) over five-year Treasury.

[16]It is possible that during this time period, and the event surrounding the conservatorship, there were changes in both the probability of default and the expected loss given default, which might affect CDS spreads and bond spreads differentially.

Table 1. Fannie Mae Swap-OAS Regressions with GSE Bond Spreads

	(1) OLS	(2) IV	(3) OLS	(4) IV
Bond Spread	0.87***	0.93***	0.43***	0.39***
	(0.04)	(0.05)	(0.06)	(0.09)
Total MBS Purchases	67.81***	60.15***	34.34***	20.71*
	(11.74)	(12.81)	(10.02)	(12.28)
OAS (t−1)			0.54***	0.57***
			(0.05)	(0.08)
Number of Observations	179	169	178	168
R^2	0.75	0.77	0.84	0.84

Notes: This table shows the results from regression (4). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the swap-OAS (Bloomberg ticker: NOASFNCL.IND). The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In columns 2 and 4 we instrument for the bond spread series with (i) the level of the Case-Shiller house-price index, (ii) the month-on-month change in this index, and (iii) the Moody's AAA bond index. Standard errors are in parentheses. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

higher-frequency series, we take the average of the observations for that week. The estimation period is the beginning of 2007 through June 2010.

In table 1 we report the regression results which show the impact of MBS purchases on the swap-OAS using the GSE debt spread as the control variable. We ran regression equation (4), with the swap-OAS as the dependent variable:

$$OAS_t = \alpha + \beta_1 * GSE\_Spread_t + \beta_2 * Total\_MBS\_Purchases_t + \varepsilon_t. \tag{4}$$

Recall that we do not need to proxy for prepayment risk, since this is already removed from the OAS series. In columns 1 and 2 we show the OLS and instrumental variable results. In columns 3 and 4 we also include the lagged value of the OAS series, to allow for

Table 2. Fannie Mae Swap-OAS Regressions with Subordinated Debt Spreads

	(1) OLS	(2) IV	(3) OLS	(4) IV
Bond Spread	0.24***	0.25***	0.08***	0.09***
	(0.01)	(0.01)	(0.02)	(0.03)
Total MBS Purchases	22.81**	7.81	3.84	−1.32
	(10.85)	(11.46)	(8.56)	(9.49)
OAS ($t-1$)			0.67***	0.63***
			(0.06)	(0.09)
Number of Observations	178	168	177	167
R^2	0.72	0.75	0.84	0.84

Notes: This table shows the results from regression (4) with the subordinated debt spread replacing the GSE bond spread. The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the swap-OAS (Bloomberg ticker: NOASFNCL.IND). The "Bond Spread" control variable captures the spread between Fannie Mae's Subordinated Benchmark Note series and five-year Treasuries (Bloomberg ticker: FNMA 4.625 05/01/13 Corp). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In columns 2 and 4 we instrument for the bond spread series with (i) the level of the Case-Shiller house-price index, (ii) the month-on-month change in this index, and (iii) the Moody's AAA bond index. Standard errors are in parentheses. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

the possibility that the effects of the purchases are distributed over time.

Observe in table 1 that the OAS moves closely with the bond spread, just as theory would predict: The OAS increases when the perceived probability of default increases. However, as measured by the coefficient on the MBS purchase volume, the impact of the purchases on the OAS was either significantly positive or insignificantly different from zero. In this specification there is no evidence that the increase in the MBS purchases led to a reduction in mortgage interest rate spreads as measured by this conventional OAS measure, once changes in default risk are controlled for.

Table 2 is analogous to table 1 except that we control for default risk using the subordinated debt series rather than the GSE debt spread. Again, the coefficients on the total volume of MBS purchased are either positive or statistically insignificantly different from zero.

Another possible specification includes a dummy variable for whether or not there was an MBS purchase program along with the variable for the volume of purchases:

$$OAS_t = \alpha + \beta_1 * GSE\_Spread_t + \beta_2 * Total\_MBS\_Purchases_t + \beta_3 * I_{\{Program\ Event\},t} + \varepsilon_t. \quad (5)$$

The results are shown in table 3, which reports the effects of four different "program event" dummy variables. In each regression the dummy is set to 0 at the start of the sample period and then increased to 1 at a later date. In column 1 the dummy is set to 1 starting in September 2008, when the Treasury started buying MBS and Fannie and Freddie were taken into government conservatorship. In column 2 the dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 we also consider the impact of the announcement of the program's expansion by the Federal Reserve on March 18, 2009. On this date it was announced that the Federal Reserve would more than double the size of the program, from $500 billion to $1.25 trillion.

The estimated coefficients in table 3 do not indicate that either the program's existence or the volume of purchases had a statistically significant negative effect on the swap-OAS. The coefficients are insignificant or positive.

To understand why the MBS program's volume or existence do not pick up an effect on the swap-OAS, it is useful to consider the residuals of regressions of the OAS series on the bond spread series (the risk indicator), without including the MBS explanatory variables. Figure 5 shows the residuals from such a regression for the swap-OAS series, along with the actual and predicted swap-OAS series over the sample period. Notice that the residuals through this whole period remain fairly evenly spread around zero. Movements in prepayment risk (as measured by swap-OAS) and default risk (as

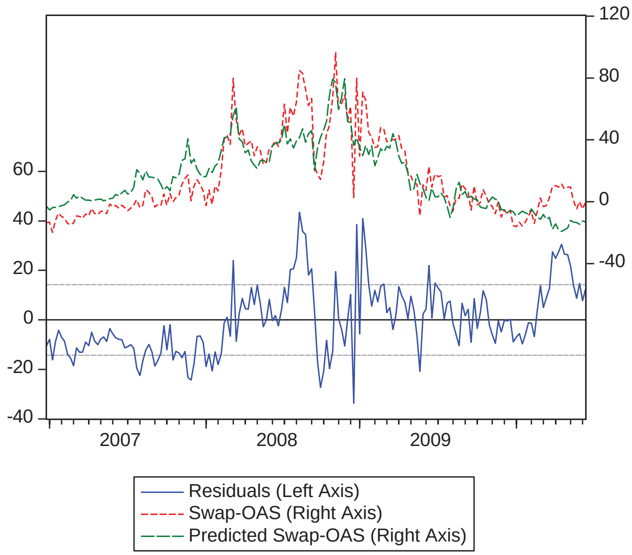
Table 3. Fannie Mae Swap-OAS with Program Dummies

	(1) Swap-OAS (Bloomberg)	(2) Swap-OAS (Bloomberg)	(3) Swap-OAS (Bloomberg)	(4) Swap-OAS (Bloomberg)
Bond Spread	0.88***	0.85***	0.85***	0.84***
	(0.05)	(0.04)	(0.04)	(0.04)
Total MBS Purchases	69.73***	24.58	31.22	63.47**
	(20.46)	(20.21)	(19.13)	(25.67)
MBS Treasury Dummy	−0.39			
	(3.45)			
MBS Federal Reserve Dummy		10.11***		
		(3.88)		
MBS Federal Reserve Announce			8.41**	11.14***
			(3.50)	(3.78)
MBS Federal Reserve Announce Expansion				−10.57*
				(5.66)
Number of Observations	179	179	179	179
R^2	0.75	0.76	0.75	0.76

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the swap-OAS (Bloomberg ticker: NOASFNCL.IND). The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

measured by agency debt spreads) account for the major movements in mortgage spreads. Little remains to be explained by the MBS purchases. This is the reason why the coefficient on MBS purchases is very small in the swap-OAS regressions.

Figure 5. Residual Analysis of Swap-OAS

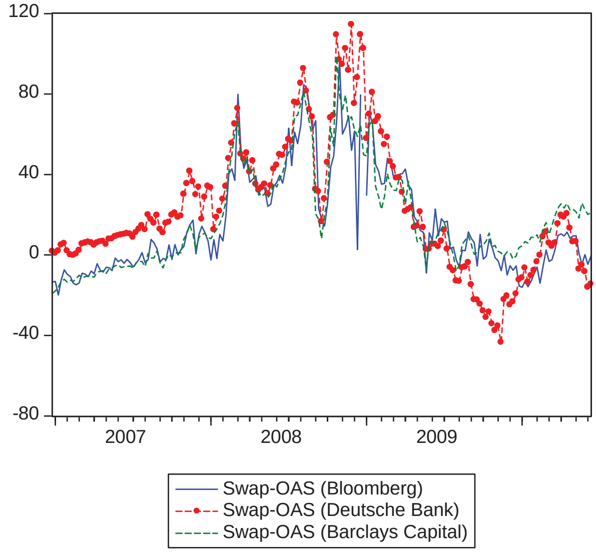


Notes: The swap-OAS line is the Bloomberg series NOASFNCL.IND. The predicted swap-OAS line shows the predicted values of a regression: $OAS_t = \alpha + \beta_1 * GSE\_Spread_t + \varepsilon_t$, where the GSE_Spread series is given by the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg series: FNMGVN5.IND). The residual series plots the residuals from the regression.

One possible concern with the analyses presented above is that the option-adjusted spread relies on the quality of the Bloomberg prepayment and interest rate models used to construct the OAS.[17] Below we present two robustness checks to the previous analysis. In the first robustness check we use two alternative swap-OAS series

[17] One might be concerned about this since during the crisis OAS calculations became particularly difficult to perform. Falling house prices and resulting negative home equity lowered the probability of refinancing, as did the tightening of lending standards and the market exit of a number of important mortgage lenders. This suggests that models that were not adequately updated would likely overstate the value of the prepayment option. In addition, during the crisis the dynamics of the swap curve and the Treasury curve might have changed, complicating the use of interest rate models that were calibrated to the pre-crisis economy.

Figure 6. Comparing Three Sources of Swap-OAS Spread



Notes: This figure compares the three swap-OAS series. The solid line represents an OAS series for Fannie Mae thirty-year current-coupon MBS computed by Bloomberg (Bloomberg ticker: NOASFNCL.IND). The dashed line with circles represents a swap-OAS series for a portfolio of agency MBS compiled by Deutsche Bank (Bloomberg ticker: Deutsche DBIQ US TBA MBS OAS Libor). The dashed line represents an OAS series for Fannie Mae thirty-year current-coupon MBS computed by Barclays Capital (retrieved through Barclays Live).

as dependent variables to remove prepayment risk. These alternative series were constructed using different interest rate and prepayment models. We use (i) an OAS series for thirty-year Fannie Mae current-coupon MBS computed by Barclays Capital and (ii) an OAS series for a monthly rebalanced index of agency MBS, compiled by Deutsche Bank (Bloomberg ticker: DBIQ US TBA MBS OAS Libor).[18] Figure 6 plots the two alternative OAS series and the Bloomberg OAS series. Up to the end of 2009, the Barclays and the Bloomberg OAS series co-moved very closely, suggesting that the

[18]This index includes MBS from Fannie Mae, Freddie Mac, and Ginnie Mae with durations of fifteen years or thirty years. It is described at https://index.db.com/htmlPages/MBS_Index_Guide_V.pdf.

models used by Bloomberg and Barclays Capital were rather similar (both series construct OAS for thirty-year Fannie Mae current-coupon MBS). However, during the last few months of the MBS program, the OAS series computed by Barclays rose significantly more, which implies that their model valued the prepayment option less than the Bloomberg model. The Deutsche Bank series does not capture the OAS of a single security, but of an index of MBS. Before the crisis, this OAS was higher than that of the thirty-year Fannie Mae current-coupon MBS. During the crisis, the co-movement with the MBS index provided by Bloomberg increased significantly.

In table 4 we repeat the key regressions from table 3 using the Barclays swap-OAS series. Notice that the coefficient on the total volume of MBS purchased by the Federal Reserve and Treasury is statistically significant and *positive*. This finding relative to the Bloomberg OAS-spread regressions is not surprising, since the Barclays OAS series increased at a faster rate than the Bloomberg OAS series during the first months of 2010, while the Federal Reserve continued to purchase more MBS during that period. The coefficients on the program announcements are negative; however, they are very small, and the net impact of the announcement effect and the volume effect is positive. Thus, results using this alternative OAS series do not provide any evidence that the Federal Reserve's MBS purchase program had an impact in reducing mortgage spreads, after controlling for prepayment risk and default risk.

In table 5 we present a similar set of regressions, using the OAS series provided by Deutsche Bank as the dependent variable. As before, none of the specifications suggest a significant reduction in option-adjusted spreads of agency MBS as a result of either the existence of the program or its volume, after we have controlled for changes in prepayment risk and default risk. This is highly consistent with the results found in table 4 using the Barclays OAS series.

A second set of robustness checks considers regressions where the secondary MBS market spread is the left-hand-side variable, without attempting to control for changes in prepayment risk. One interpretation of this is that the value of the prepayment option is assumed to be zero.[19] In addition, if the Federal Reserve's actions did

[19] Given that we argued in footnote 17 that model misspecification most likely led to an overvaluation of the prepayment option, this specification can provide a bound on the error resulting from valuing this option incorrectly.

Table 4. Fannie Mae Swap-OAS from Barclays Capital

	(1) Swap-OAS (Barclays)	(2) Swap-OAS (Barclays)	(3) Swap-OAS (Barclays)	(4) Swap-OAS (Barclays)
Bond Spread	0.97***	0.92***	0.90***	0.89***
	(0.05)	(0.04)	(0.04)	(0.04)
Total MBS Purchases	153.41***	150.74***	115.35***	195.62***
	(20.03)	(20.33)	(19.21)	(24.80)
MBS Treasury Dummy	−7.80**			
	(3.36)			
MBS Federal Reserve Dummy		−8.23**		
		(3.90)		
MBS Federal Reserve Announce			0.05	6.32*
			(3.49)	(3.54)
MBS Federal Reserve Announce Expansion				−25.79***
				(5.43)
Number of Observations	180	180	180	180
R^2	0.74	0.74	0.73	0.76

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the OAS of Fannie Mae thirty-year current-coupon MBS as computed by Barclays Capital. The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

Table 5. Swap-OAS from Deutsche Bank

	(1) Swap-OAS (Deutsche Bank)	(2) Swap-OAS (Deutsche Bank)	(3) Swap-OAS (Deutsche Bank)	(4) Swap-OAS (Deutsche Bank)
Bond Spread	0.84***	0.89***	0.84***	0.82***
	(0.07)	(0.06)	(0.06)	(0.04)
Total MBS Purchases	−38.93	2.74	−66.46***	83.34***
	(26.98)	(27.30)	(25.17)	(30.24)
MBS Treasury Dummy	3.88			
	(4.53)			
MBS Federal Reserve Dummy		−5.34		
		(5.24)		
MBS Federal Reserve Announce			10.65**	22.36***
			(4.57)	(4.33)
MBS Federal Reserve Announce Expansion				−48.14***
				(6.63)
Number of Observations	180	180	180	180
R^2	0.72	0.72	0.73	0.79

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the OAS for a monthly rebalanced index, compiled by Deutsche Bank, that tracks the MBS TBA Market (Bloomberg ticker: DBIQ US TBA MBS OAS Libor). The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

contribute to a decline in prepayment risk, we would like to measure this contribution to an overall decline in mortgage rates. In table 6 the specific dependent variable is the spread of the secondary market yield of thirty-year Fannie Mae MBS over the ten-year swap rate.

Table 6. Secondary Market Spread over Swap Rates

	(1) Secondary Market Spread	(2) Secondary Market Spread	(3) Secondary Market Spread	(4) Secondary Market Spread
Bond Spread	0.67***	0.74***	0.71***	0.71***
	(0.08)	(0.07)	(0.07)	(0.07)
Total MBS Purchases	−20.96	−49.65	−37.93	21.18
	(32.11)	(31.96)	(29.89)	(40.45)
MBS Treasury Dummy	12.89**			
	(5.40)			
MBS Federal Reserve Dummy		21.35***		
		(6.13)		
MBS Federal Reserve Announce			18.28***	22.89***
			(5.43)	(5.79)
MBS Federal Reserve Announce Expansion				−18.99**
				(8.86)
Number of Observations	180	180	180	180
R^2	0.51	0.53	0.53	0.54

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the spread of Fannie Mae thirty-year current-coupon MBS (Bloomberg ticker: MTGEFNCL.IND) over ten-year swap rates. The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

The results are consistent with the findings using the option-adjusted spreads: The program dummies are positive, and the volume of purchases is never statistically significant. Again, after attempting to control for default risk using the bond spreads, it does not appear

Table 7. Fannie Mae Treasury-OAS Regressions with GSE Bond Spreads

	(1) OLS	(2) IV	(3) OLS	(4) IV
Bond Spread	0.89***	0.80***	0.40***	0.24***
	(0.06)	(0.06)	(0.06)	(0.09)
Total MBS Purchases	−143.73***	−184.18***	−57.70***	−68.55***
	(15.75)	(17.44)	(14.94)	(16.95)
OAS (t−1)			0.58***	0.66***
			(0.06)	(0.07)
Number of Observations	180	170	180	170
R^2	0.83	0.83	0.89	0.88

Notes: This table shows the results from regression (4). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the Treasury-OAS (Bloomberg ticker: MOASFNCL.IND). The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In columns 2 and 4 we instrument for the bond spread series with (i) the level of the Case-Shiller house-price index, (ii) the month-on-month change in this index, and (iii) the Moody's AAA bond index. Standard errors are in parentheses. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

as if the program significantly lowered secondary market spreads of agency-insured MBS.

3.2 Spreads over Treasury Rates

Tables 7 and 8 consider the same regressions as tables 1 and 2 except that Treasury-OAS replaces swap-OAS as the dependent variable. Here the sign of the coefficient on the MBS purchase volume shifts from positive to negative and statistically significant, indicating that the purchases have a negative effect on the Treasury-OAS. According to the estimated regression coefficient in column 2, a purchase of $500 billion worth of MBS (approximately 10 percent of the market) is associated with a reduction in the Treasury-OAS of 18.4 basis points.

Table 8. Fannie Mae Treasury-OAS Regressions with Subordinated Debt Spreads

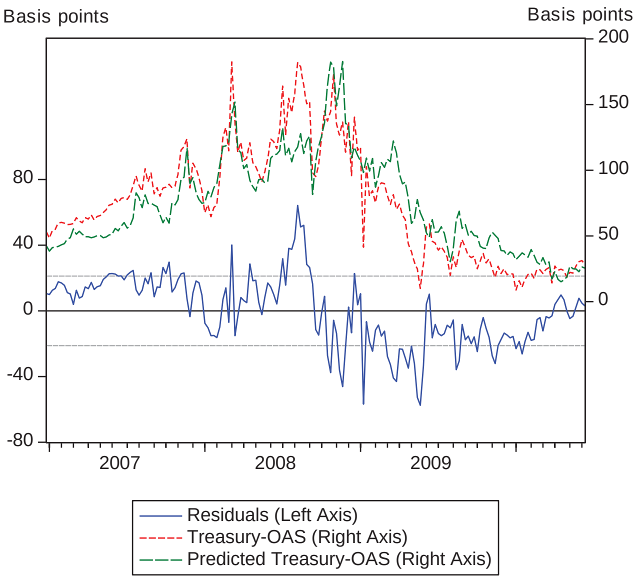
	(1) OLS	(2) IV	(3) OLS	(4) IV
Bond Spread	0.18***	0.17***	0.03*	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Total MBS Purchases	−216.02***	−247.82***	−56.03***	−62.89***
	(17.96)	(19.74)	(16.05)	(18.15)
OAS (t−1)			0.77***	0.79***
			(0.05)	(0.06)
Number of Observations	179	169	179	169
R^2	0.72	0.71	0.87	0.87

Notes: This table shows the results from regression (4). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the Treasury-OAS (Bloomberg ticker: MOASFNCL.IND). The "Bond Spread" control variable captures the spread between Fannie Mae's Subordinated Benchmark Note series and five-year Treasuries (Bloomberg ticker: FNMA 4.625 05/01/13 Corp). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In columns 2 and 4 we instrument for the bond spread series with (i) the level of the Case-Shiller house-price index, (ii) the month-on-month change in this index, and (iii) the Moody's AAA bond index. Standard errors are in parentheses. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

To better understand this estimated effect of the program, figure 7 (which is analogous to figure 5 for the swap-OAS) shows the residuals from the regression of the Treasury-OAS on the default risk indicator. Here we see that the residuals are below zero for almost all of 2009, which is what is being picked up by the MBS purchase coefficient. However, note that the residuals show little trend movement throughout 2009, as the Federal Reserve's and the Treasury's MBS stock continuously grew in size. If the actual volume of purchases was a partial driving factor, we would expect residuals to become significantly more negative over time, as purchases expanded.

Rather, it appears as if there was a single downward shift in residuals without a further effect from conducting actual purchases. This suggests that the specification that includes program dummies might be superior. Table 9 introduces the same program dummies as table 3, using the Treasury-OAS as the dependent variable. The

Figure 7. Residual Analysis of Treasury-OAS



Notes: The Treasury-OAS line is the Bloomberg series MOASFNCL.IND. The predicted swap-OAS line shows the predicted values of a regression: $OAS_t = \alpha + \beta_1 * GSE\_Spread_t + \varepsilon_t$, where the GSE_Spread series is given by the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg series: FNMGVN5.IND). The residual series plots the residuals from the regression.

actual volume of the MBS held by the Federal Reserve and Treasury appears to have no statistically significant effect on the Treasury-OAS. However, unlike the swap-OAS regressions, the coefficient on the dummy variables in these regressions indicates an effect of the existence or the announcement of the MBS purchase program. The estimated coefficients imply a negative effect of about 30 basis points on the Treasury-OAS.

To examine the robustness of this finding, we looked at the secondary market spread over Treasuries. As shown in table 10, a regression with the spread of the Fannie Mae secondary market rate over constant-maturity ten-year Treasury rates also shows a statistically significant negative effect of the announcement of the program of about 30 basis points, without a significant further effect due to

Table 9. Fannie Mae Treasury-OAS with Program Dummies

	(1) Treasury-OAS (Bloomberg)	(2) Treasury-OAS (Bloomberg)	(3) Treasury-OAS (Bloomberg)	(4) Treasury-OAS (Bloomberg)
Bond Spread	1.16***	0.96***	0.99***	0.98***
	(0.06)	(0.05)	(0.05)	(0.05)
Total MBS Purchases	5.30	18.41	−20.53	54.05*
	(23.54)	(23.25)	(22.94)	(30.34)
MBS Treasury Dummy	−30.71***			
	(3.95)			
MBS Federal Reserve Dummy		−37.94***		
		(4.46)		
MBS Federal Reserve Announce			−28.31***	−22.49***
			(4.17)	(4.34)
MBS Federal Reserve Announce Expansion				−23.97***
				(6.65)
Number of Observations	180	180	180	180
R^2	0.87	0.88	0.86	0.87

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the swap-OAS (Bloomberg ticker: NOASFNCL.IND). The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

Table 10. Secondary Market Spread over Treasuries

	(1) Secondary Market Spread	(2) Secondary Market Spread	(3) Secondary Market Spread	(4) Secondary Market Spread
Bond Spread	1.14***	0.92***	0.95***	0.94***
	(0.06)	(0.05)	(0.05)	(0.05)
Total MBS Purchases	9.02	−14.83	−33.06	31.78
	(21.99)	(23.62)	(22.30)	(29.72)
MBS Treasury Dummy	−30.78***			
	(3.69)			
MBS Federal Reserve Dummy		−29.37***		
		(4.53)		
MBS Federal Reserve Announce			−24.66***	−19.59***
			(4.05)	(4.25)
MBS Federal Reserve Announce Expansion				−20.83***
				(6.51)
Number of Observations	180	180	180	180
R^2	0.88	0.87	0.86	0.87

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the spread of Fannie Mae thirty-year current-coupon MBS (Bloomberg ticker: MTGEFNCL.IND) over ten-year U.S. Treasury yields. The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

the volume of the program. This effect is very similar to the effect detected on the Treasury-OAS.[20]

[20] This suggests that any decline in the Treasury-OAS can be attributed more to changes in the default risk than to an increase in the value of the prepayment option.

3.3 Did the MBS Program Make the Implicit Guarantee More Explicit?

As another robustness check, we examined whether the absence of size effects and the presence of program effects on Treasury-OAS might be due to the program's mere existence signaling to the market that federal government guarantees of the GSEs had become more likely. If investors believed that the government would always bail out Fannie and Freddie, despite the lack of explicit "full faith and credit" insurance, mortgage spreads over Treasuries would not have increased in 2007 and 2008 nor have remained high after the federal government takeover. The fact that spreads were positive suggests that market participants attached some likelihood to the government not bailing out Fannie and Freddie (in addition to some differences in the liquidity of the two securities). By directly purchasing GSE debt and GSE-insured MBS, the Federal Reserve increased its own financial exposure to the GSEs, increasing the perceived strength of the guarantee. For a discussion of the public's perception of U.S. government guarantees for GSEs, see Passmore (2005).

To try and separate the impact that these "implicit guarantee" effects had on the OAS from possible effects related to a provision of liquidity to mortgage markets, we analyze the development of the option-adjusted spread on MBS that are guaranteed by Ginnie Mae. Ginnie Mae securities are the only MBS that are explicitly guaranteed by the full faith and credit of the U.S. government. If Fannie Mae OAS declined significantly more following the announcement of the MBS purchase program than Ginnie Mae OAS, then this is evidence for an "implicit guarantee" explanation of any observed decline in spreads.

In table 3 we analyzed the swap-OAS and in table 9 the Treasury-OAS of Fannie Mae securities. In tables 11 and 12 we repeat the same regressions but use the OAS on Ginnie Mae securities as the dependent variable.[21] As was the case with the swap-OAS on Fannie Mae MBS in table 3, when analyzing the swap-OAS of Ginnie Mae

[21]We use the NOASGNSF.IND series from Bloomberg for the swap-OAS on thirty-year Ginnie Mae–insured MBS, and the MOASGNSF.IND series for the Treasury-OAS series.

Table 11. Ginnie Mae Swap-OAS with Program Dummies

	(1) Swap-OAS	(2) Swap-OAS	(3) Swap-OAS	(4) Swap-OAS
Bond Spread	0.78***	0.87***	0.85***	0.85***
	(0.06)	(0.05)	(0.05)	(0.05)
Total MBS Purchases	24.25	12.22	12.34	−16.61
	(22.73)	(22.53)	(20.98)	(28.24)
MBS Treasury Dummy	16.19***			
	(3.83)			
MBS Federal Reserve Dummy		21.19***		
		(4.33)		
MBS Federal Reserve Announce			20.80***	18.34***
			(3.84)	(4.15)
MBS Federal Reserve Announce Expansion				9.50
				(6.23)
Number of Observations	179	179	179	179
R^2	0.71	0.72	0.73	0.73

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the swap-OAS for Ginnie Mae securities (Bloomberg ticker: NOASGNSF.IND). The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

MBS, we find an incorrectly signed effect. The coefficients suggest that the program announcement, program start, and program volume all contributed to an *increase* in the spread, while the effects of program volume are not statistically significant. When looking at the effects on the Treasury-OAS of Ginnie Mae MBS, we find effects

Table 12. Ginnie Mae Treasury-OAS with Program Dummies

	(1) Treasury-OAS	(2) Treasury-OAS	(3) Treasury-OAS	(4) Treasury-OAS
Bond Spread	0.93***	0.90***	0.90***	0.90***
	(0.06)	(0.05)	(0.05)	(0.05)
Total MBS Purchases	−73.79***	−54.45**	−71.85***	−66.20**
	(23.13)	(23.14)	(21.80)	(29.88)
MBS Treasury Dummy	−6.25			
	(3.89)			
MBS Federal Reserve Dummy		−11.62***		
		(4.44)		
MBS Federal Reserve Announce			−7.42*	−6.98
			(3.95)	(4.28)
MBS Federal Reserve Announce				−1.81
				(6.55)
Number of Observations	180	180	180	180
R^2	0.85	0.85	0.85	0.85

Notes: This table shows the results from regression (5). The observations are at a weekly frequency between 2007 and June 2010. The dependent variable is the Treasury-OAS for Ginnie Mae securities (Bloomberg ticker: MOASGNSF.IND). The "Bond Spread" control variable captures the spread between five-year Fannie Mae bonds and U.S. Treasury active (on-the-run) securities (Bloomberg ticker: FNMGVN5.IND). "Total MBS Purchases" captures the stock of GSE-insured MBS held by the Federal Reserve and Treasury as a percentage of the total market of about $5 trillion. In column 1 the program dummy is set to 1 starting in September 2008, when the Treasury started buying MBS. In column 2 the program dummy is set to 1 starting in January 2009, when the Federal Reserve purchases of MBS started. In column 3 the program dummy is set to 1 starting with the announcement of the Federal Reserve's MBS purchase program on November 25, 2008. In column 4 the additional "MBS Federal Reserve Announce Expansion" dummy is set to 1 at the MBS program expansion announcement on March 18, 2009. Significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$).

that are between one-third and one-half the size of the effect on the Fannie Mae Treasury-OAS.[22] The results thus suggest that at

[22] These results survive in a specification where we drop the total volume of MBS purchased from the regression.

least 50 percent of the observed fall in Treasury-OAS on mortgage-backed securities guaranteed by Fannie Mae could be attributed to the "implicit guarantee" effect. This leaves at most a decline of about 15 basis points to be explained by the MBS purchase program.

If this "implicit guarantee" effect is indeed a key channel through which the MBS purchases and GSE debt purchases affected mortgage spreads,[23] a significantly more straightforward way to achieve the same goal would have been to extend, formally and explicitly, the full faith and credit of the United States to Fannie and Freddie, in a similar fashion as it is already extended to Ginnie Mae. Moreover, if the implicit guarantee was the channel, then similar MBS purchase programs used in the future are not likely to have any impact on spreads.

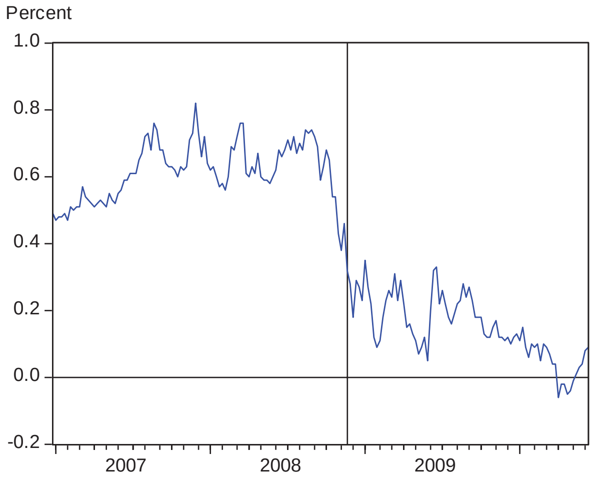
4. A Shift in the Swap Spread

The results in the previous section reveal a strong positive effect of risk factors and no negative effect of the *volume* of MBS purchases on mortgage spreads. These results are robust to alternative measures and specifications. The results also reveal a marked difference in the estimated effect of the MBS program's *existence* or *announcement* on spreads over swaps rates versus spreads over Treasury rates: Program dummies show no negative effect of the program on the swap-OAS and about a 30-basis-point negative effect on the Treasury-OAS. This result is also robust to alternative measures and specifications.

This difference in the estimated dummy coefficients in the regression equation for spreads over swaps versus spreads over Treasury rates implies certain relative movements of swaps and Treasury rates during this period. In particular, it implies that the spread between swaps rates and Treasury rates—commonly referred to as the *swap spread*—should have narrowed during this period. To show this simply, we can abstract from maturity differences or the term structure and let M = mortgage rate, S = swap rate, and T = Treasury rate. Then the two mortgage spreads discussed in the previous section are M-S and M-T, and the swap spread is S-T. Our empirical

[23] An "implicit guarantee" channel might also have contributed to the decline in the bond spreads, which we use to control for the default risk of the MBS.

Figure 8. Weekly Averages of the Ten-Year Swap Spread



Notes: This figure shows weekly averages of the difference between the ten-year swap rate and the ten-year constant-maturity Treasury rate. The vertical line is drawn for the week ending November 15, 2008.

results show that the M-S spread was unchanged during the period of the program (after controlling for prepayment and default risk), while the M-T spread decreased. So the implication is that (M-T) – (M-S) decreased, which means of course that the swap spread S-T decreased.[24]

In fact, the swap spreads did decrease during this period. We examined the one-year, two-year, five-year, and ten-year swap spreads. Figure 8 shows the ten-year swap spread, or the difference between the ten-year swap rate and the ten-year constant-maturity Treasury rate. Clearly there was a significant downshift in the swap spread during this time period. The spread averaged about 0.5 percent from 2005 through 2007 and about 0.1 percent from the start of 2009 through June 2011. The story is similar for the swap spreads at

[24]To consider the whole term structure, you can use the derivation of the OAS in equation (3). The interest rate (r_{it}) used for computing the OAS (θ) is based on the LIBOR swap curve in the case of the swap-OAS, while it is based on the Treasury yield curve in the case of the Treasury-OAS. The difference between these two curves is due to differences in the swap spreads at various maturities.

other maturities, though the shorter maturities increased by a larger amount during the panic in late September and early October 2008 before decreasing.[25]

The decline in the swap spread shown in figure 8 is well known to traders and investors in the swap and Treasury markets. The most commonly cited explanation[26] for the decline is the huge increase in Treasury borrowing relative to private-sector borrowing as the Federal deficit increased sharply when the economy went into a downturn in late 2008. This increased the demand for Treasury borrowing and decreased the demand for private-sector borrowing; hence, according to this explanation, the spread between swap rates and Treasury rates narrowed. In support of this explanation, the Treasury Borrowing Advisory Committee (2010) reported that until October 2008, the Treasury had been adding incrementally to coupon auction sizes. In October 2008, the Treasury surprised the market with $40 billion of 2015/18 issues, which was followed by a rapid rise in coupon issuance for a full year.

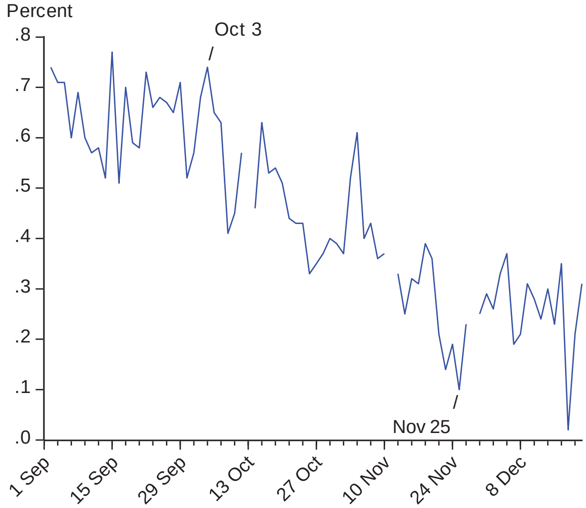
Given the algebraic link between the three spreads (M-S, M-T, and S-T), at least two possible explanations for the decline in the swap spread (S-T) thus emerge from our analysis. The first possible explanation—discussed in the previous section of this paper—is that the MBS program reduced mortgage spreads over Treasuries but did not reduce mortgage spreads over swaps and thereby led to a decline in the swap spread. The second explanation—discussed in the Treasury Borrowing Advisory Committee report (2010) and elsewhere—is that a large increase in the supply of Treasury debt drove down the swap spread and thereby created a differential between mortgage spreads over swap rates and mortgage spreads over Treasury rates.

There are potentially important timing differences which might help to distinguish between the two explanations. For example, the second explanation implies that the shift in the swap spreads would occur in October 2008 when Treasury issuance rose, while the first

[25] During the panic, the TED spread (three-month LIBOR over three-month Treasury-bill rates) and the LIBOR overnight index swap (OIS) spread were also spiking. See Smith (2010) and Taylor and Williams (2009) for a discussion of movements in LIBOR OIS around the period of the panic.

[26] See, for example, the quarterly Treasury Borrowing Advisory Committee report of May 2010.

Figure 9. Daily Observations on the Ten-Year Swap Spread in Late 2008



Notes: This figure shows daily observations on the difference between the ten-year swap rate and the ten-year constant-maturity Treasury rate.

explanation implies that the swap shift would begin at the time of the MBS program announcement or startup. In figure 8 we have drawn a vertical line at the week ending November 15, which was before the November 25 announcement of the Federal Reserve's purchase program. In figure 9 we show daily observations on that same swap spread. Most of the movement in the spread occurred before the announcement of the program by the Federal Reserve. While this provides some evidence in favor of the second hypothesis, rigorous testing between these two explanations will require additional research, including further specifying and exploring the second explanation, which is beyond the scope of this paper.

5. Conclusion

In this paper we endeavored to estimate the quantitative impact of the Federal Reserve's mortgage-backed securities purchase program on mortgage interest spreads using a multivariate statistical framework which takes into account other possible influences on spreads. We controlled for two other possible influences on mortgage spreads: changes in prepayment risk and changes in default risk. Our results can be summarized as follows:

- Using conventional option-adjusted spreads (OAS) from Bloomberg based on LIBOR swaps to control for prepayment risks, it is difficult to detect a significant effect of the MBS purchases. Movements in prepayment risk and particularly movements in default risk explain virtually all of the movements in mortgage spreads, as captured by the OAS relative to the swap curve. We find similar results when using other swap-OAS series compiled by Barclays and Deutsche Bank, as well as when considering the secondary market MBS spread without controlling for possible changes in prepayment risk.
- A statistically significant effect on mortgage spreads—about 30 basis points—can be found if one uses an alternative measure of OAS based on the Treasury yield curve, but even with this measure the volume of purchases has no effect over and above the mere announcement or existence of the program. In other words, the impact has not increased with the additional purchases of MBS since the start of the program. We find a similar effect when we consider the secondary market MBS spread of MBS over Treasuries, without attempting to control for prepayment risk.
- When also analyzing the impact on the OAS of MBS guaranteed by Ginnie Mae, which has the U.S. government's explicit full faith and credit guarantee, we find evidence for the suggestion that about 50 percent of the 30-basis-points decline in Treasury-OAS for MBS guaranteed by Fannie Mae can be attributed to what we call the "implicit guarantee" effect. This suggests that about 15 basis points of the decline in Treasury-OAS can be explained by increased liquidity in agency-insured MBS markets.
- Finally we showed that the estimated negative impact on the Treasury-OAS compared with the estimated zero impact on the swap-OAS implies a downward shift in the swap spread or the difference between swap rates and Treasury rates during

this period. Such a shift did indeed occur and has been noted by market participants, who have offered an explanation unrelated to the MBS purchase program. While timing differences provide support for this alternative explanation, further research is required to discriminate rigorously between these hypotheses. We hope that the information in this study will be of value in such research.

Analyzing the effectiveness of the MBS purchase program is very difficult. The creation of adequate counterfactuals is complicated by the simultaneous government interventions in a large number of markets. Furthermore, the conservatorship status of the GSEs has contaminated many of the relevant GSE default risk proxies that are most important to control for when analyzing the development of spreads on GSE-insured MBS. Our analysis has used a variety of different approaches to proxy for this risk, each with its own problems. Nevertheless, on balance this paper suggests that the impact of the Federal Reserve's MBS purchase program on mortgage spreads has been small and uncertain, once the effects of default risk and prepayment risk have been taken into account.

While this paper is unlikely to be the final word on the program's effectiveness, our empirical results thus raise questions about the ability of central banks to conduct price-keeping operations reliably by increasing and decreasing asset purchases in particular markets. They also raise doubts about the benefits in terms of lower mortgage interest rates of further increases in the size of the Federal Reserve's MBS portfolio or about the costs in terms of higher interest rates of gradually reducing the size of that portfolio.

References

Aït-Sahalia, Y., J. Andritzky, A. Jobst, S. Nowak, and N. Tamirisa. 2010. "Market Response to Policy Initiatives during the Global Financial Crisis." NBER Working Paper No. 15809 (March).

Belikoff, A., K. Levin, H. Stein, and X. Tian. 2010. "Analysis of Mortgage Backed Securities: Before and After the Crisis." Bloomberg LP, Version 2.2.

Bernanke, B. 2009. "The Crisis and the Policy Response." Speech at the Stamp Lecture, London School of Economics, London, England, January 13.

Boudoukh, J., M. Richardson, R. Stanton, and R. Whitelaw. 1999. "The Pricing and Hedging of Mortgage-Backed Securities: A Multivariate Density Estimation Approach." In *Advanced Fixed-Income Valuation Tools*, ed. N. Jegadeesh and B. Tuckman, Chapter 9. John Wiley Publishing.

Brigo, D., and F. Mercurio. 2006. *Interest Rate Models: Theory and Practice.* 2nd Edition. Springer Finance.

Ciorciari, J., and J. Taylor, eds. 2009. *The Road Ahead for the Federal Reserve,* by George Shultz, Allan Meltzer, Peter Fisher, Donald Kohn, James Hamilton, John Taylor, Myron Scholes, Darrell Duffie, Andrew Crockett, Michael Halloran, Richard Herring, and John Ciorciari. Stanford, CA: Hoover Press.

Deng, Y., J. M. Quigley, and R. Van Order. 2000. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica* 68 (2): 275–307.

Duygan-Bump, B., P. Parkinson, E. Rosengren, G. Suarez, and P. Willen. 2010. "How Effective Were the Federal Reserve Emergency Liquidity Facilities? Evidence from the Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility." Working Paper No. QAU10-3, Federal Reserve Bank of Boston (April 29).

Fabozzi, F., and S. Mann. 2001. *Introduction to Fixed Income Analytics.* New Hope, PA: Frank J. Fabozzi Associates.

Fannie Mae. 2001. "2001 Annual Report." Statement by CFO Timothy Howard.

Fuster, A., and P. Willen. 2010. "$1.25 Trillion Is Still Real Money: Some Facts about the Effects of the Federal Reserve's Mortgage Market Investments." Discussion Paper No. 10-4, Federal Reserve Bank of Boston.

Gagnon, J., M. Raskin, J. Remanche, and B. Sack. 2011. "The Financial Market Effects of the Federal Reserve's Large-Scale Asset Purchases." *International Journal of Central Banking* 7 (1): 3–44.

Hancock, D., and W. Passmore. 2011. "Did the Federal Reserve's MBS Purchase Program Lower Mortgage Rates?" FEDS Working Paper No. 2011-01, Board of Governors of the Federal Reserve System.

Krishnamurthy, A. 2010. "How Debt Markets Have Malfunctioned in the Crisis." *Journal of Economic Perspectives* 24 (1): 3–28.

Kupiec, P., and A. Kah. 1999. "On the Origin and Interpretation of OAS." *Journal of Fixed Income* 9 (3): 82–92.

Passmore, W. 2005. "The GSE Implicit Subsidy and the Value of Government Ambiguity." FEDS Working Paper No. 2005-05, Board of Governors of the Federal Reserve System.

Rudebusch, G. 2010. "Macro-Finance Models of Interest Rates and the Economy." Working Paper No. 2010-01, Federal Reserve Bank of San Francisco.

Sack, B. 2009. "The Fed's Expanded Balance Sheet." Remarks at the Money Marketeers of New York University, December 2. Available at http://www.newyorkfed.org/newsevents/speeches/2009/sac091202.html.

Smith, J. 2010. "The Term Structure of Money Market Spreads during the Financial Crisis." Working Paper, Stanford University.

Taylor, J., and J. Williams. 2009. "A Black Swan in the Money Market." *American Economic Journal: Macroeconomics* 1 (1): 58–83.

Treasury Borrowing Advisory Committee. 2010. Presentation to the quarterly meeting, May 4. Available at http://www.treasury.gov/resource-center/data-chart-center/quarterly-refunding/Documents/dc-2010-q2.pdf.

Windas, T. 1996. *An Introduction to Option-Adjusted Spread Analysis*. Revised Edition. Princeton, NJ: Bloomberg Press.

Discretionary Fiscal Policies over the Cycle: New Evidence Based on the ESCB Disaggregated Approach*

Luca Agnello[a] and Jacopo Cimadomo[b]
[a]Banque de France, Service d'Etude des Politiques de Finances Publiques
[b]European Central Bank, Fiscal Policy Division

This paper explores how discretionary fiscal policies on the revenue side of the government budget have reacted to economic fluctuations in European Union countries. For this purpose, it uses data on legislated revenue changes and structural indicators provided twice per year by national central banks of European Union countries in the ESCB framework for analyzing fiscal policy. Results suggest that, overall, legislated changes in taxes and social security contributions have responded in a strongly procyclical way to the business cycle, while commonly used cyclical-adjustment methods point to acyclicality.

JEL Codes: E62, E65, H20.

 We are grateful to the ESCB's Working Group on Public Finance (WGPF) for authorizing the use of data from the "disaggregated approach" for this study. These data are confidential, and therefore they are not reported herein. We are particularly grateful to Paul Hiebert, Richard Morris, and Christiane Nickel for their input in this paper. We also would like to thank two anonymous referees, the editor Giancarlo Corsetti, and the participants at the WGPF meeting of March 2009, and at an ECB seminar, for helpful comments and suggestions. In particular, we would like to thank António Afonso, Sebastian Hauptmeier, Fédéric Holm-Hadulla, Gerrit Koester, Michele Lenza, Bernhard Manzke, Alex Muscatelli, Laurent Paul, Javier J. Pérez, Philipp Rother, Dan Stegarescu, Ad van Riet, Karsten Wendorff, and Guido Wolswijk. The opinions expressed herein are those of the authors and do not necessarily reflect those of the European Central Bank, the Banque de France, or the Eurosystem. Author e-mails: luca.agnello@banque-france.fr and jacopo.cimadomo@ecb.int.

1. Introduction

This paper provides new evidence on the cyclical stance of fiscal policy in European Union (EU) countries, over the period 1998 through 2008. In particular, it explores whether EU governments have adopted countercyclical revenue-based fiscal policies that helped to smooth the business cycle, or whether they have acted in a procyclical way, for instance, by raising taxes during slowdowns while implementing expansionary policies in booming times, thereby exacerbating swings of the cycle.

The paper focuses on *legislated changes* in taxes and social security contributions, as defined by the European System of Central Banks (ESCB) disaggregated approach (see Kremer et. al 2006). In fact, as stressed in this framework, changes in laws that have a budgetary impact should reflect the discretionary component of fiscal policy more effectively than commonly used cyclically adjusted fiscal indicators. The latter might be affected by a number of limitations, mostly related to the filtering techniques applied to net out the automatic impact of the cycle from headline fiscal figures (see, e.g., Canova 1998; Jaeger and Schuknecht 2004; Darby and Melitz 2008; Larch and Turrini 2009).

Measures of legislated changes in taxes and social security contributions used in this paper are based on information collected within the ESCB. In detail, a database consisting of legislative changes affecting government revenues is based on estimates by the national central banks (NCBs) of the twenty-seven European Union countries compiled at a biannual frequency. This database comprises *qualitative* information on fiscal measures judged to have an impact of a relevant magnitude on government finances, including the estimated duration of these measures. The database also reports *quantitative* information, which relates to the budgetary impact of measures. In most cases this is based on official government estimates (if existing), but ESCB fiscal experts' judgment may deviate from official estimates in some cases, e.g., if there is reason to believe that the official estimates are overly optimistic or pessimistic.[1]

While information concerning the estimated impact of legislative changes is typically provided in budget documents, long time-series

[1] In section 3, some specific examples of changes in legislation, and of cases in which the NCBs' assessment deviates from the official one, are provided.

and harmonized cross-country data are generally not available from other sources. The European Commission, in its yearly publication *Public Finances in EMU*, reports estimates on discretionary legislated measures for EU-27 countries.[2] However, this has only been done since 2004, and quantitative impacts are only reported for the most important measures. Moreover, there is generally no distinction between the (permanent vs. temporary) nature of measures. These features clearly limit the scope for using the information provided in the Commission's reports for a panel regression analysis, as proposed in this paper. At the same time, we have carried out an analysis which confirms that ESCB data and the information provided by the Commission in its reports—when overlapping in terms of country and time coverage—are highly correlated.[3] A data set of legislative changes has been also collected by Romer and Romer (2009a, 2009b, 2010), but only for the United States.[4]

In sum, for what concerns changes in legislation on the revenue side, the ESCB data set is—to our knowledge—unique in its kind for EU countries, given its (relatively) large country and time coverage and the amount of descriptive information provided. At the same time, like for example for the measurement of the output gap, the actual impact of legislative changes is unobservable and any estimates thereof are subject to a certain degree of uncertainty.

The ESCB data set of legislated changes is used to estimate the *cyclical sensitivity* of revenue-based discretionary fiscal measures, based on the estimation of fiscal policy rules linking these measures of legislated changes to the output gap and other control variables.

---

[2]This report has been published since 2000. However, estimates for discretionary measures are available only as of the 2004 edition. Past editions of *Public Finances in EMU* are available at http://ec.europa.eu/economy_finance/publications/european_economy/public_finances_emu_en.htm.

[3]This is also explained by the fact that both approaches are primarily based on the same underlying source, i.e., official budget documents.

[4]The authors use a "narrative record" of presidential speeches, executive-branch documents, Congressional reports, and budget laws to identify the size, timing, and principal motivation for all major post-war tax policy actions. Based on this data, the authors uncover that tax increases have strong contractionary effects on output in the United States. The data reported in tables from the ESCB disaggregated approach also provide a description of the size, timing, and nature of legislated revenue changes that took place in EU countries. However, at the current juncture, these data do not allow to identify also the motivation behind each policy decision (e.g., cyclical stabilization, deficit control, or other "exogenous" reasons) as in the data set of Romer and Romer.

In particular, the cyclical reaction of discretionary changes in overall taxes and social security contributions is analyzed, together with the cyclical behavior of four subcategories of government revenues: direct taxes paid by households, direct taxes paid by enterprises, indirect taxes, and social security contributions. Finally, results from these baseline regressions are compared with what is obtained using fiscal indicators that are cyclically adjusted according to the ESCB disaggregated approach.

Based on this analysis, it emerges that, in line with existing studies, changes in cyclically adjusted taxes and social security contributions seem to be disconnected from the economic cycle in EU countries over the last ten years. However, when legislated measures are used instead of cyclically adjusted indicators, it is found that, overall, revenue changes have responded in a *strongly procyclical* way to the business cycle. In particular, this result appears to be the outcome of a significant procyclical reaction of direct taxes payable by households and of social security contributions. More specifically, there is evidence that direct taxes paid by households have been significantly reduced during economic expansions. As for indirect taxes, while estimates based on cyclically adjusted indicators suggest procyclicality, changes in indirect tax legislation seem to be acyclical. In addition, indirect taxes are the only revenue subcomponent that appears to have been used to stabilize the government debt-to-GDP ratio.

The rest of this paper is organized as follows. In section 2, a review of the empirical literature that explores the cyclical sensitivity of discretionary fiscal policies, based on cyclically adjusted indicators, is provided. Section 3 describes how the data set used in this analysis is constructed using information provided by NCBs. In section 4, the fiscal policy reaction functions used in the analysis are described. In section 5, the main results from the analysis are shown and discussed. Section 6 is devoted to a series of robustness checks, and section 7 concludes.

2. Review of the Literature Based on Cyclically Adjusted Indicators

The bulk of the macroeconomic literature typically employs "cyclically adjusted" indicators as a proxy for the discretionary component of fiscal policy. Based on cyclically adjusted indicators regularly published by the main international institutions, a number of policy and academic papers have investigated issues related to discretionary fiscal policies (see, for example, Alesina and Perotti 1995; Giavazzi and Pagano 1996).

In particular, as regards the issue of the policy responsiveness to economic fluctuations—which is often referred to as "cyclical sensitivity"—the standard approach in the fiscal policy literature is to estimate fiscal policy rules linking (levels or changes of) cyclically adjusted fiscal variables, as measures of the discretionary component of fiscal policy, to business-cycle indicators (generally, the output gap) and other explanatory variables, such as, in particular, the public debt (see, e.g., Taylor 2000; Auerbach 2002).[5]

While the empirical evidence from this literature is quite mixed, a relatively large consensus has emerged on the fact that discretionary fiscal policies seem to be substantially *acyclical* in EU countries, especially as far as *government revenues* are concerned. In particular, Galí and Perotti (2003) estimate fiscal policy rules for a panel of euro-area countries and find that cyclically adjusted budget balances have not reacted to the business cycle in the period that followed the entry into force of the Maastricht Treaty in 1992. When they focus on the breakdown between (cyclically adjusted) spending and revenue, they show that both these components seem to have been acyclical over the period considered. Similar results are found by Turrini (2008), who, however, highlights that the substantial neutrality of the cyclically adjusted budget balance over the cycle is likely to stem

[5]Within this literature, most papers employ "ex post" data—i.e., observations in a revised form taken from the latest available vintage—to study "actual" or "realized" policies. Recently, a new strand of literature focuses on fiscal policy plans, or "intentions," rather than actual policies. In particular, Cimadomo (2011) proposes to estimate fiscal policy rules based on an information set which closely mimics the one available to fiscal policymakers at the time of budgeting. Budget plans reported at the time of budgetary decisions, and other real-time macro indicators, are used in this framework. It is found that ex ante fiscal plans are countercyclical, especially in expansions, whereas ex post data point to acyclicality (for related studies, see Giuliodori and Beetsma 2008; Beetsma and Giuliodori 2010). At the same time, other papers focus on actual policies and incorporate revised information as regards fiscal indicators and real-time data for the output gap and other explanatory variables (see Forni and Momigliano 2005; Golinelli and Momigliano 2006).

from the combined effect of a significant countercyclical response of revenue compensated by a procyclical reaction of expenditure. More recently, Fatás and Mihov (2009) have shown that cyclically adjusted spending has become more procyclical in euro-area countries, after the adoption of the single currency, whereas cyclically adjusted taxes have become more countercyclical. Many other studies employing cyclically adjusted budget balances, and revised data, tend to confirm that discretionary fiscal policies are acyclical in EU and euro-area countries (see, e.g., Ballabriga and Martinez-Mongay 2002; OECD 2003; European Commission 2006; Wyplosz 2006).

While structural indicators represent a useful benchmark to evaluate the fiscal policy stance, especially in the absence of more detailed information on changes in legislation passed by parliaments, they are clearly subject to a number of limitations.[6] These latter are notably related to the substantial uncertainty inherent in the cyclical-adjustment procedure. In particular, there is a certain degree of arbitrariness in the selection of the statistical smoothing technique used to extract the cyclical component from the unadjusted level of the budget and budgetary categories. In addition, while standard methods imply that elasticities of budgetary components with respect to output are treated as constant, empirical evidence suggests that they can vary over time and that they might be characterized by high volatility (see Eschenbach and Schuknecht 2004; Jaeger and Schuknecht 2004). Moreover, additional difficulties may stem from the (unconsensual) definition of temporary measures, as long as structural indicators are considered. Finally, it cannot be excluded that not only do unemployment benefits react to the cycle (as generally assumed), but other categories of social spending (such as, for example, age- and health-related expenditure as well as incapacity and sick benefits) also move in response to cyclical fluctuations (on this point, see in particular Darby and Melitz 2008). Therefore, as discussed by some authors (see, for example, Chalk 2002; Larch and Salto 2005), commonly used cyclically adjusted fiscal indicators may provide inaccurate measures of discretionary policies, and empirical results based on these indicators should be interpreted cautiously. Against this background, as also suggested

[6] See in particular Larch and Turrini (2009) for an extensive account of downsides of cyclically adjusted indicators.

by Kremer et. al (2006), data on legislated policy changes should reflect better the discretionary stance of fiscal policy.

3. A Data Set for Discretionary Fiscal Policies in the EU

The starting point of this study is the identification of legislated revenue changes as measures of the discretionary component of fiscal policy. To that aim, data from the ESCB's disaggregated approach are used. Data include discretionary measures that have been approved by the respective national parliament and that are assessed to have a sizable impact on government finances. Both qualitative information (i.e., the description of each measure) and quantitative information (i.e., the estimated budgetary impact and its duration) is provided. While the data regularly collected within the ESCB is confidential, in Kremer et al. (2006) the disaggregated approach is documented and explained, and information on the impact of measures over the period 1998–2004 for some countries is reported. Here, two examples from that earlier paper are briefly discussed as illustrative cases. The first is a major tax reform implemented in the Netherlands in 2001.[7] That tax reform (named "Wet inkomstenbelasting 2001"[8]) implied a shift from direct to indirect tax revenues and, according to ESCB calculations, had a significant negative impact on government revenues. Revenues from social contributions and direct taxes were estimated to be reduced by about 2.4 percent of GDP in the year of the reform. At the same time, the VAT rate was increased from 17.5 percent to 19 percent, and energy taxes were also increased. This elicited additional revenues from indirect taxation amounting to around 0.6 percent of GDP. Thus, overall, the impact on tax revenues of these changes was assessed as amounting to 1.8 percent of GDP in 2001.

In most cases the ESCB data concerning legislative changes is based on official government estimates provided during the legislative process. However, ESCB fiscal experts' judgment may deviate

[7]See Kremer et al. (2006, p. 34).

[8]The reform was approved by the Dutch parliament on May 22, 2000 through the State law number BWBR0011353, which came into force on January 1, 2001.

from official estimates in some cases (e.g., if there is reason to believe that the official estimate is overly optimistic or pessimistic). For example, again citing Kremer et al. (2006), revenue developments in Italy in the years 1998–2004 were influenced by the 1998 tax reform which, inter alia, introduced a new regional tax on productive activities (the "IRAP" tax). While according to official estimates released when the reform was introduced, the IRAP tax was expected to have a neutral effect on total revenue, ESCB fiscal experts assessed that the reform induced a negative budgetary impact close to 0.5 percent of GDP, cumulatively over the considered period. Overall, the tax reform was assessed as having implied reductions in social security contributions (estimated at 2.1 percent of GDP) and direct taxes payable by corporations (0.9 percent of GDP), only partly offset by the increase in indirect taxes (2.5 percent of GDP), where the new tax was classified.[9]

All in all, information on legislated changes has been combined to compile two complementary data sets. The first one, based on the so-called disaggregated approach (DA) tables, as described below, incorporates data on legislated revenue changes and cyclically adjusted indicators for nineteen EU countries. The second—compiled on the basis of DA tables and on further complementary information from NCBs—comprises only legislated revenue changes, but for the larger set of twenty-seven EU countries.[10]

Two facts should be stressed. First, while structural indicators for government expenditure are also reported in DA tables, estimates

[9]See Kremer et al. (2006, pp. 29–31). See also this paper for other examples of changes in legislation.

[10]While cyclically adjusted indicators are produced by different international institutions, only the European Commission produces comparable data on the budgetary impact of the main changes in legislation, for each of the EU-27 countries. These data are reported in the yearly publication *Public Finances in EMU*. However, the coverage of the Commission's data set is significantly smaller than that of the ESCB's data set, given that discretionary measures started to be reported only as of 2004 and that for some measures the budgetary impact is not reported. Nevertheless, we have manually collected the Commission's data from past editions of *Public Finances in EMU*, from 2004 to 2008. Data had to be collected from PDF files for each country, as they are not included in any Commission's electronic database. The estimated correlation coefficient between the Commission's and the ESCB's data on legislative changes is high (0.87) and statistically significant, indicating that the two institutions tend to adopt a broadly similar assessment regarding the budgetary impact of revenue measures.

for the budgetary impact of legislative measures on the spending side are not assessed. This reflects the very different nature of most government spending to most revenues. In fact, a benchmark path that would be followed by revenue items in the absence of discretionary measures can be identified in a (relatively) easy way, on the basis of growth assumptions on the relevant tax bases. By contrast, there is no predetermined path for most (especially non-social) government spending, which instead has to be set annually in the budget. In addition, discretionary spending actions taken at the administrative level are often even more important than parliamentary decisions, and keeping track of all of them is virtually impossible (see also Kremer et al. 2006).

Second, this paper focuses on *permanent* (legislated) discretionary measures on the revenue side of the government budget. An analysis of temporary measures would also be interesting, given that these measures are often used for countercyclical actions. Unfortunately, however, the ESCB data concerning temporary measures is (partly) aggregated to the level of total revenues, rather than being reported for individual tax categories) (see also table 12 in the appendix for an example of a DA table for Italy, as reported in Kremer et al. 2006). This clearly limits the scope for such type of analysis.

In the following, an overview of the information provided in DA tables, and the construction of the data set, is described in more detail.

3.1 The "Disaggregated Approach" Tables

DA tables are compiled in a standardized form according to the ESCB disaggregated approach.[11] These tables report changes in cyclically adjusted total expenditure and revenue, along with the breakdown of the latter into its four main categories (see table 12 in the appendix). In particular, structural ratios are computed for (i) *direct taxes payable by corporations* (ΔR^{dte}); (ii) *direct taxes payable by households* (ΔR^{dth}); (iii) *indirect taxes* (ΔR^{itx}); and (iv) *social contributions* (ΔR^{sct}). The sum of these four categories is defined as

[11] See Kremer et al. (2006) for a detailed description of this methodology.

overall taxes and social contributions, simply labeled *overall revenue* (ΔR^{txs}) in this paper.[12]

As regards the cyclical-adjustment procedure, the approach developed by Bouthevillain et al. (2001) is followed. In particular, cyclically adjusted revenue and expenditure categories are adjusted individually by applying specific elasticities to the deviation of their respective macroeconomic bases from an estimated trend. Specifically, for each budgetary category considered, labeled as j, the structural level X^j is computed as

$$X^j = x^j_u - x^j_c - x^j_T, \quad (1)$$

where x^j_u is the unadjusted level of the budget item j, x^j_c is its cyclical component, and x^j_T is the temporary measures.[13] In this context, all variables are expressed as ratios of nominal trend GDP, as estimated through the Hodrick-Prescott methodology. DA tables span a period typically ranging from $t-8$ until $t+2$. For example, the spring 2008 vintage includes data from 2000 through 2010.

Importantly, DA tables also incorporate data on legislated changes, defined as $\Delta\ell^j$ hereafter, for each budgetary category j.[14] The impact of legislated changes in DA tables is expressed as a ratio of trend GDP. To be noted, in this context the difference between changes in cyclically adjusted indicators and legislated changes for

[12]Overall taxes and social contributions might differ from total revenue, as the former do not include non-tax-related revenue—as, for example, revenue from EU funds.

[13]Standard approaches for cyclical adjustment typically assume that the relevant tax base is proxied by the output gap. By contrast, the cyclical-adjustment approach developed by Bouthevillain et al. (2001) allows to take into account "composition effects" arising from unbalanced growth. This implies that various macroeconomic bases for government revenue and expenditure might vary in different phases of the cycle or exhibit fluctuations of different magnitude. At the same time, this approach is subject to limitations similar to the alternative methods. In particular, tax bases are defined as deviations from trends, which are estimated based on statistical filtering tecniques (Hodrick-Prescott filter). In addition, elasticities of budgetary categories to the tax bases are assumed to be constant over time (see also Larch and Turrini 2009).

[14]The "first-difference" notation $\Delta\ell^j$ to indicate the fiscal impact of legislated changes is a simple convention, as in reality "levels" of legislation (and their implied budgetary effects) are unobserved. On the other hand, the budgetary impact of a change in legislation can be estimated.

each revenue category j is accounted for by three separate factors: (i) fiscal drag (fd^j), (ii) decoupling of the tax base from GDP (de^j), and (iii) a residual component (re^j).[15]

3.2 Data Availability and Data Construction

Two complementary data sets have been constructed, based on DA tables and on other information from NCBs:

(i) The first data set incorporates an unbalanced panel for nineteen EU countries, over the period 1998–2008, including changes in *both* structural indicators (ΔR^j) and legislated measures ($\Delta \ell^j$). Sources of this data set are only DA tables.

(ii) The second data set comprises an (unbalanced) panel covering twenty-seven EU countries, where only data on legislated measures are incorporated. This data set is constructed based on DA tables and other complementary information from NCBs.

The analysis is based on the spring 2008 vintage of DA tables, but information from earlier vintages has also been used, in case of missing values. Table 13 in the appendix reports the time and country coverage of these two data sets. To be noted, the first data set does not include eight EU countries (Bulgaria, Denmark, Estonia, Hungary, Romania, Slovakia, Sweden, and the United Kingdom)

[15]These three factors, which account for changes in structural revenue ratios (ΔR) beyond the ones explained by legislation changes, are defined as follows: (i) The *fiscal drag* refers to the increase in average tax rates in a progressive income tax scheme that stems from an increase in nominal income, due to inflation or real growth. As such, the fiscal drag may affect structural revenues, even in the absence of legislation changes. (ii) The so-called *decoupling of the tax base from GDP* refers to the possibility that (structural) revenue ratios to nominal (trend) GDP might change even when the elasticity with respect to the macroeconomic base amounts to unity, and even when legislation is unmodified. This may occur when the (trend) growth rate of the tax base deviates from the (trend) growth rate of nominal GDP. (iii) Changes in the structural revenue ratios not explained by the two factors decribed above, and by legislation changes, are denoted as *residuals*. A more thorough description of these components and details concerning their computation can be found in Kremer et al. (2006).

Table 1. Correlation Matrix of Overall Taxes and Social Contributions, and Four Subcomponents Computed According to the ESCB's Disaggregated Approach

	ΔR^{txs}	fd^{txs}	de^{txs}	$\Delta \ell^{txs}$	rs^{txs}
ΔR^{txs}	1.00				
fd^{txs}	−0.10	1.00			
de^{txs}	0.12	0.26	1.00		
$\Delta \ell^{txs}$	0.37	−0.23	−0.32	1.00	
rs^{txs}	0.75	−0.19	−0.06	−0.05	1.00

Notes: Variables refer to overall taxes and social security contributions, net of temporary measures (txs). ΔR are changes in cyclically adjusted values, net of temporary measures. These changes are explained by four components, namely fd: fiscal drag, de: decoupling of the tax base from GDP, legislated changes, and rs: residual component. Values reported are the simple averages of single countries' correlations, for the panel EU-19 over the period 1998–2008.

based on the fact that only legislated changes, but not cyclically adjusted revenue indicators, are available for these countries.

The sample used in the empirical analysis ends in 2008, given that in projection years actual output tends to converge to the potential one for the majority of countries considered (as it typically occurs at the end of the forecast horizon of projection exercises). Therefore, an analysis which focuses on the cyclical stance of fiscal policy for these projection years would not make much sense, given that the output gap tends to shrink over these years.

From a preliminary descriptive analysis of data on legislated changes collected as documented above, some interesting evidence emerges. In particular, focusing on observations averaged over all countries considered, table 1 presents the correlation matrix of cyclically adjusted changes of total taxes and social contributions (ΔR^{txs}) and its driving factors as defined by the disaggregated approach. We note that unsystematic events, as captured by the residual component rs^{txs}, seem to have been important in explaining structural developments of taxes and social contributions. Looking at the impact of the other three factors, it turns out that legislated changes display a positive and relatively high correlation with changes in structural revenues, although such correlation is lower than the one between residuals and structural revenues.[16] Focusing on each revenue category j (see table 2), it emerges that the correlation of legislated changes with structural indicators is higher for social security contributions (0.72), followed by direct taxes paid by households (0.48), indirect taxes (0.41), and direct taxes paid by enterprises (0.33). Except for social security contributions, most of the variation in structural indicators is explained by the residual component.[17]

Table 3 reports the variance-covariance matrix of legislated changes across each revenue category j. As shown, legislated changes display low cross-covariances, suggesting that revenue decisions seem to have been made independently from each other. The variance-covariance matrix also allows to derive the variance decomposition of overall taxes and social contributions. This is equal to the sum of variances of the four components, plus two times all the cross-covariances (these are, however, low, and are ignored for simplicity). It emerges that the bulk of the variability of the total change in legislation is mainly due to changes of direct taxes payable by households, social security contributions, and indirect taxes, whereas changes in

[16] A variance-decomposition analysis suggests that legislated measures explain a rather sizable share of variation in total cyclically adjusted revenues. This is in particular driven by legislated changes for direct taxes by households and social security contributions (results are not reported but are available from the authors).

[17] There are two main possible causes of residuals. One possible cause is that the impact of legislative changes is misestimated. In this case, however, we would expect to observe particularly large residuals in years in which the impact of tax reforms is particularly large. In general, this is not observed in the data. In addition, we would have a large and negative correlation between the estimated impact of legislative changes and residuals. However, table 2 shows that such correlation is negligible. The other cause—and in our view the main cause—is that the residual reflects differences between the underlying cyclical-adjustment model (i.e., tax bases and elasticities), which underpins the estimated impact of fiscal drag and decoupling, and the true (more complex) tax base. In this case, we would expect that the role of the residual in explaining structural changes in tax revenues would be more important for those tax categories for which we are less able to model the true tax base. This indeed seems to be the case. As can be seen in table 2, for example, the correlation between residuals and changes in structural revenues is highest for direct taxes payable by enterprises, the tax base for which is notoriously difficult to proxy given the complex nature of business taxation. It is lowest for social contributions, the base for which is much easier to proxy. For a more in-depth discussion of the "residual," see Morris et al. (2009).

Table 2. Correlation Matrices of Each Revenue Category, and Four Subcomponents Computed According to the ESCB's Disaggregated Approach

	ΔR^{dte}	fd^{dte}	de^{dte}	$\Delta\ell^{dte}$	rs^{dte}
ΔR^{dte}	1.00				
fd^{dte}	0.13	1.00			
de^{dte}	0.07	0.00	1.00		
$\Delta\ell^{dte}$	0.33	0.20	−0.04	1.00	
rs^{dte}	0.89	−0.07	−0.03	−0.10	1.00
	ΔR^{dth}	fd^{dth}	de^{dth}	$\Delta\ell^{dth}$	rs^{dth}
ΔR^{dth}	1.00				
fd^{dth}	−0.14	1.00			
de^{dth}	0.10	0.19	1.00		
$\Delta\ell^{dth}$	0.48	−0.28	−0.14	1.00	
rs^{dth}	0.72	−0.21	−0.13	−0.14	1.00
	ΔR^{itx}	fd^{itx}	de^{itx}	$\Delta\ell^{itx}$	rs^{itx}
ΔR^{itx}	1.00				
fd^{itx}	0.00	1.00			
de^{itx}	0.17	0.02	1.00		
$\Delta\ell^{itx}$	0.41	0.32	−0.14	1.00	
rs^{itx}	0.79	−0.32	−0.01	−0.16	1.00
	ΔR^{sct}	fd^{sct}	de^{sct}	$\Delta\ell^{sct}$	rs^{sct}
ΔR^{sct}	1.00				
fd^{sct}	0.08	1.00			
de^{sct}	0.17	−0.03	1.00		
$\Delta\ell^{sct}$	0.72	0.01	−0.06	1.00	
rs^{sct}	0.61	−0.11	−0.32	0.10	1.00

Notes: Variables refer to direct taxes payable by enterprises (dte), direct taxes by households (dth), indirect taxes (itx), and social security contributions (sct). ΔR are changes in cyclically adjusted values, net of temporary measures. These changes are explained by four components, namely fd: fiscal drag, de: decoupling of the tax base from GDP, legislated changes, and rs: residual component. Values reported are the simple averages of single countries' correlations, for the panel EU-19 over the period 1998–2008.

Table 3. Variance-Covariance Matrix of Legislated Changes

	$\Delta\ell^{txs}$	$\Delta\ell^{dth}$	$\Delta\ell^{dte}$	$\Delta\ell^{itx}$	$\Delta\ell^{sct}$
$\Delta\ell^{txs}$	0.261				
$\Delta\ell^{dth}$	0.109	0.102			
$\Delta\ell^{dte}$	0.033	0.013	0.034		
$\Delta\ell^{itx}$	0.051	−0.001	−0.021	0.107	
$\Delta\ell^{sct}$	0.067	−0.005	0.007	−0.034	0.099

Notes: Variables refer to legislated changes $\Delta\ell$ in overall taxes and social security contributions (*txs*); direct taxes paid by households (*dth*); direct taxes paid by enterprises (*dte*); indirect taxes (*itx*); and social security contributions (*sct*). Sample: 27-EU countries over the period 1998–2008.

direct taxes payable by corporations explain less. This suggests that the former three policy instruments seem to have been predominantly used by European fiscal policymakers in setting their revenue policies.

4. Empirical Strategy

In order to assess the cyclical sensitivity of legislated changes in the EU during the period 1998–2008, and to compare results from regressions based on cyclically adjusted indicators as obtained from the disaggregated approach (for the restricted set of nineteen EU countries), we run separately the following panel regressions:

$$\Delta R^j_{i,t} = \beta Y_{i,t} + \gamma \widetilde{R}^j_{i,t-1} + \phi B_{i,t-1} + \mathbf{Z}'_{i,t}\boldsymbol{\delta} + \alpha_i + \varepsilon_{i,t} \qquad (2)$$

$$\Delta \ell^j_{i,t} = \beta Y_{i,t} + \gamma \widetilde{R}^j_{i,t-1} + \phi B_{i,t-1} + \mathbf{Z}'_{i,t}\boldsymbol{\delta} + \alpha_i + \varepsilon_{i,t} \qquad (3)$$

for $i = 1, \ldots . N$ and $t = 1, \ldots .., T_i$, where N is the number of countries included in our sample and T_i is the number of observations available for each country.

We estimate model (2) and (3) for the aggregate series of overall taxes and social contributions, and for each individual revenue-side budgetary category j. As discussed above, both of the discretionary fiscal policy indicators considered, $\Delta R^j_{i,t}$ and $\Delta\ell^j_{i,t}$, are expressed as a percentage of nominal trend GDP.

The explanatory variable $\widetilde{R}^j_{i,t-1}$ measures the level of the *cyclically adjusted* component of each budgetary category j. It is computed as the difference between the unadjusted volume of each budgetary item j and its cyclical component divided by the nominal trend GDP (as estimated by NCBs). To be noted, $\widetilde{R}^j_{i,t-1}$ includes temporary measures, while $R^j_{i,t-1}$ does not.[18] The regressor $\widetilde{R}^j_{i,t-1}$ captures persistence of budgetary policy decisions. In fact, it is reasonable to assume that the initial level of each revenue category affects the way discretionary policies for that item are set for the following year. For example, it can be expected that the higher the initial level of taxation (in structural terms), the higher the size of the downward adjustment implemented. According to this hypothesis, we expect that the coefficient γ ranges between -1 and 0 (i.e., $-1 < \gamma < 0$). $\widetilde{R}^j_{i,t-1}$ is also used as a proxy to capture persistence in equation (3) since, clearly, it is not possible to reconstruct levels of "legislation."[19] In addition, the inclusion of $\widetilde{R}^j_{i,t-1}$ in equation (3) facilitates comparisons with model (2), as the set of control variables is exactly the same in the two equations.

The economic cycle is represented by the output gap, labeled $Y_{i,t}$. This is the key variable used to measure the systematic response of fiscal discretionary measures to cyclical conditions. $Y_{i,t}$ is computed by NCBs as the difference between nominal GDP and nominal trend GDP (estimated through the Hodrick-Prescott filter), over trend GDP. A positive value of β indicates countercyclicality; i.e., taxes or social contributions increase during economic booms and decrease during slowdowns. On the other hand, a negative value of β points to

[18]In principle, levels of the structural ratio indicators as computed in equation (1) should be used as regressors. Unfortunately, the lack of information concerning temporary measures adopted by national governments and impacting on each budgetary item makes it impossible to reconstruct levels of the variable $R^j_{i,t}$. For this reason, the variable $\widetilde{R}^j_{i,t}$ is used as a proxy. This variable is computed without netting out the temporary measures.

[19]An experiment based on the inclusion of lagged values of $\Delta\ell$ has also been carried out. Unfortunately, however, these controls enter insignificantly in the regression equations. In addition, empirical fiscal policy rules are generally specified such that the initial state of public finances, as represented by fiscal variables included on the right-hand side of the regression equation, are expressed in levels rather than in differences (see, e.g., Galí and Perotti 2003). This is consistent with the idea that fiscal policymakers are more concerned about levels of deficits, revenues, and expenditure, rather than their growth rates.

procyclicality, given that discretionary fiscal policies become looser during expansions and tighter during downturns.

Following several other papers (see, e.g., Bohn 1998; Ballabriga and Martinez-Mongay 2002; Wyplosz 2002; Favero 2003), we also incorporate a "debt stabilization" motive by adding as a regressor the general government debt (relative to nominal trend GDP) outstanding at the time of the budget decision, denoted as $B_{i,t-1}$.[20] A positive sign of ϕ is expected because more indebtedness may lead to more concern about fiscal sustainability and induce the governments to adopt more stringent discretionary measures.

The vector $\mathbf{Z}_{i,t}$ includes the following set of additional variables:

- $NONACTIVE_{i,t}$ accounts for "*population aging*" effects on the fiscal policymaking process (see, e.g., Beetsma and Giuliodori 2010). This variable is computed as the share of the population that is not of working age (i.e., the number of fifteen-year or younger plus the number of sixty-five or older).[21]
- $ELECT_{i,t}$ acts as a control for the possible influence of the electoral cycle, and takes a value of one (zero) if year t is (not) an election year.[22]
- $EMU15_i$ is a dummy variable that equals one for years in which EMU countries have joined the European Monetary Union, and zero otherwise. The idea is to check whether the fiscal policy stance differs between countries belonging to the EMU area and other EU countries. In particular, we expect that accession countries have incentives to pursue tighter fiscal policies in order to consolidate public finances, meet the convergence criteria, and benefit from joining the single currency.

[20] Data on nominal debt are taken from World Economic Outlook (WEO).

[21] Source: AMECO data set.

[22] Recent work shows that elections may play a role in explaining the fiscal stance in euro-area countries in the past decades (see, e.g., Debrun et al. 2008; Golinelli and Momigliano 2009). Data on election years are taken from the web site of the International Institute for Democracy and Electoral Assistance (www.idea.int) and from the Election Resources on the Internet web site http://electionresources.org.

- T_i, is a linear trend that is intended to capture common third factors driving discretionary fiscal behavior of all countries that are unlikely due to common economic circumstances. In practice, we account for the possibility that both $\Delta R_{i,t}$ and $\Delta \ell_{i,t}$ might evolve according to a deterministic trend (see, e.g., Fatàs and Mihov 2003; Afonso, Agnello, and Furceri 2010). Finally, α_i denote the country-specific effects, and $\varepsilon_{i,t}$ are the error terms.

A problem related to estimation of model (2)–(3) concerns the potential reverse causality between $Y_{i,t}$ and our dependent variables. Therefore, this issue is tackled by estimating the panel regressions through the instrumental variables (IV/2SLS) method (see, e.g., Jaimovich and Panizza 2007). In particular, we instrument $Y_{i,t}$ by using its lagged value, $Y_{i,t-1}$—or its second and third lag, depending on the Hansen's J-test for overidentifying restrictions—and the lagged value of the output gap of the United States (see, e.g., Galí and Perotti 2003).[23] In addition, we will estimate our models also using the fixed-effects OLS estimator. In fact, although OLS estimates are in principle biased if the output gap reacts to fiscal policy actions, it cannot be discarded that the reversal causality is weak given that we are considering subcomponents of the budget balance, whose influence on the economic activity might be feeble. In addition, OLS estimates serve as a useful benchmark, given that they are not affected by the choice of instruments.

5. Baseline Results

First, results for the aggregate indicator, as represented by overall taxes and social contributions, and panel regressions (2)–(3), are reported in table 4. In estimates reported in table 4, and in all the following tables, standard errors have been corrected for heteroskedasticity and autocorrelation in residuals. Columns 1 and 2 of that table present the IV/2SLS and OLS estimates of regression

[23]The U.S. output gap is used as instrument, given that a business-cycle variable which is likely to be correlated (for reasons other than the existence of coordinated fiscal policies) with the EU-countries-specific output gap is needed.

Table 4. Estimates for Overall Taxes and Social Contributions (*txs*): EU-19 and EU-27

Dependent Variable	ΔR^{txs}		$\Delta \ell^{txs}$		$\Delta \ell^{txs27}$	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	IV/2SLS	*OLS*	*IV/2SLS*	*OLS*	*IV/2SLS*	*OLS*
$Output\ gap_t$	−0.086	0.036	−0.118***	−0.107***	−0.156***	−0.094***
	[0.076]	[0.049]	[0.033]	[0.023]	[0.041]	[0.021]
$\widetilde{R}^{txs}_{t-1}$	−0.271***	−0.294***	−0.056	−0.058*	−0.054	−0.064**
	[0.041]	[0.072]	[0.037]	[0.030]	[0.034]	[0.029]
$Debt_{t-1}$	0.007	0.009	0.002	0.002	−0.001	0.000
	[0.010]	[0.010]	[0.005]	[0.006]	[0.005]	[0.005]
EMU15	−1.379*	−1.451***	−0.486*	−0.492**	−0.458*	−0.496**
	[0.716]	[0.541]	[0.256]	[0.222]	[0.257]	[0.229]
T	0.062**	0.068***	0.014	0.015	0.012	0.015
	[0.025]	[0.023]	[0.014]	[0.014]	[0.014]	[0.013]
$NONACTIVE_t$	−0.310***	−0.288**	0.044	0.046	0.026	0.042
	[0.113]	[0.130]	[0.073]	[0.062]	[0.076]	[0.062]
$ELECT_t$	−0.021	−0.012	0.087	0.088	0.079	0.082
	[0.146]	[0.135]	[0.065]	[0.070]	[0.059]	[0.064]
Obs.	189	189	189	189	222	222
Number of Countries	19	19	19	19	27	27
R-squared	0.23	0.26	0.22	0.22	0.17	0.20
Hansen Statistic[a]	2.93		2.87		0.04	
(P-value)	(0.09)		(0.24)		(0.85)	

Notes: *, ***, and *** are significant at 10, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses.

(2), which include as dependent variables structural taxes and social contribution, for the panel of nineteen EU countries. For this regression and the following ones, the Hansen's test fails to reject the null hypothesis of overidentifying restrictions, indicating the validity of the chosen instruments. The estimated coefficient on output gap, $\hat{\beta}$, is not statistically significant, suggesting that changes in revenue items have behaved in an acyclical way. When legislated changes are considered, results on cyclical sensitivity turn out to be remarkably different. In fact, as shown in columns 3 and 4 (for the EU-19 sample) and columns 5 and 6 (for the EU-27 sample), it is found that changes in taxes and social contribution laws are highly responsive to cyclical conditions. In particular, the negative sign of the parameter $\hat{\beta}$ indicates that changes in legislation have been implemented in a *procyclical way; i.e., governments have tended to reduce tax pressure and cut social contributions during booms while increase revenues in recessions.* The effects are sizable: an increase (decrease) of 1 percentage point in $Y_{i,t}$ induces policymakers to cut (raise) taxes and social contributions by about 0.12 to 0.16 percent (of nominal trend GDP), thereby contributing to a further overheating (contraction) of the economic activity.[24]

Regarding the other parameters, the estimated coefficients $\hat{\gamma}$ tend to be significant and have the expected negative sign. This implies that regardless of the measure considered, discretionary fiscal policy actions show a certain degree of persistence. The coefficients on debt, $\hat{\phi}$, are not statistically significant. Nevertheless, this result does not necessarily imply that fiscal decisions are not sensitive to debt developments, as it cannot be excluded that debt stabilization is achieved mainly through expenditure-based policies. Analyzing the significance of controls $\mathbf{Z}_{i,t}$, it is found that the variable *EMU15* is always significant and has a negative sign, in line with the possible interpretation proposed above. In particular, it is likely that the incentives to join the monetary union have induced accession

[24]In addition to these effects, procyclicality of fiscal policies may also affect the size of fiscal multipliers. For example, based on a fully fledged open-economy DSGE model and focusing on government spending, Corsetti, Meier, and Müller (2010) show that procyclicality in government expenditure tends to reduce both domestic and cross-border multipliers.

Table 5. Estimates for Overall Taxes and Social Contributions: EMU Countries

	ΔR^{txs}		$\Delta \ell^{txs}$	
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**
	IV/2SLS	*OLS*	*IV/2SLS*	*OLS*
$Output\ gap_t$	0.049	0.072	−0.178**	−0.093**
	[0.100]	[0.079]	[0.078]	[0.046]
$\widetilde{R}^{txs}_{t-1}$	−0.288***	−0.281***	−0.111*	−0.143***
	[0.069]	[0.070]	[0.066]	[0.051]
$Debt_{t-1}$	0.006	0.004	−0.002	0.002
	[0.015]	[0.013]	[0.007]	[0.009]
T	0.057*	0.064**	0.008	0.023
	[0.031]	[0.031]	[0.020]	[0.023]
$NONACTIVE_t$	−0.064	−0.085	−0.004	−0.077
	[0.235]	[0.214]	[0.175]	[0.159]
$ELECT_t$	−0.209**	−0.196	−0.007	0.001
	[0.090]	[0.140]	[0.054]	[0.089]
Obs.	124	127	125	127
Number of Countries	13	15	13	15
R-squared	0.19	0.19	0.19	0.22
Hansen Statistic[a]	5.60		0.74	
(P-value)	(0.13)		(0.39)	

Notes: *, ***, and *** are significant at 10, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses.

countries to increase taxation and consolidate public finances with a view of joining the EMU. The demographic variable *NONACTIVE* seems to play a role only for $\Delta R^{txs}_{i,t}$, while it is never statistically significant when legislated changes are considered. Finally, it is found that the electoral cycle does not seem to be relevant in determining discretionary fiscal decisions.[25]

As an additional set of results, table 5 presents estimates for EMU countries. By comparing results with those obtained using the

[25]See also Turrini (2008) for similar results on the electoral cycle.

full sample (and reported in table 4), previous conclusions remain quantitatively and qualitatively unaltered. In fact, there seem to be no sizable differences between EU-19, EU-27, and EMU countries in the responsiveness of legislated changes to the economic cycle.

5.1 Estimating the Cyclical Sensitivity of Revenue-Side Categories

This section is devoted to examining whether the cyclical sensitivity varies across revenue categories and investigating to what extent each of these budgetary items determines the procyclical nature of revenue-side discretionary fiscal policies. To that end, panel regressions (2)–(3) are reestimated for each of the four revenue budgetary items considered. Results are reported in tables 6–9. As in tables 4 and 5, both IV and OLS estimates are provided.

Focusing on the cyclical sensitivity parameter, it is found that results for direct taxes payable by households and social contributions are qualitatively the same as those obtained for overall taxes and social contributions. In fact, from tables 6 and 9, it can be noticed that while their cyclically adjusted components ($\Delta R_{i,t}^{dth}$ and $\Delta R_{i,t}^{sct}$) are not responsive to the economic cycle, legislation changes ($\Delta \ell_{i,t}^{dth}$ and $\Delta \ell_{i,t}^{sct}$) behave in a *procyclical way*. However, some differences emerge as regards the size of the coefficients $\hat{\beta}$. As shown in columns 3 through 6 of table 6, it seems that legislated changes impacting on *taxes payable by households are highly procyclical*. In particular, an increase of 1 percentage point in the output gap induces policymakers to cut these taxes by about 0.10 to 0.11 percent of nominal trend GDP. On the other side, the degree of procyclicality of legislated changes associated to social contributions is lower, as suggested by results in columns 3 through 6 of table 9). In fact, an increase of 1 percentage point in the output gap reduces social contributions by around 0.04 percent. This effect, although rather small, is still statistically significant.

Results from regression estimated based on taxes by corporations and indirect taxes point to different conclusions. In particular, while results on fiscal rules estimated using $\Delta R_{i,t}^{dte}$ point clearly towards countercyclicality (see columns 1 and 2 of table 7), estimates based

Table 6. Estimates for Direct Taxes Payable by Households (dth)

Dependent Variable	ΔR^{dth} (1)	ΔR^{dth} (2)	$\Delta \ell^{dth}$ (3)	$\Delta \ell^{dth}$ (4)	$\Delta \ell^{dth27}$ (5)	$\Delta \ell^{dth27}$ (6)
	IV/2SLS	*OLS*	*IV/2SLS*	*OLS*	*IV/2SLS*	*OLS*
$Output\ gap_t$	0.036	−0.025	−0.115***	−0.060***	−0.107***	−0.050***
	[0.031]	[0.023]	[0.032]	[0.013]	[0.030]	[0.012]
$\widetilde{R}^{dth}_{t-1}$	−0.405***	−0.372***	−0.106	−0.136	−0.116	−0.142*
	[0.114]	[0.124]	[0.096]	[0.088]	[0.089]	[0.081]
$Debt_{t-1}$	−0.001	−0.002	−0.004	−0.003	−0.005	−0.003
	[0.007]	[0.006]	[0.005]	[0.004]	[0.005]	[0.004]
EMU15	−0.444*	−0.402*	0.020	−0.018	0.017	−0.023
	[0.246]	[0.239]	[0.111]	[0.131]	[0.113]	[0.134]
T	0.007	0.005	−0.017	−0.016	−0.017	−0.015
	[0.020]	[0.014]	[0.014]	[0.013]	[0.014]	[0.013]
$NONACTIVE_t$	−0.084	−0.102	−0.014	0.003	−0.017	0.003
	[0.081]	[0.068]	[0.053]	[0.039]	[0.052]	[0.039]
$ELECT_t$	−0.009	−0.010	−0.014	−0.012	0.001	0.002
	[0.065]	[0.058]	[0.046]	[0.045]	[0.043]	[0.041]
Obs.	189	189	190	190	222	222
Number of Countries	19	19	19	19	27	27
R-squared	0.17	0.21	0.10	0.16	0.07	0.14
Hansen Statistic[a]	7.89		1.40		1.21	
(P-value)	(0.10)		(0.24)		(0.27)	

Notes: *, ***, and *** are significant at 10, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses.

Table 7. Estimates for Direct Taxes Payable by Corporations (*dte*)

Dependent Variable	ΔR^{dte}		$\Delta \ell^{dte}$		$\Delta \ell^{dte27}$	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	IV/2SLS	*OLS*	*IV/2SLS*	*OLS*	*IV/2SLS*	*OLS*
$Output\ gap_t$	0.095**	0.111***	0.005	0.008	0.005	0.008
	[0.044]	[0.024]	[0.016]	[0.010]	[0.017]	[0.009]
$\widetilde{R}^{dte}_{t-1}$	−0.327***	−0.339***	−0.085***	−0.087***	−0.085**	−0.086***
	[0.073]	[0.078]	[0.032]	[0.032]	[0.035]	[0.030]
$Debt_{t-1}$	0.000	0.000	−0.004	−0.004	−0.004	−0.004
	[0.006]	[0.006]	[0.003]	[0.003]	[0.002]	[0.002]
$EMU15$	−0.144	−0.15	−0.108	−0.109	−0.111	−0.111
	[0.238]	[0.254]	[0.104]	[0.081]	[0.107]	[0.082]
T	0.029**	0.030**	0.004	0.004	0.004	0.004
	[0.013]	[0.012]	[0.006]	[0.006]	[0.007]	[0.006]
$NONACTIVE_t$	−0.053	−0.05	0.005	0.006	0.006	0.006
	[0.063]	[0.052]	[0.022]	[0.023]	[0.025]	[0.023]
$ELECT_t$	−0.066	−0.064	0.000	0.000	−0.004	−0.004
	[0.059]	[0.065]	[0.029]	[0.030]	[0.027]	[0.027]
Obs.	189	189	190	190	222	222
Number of Countries	19	19	19	19	27	27
R-squared	0.22	0.22	0.10	0.10	0.10	0.10
Hansen Statistic[a]	1.69		0.22		0.26	
(P-value)	(0.19)		(0.64)		(0.61)	

Notes: *, ***, and *** are significant at 0, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses.

Table 8. Estimates for Indirect Taxes (*itx*)

Dependent Variable	ΔR^{itx}		$\Delta \ell^{itx}$		$\Delta \ell^{itx27}$	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	IV/2SLS	*OLS*	*IV/2SLS*	*OLS*	*IV/2SLS*	*OLS*
$Output\ gap_t$	−0.122***	−0.049	−0.014	−0.012	−0.016	−0.014
	[0.035]	[0.032]	[0.016]	[0.014]	[0.015]	[0.013]
$\widetilde{R}^{itx}_{t-1}$	−0.301***	−0.289***	−0.037*	−0.035**	−0.039*	−0.036**
	[0.033]	[0.077]	[0.020]	[0.015]	[0.020]	[0.015]
$Debt_{t-1}$	0.008	0.009	0.009**	0.009**	0.008**	0.008**
	[0.005]	[0.006]	[0.004]	[0.004]	[0.004]	[0.003]
EMU15	−0.639	−0.720**	−0.347**	−0.350**	−0.344**	−0.346**
	[0.388]	[0.316]	[0.137]	[0.144]	[0.145]	[0.145]
T	0.007	0.011	0.009	0.009	0.008	0.008
	[0.014]	[0.017]	[0.010]	[0.013]	[0.011]	[0.013]
$NONACTIVE_t$	−0.234***	−0.204*	−0.058	−0.054	−0.059	−0.056
	[0.078]	[0.109]	[0.047]	[0.064]	[0.048]	[0.064]
$ELECT_t$	−0.041	−0.035	−0.013	−0.01	−0.018	−0.016
	[0.085]	[0.087]	[0.036]	[0.040]	[0.034]	[0.037]
Obs.	189	189	188	190	220	222
Number of Countries	19	19	19	19	27	27
R-squared	0.22	0.25	0.08	0.08	0.08	0.08
Hansen Statistic[a]	0.14		8.43		6.83	
(P-value)	(0.71)		(0.08)		(0.15)	

Notes: *, ***, and *** are significant at 10, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses.

Table 9. Social Contribution (*sct*)

Dependent Variable	ΔR^{sct} (1)	ΔR^{sct} (2)	$\Delta \ell^{sct}$ (3)	$\Delta \ell^{sct}$ (4)	$\Delta \ell^{sct27}$ (5)	$\Delta \ell^{sct27}$ (6)
	IV/2SLS	*OLS*	*IV/2SLS*	*OLS*	*IV/2SLS*	*OLS*
$Output\ gap_t$	−0.002	0.012	-0.043^{*}	-0.030^{*}	-0.041^{**}	-0.028^{*}
	[0.036]	[0.027]	[0.022]	[0.016]	[0.020]	[0.015]
$\widetilde{R}^{sct}_{t-1}$	-0.518^{***}	-0.519^{***}	-0.270^{**}	-0.272^{**}	-0.269^{**}	-0.269^{**}
	[0.109]	[0.157]	[0.133]	[0.118]	[0.126]	[0.110]
$Debt_{t-1}$	0.006	0.006	0.003	0.003	0.003	0.003
	[0.010]	[0.006]	[0.007]	[0.005]	[0.007]	[0.004]
EMU15	−0.041	−0.055	0.087	0.074	0.083	0.071
	[0.113]	[0.120]	[0.085]	[0.076]	[0.084]	[0.075]
T	0.016	0.017	0.013	0.014	0.013	0.014
	[0.020]	[0.014]	[0.015]	[0.011]	[0.016]	[0.011]
$NONACTIVE_t$	0.017	0.022	0.054	0.058	0.054	0.059
	[0.096]	[0.066]	[0.063]	[0.045]	[0.064]	[0.045]
$ELECT_t$	0.083^{*}	0.084	0.097^{***}	0.098^{**}	0.086^{***}	0.087^{**}
	[0.046]	[0.064]	[0.032]	[0.041]	[0.029]	[0.037]
Obs.	189	189	190	190	222	222
Number of Countries	19	19	19	19	27	27
R-squared	0.27	0.27	0.24	0.25	0.23	0.23
Hansen Statistic[a]	0.73		2.73		2.31	
(P-value)	(0.39)		(0.44)		(0.51)	

Notes: *, ***, and *** are significant at 10, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses.

on legislated changes on taxes payable by enterprises always indicate acyclicality.

Turning to indirect taxes, results from table 8 show that changes in cyclically adjusted indirect taxes respond in a procyclical way to the cycle, while when legislated changes are considered, results indicate that they are not responsive to economic developments.

Overall, these findings suggest that the observed acyclicality of the changes in the cyclically adjusted overall taxes and social contributions seems to be explained by acyclicality of both taxes payable by households and social contributions and, on the other hand, by the possibility that the countercyclicality found for taxes payable by corporations offsets the procyclical behavior of changes in indirect taxes. At the same time, as regards legislated revenue changes, *their procyclical behavior can be interpreted as driven by a significant procyclicality of changes in taxes payable by households and (though to a minor extent) social contributions.*

Finally, by analyzing the statistical significance of the other explanatory variables in each specification, it emerges that, with the exception of legislated changes in direct taxes payable by households, the coefficient $\hat{\gamma}$ is always significant and with the expected negative sign. On the contrary, the parameter $\hat{\phi}$ has been found to be significant and with the expected positive sign only in the regressions of legislated changes on indirect taxes. This suggests that *indirect taxes are the only revenue subcomponent that seems to have been used to stabilize the debt-output ratio.* One possible interpretation of this is that increases in indirect taxes can be phased in rapidly. In addition, they are generally perceived to be less problematic, from a "social" point of view, than, for instance, increases in direct taxes on households or corporations. For these reasons, in situations in which the sustainability of public finances is perceived to be at risk, governments might decide to rely primarily on these measures for urgent consolidation interventions. For the same reasons, as a response to the severe deterioration of public finances that followed the 2008–09 crisis, many governments adopted consolidation packages in which increases in indirect taxation had a prominent role.

Regarding the importance of control variables, $\mathbf{Z}_{i,t}$, it is found that only for indirect taxes, and in the equation of cyclically adjusted taxes payable by households, the dummy *EMU15* is statistically significant.

5.2 Testing for Asymmetries

In this section the possible presence of non-linearity in the way discretionary fiscal policies react to the evolution of the economic cycle and to debt accumulation is explored.

In order to test whether there is a significant difference in the policymaker's response along the business cycle, the notion of "*good times*" as periods of positive output gap and "*bad times*" as years in which the output gap is negative, is used.[26] Asymmetries in the discretionary fiscal response may also be at play as regards the level of government debt. In the framework of the European Monetary Union (EMU) and the Stability Growth Pact (SGP), for instance, fiscal authorities may pursue more sustainable policies, attempting to reduce government debt, when the 60 percent ceiling is approached or exceeded. More generally, it can be expected that governments are more concerned about the sustainability of public finances when the government debt is high rather than when it is low.

In order to account for asymmetric effects, the following panel regressions are estimated:

$$\Delta R^j_{i,t} = \beta^+ Y_{i,t} d_1 + \beta^- Y_{i,t}(1-d_1) + \gamma \widetilde{R}^j_{i,t-1} + \phi^+ B_{i,t-1} d_2 + \phi^- B_{i,t-1}(1-d_2) + \mathbf{Z}'_{i,t}\boldsymbol{\delta} + \alpha_i + \varepsilon_{i,t} \quad (4)$$

$$\Delta \ell^j_{i,t} = \beta^+ Y_{i,t} d_1 + \beta^- Y_{i,t}(1-d_1) + \gamma \widetilde{R}^j_{i,t-1} + \phi^+ B_{i,t-1} d_2 + \phi^- B_{i,t-1}(1-d_2) + \mathbf{Z}'_{i,t}\boldsymbol{\delta} + \alpha_i + \varepsilon_{i,t}, \quad (5)$$

where

$$d_1 = \begin{cases} 1 & \text{if } Y_{i,t} > 0 \\ 0 & \text{if } Y_{i,t} \leq 0 \end{cases} \qquad d_2 = \begin{cases} 1 & \text{if } B_{i,t} > 60 \\ 0 & \text{if } B_{i,t} \leq 60. \end{cases}$$

The variables $Y_{i,t}$ and $B_{i,t-1}$ from the baseline model (3) have been replaced in equations (4)–(5) with two sets of regressors. The first set

[26] The same approach is followed by many other papers on fiscal policy rules (see, e.g., Gavin and Perotti 1997; OECD 2003; European Commission 2004; Turrini 2008).

of regressors is constructed by interacting $Y_{i,t}$ with a dummy indicator which equals one (zero) if the output gap is positive (negative) and their associated coefficients are β^+ and β^-, respectively. The second one is computed by interacting $B_{i,t-1}$ with a dummy indicator which equals one (zero) when the debt at time $t-1$ is above (below) the 60 percent ceiling. The coefficients associated to these regressors are denoted as ϕ^+ and ϕ^-.

Tables 14 and 15 (see the online appendix at www.ijcb.org) report estimates of the panel models (4) and (5), respectively, when we use the sample of nineteen EU countries. In addition, table 16 of the online appendix shows results from model (5) when we consider the group of twenty-seven EU countries. Specifically, for each budget category, we report the estimates from three possible nested specifications of our models. In particular, the first column shows estimates accounting only for asymmetric responses of discretionary fiscal policy indicators to the economic cycle. The second one considers the possibility of asymmetries linking discretionary measures to debt developments. Finally, the third column reports estimates from the nested model, where the reaction to both the output gap and debt development can be characterized by two regimes.

To formally test for asymmetric effects in these models, some Wald tests are proposed. The first one tests the null hypothesis $\beta^+ = \beta^-$, i.e., equality of coefficients associated with positive and negative output gaps. The second one tests the linear restriction $\beta^+ = \beta^- = 0$, i.e., equal coefficients, and jointly not statistically different from zero. These two tests are also carried out for the parameters ϕ^+ and ϕ^-, which measure the reaction of discretionary fiscal policies to debt developments when the debt ratio is respectively above and below the 60 percent threshold. Results are reported at the bottom of each table.

Table 14 indicates that when we consider cyclically adjusted indicators as a measure of discretionary fiscal policy, changes in taxes and social contributions seem to be disconnected from the economic cycle and do not behave in an asymmetric way (in fact, both the parameter estimates $\hat{\beta}^+$ and $\hat{\beta}^-$ are not statistically significant, and the Wald test rejects the null that $\hat{\beta}^+ = \hat{\beta}^-$). Analyzing the behavior of the revenue-side budgetary subcomponents, we find that changes in direct taxes by households and social contributions remain acyclical in each phase of the cycle. At the same time, changes in taxes payable by enterprises have responded in a countercyclical way, especially during bad times ($\hat{\beta}^- = 0.15$) rather than good times ($\hat{\beta}^+ = 0.08$). However, the hypothesis that these two coefficients are statistically different is rejected by the reported Wald statistic. Finally, asymmetric effects due to debt development do not seem to exist.

Tables 15 and 16 show that legislated changes in taxes and social contributions occur in a strongly procyclical way in "good times" and in an acyclical way during "bad times." This evidence holds in particular for the sample of nineteen countries, for which the Wald test rejects (at the 10 percent level) the null hypothesis that $\beta^+ = \beta^-$. At the same time, as regards the panel of twenty-seven EU countries, while the coefficient associated with positive output gaps is again negative and significant at the 99 percent level, and $\hat{\beta}^-$ is not significantly different from zero, there is no strong statistical evidence that the two parameters are different from each other (although the Wald statistic associated is very close to the 10 percent critical level).

As regards the legislation process related to each single revenue category, changes in taxes payable by households seem to be strongly procyclical when the output gap is positive. In particular, while the parameter $\hat{\beta}^+$ is equal to around -0.10 for the subgroup of nineteen countries, it decreases to -0.07 for the overall country sample. A possible interpretation of this result is that the fiscal position of governments tends to improve during economic upswings, at least partially on account of the fiscal drag that operates in progressive tax systems (e.g., in European ones). As a consequence, EU governments may have reacted to these developments by lowering personal income taxes, as a compensation of losses in disposable income due to the fiscal drag.

Our results also indicate that asymmetries related to the reaction of legislative changes to debt developments do not seem to be at play. In fact, although the coefficients ϕ^+ and ϕ^-, in the equation of legislated change in indirect taxes, are both significant and positive (as expected), their difference is not statistically different from zero.

Finally, estimates indicate that asymmetries due to economic and debt development do not seem to occur simultaneously.

6. Robustness Checks

The robustness of the empirical findings presented above is explored in different directions. First, we test whether our results are robust to various alternative business-cycle indicators. In particular, we reestimate our baseline models (2)–(3) by replacing the output gap as reported by NCBs with the one published by the European Commission in its AMECO database. In addition, we check the sensitivity of our results by using the simple growth rate of real GDP as an alternative measure of the cycle. In fact, it cannot be excluded that fiscal policymakers do not base their decisions on the output gap—due, for example, to the unreliability of potential output estimates—and might simply use real GDP growth as a proxy for the state of the economy. Estimates of models (2)–(3) for overall taxes and social contributions obtained by these alternative measures of cyclical conditions are reported in table 10, along with our baseline estimates. Results from the use of the AMECO output gap clearly point to results in line with the baseline: overall taxes and social contributions appear to be procyclical when legislated changes are considered, but acyclical when structural indicators are used (columns 2, 5, and 8). When GDP growth is used instead, regressions based on structural indicators yield to a statistically insignificant $\hat{\beta}$. However, when legislated revenue changes are incorporated, $\hat{\beta}$ becomes negative, indicating again procyclicality, as in the baseline case (columns 3, 6, and 9).

A further robustness test consists of examining the effects of including in equations (2) and (3) additional control variables. Specifically, we are interested in testing whether the omission of variables potentially correlated to ΔR^{txs} or $\Delta \ell^{txs}$ may impact significantly on the relation between discretionary fiscal policy behavior and economic fluctuations. In particular, we control for changes in total and primary government spending and changes in the inflation rate. In addition, we check whether the ideology behind national political parties in power influences the degree of fiscal policy discretion.

By controlling for changes in government spending, we account for the possibility that changes in taxes and social contributions may

Table 10. Estimates Using Different Measures of Business Cycle

	ΔR^{txs}			$\Delta \ell^{txs}$			$\Delta \ell^{txs27}$		
	$Outgap_t$		$Growth_t$	$Outgap_t$		$Growth_t$	$Outgap_t$		$Growth_t$
	BASE-LINE- (1)	AMECO (2)	(3)	BASE-LINE- (4)	AMECO (5)	(6)	BASE-LINE- (7)	AMECO (8)	(9)
β	−0.086	−0.076	0.151	−0.118***	−0.122***	−0.113*	−0.156***	−0.112***	−0.088
	[0.076]	[0.053]	[0.118]	[0.033]	[0.036]	[0.067]	[0.041]	[0.033]	[0.062]
$\widetilde{R}_{t-1}$	−0.271***	−0.276***	−0.272***	−0.056	−0.060*	−0.090**	−0.054	−0.067*	−0.089**
	[0.041]	[0.040]	[0.051]	[0.037]	[0.036]	[0.042]	[0.034]	[0.036]	[0.040]
$Debt_{t-1}$	0.007	0.006	−0.003	0.002	0.000	0.012**	−0.001	−0.001	0.009
	[0.010]	[0.010]	[0.014]	[0.005]	[0.005]	[0.006]	[0.005]	[0.005]	[0.005]
$EMU15$	−1.379*	−1.352*	−1.447**	−0.486*	−0.431	−0.542*	−0.458*	−0.438	−0.536*
	[0.716]	[0.711]	[0.668]	[0.256]	[0.279]	[0.288]	[0.257]	[0.292]	[0.295]
T	0.062**	0.064**	0.059**	0.014	0.016	0.025	0.012	0.017	0.021
	[0.025]	[0.026]	[0.027]	[0.014]	[0.015]	[0.016]	[0.014]	[0.015]	[0.015]
$NONACTIVE_t$	−0.310***	−0.332***	−0.191	0.044	0.005	−0.013	0.026	−0.001	0.002
	[0.113]	[0.115]	[0.174]	[0.073]	[0.084]	[0.092]	[0.076]	[0.085]	[0.091]
$ELECT_t$	−0.021	−0.023	0.020	0.087	0.082	0.070	0.079	0.076	0.069
	[0.146]	[0.146]	[0.148]	[0.065]	[0.062]	[0.068]	[0.059]	[0.057]	[0.061]
Obs.	189	189	189	189	189	189	222	222	222
Number of Countries	19	19	19	19	19	19	27	27	27
Centered R-squared	0.23	0.23	0.28	0.22	0.21	0.07	0.17	0.20	0.09
Hansen Statistic[a]	2.93	1.95	1.72	2.87	0.75	5.18	0.04	0.91	5.83
(P-value)	(0.09)	(0.16)	(0.19)	(0.24)	(0.39)	(0.02)	(0.85)	(0.34)	(0.02)

Notes: *, ***, and *** are significant at 10, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses. Growth rate instrumented using its own lagged values.

be driven by changes in expenditures.[27] In order to test for such hypothesis, we use two measures of spending changes: the change in the total expenditure and the changes in the primary expenditure (i.e., total expenditure less interest payments). Both series are taken from the AMECO database and are divided by nominal trend GDP. Lagged values of these series are used since a reverse causality (expenditure are adjusted to respond to revenue dynamics) cannot be a priori excluded.

In addition, experiments based on the inclusion of changes in the inflation rate are carried out. This may reflect the possibility that taxes, especially the indirect ones (VAT), may be lowered to mitigate the impact of increases in prices on consumption goods. Similarly to spending, lagged values of inflation changes are used, given that taxes may impact on prices in the same period.

Finally, the possibility that the "political orientation" of coalitions in power in the countries considered may affect the way discretionary fiscal policies are decided is tested. To that aim, the information from the Dataset of Political Indicators (DPI), provided by the World Bank, is used. In particular, the variable *EXECRCL* distinguishes between right, center, and left parties according to the government orientation with respect to economic policy. Based on this information, for each of the twenty-seven countries, the succession of political parties in charge since 1998 is reconstructed. Specifically, three dummy variables which enter as regressors in the baseline model are introduced, namely (i) *Right* takes the value of one during the years in which parties in charge can be defined as "conservative," Christian democratic, or right-wing, and zero otherwise; (ii) *Center* takes the value of one when the parties can best be described as "centrist," and zero otherwise; and finally (iii) *Left* takes value of one for parties that are defined as communist, socialist, social democratic, or left-wing, and zero otherwise.

As table 11 shows, none of these additional variables alter significantly the baseline results. In particular, spending developments do not seem to be associated with revenue decisions. Inflation appears

[27]As an alternative, one may also want to include the government deficit, instead of spending, as an additional control variable. However, the deficit clearly incorporates revenues, and therefore problems related to endogeneity are very likely to arise.

Table 11. Testing for Omitted Controls

	ΔR^{txs}	$\Delta \ell^{txs}$	ΔR^{txs}	$\Delta \ell^{txs}$	ΔR^{txs}	$\Delta \ell^{txs}$	ΔR^{txs}	$\Delta \ell^{txs}$
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Output gap$_t$	−0.087	−0.118***	−0.086	−0.117***	−0.076	−0.120***	−0.092	−0.120***
	[0.072]	[0.032]	[0.072]	[0.032]	[0.076]	[0.032]	[0.079]	[0.033]
$\widetilde{R}_{t-1}$	−0.275***	−0.057*	−0.275***	−0.055	−0.290***	−0.053	−0.285***	−0.073*
	[0.045]	[0.034]	[0.062]	[0.037]	[0.048]	[0.036]	[0.048]	[0.039]
$Debt_{t-1}$	0.006	0.002	0.005	0.002	0.009	0.002	0.008	0.003
	[0.010]	[0.005]	[0.010]	[0.005]	[0.011]	[0.005]	[0.010]	[0.005]
EMU15	−1.403*	−0.492**	−1.408**	−0.484*	−1.319*	−0.493*	−1.273*	−0.365*
	[0.719]	[0.235]	[0.656]	[0.251]	[0.693]	[0.268]	[0.756]	[0.211]
T	0.063**	0.014	0.064**	0.014	0.068**	0.013	0.068**	0.020
	[0.025]	[0.013]	[0.025]	[0.014]	[0.028]	[0.015]	[0.026]	[0.013]
$NONACTIVE_t$	−0.332***	0.039	−0.330**	0.046	−0.368***	0.050	−0.334***	0.008
	[0.126]	[0.069]	[0.134]	[0.073]	[0.128]	[0.076]	[0.114]	[0.085]
$ELECT_t$	−0.027	0.086	−0.026	0.088	−0.024	0.088	−0.029	0.067
	[0.140]	[0.066]	[0.135]	[0.066]	[0.140]	[0.064]	[0.147]	[0.062]

(continued)

Table 11. (Continued)

Dependent Variable	ΔR^{txs} (1)	$\Delta \ell^{txs}$ (2)	ΔR^{txs} (3)	$\Delta \ell^{txs}$ (4)	ΔR^{txs} (5)	$\Delta \ell^{txs}$ (6)	ΔR^{txs} (7)	$\Delta \ell^{txs}$ (8)
$\Delta Spending^{(1)}_{t-1}$	−0.033	−0.008						
	[0.048]	[0.020]						
$\Delta Spending^{(2)}_{t-1}$			−0.036	0.003				
			[0.050]	[0.022]				
$\Delta Inflation_{t-1}$					−0.060	0.007		
					[0.050]	[0.021]		
Left							0.024	0.228
							[0.299]	[0.304]
Center							−0.305	−0.199
							[0.310]	[0.313]
Right							−0.057	0.189
							[0.142]	[0.265]
Obs.	189	189	189	189	189	189	189	189
Number of Countries	19	19	19	19	19	19	19	19
R-squared	0.23	0.22	0.23	0.22	0.25	0.22	0.23	0.25
Hansen Statistic[a]	2.30	3.51	3.06	3.17	4.86	2.75	2.75	3.12
(P-value)	(0.13)	(0.17)	(0.08)	(0.21)	(0.03)	(0.25)	(0.10)	(0.21)

Notes: *, ***, and *** are significant at 10, 5, and 1 percent, respectively. Robust HAC standard error is in square brackets.
[a]Overidentifying restrictions test and p-value are in parentheses.
(1) Excluding interest rates.
(2) Total expenditure.

to be not significant in the regression for ΔR^{txs} and $\Delta \ell^{txs}$. Finally, the political orientation of ruling parties plays no role in accounting for revenue developments. In all these experiments, the core findings related to the coefficient on the output gap are unaffected, pointing to the fact the baseline model seems to be well specified.

7. Conclusion

This paper proposes a new approach to study the cyclical sensitivity of discretionary revenue policies in EU countries, based on measures of legislated revenue changes as provided by EU national central banks in the framework of the "ESCB disaggregated approach." The findings from this analysis, which focuses on the period 1998–2008, can be summarized as follows.

First, it emerges that legislated changes on the revenue side of the government budget seem to have been used in a strongly procyclical way, in contrast to what emerges based on cyclically adjusted indicators, which point to acyclicality. This result holds for the whole set of EU countries, but also when countries belonging to the EMU are analyzed separately.

Second, the observed procyclicality of discretionary revenue appears to be mainly driven by direct taxes paid by households and—to a minor extent—social security contributions. In particular, the former have been significantly reduced during booming economic times, thus probably contributing to overheating economies which were already in an expansionary phase. These developments might be driven by deliberate decisions of governments to lower personal income taxes in booming times, as a (at least partial) compensation of losses in disposable income generated by the fiscal drag, in progressive tax systems.

Finally, indirect taxes are the only revenue item that seems to react to the government debt-to-GDP ratio.

These findings suggest that a sound identification of the discretionary component of fiscal policy is of key importance to appropriately assess how fiscal policies have been implemented over the economic cycle. This is, in particular, relevant for the EU fiscal surveillance process where, currently, cyclically adjusted indicators have a prominent role in the analysis of fiscal policymaking.

Appendix

Table 12. Changes in the Structural Fiscal Components (as Percentage of Trend GDP) from a DA Table for Italy, as Reported in Kremer et al. (2006)

Increasing +, Decreasing −	1998	1999	2000	2001	2002	2003	2004	1998–2004
Unadjusted Balance	0.3	1.4	1.5	−2.4	−1.7	−1.2	1.1	−1.0
Cyclical Component	0.8	0.4	0.6	1.0	−0.2	−1.4	−1.2	0.1
Temporary Measures	0.0	0.0	0.6	−0.9	0.3	0.0	0.2	0.2
Balance	−0.5	1.0	0.3	−2.6	−1.8	0.3	2.0	−1.3
Interest Payments	−0.3	−0.3	−0.6	−0.5	−0.4	−0.2	−0.1	−2.4
Due to Changes in Average Interest Rate	0.0	−0.1	−0.2	−0.2	−0.3	−0.2	−0.1	−1.2
Due to Changes in Debt Level	−0.2	−0.2	−0.4	−0.3	−0.1	0.0	0.1	−1.1
Primary Balance	−0.8	0.7	−0.3	−3.1	−2.2	0.1	1.9	−3.7
Total Revenue	−0.7	1.3	−0.2	−1.9	−1.7	0.1	1.4	−1.7
Direct Taxes Payable by Corporations	0.0	0.0	0.0	−0.1	−0.6	−0.5	0.2	−1.1
Fiscal Drag	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Decoupling of Base from GDP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	−0.2
Legislation Changes	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2
Residual	0.0	0.0	0.0	−0.1	−0.6	−0.5	0.0	−1.1
Direct Taxes Payable by Households	−0.3	0.0	0.0	−0.6	0.5	0.0	−0.1	−0.5
Fiscal Drag	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
Decoupling of Base from GDP	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.4
Legislation Changes	−0.2	−0.1	−0.2	−0.6	0.3	−0.2	0.0	−1.1
Residual	−0.2	0.0	0.1	−0.1	0.1	0.1	−0.1	0.0
Memo Item: Included in Expenditure	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.2

(continued)

Table 12. (Continued)

Increasing +, Decreasing −	1998	1999	2000	2001	2002	2003	2004	1998–2004
Social Contributions	−0.2	0.8	−0.1	−2.4	−0.7	0.7	0.8	−1.0
Fiscal Drag	−0.2	−0.2	−0.3	−0.3	−0.2	−0.2	−0.1	−1.4
Decoupling of Base from GDP	0.3	0.3	0.1	0.2	0.1	0.0	0.0	0.9
Legislation Changes	−0.1	0.2	0.0	−1.8	−0.1	0.8	0.3	−0.8
Residual	−0.2	0.5	0.1	−0.5	−0.4	0.1	0.7	0.3
Memo Item: Included in Expenditure	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Indirect Taxes	0.0	0.5	−0.1	0.4	−0.4	0.1	0.4	1.0
Fiscal Drag	0.0	0.0	−0.1	−0.1	−0.1	0.0	0.0	−0.3
Decoupling of Base from GDP	0.0	0.0	−0.1	−0.1	−0.1	−0.1	0.0	−0.3
Legislation Changes	0.0	0.1	0.0	0.6	−0.2	0.3	0.2	0.9
Residual	0.0	0.4	0.0	0.0	0.0	−0.1	0.3	0.7
Taxes and Social Contributions Overall	−0.4	1.4	−0.4	−2.6	−1.3	0.3	1.4	−1.6
Fiscal Drag	−0.2	−0.2	−0.3	−0.4	−0.2	−0.2	−0.1	−1.5
Decoupling of Base from GDP	0.4	0.4	0.1	0.1	0.0	−0.1	−0.1	0.8
Legislation Changes	−0.3	0.2	−0.2	−1.8	−0.1	0.8	0.6	−0.8
Residual	−0.3	1.0	0.1	−0.6	−0.9	−0.3	1.0	−0.1
Memo Item: Included in Expenditure	0.1	0.1	0.0	0.0	0.1	0.0	−0.1	0.2

(continued)

Table 12. (Continued)

Increasing +, Decreasing −	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**1998–2004**
Non-Tax-Related Revenue								
of which EU	−0.3	−0.1	0.2	0.7	−0.4	−0.2	0.0	−0.1
Total Primary Expenditure	0.1	0.7	0.1	1.2	0.5	0.0	−0.6	2.0
Social Payments	−0.2	−0.2	−0.1	0.0	0.3	0.2	−0.1	−0.2
of which old-age pensions	0.0	0.0	−0.1	−0.1	0.0	0.0	0.0	−0.1
of which unemployment benefits	0.0	−0.1	−0.1	0.0	−0.1	−0.1	−0.1	−0.3
of which social transfers in kind	0.1	0.1	0.1	0.2	0.5	0.2	0.1	1.3
Subsidies	−0.2	0.1	−0.1	0.0	0.0	−0.1	0.0	−0.2
of which EU	−0.1	0.0	0.0	−0.1	0.0	0.0	0.0	−0.2
Compensation of Employees	0.0	0.2	−0.1	−0.1	0.1	0.1	0.0	0.2
Intermediate Consumption	0.1	0.3	0.00	0.8	0.1	0.0	−0.1	1.1
Government Investment	0.1	0.1	0.1	0.3	0.2	−0.2	−0.3	0.3
Other	0.3	0.2	0.3	0.2	−0.1	0.1	−0.1	0.8
of which EU	0.2	−0.1	0.0	−0.1	−0.2	0.0	0.1	0.0
Memorandum Items								
Health Care	0.0	0.1	0.2	0.2	0.5	0.2	0.1	1.3
Trend Growth of Real GDP	3.0	2.8	2.4	2.1	1.7	1.6	1.5	
Change in GDP Deflator	1.7	1.6	3.9	5.2	3.8	2.5	0.9	
Change in Public Employees	3.1	0.9	0.8	1.8	2.1	0.4	−0.7	

Table 13. Data Set Description (countries and time coverage)

Variables	Data Set 1 (EU-19 Countries) $\Delta R^j, \Delta \ell^j$	Data Set 2 (EU-27 Countries) $\Delta \ell^j$
Austria	00–08	00–08
Belgium	98–08	98–08
Bulgaria		06–08
Cyprus	00–08	00–08
Czech Rep.	98–07	98–08
Denmark		04–08
Estonia		04–08
Finland	98–08	98–08
France	98–08	98–08
Germany	98–08	98–08
Greece	00–08	00–08
Hungary		04–08
Ireland	00–08	00–08
Italy	00–08	98–08
Latvia	98–08	98–08
Lithuania	98–08	98–08
Luxembourg	00–08	00–08
Malta	04–08	04–08
Netherlands	00–08	98–08
Poland	99–08	99–08
Portugal	00–08	00–08
Romania		06–08
Slovakia		04–08
Slovenia	00–08	00–08
Spain	96–08	96–08
Sweden		04–08
United Kingdom		04–08

Notes: The first data set incorporates data on (i) fiscal drag (fd^j), (ii) decoupling of the tax base from GDP (de^j), and (iii) residual component (re^j) and legislated changes ($\Delta\ell^j$). The sum of these four components is equal to the change in structural revenues (ΔR^j). The second data set incorporates only data on $\Delta\ell^j$.

References

Afonso, A., L. Agnello, and D. Furceri. 2010. "Fiscal Policy Responsiveness, Persistence and Discretion." *Public Choice* 145 (3): 503–30.

Alesina, A., and R. Perotti. 1995. "Fiscal Expansions and Fiscal Adjustments in OECD Countries." NBER Working Paper No. 5214.

Auerbach, A. 2002. "Is There a Role for Discretionary Fiscal Policy?" NBER Working Paper No. 9306.

Ballabriga, F., and C. Martinez-Mongay. 2002. "Has EMU Shifted Policy?" Economic Papers of the European Commission, Directorate-General for Economic and Financial Affairs, No. 166.

Beetsma, R., and M. Giuliodori. 2010. "Fiscal Adjustment to Cyclical Developments in the OECD: An Empirical Analysis Based on Real-Time Data." *Oxford Economic Papers* 62 (3): 419–41.

Bohn, H. 1998. "The Behavior of U.S. Public Debt and Deficits." *Quarterly Journal of Economics* 113 (3): 949–63.

Bouthevillain, C., P. Cour-Thimann, G. van den Dool, P. Hernández de Cos, G. Langenus, M. Mohr, S. Momigliano, and M. Tujula. 2001. "Cyclically Adjusted Budget Balances: An Alternative Approach." ECB Working Paper No. 77.

Canova, F. 1998. "Detrending and Business Cycle Facts." *Journal of Monetary Economics* 41 (3): 475–512.

Chalk, N. A. 2002. "Structural Balances and All That: Which Indicators to Use in Assessing Fiscal Policy." IMF Working Paper No. 02/101.

Cimadomo, J. 2011. "Fiscal Policy in Real Time." Forthcoming in *Scandinavian Journal of Economics.*

Corsetti, G., A. Meier, and G. Müller. 2010. "Cross-Border Spillovers from Fiscal Stimulus." *International Journal of Central Banking* 6 (1): 5–37.

Darby, J., and J. Melitz. 2008. "Social Spending and Automatic Stabilisers in the OECD." *Economic Policy* 23 (56): 715–56.

Debrun, X., L. Moulin, A. Turrini, J. Ayuso-i-Casals, and M. Kumar. 2008. "Tied to the Mast? National Fiscal Rules in the European Union." *Economic Policy* 23 (54): 297–362.

Eschenbach, F., and L. Schuknecht. 2004. "Budgetary Risks from Real Estate and Stock Markets." *Economic Policy* 19 (39): 313–46.

European Commission. 2004. "Public Finances in EMU — 2004." *European Economy* No. 3.

———. 2006. "Public Finances in EMU — 2006 — Part IV, Fiscal Policy in Good Times." *European Economy* No. 3.

Fatás, A., and I. Mihov. 2003. "The Case for Restricting Fiscal Policy Discretion." *Quarterly Journal of Economics* 118 (4): 1419–47.

———. 2009. "The Euro and Fiscal Policy." NBER Working Paper No. 14722.

Favero, C. 2003. "How Do European Monetary and Fiscal Authorities Behave?" In *Monetary and Fiscal Policies in EMU,* ed. M. Buti (chapter 7). Cambridge University Press.

Forni, L., and S. Momigliano. 2004. "Cyclical Sensitivity of Fiscal Policies Based on Real-Time Data." *Applied Economics Quarterly* 50 (3): 299–326.

Galí, J., and R. Perotti. 2003. "Fiscal Policy and Monetary Integration in Europe." *Economic Policy* 18 (37): 533–72.

Gavin, M., and R. Perotti. 1997. "Fiscal Policy and Saving in Good Times and Bad Times." In *Promoting Savings in Latin America,* ed. R. Hausman and H. Reisen, 149–67. IDB and OECD.

Giavazzi, F., and M. Pagano. 1996. "Non-Keynesian Effects of Fiscal Policy Changes: International Evidence and the Swedish Experience." NBER Working Paper No. 5332.

Giuliodori, M., and R. Beetsma. 2008. "On the Relationship between Fiscal Plans in the European Union: An Empirical Analysis Based on Real-Time Data." *Journal of Comparative Economics* 36 (2): 221–42.

Golinelli, R., and S. Momigliano. 2006. "Real-Time Determinants of Fiscal Policies in the Euro Area." *Journal of Policy Modeling* 28 (9): 943–64.

———. 2009. "The Cyclical Reaction of Fiscal Policies in the Euro Area: The Role of Modelling Choices and Data Vintages." *Fiscal Studies* 30 (1): 39–72.

Jaeger, A., and L. Schuknecht. 2004. "Boom-Bust Phases in Asset Prices and Fiscal Policy Behaviour." IMF Working Paper No. 04/54.

Jaimovich, D., and U. Panizza. 2007. "Procyclicality or Reverse Causality?" Inter-American Development Bank Working Paper No. 599.

Kremer, J., C. Rodrigues Braz, T. Brosens, G. Langenus, S. Momigliano, and M. Spolander. 2006. “A Disaggregated Framework for the Analysis of Structural Developments in Public Finances.” ECB Working Paper No. 579.

Larch, M., and M. Salto. 2005. “Fiscal Rules, Inertia and Discretionary Fiscal Policy.” *Applied Economics* 37 (10): 1135–46.

Larch, M., and A. Turrini. 2009. “The Cyclically Adjusted Budget Balance in EU Fiscal Policymaking. Love at First Sight Turned into a Mature Relationship.” MPRA Paper No. 20594.

Morris, R., C. Rodrigues Braz, F. De Castro, S. Jonk, J. Kremer, S. Linehan, M. R. Marino, C. Schalck, and O. Tkacevs. 2009. “Explaining Government Revenue Windfalls and Shortfalls: An Analysis for Selected EU Countries.” ECB Working Paper No. 1114.

OECD (Organisation for Economic Cooperation and Development). 2003. “Fiscal Stance over the Cycle: Evidence from Panel Analysis.” Appendix to chapter IV, “Fiscal Stance over the Cycle: The Role of Debt, Institutions, and Budget Constraints.” *Economic Outlook* 74 (December): 137–41.

Romer, C., and D. H. Romer. 2009a. “A Narrative Analysis of Post-War Tax Changes.” Appendix to “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks” (June).

———. 2009b. “Do Tax Cuts Starve the Beast? The Effect of Tax Changes on Government Spending.” *Brookings Papers on Economic Activity* (Spring): 139–200.

———. 2010. “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks.” *American Economic Review* 100 (3): 763–801.

Taylor, J. B. 2000. “Reassessing Discretionary Fiscal Policy.” *Journal of Economic Perspectives* 14 (3): 21–36.

Turrini, A. 2008. “Fiscal Policy and the Cycle in the Euro Area: The Role of Government Revenue and Expenditure.” Economic Papers of the European Commission, Directorate-General for Economic and Financial Affairs, No. 323.

Wyplosz, C. 2002. “Fiscal Discipline in EMU: Rules or Institutions?” Mimeo, Graduate Institute for International Studies, Geneva.

———. 2006. “European Monetary Union: The Dark Sides of a Major Success.” *Economic Policy* 21 (46): 208–61.

Online Appendix

Table 14. Asymmetric Behavior of Cyclically Adjusted Measures (nineteen countries)

	ΔR^{txs}			ΔR^{dth}			ΔR^{dte}			ΔR^{itx}			ΔR^{sct}		
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
Output gap$_t(\beta)$	—	0.041	—	—	−0.025	—	—	0.112***	—	—	−0.048	—	—	0.014	—
	—	[0.048]	—	—	[0.023]	—	—	[0.024]	—	—	[0.029]	—	—	[0.027]	—
Output gap$_t^+(\beta^+)$	−0.027	—	−0.012	−0.022	—	−0.023	0.081*		0.084*	−0.093*	—	−0.086	0.015	—	0.022
	[0.108]	—	[0.107]	[0.041]	—	[0.042]	[0.046]		[0.046]	[0.055]	—	[0.055]	[0.043]	—	[0.042]
Output gap$_t^-(\beta^-)$	0.128	—	0.117	−0.029	—	−0.028	0.152**		0.149**	0.012	—	0.007	0.007	—	0.002
	[0.119]	—	[0.118]	[0.045]	—	[0.045]	[0.061]		[0.061]	[0.072]	—	[0.072]	[0.052]	—	[0.051]
$\tilde{R}_{t-1}$	−0.296***	−0.304***	−0.305***	−0.372***	−0.372***	−0.372***	−0.334***	−0.340***	−0.336***	−0.294***	−0.293***	−0.297***	−0.519***	−0.537***	−0.536***
	[0.072]	[0.068]	[0.068]	[0.125]	[0.125]	[0.125]	[0.079]	[0.078]	[0.079]	[0.047]	[0.046]	[0.047]	[0.159]	[0.156]	[0.157]
$Debt_{t-1}(\Phi)$	0.009	—	—	−0.002	—	—	0.001	—	—	0.009	—	—	0.006	—	—
	[0.010]	—	—	[0.006]	—	—	[0.005]	—	—	[0.008]	—	—	[0.006]	—	—
$Debt_{t-1}^+(\Phi^+)$	—	0.000	0.001	—	−0.001	−0.001	—	−0.001	−0.001	—	0.004	0.005	—	0.002	0.002
	—	[0.011]	[0.011]	—	[0.006]	[0.006]	—	[0.007]	[0.007]	—	[0.008]	[0.008]	—	[0.006]	[0.006]
$Debt_{t-1}^-(\Phi^-)$	—	−0.008	−0.007	—	−0.001	−0.001	—	−0.003	−0.003	—	0.000	0.001	—	−0.002	−0.002
	—	[0.013]	[0.013]	—	[0.007]	[0.007]	—	[0.008]	[0.008]	—	[0.010]	[0.010]	—	[0.007]	[0.007]
EMU15	−1.456***	−1.336***	−1.344***	−0.402*	−0.409*	−0.408*	−0.157	−0.129	−0.137	−0.723***	−0.669**	−0.674**	−0.055	−0.003	−0.002
	[0.537]	[0.490]	[0.487]	[0.240]	[0.235]	[0.236]	[0.247]	[0.252]	[0.245]	[0.269]	[0.270]	[0.271]	[0.120]	[0.150]	[0.149]
T	0.064***	0.055**	0.053**	0.005	0.005	0.005	0.028**	0.028**	0.026*	0.008	0.005	0.003	0.017	0.011	0.011
	[0.023]	[0.024]	[0.025]	[0.014]	[0.014]	[0.014]	[0.013]	[0.013]	[0.014]	[0.015]	[0.016]	[0.016]	[0.014]	[0.015]	[0.015]
$NONACTIVE_t$	−0.280**	−0.307**	−0.300**	−0.103	−0.102	−0.102	−0.045	−0.053	−0.048	−0.200***	−0.212***	−0.207***	0.022	0.013	0.012
	[0.131]	[0.131]	[0.133]	[0.068]	[0.068]	[0.068]	[0.051]	[0.052]	[0.052]	[0.068]	[0.068]	[0.068]	[0.067]	[0.065]	[0.065]
$ELECT_t$	−0.020	0.007	0.000	−0.010	−0.012	−0.012	−0.068	−0.060	−0.064	−0.04	−0.025	−0.03	0.084	0.093	0.094
	[0.130]	[0.135]	[0.130]	[0.058]	[0.059]	[0.059]	[0.065]	[0.066]	[0.066]	[0.090]	[0.089]	[0.090]	[0.063]	[0.065]	[0.064]
H_0: $\beta^+ = \beta^-$	0.57	—	0.40	0.01	—	0.01	0.55	—	0.47	0.86	—	0.67	0.01	—	0.07
H_0: $\beta^+ = \beta^- = 0$	0.48	—	0.73	0.64	—	0.64	10.6***	—	10.6***	1.85	—	1.67	0.10	—	0.16
H_0: $\Phi^+ = \Phi^-$	—	3.94**	3.68*	—	0.08	0.07	—	0.85	0.72	—	2.21	2.00	—	2.73*	2.77*
H_0: $\Phi^+ = \Phi^- = 0$	—	2.37	2.28	—	0.09	0.08	—	0.53	0.48	—	1.77	1.72	—	1.68	1.68
Obs.	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189
Centered R-squared	0.26	0.28	0.28	0.21	0.21	0.21	0.23	0.23	0.23	0.25	0.26	0.26	0.27	0.28	0.28

Notes: *, **, and *** indicate OLS estimates significant at 10, 5, and 1 percent, respectively. Robust standard errors are in square brackets.

Table 15. Asymmetric Behavior of Legislated Revenue Changes (nineteen countries)

	$\Delta\ell^{txs}$			$\Delta\ell^{dth}$			$\Delta\ell^{dte}$			$\Delta\ell^{itx}$			$\Delta\ell^{sct}$		
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
Output $gap_t(\beta)$	—	−0.109***	—	—	−0.060***	—	—	0.008	—	—	−0.013	—	—	−0.029*	—
	—	[0.023]	—	—	[0.013]	—	—	[0.010]	—	—	[0.016]	—	—	[0.016]	—
Output $gap_t^+(\beta^+)$	−0.176***	—	−0.181***	−0.097***	—	−0.100***	−0.009	—	−0.010	−0.024	—	−0.027	−0.034	—	−0.031
	[0.050]	—	[0.051]	[0.028]	—	[0.029]	[0.020]	—	[0.020]	[0.031]	—	[0.031]	[0.028]	—	[0.028]
Output $gap_t^-(\beta^-)$	−0.008	—	−0.005	−0.007	—	−0.005	0.031	—	0.031	0.005	—	0.008	−0.026	—	−0.027
	[0.053]	—	[0.053]	[0.026]	—	[0.026]	[0.026]	—	[0.026]	[0.040]	—	[0.040]	[0.031]	—	[0.031]
$\tilde{R}_{t-1}$	−0.059*	−0.055*	−0.056*	−0.132	−0.136	−0.133	−0.084***	−0.087***	−0.084**	−0.036	−0.033	−0.034	−0.272**	−0.278**	−0.278**
	[0.030]	[0.030]	[0.030]	[0.087]	[0.087]	[0.086]	[0.032]	[0.032]	[0.032]	[0.026]	[0.026]	[0.026]	[0.119]	[0.120]	[0.121]
$Debt_{t-1}(\Phi)$	0.003	—	—	−0.003	—	—	−0.003	—	—	0.009**	—	—	0.003	—	—
	[0.006]	—	—	[0.004]	—	—	[0.002]	—	—	[0.004]	—	—	[0.005]	—	—
$Debt_{t-1}^+(\Phi^+)$	—	0.004	0.005	—	−0.002	−0.002	—	−0.003	−0.003	—	0.011**	0.011**	—	0.002	0.002
	—	[0.006]	[0.006]	—	[0.004]	[0.004]	—	[0.002]	[0.002]	—	[0.005]	[0.005]	—	[0.004]	[0.004]
$Debt_{t-1}^-(\Phi^-)$	—	0.006	0.008	—	−0.001	0.000	—	−0.003	−0.003	—	0.013**	0.013**	—	0.000	0.000
	—	[0.007]	[0.007]	—	[0.004]	[0.004]	—	[0.003]	[0.003]	—	[0.005]	[0.005]	—	[0.005]	[0.005]
EMU15	−0.498**	−0.521**	−0.532**	−0.024	−0.031	−0.040	−0.113	−0.109	−0.114	−0.351**	−0.371**	−0.373**	0.074	0.094	0.093
	[0.218]	[0.216]	[0.211]	[0.131]	[0.121]	[0.121]	[0.078]	[0.084]	[0.080]	[0.151]	[0.151]	[0.152]	[0.077]	[0.088]	[0.089]
T	0.010	0.018	0.014	−0.018	−0.014	−0.016	0.003	0.004	0.003	0.008	0.011	0.01	0.014	0.012	0.012
	[0.013]	[0.014]	[0.014]	[0.013]	[0.012]	[0.012]	[0.006]	[0.006]	[0.006]	[0.009]	[0.009]	[0.009]	[0.011]	[0.011]	[0.011]
$NONACTIVE_t$	0.054	0.050	0.061	0.009	0.005	0.011	0.008	0.006	0.009	−0.053	−0.051	−0.049	0.059	0.055	0.055
	[0.061]	[0.064]	[0.063]	[0.038]	[0.039]	[0.038]	[0.023]	[0.023]	[0.023]	[0.038]	[0.038]	[0.038]	[0.045]	[0.045]	[0.044]
$ELECT_t$	0.080	0.083	0.074	−0.016	−0.014	−0.019	−0.002	0.000	−0.002	−0.012	−0.014	−0.016	0.098**	0.102**	0.102**
	[0.070]	[0.071]	[0.070]	[0.044]	[0.047]	[0.046]	[0.030]	[0.030]	[0.030]	[0.050]	[0.050]	[0.050]	[0.040]	[0.042]	[0.042]
H_0: $\beta^+ = \beta^-$	3.27*		3.56*	3.49*	—	3.66**	0.96	—	0.99	0.21	—	0.30	0.03	—	0.00
H_0: $\beta^+ = \beta^- = 0$	10.48***		10.5***	9.88***	—	9.63***	0.76	—	0.76	0.37	—	0.44	1.73	—	1.65
H_0: $\Phi^+ = \Phi^-$	—	0.85	1.19	—	0.68	0.89	—	0.00	0.02	—	1.31	1.39	—	0.97	0.95
H_0: $\Phi^+ = \Phi^- = 0$	—	0.52	0.76	—	0.51	0.57	—	1.30	1.19	—	2.83**	2.89**	—	0.57	0.57
Obs.	189	189	189	190	190	190	190	190	190	190	190	190	190	190	190
Centered R-squared	0.23	0.22	0.24	0.17	0.17	0.18	0.11	0.10	0.11	0.08	0.09	0.09	0.25	0.25	0.25

Notes: *, **, and *** indicate OLS estimates significant at 10, 5, and 1 percent, respectively. Robust standard errors are in square brackets.

Table 16. Asymmetric Behavior of Legislated Revenue Changes (overall sample)

	$\Delta\ell^{txs27}$			$\Delta\ell^{dth27}$			$\Delta\ell^{dte27}$			$\Delta\ell^{itx27}$			$\Delta\ell^{sct27}$		
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
Output gap$_t(\beta)$	—	−0.094***	—	—	−0.050***	—	—	0.008	—	—	−0.015	—	—	−0.028*	—
	—	[0.021]	—	—	[0.012]	—	—	[0.009]	—	—	[0.014]	—	—	[0.015]	—
Output gap$_t^+(\beta^+)$	−0.143***	—	−0.145***	−0.068***	—	−0.070***	−0.005	—	−0.005	−0.037	—	−0.038	−0.027	—	−0.025
	[0.042]	—	[0.042]	[0.025]	—	[0.025]	[0.016]	—	[0.016]	[0.026]	—	[0.026]	[0.024]	—	[0.024]
Output gap$_t^-(\beta^-)$	−0.017	—	−0.015	−0.023	—	−0.021	0.027	—	0.027	0.02	—	0.022	−0.031	—	−0.033
	[0.048]	—	[0.048]	[0.025]	—	[0.025]	[0.024]	—	[0.024]	[0.036]	—	[0.036]	[0.028]	—	[0.028]
$\tilde{R}_{t-1}$	−0.066**	−0.063**	−0.064**	−0.142*	−0.143*	−0.143*	−0.085***	−0.086***	−0.085***	−0.038	−0.034	−0.036	−0.269**	−0.273**	−0.273**
	[0.030]	[0.029]	[0.029]	[0.081]	[0.081]	[0.081]	[0.030]	[0.030]	[0.030]	[0.024]	[0.024]	[0.024]	[0.112]	[0.111]	[0.112]
$Debt_{t-1}(\Phi)$	0.001	—	—	−0.003	—	—	−0.003*	—	—	0.008**	—	—	0.003	—	—
	[0.005]	—	—	[0.004]	—	—	[0.002]	—	—	[0.004]	—	—	[0.004]	—	—
$Debt_{t-1}^+(\Phi^+)$	—	0.002	0.003	—	−0.002	−0.002	—	−0.004*	−0.003	—	0.010**	0.010**		0.002	0.002
	—	[0.005]	[0.005]	—	[0.004]	[0.004]	—	[0.002]	[0.002]	—	[0.004]	[0.004]		[0.004]	[0.004]
$Debt_{t-1}^-(\Phi^-)$	—	0.003	0.004	—	−0.001	−0.001	—	−0.004	−0.003	—	0.011**	0.012**		0.000	0.000
	—	[0.006]	[0.006]	—	[0.004]	[0.004]	—	[0.002]	[0.002]	—	[0.005]	[0.005]		[0.004]	[0.004]
EMU15	−0.501**	−0.517**	−0.525**	−0.026	−0.035	−0.040	−0.114	−0.111	−0.115	−0.349**	−0.366**	−0.370**	0.071	0.089	0.089
	[0.227]	[0.225]	[0.221]	[0.135]	[0.125]	[0.126]	[0.079]	[0.084]	[0.081]	[0.142]	[0.143]	[0.143]	[0.075]	[0.085]	[0.085]
T_i	0.012	0.017	0.014	−0.016	−0.014	−0.015	0.004	0.004	0.004	0.007	0.01	0.009	0.014	0.012	0.012
	[0.013]	[0.013]	[0.013]	[0.013]	[0.012]	[0.012]	[0.006]	[0.006]	[0.006]	[0.008]	[0.008]	[0.008]	[0.011]	[0.011]	[0.011]
$NONACTIVE_{i,t}$	0.049	0.045	0.053	0.006	0.005	0.008	0.008	0.006	0.009	−0.053	−0.053	−0.05	0.059	0.056	0.056
	[0.061]	[0.063]	[0.062]	[0.038]	[0.039]	[0.038]	[0.023]	[0.023]	[0.023]	[0.036]	[0.036]	[0.036]	[0.045]	[0.045]	[0.045]
$ELECT_{i,t}$	0.079	0.079	0.076	0.001	0.000	−0.002	−0.004	−0.004	−0.005	−0.017	−0.019	−0.02	0.087**	0.089**	0.090**
	[0.064]	[0.064]	[0.064]	[0.041]	[0.043]	[0.043]	[0.027]	[0.027]	[0.027]	[0.044]	[0.044]	[0.044]	[0.037]	[0.038]	[0.038]
H_0: $\beta^+ = \beta^-$	2.49	—	2.67*	1.09	—	1.20	0.84	—	0.86	1.05	—	1.20	0.01	—	0.03
H_0: $\beta^+ = \beta^- = 0$	9.96***	—	9.88***	8.17***	—	7.92***	0.72	—	0.73	1.02	—	1.11	1.92	—	1.94
H_0: $\Phi^+ = \Phi^-$	—	0.50	0.67	—	0.56	0.65	—	0.00	0.00	—	1.22	1.38	—	0.94	0.95
H_0: $\Phi^+ = \Phi^- = 0$	—	0.26	0.37	—	0.48	0.49	—	1.58	1.42	—	2.80*	2.98**	—	0.53	0.53
Obs.	222	222	222	222	222	222	222	222	222	222	222	222	222	222	222
Centered R-squared	0.21	0.20	0.22	0.15	0.15	0.15	0.11	0.10	0.11	0.09	0.09	0.09	0.23	0.24	0.24

Notes: *, **, and *** indicate OLS estimates significant at 10, 5, and 1 percent, respectively. Robust standard errors are in square brackets.

Traditional versus New Keynesian Phillips Curves: Evidence from Output Effects*

Werner Roeger[a] and Bernhard Herz[b]
[a]European Commission
[b]University of Bayreuth

We identify a crucial difference between the backward-looking and forward-looking Phillips curve concerning the real output effects of monetary policy shocks. The backward-looking Phillips curve predicts a strict intertemporal trade-off in the case of monetary shocks: a positive short-run response of output is followed by a period in which output is below baseline and the cumulative output effect is exactly zero. In contrast, the forward-looking model implies a positive cumulative output effect. The empirical evidence on the cumulated output effects of money is consistent with the forward-looking model. We also use this method to determine the degree of forward-looking price setting.

JEL Codes: E31, E32, E40.

1. Introduction

The dynamic effects of aggregate demand on output and inflation are still an open question. Even after decades of investigation, this issue is still highly controversial, with only a few definitive answers available. At stake in this discussion are the nature of the business cycle and the appropriate conduct of monetary policy, among others. The discussion has typically been framed within a Phillips-curve

*The views expressed in this paper are those of the authors and not necessarily those of the European Commission. The authors thank seminar participants at Halle, Deutsche Bundesbank, Annual Conference of German Economic Society, and ICMIAF for useful comments as well as Peter McAdam for stimulating discussions. Very helpful comments by two anonymous referees are gratefully acknowledged. Author e-mails: Bernhard.herz@uni-bayreuth.de; Werner.roeger@ec.europa.eu.

setup. The traditional Phillips curve relates inflation to some cyclical indicator and lagged values of inflation. It implies that inflation is a backward-looking phenomenon, produced by adaptive expectations or by price-setting behavior based on a backward-looking rule of thumb. While the traditional Phillips curve is subject to considerable theoretical criticism, the empirical evidence indicates that it describes the post-war inflation in the United States and Europe reasonably well; see, e.g., Rudebusch and Svensson (1999) and Galí, Gertler, and López-Salido (2001).

The so-called New Keynesian Phillips curve relates inflation to the output gap and a "cost-push" effect influenced by expected inflation. Obviously, inflation then is a forward-looking phenomenon caused by staggered nominal price setting as developed by Taylor (1979) and Calvo (1983) or quadratic price adjustment cost (Rotemberg 1982). This model is now widely used in the theoretical analysis of monetary policy (Clarida, Galí, and Gertler, 1999) and has been portrayed as "the closest thing there is to a standard specification" (McCallum 1997). While the new Phillips curve is attractive on theoretical terms, it has not proved to be a simple task to reconcile it with the data. In particular, the relevance of the forward-looking term has been subject to criticism by Fuhrer and Moore (1995), Fuhrer (1997), and Rudd and Whelan (2005).[1]

As an alternative, a number of researchers have investigated a hybrid form of the backward-looking and the forward-looking Phillips curve, e.g., Fuhrer and Moore (1995), Brayton et al. (1997), Roberts (1997), and Christiano, Eichenbaum, and Evans (2005). Empirically, the hybrid Phillips curve also has only limited success. For example, Chadha, Masson, and Meredith (1992), Fuhrer (1997), and Roberts (2001) argue that this specification is not able to replicate inflation dynamics at the quarterly frequency. Evidence from Galí and Gertler (1999), Galí, Gertler, and López-Salido (2001), McAdam and Willman (2004), and Christiano, Eichenbaum, and Evans (2005) is more favorable to the hybrid Phillips curve. However,

[1]McAdam and Willmann (2010) propose a new form of state-dependent Calvo price-setting signal dependent on inflation and aggregate competitiveness which allows to derive a New Keynesian Phillips curve expressed in terms of the actual levels of variables and thus is not regime dependent.

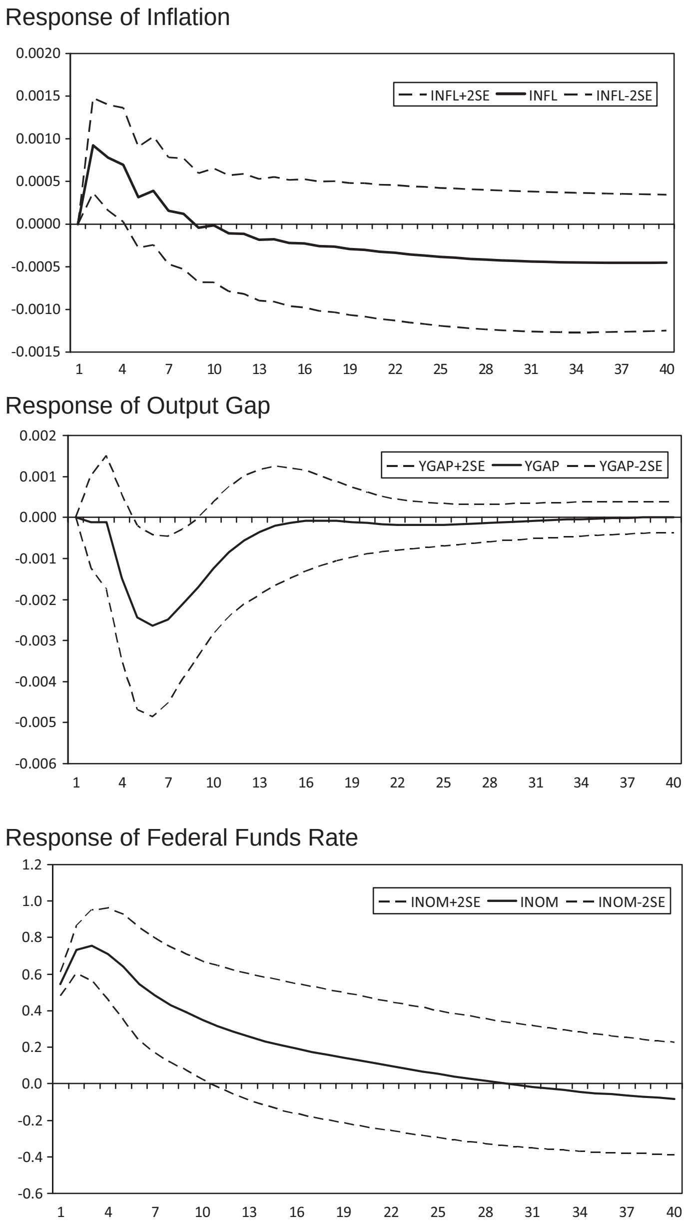
Rudd and Whelan (2006) have recently questioned the relevance of the forward term in the hybrid specification as well.

These empirical tests have so far exclusively focused on the price equation itself and have ignored other dimensions in which the two models—namely the Phillips curve with only backward-looking dynamics and the Phillips curve with (partly) forward-looking inflation expectations—make different predictions. One such dimension is the output response to a monetary policy shock. While both models imply that money is neutral in the long run, they make strikingly different predictions concerning the adjustment of output following a temporary monetary policy shock. Let us consider the effect of a one-period positive shock to the policy rate. In the backward-looking model this leads to a negative output effect due to the fall in demand, a decline in inflation in the current period, and a downward adjustment of inflation expectations in the next period, when the negative demand shock has already disappeared and actual inflation adjusts upwards. This positive inflation surprise leads to an increase of output above its potential during the adjustment process. Taken together, these positive output effects compensate for the initial output loss. In the forward-looking model, in contrast, firms correctly anticipate a rising inflation after the initial drop and respond by lowering prices insufficiently. Therefore, prices are too high over the entire adjustment period, which keeps output below potential and implies that the cumulated output effects are negative. This behavior of output as a response to temporary monetary policy shocks is in accordance with numerous VAR (vector autoregression) studies (e.g., Sims and Zha 2006, Castelnuovo 2009). Figure 1 illustrates the VAR evidence for the United States by showing the impulse-response functions of the output gap, inflation, and the federal funds rate to a monetary policy shock.

As will be shown in this paper, this feature of impulse responses for output is consistent with highly forward-looking price setting and cannot be generated with backward-looking Phillips curves. This dynamic adjustment pattern thus constitutes important empirical evidence in favor of the New Keynesian Phillips curve and against backward-looking specifications.

In this paper we propose to test the purely backward-looking Phillips curve and the purely forward-looking Phillips curve against a hybrid Phillips curve via their implications for cumulative output

Figure 1. Impulse Response to a Monetary Policy Shock



Notes: Sample is the United States, 1970:Q1–2009:Q4. Identification is via Cholesky decomposition using a lower triangular matrix with inflation, output gap, and the federal funds rate. VAR is estimated with four lags. The output gap has been calculated by detrending the logarithm of real U.S. GDP using an HP filter with lamda = 1600.

effects of monetary policy shocks. Compared with existing tests, our procedure is innovative in the sense that we have found a way of testing both versions of the Phillips curve against the hybrid model, while previous procedures test the forward-looking model against the backward-looking model without testing the traditional Phillips curve. Given the price puzzle with standard VARs which is also present in figure 1 and discussions about misspecification of VAR models (see Castelnuovo 2009 for a recent discussion), we do not test the zero cumulative effect within a VAR framework but resort to the approach suggested by Romer and Romer (2004), who carefully construct measures of monetary policy shocks. We estimate the output reaction to Romer and Romer monetary shocks under the hypothesis of long-run monetary neutrality. For our purpose of comparing the forward-looking and the backward-looking model with each other, this restriction suggests itself naturally, since it is implied by both hypotheses. As shown below, this hypothesis cannot be rejected by the data. Our tests focus on cumulative properties of the output response, the distinguishing feature between the traditional and the New Keynesian Phillips curve. We find that the empirical evidence is clearly more consistent with the forward-looking model.

The paper proceeds as follows. Section 2 derives the output implications of the backward-looking and the forward-looking Phillips curve—in particular, the adjustment process of output to a monetary shock. Section 3 presents empirical estimates and tests. Section 4 concludes.

2. The Model: Output Effects of Forward- and Backward-Looking Price Setting

2.1 The Traditional Phillips Curve

In a first step we derive that the backward-looking Phillips curve implies long-run monetary neutrality; i.e., a permanent increase in the money stock does not change the level of output in the long run. The traditional Phillips-curve approach is analyzed in the following standard macroeconomic model with aggregate supply,

$$\pi_t = \pi_{t-1} + \kappa x_t + u_t, \tag{1}$$

aggregate demand,

$$x_t = -\lambda(i_t - \pi_{t+1}) + e_t, \tag{2}$$

monetary policy (interest rate) rule,

$$i_t = \alpha_\pi \pi_t + \alpha_x x_t + v_t, \tag{3}$$

and money demand,

$$m_t - p_t = -\theta_i i_t + \theta_x x_t + z_t. \tag{4}$$

x denotes real output, m a nominal money aggregate, and p the price level. All of these variables are in logs. i is the nominal interest rate and π denotes the inflation rate. u, e, ν, and z denote i.i.d. shocks in aggregate supply, aggregate demand, the interest rate rule, and money demand, respectively. Setting the non-policy shocks equal to zero and solving equations equations (1), (2), and (3) for inflation yields

$$\left[1 - \frac{1 + \lambda\alpha_x + \kappa\lambda\alpha_\pi}{\kappa\lambda} L + \frac{1 + \lambda\alpha_x}{\kappa\lambda} L^2\right] \pi_{t+1} = v_t. \tag{5}$$

Using standard calculus, we can show for the roots of the characteristic equation that one root is smaller and one root is larger than one. Without loss of generality, we define τ_1 to be the larger root, i.e., $\tau_1 > 1$, and τ_2 to be the smaller root, $\tau_2 < 1$. Equation (5) can then be rewritten as

$$\pi_t = \tau_2 \pi_{t-1} - \left[\sum_{i=1}^{\infty} \left(\frac{1}{\tau_1}\right)^{\mathrm{i}} v_{t+i}\right]. \tag{6}$$

Taking into account $\pi_{t-1} = 0$, the effect of a temporary interest rate shock ν_t on period i inflation is

$$\pi_{t+i} = -\tau_2^i v_t. \tag{7}$$

Thus, the long-run effect of the interest rate shock on inflation is zero; i.e., a one-time change in the interest rate does not affect inflation in the long term:

$$\lim_{i\to\infty} \frac{d\pi_{t+i}}{dv_t} = 0. \tag{8}$$

How does this result relate to long-run monetary neutrality? From the traditional Phillips curve, equation (1), it follows that only changes in inflation affect output. Together with equation (7), this yields

$$x_{t+i} = \frac{1}{\kappa}(\pi_{t+i} - \pi_{t+i-1}) = \frac{1}{k}\tau_2^{i-1}(1-\tau_2)v_t. \tag{9}$$

Accordingly, the temporary interest rate shock does not affect output in the long run:

$$\lim_{i\to\infty} \frac{\partial x_{t+i}}{\partial v_t} = 0. \tag{10}$$

To derive the monetary effects of an interest rate shock, we combine equations (3) and (4) to get

$$m_t = p_t - \theta_i \alpha_\pi \pi_t + (\theta_x - \theta_i \alpha_\pi) x_t - \theta_i v_t. \tag{11}$$

Making use of

$$p_{t+n} = \sum_{i=0}^{n} \pi_{t+i} = \sum_{i=0}^{n} p_{t+i} - p_{t+i-1} = \sum_{i=0}^{n} -\left(\frac{1}{\tau_1}\right)^{i+1} v_t \tag{12}$$

and equation (11), the long-run effect of a temporary interest rate shock on the price level and the money stock is given by

$$\lim_{n\to\infty} p_{t+n} = -\frac{1}{1-\tau_1} v_t \tag{13}$$

and

$$\lim_{n\to\infty} m_{t+n} = -\frac{1}{1-\tau_1} v_t, \tag{14}$$

respectively.

A temporary negative interest rate shock causes a permanent increase in the money stock, equation (14), while it does not affect output in the long run, equation (10), so that monetary neutrality holds under the traditional Phillips curve:

$$\lim_{t\to\infty} \frac{dx_t}{dv_t} \bigg/ \frac{dm_t}{dv_t} = 0. \tag{15}$$

What are the cumulative output effects of the temporary interest rate shock? We get this cumulative output effect by summing up the output effects (equation (9)) during the adjustment process following the temporary interest rate shock, which yields

$$\sum_{i=0}^{\infty} x_{t+i} = \sum_{i=0}^{\infty} \frac{1}{\kappa}(\pi_{t+i} - \pi_{t+i-1})$$
$$= (v_t - 0) + (v_{t+1} - v_t) + (v_{t+2} - v_{t+1}) + \cdots = 0. \quad (16)$$

This zero cumulative output effect of a temporary interest rate shock is related to the well-known superneutrality property of the traditional Phillips curve.

In summary, a standard macroeconomic model with a traditional Phillips curve implies that a permanent increase in the level of money

- does not affect output in the long run (long-run neutrality) and
- entails a cumulative output effect of zero during the adjustment process.

2.2 The New Keynesian Phillips Curve

What are the implications of the New Keynesian Phillips curve for monetary neutrality? We analyze the New Keynesian Phillips curve in the following standard macroeconomic model with aggregate supply,

$$\pi_t = \beta\pi_{t+1} + \kappa x_t + u_t, \quad (17)$$

and aggregate demand, monetary policy rule, and money demand as in the previous section, i.e., aggregate demand,

$$x_t = -\lambda(i_t - \pi_{t+1}) + e_t, \quad (2)$$

monetary policy (interest rate) rule,

$$i_t = \alpha_\pi \pi_t + \alpha_x x_t + v_t, \quad (3)$$

and money demand,

$$m_t - p_t = -\theta_i i_t + \theta_x x_t + z_t. \tag{4}$$

In addition to the previous notation, β is the discount rate with $0 < \beta \leq 1$.

Combining equations (2), (3), and (17) yields a first-order difference equation for inflation:

$$\pi_{t+1}\left(1 - \frac{1 + \lambda\alpha_x + \kappa\lambda\alpha_\pi}{\beta + \beta\lambda\alpha_x + \kappa\lambda}\mathbf{L}\right) = \frac{\kappa\lambda}{\beta + \beta\lambda\alpha_x + \kappa\lambda}v_t \tag{18}$$

with $\delta = \dfrac{1 + \lambda\alpha_x + \kappa\lambda\alpha_\pi}{\beta(1 + \lambda\alpha_x) + \kappa\lambda} > 1$.

Solving of equation (18) yields

$$\pi_{t+1} = -\sum_{i=1}^{\infty}\left(\frac{1}{\delta}\right)^i \frac{\kappa\lambda}{\beta(1 + \lambda\alpha_x) + \kappa\lambda} v_{t+i}. \tag{19}$$

A temporary interest rate shock ν_t produces an opposite contemporaneous movement of the inflation rate,

$$\pi_t = -\mu\nu_t \tag{20}$$

with $\mu = \dfrac{1}{\delta}\dfrac{\kappa\lambda}{\beta(1 + \lambda\alpha_x) + \kappa\lambda} > 0$.

As the interest rate shock dies out, inflation returns to its steady-state value. Solving equations (17) and (20) for output yields a similar effect on output:

$$x_t = -\frac{\lambda(1 - \alpha_\pi\mu)}{1 + \lambda a_x} v_t. \tag{21}$$

Finally, the effect of the interest rate shock on the money stock can be derived by combining equations (3), (4), (20), and (21) and taking first differences:

$$\mathbf{\Delta m_t} - \pi_t = \left[\theta_i(\alpha_\pi\mu - \mathbf{1}) + \frac{\lambda(\mathbf{1} - \mu)}{\mathbf{1} + \lambda\alpha_x}(a_x\theta_i - \theta_x)\right]\mathbf{\Delta v_t}. \tag{22}$$

A one-time reduction of the interest rate—i.e., a temporary negative unit shock in the Taylor-rule equation—is then equivalent to a permanent increase in the money stock,

$$\mathbf{m_t} = \mu v_t + m_{t-1}. \tag{23}$$

Taken together, a temporary negative interest rate shock implies a contemporaneous fall in output and inflation as well as a permanent drop in the price level and a permanent increase in the money stock. As a corollary monetary neutrality holds, because the change in monetary policy with the associated permanent increase in the money stock does not affect output in the long run,

$$\frac{\frac{dx_{t+i}}{dv_t}}{\frac{dm_{t+i}}{dv_t}} = 0 \quad i \geq 1. \tag{24}$$

What are the cumulative output effects of a temporary interest rate shock, i.e., a permanent change in the money stock? According to equation (24), a temporary interest rate shock has a contemporaneous output effect only, so that the cumulative output effect is given by

$$\sum_{i=0}^{\infty} \frac{dx_{t+i}}{dv_t} = \frac{dx_t}{dv_t} = -\frac{\lambda(1-\mu)}{1+\lambda\alpha_x} < 0. \tag{25}$$

Summarizing, a temporary negative interest rate shock has a positive output effect. As output remains positive during the adjustment process, the cumulative output effect is also positive.

While the backward- and the forward-looking models both imply long-run monetary neutrality—i.e., output is not affected by a permanent rise in the level of the money stock—they differ fundamentally with respect to the transition path following such a permanent increase in the money stock. This crucial difference between the two models in terms of output response is given by equations (16) and (25). While the traditional Phillips curve implies that these cumulated output effects are zero, the New Keynesian Phillips curve predicts a positive cumulated output effect.

In order to get a first idea on the quantitative importance of the cumulative output responses under the backward-looking and forward-looking specifications, we simulate the output effect of a monetary shock on the two Phillips-curve models estimated by Galí, Gertler, and López-Salido (2001, p. 1240). In addition, we also present results for a hybrid version of the Phillips curve with 50 percent share of forward-looking price setters, which is in the neighborhood of values often estimated in empirical studies. For this purpose we modify the New Keynesian Phillips curve by introducing an additional parameter s^F, which indicates the degree of forward-looking price setting:

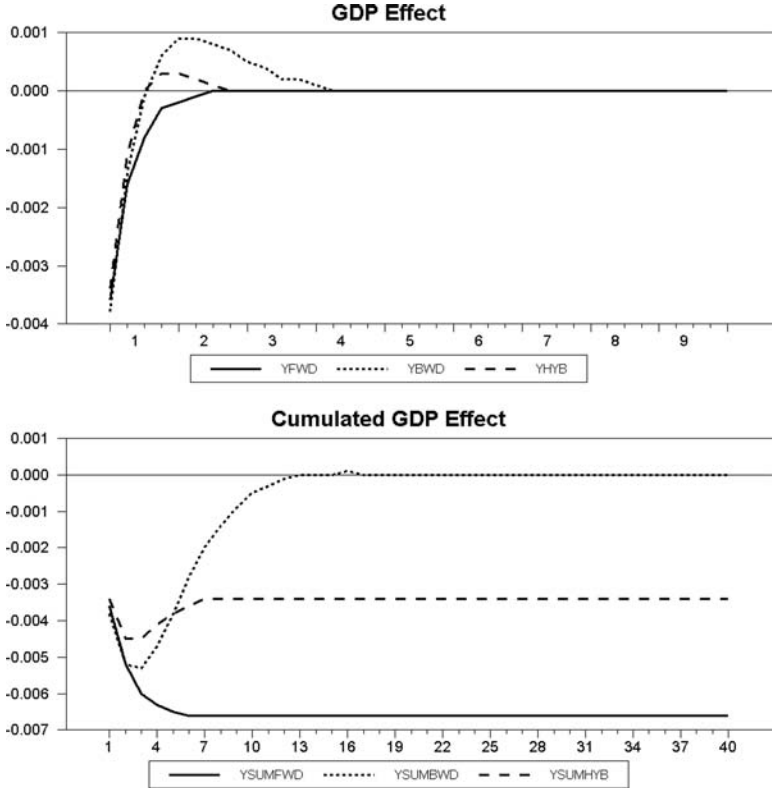
$$\pi_t = \beta(s^F \pi_{t+1} + (1 - s^F)\pi_{t-1}) + \kappa x_t + u_t \quad 0 \leq s^F \leq 1. \quad (17b)$$

In all three models, the strongest output effect of a temporary increase in the nominal interest rate occurs in the first period (figure 2). The backward-looking model generates oscillatory movements of GDP (YBWD) around the baseline. Though the output effect of a positive interest rate shock is negative initially, it turns positive after about three quarters. The cumulated sum of the output effect (YSUMBWD) is zero. In the forward-looking model, the output effect (YFWD) stays negative during the adjustment process, implying a negative cumulative output effect (YSUMFWD).[2] The hybrid model lies somewhere between these borderline cases. The interesting question that emerges in this case is how the sum of positive and negative output effects (integrals above and below the zero line) is related to the degree of forward-looking price setting.

Tables 1 and 2 give the ratio of positive and negative output effects for alternative values of the parameter s^F which indicates

[2]Notice that firms adjust prices more rapidly in the case of a permanent change in the growth rate of money than in the case of a temporary monetary growth shock. Therefore, the strong cumulative output effect of the forward-looking model in the case of a temporary monetary growth (interest rate) shock is consistent with near superneutrality in the case of a permanent monetary growth (interest rate) shock. With the same parameter values, a permanent increase in the growth rate of money by 1 percent leads to a .025 percent gain in the level of output in the long run. Notice, however, that these estimates do not include possible resource or menu costs from inflation. We derived an alternative version of the New Keynesian price equation based on a model with convex costs of price adjustment (available from the authors by request) and find that using the Galí, Gertler, and López-Salido (2001) estimates, we can identify a price adjustment cost parameter which suggests that with rates of inflation exceeding 1.5 percent, the menu costs would already exceed the markup gains of inflation.

Figure 2. Simulated Output and Cumulated Output Effects



Notes: Shock: Temporary increase of nominal interest rates. Yxx: Deviation of GDP from baseline value. YCUMxx: Cumulated deviation of GDP from baseline value.

the degree of forward-looking price setting. There is a monotonic relationship between s^F and the ratio of the sum of positive and negative output effects. Unfortunately, the relationship is not linear but (sort of) concave. Therefore, the degree of forward-looking behavior cannot be determined with very high precision, while fairly robust statements can be made whether s^F is larger or smaller than .5.

Both Phillips-curve specifications have one problem in common when confronted with VAR evidence on the adjustment of output to a monetary shock. In contrast to the VAR evidence, the peak

Table 1. Fraction of Forward-Looking Firms and Output Effects (model without habit persistence)

s^F	1	.9	.8	.7	.6	.5	.4	.3	.2	.1	0
$\frac{Sumpositive}{Sumnegative}$	0	0	0	−.005	−.008	−.26	−.50	−.72	−.85	−.94	−1

Table 2. Fraction of Forward-Looking Firms and Output Effects (model with habit persistence)

S^F	1	.9	.8	.7	.6	.5	.4	.3	.2	.1	0
$\frac{Sumpositive}{Sumnegative}$	−.006	−.007	−.008	−.013	−.026	−.46	−.71	−.88	−.96	−.99	−1

of the output effect already occurs in the first period. As shown by Fuhrer (2000), habit persistence in demand allows for a delayed peak. Therefore, we also consider an alternative aggregate demand function,

$$x_t = (1-h)E_t x_{t+1} + h x_{t-1} - \lambda(i_t - E_t \pi_{t+1}) + e_t, \tag{2b}$$

where we set the parameter h equal to .9.

Since this introduces a backward-looking element into the model, it may affect the output dynamics and, in particular, the cumulative output effect. However, it turns out (see figure 3) that habit persistence does not affect the central propositions of the two hypotheses, concerning the size of the cumulative output effect and the symmetry of positive and negative output deviations.

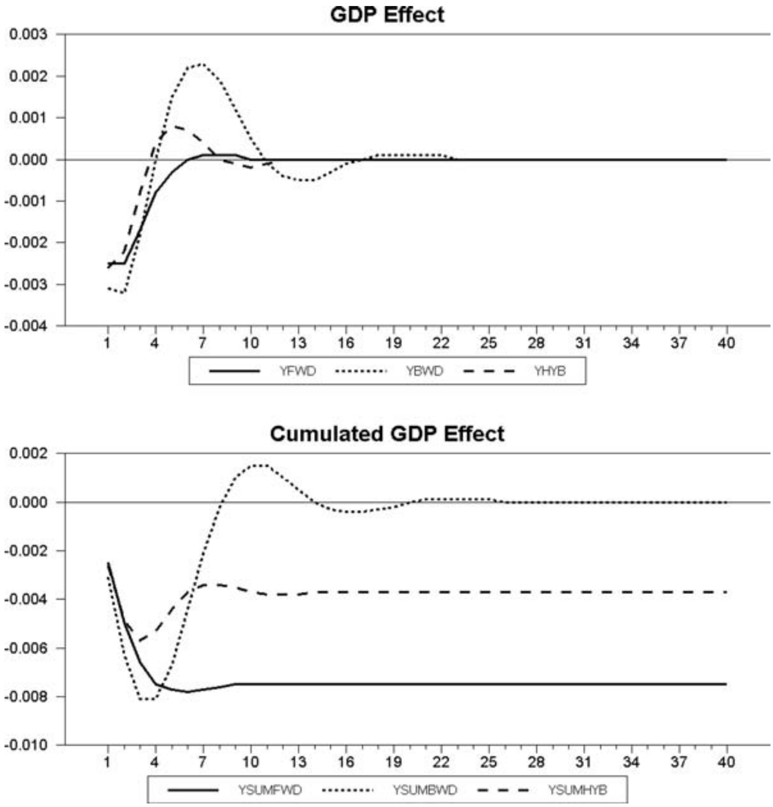
In particular, the empirical results from table 2 indicate that

- introducing habit persistence does not alter the basic prediction of the backward-looking model ($s^F = 0$) that the cumulative output effect of a monetary shock is zero, and
- the relationship between s^F and the ratio of positive and negative output effects remains concave (though slightly less).

3. Empirical Analysis

In this section we test the predictions of the forward- and backward-looking models concerning the cumulative output effects of monetary

Figure 3. Simulated Output and Cumulated Output Effects (with habit persistence)



Notes: Shock: Temporary increase of nominal interest rates. Yxx: Deviation of GDP from baseline value. YCUMxx: Cumulated deviation of GDP from baseline value.

policy shocks. The backward-looking model predicts that the sum of all output deviations is zero, while the forward-looking model predicts that a positive interest rate shock will only have negative output effects. The hybrid Phillips curve predicts a negative cumulative output effect but allows for periods with positive output effects. Using this information helps us to pin down the parameter s^F for the hybrid Phillips curve. As monetary policy shocks, we take the "new measure of monetary shocks" of Romer and Romer (2004).

They use quantitative and narrative records to infer the intended Federal Reserve interest rate policy. This series is then regressed on the Federal Reserve's internal forecasts to derive a measure that is relatively free of endogenous and anticipatory movements. Since the measure is available on a monthly frequency, we follow Romer and Romer and use the monthly measure of industrial production as our output variable. We replicate their specification by regressing the growth rate of industrial production Δx_t on its own lags, lags of the monetary policy shock measure v_{t-j}, and seasonal dummies D_{kt}:

$$\Delta x_t = a_o + \sum_{k=1}^{11} a_k D_{kt} + \sum_{i=1}^{24} b_i \Delta x_{t-i} + \sum_{j=1}^{36} c_j v_{t-j} + e_t. \qquad (26)$$

Our sample period is 1970:M1–1996:M12. Since we have shown above that both models imply monetary neutrality, we want to make sure that the data we use is consistent with this joint prediction; therefore, we also estimate the Romer and Romer specification by imposing a long-run monetary neutrality restriction $\sum_{j=1}^{36} c_j = 0$ on the parameter estimates for the monetary shocks. Tables 3 and 4 show the regression with and without neutrality restriction. The F-test yields a value of $F(1, 252) = 0.62$ with significance level 0.43; i.e., we cannot reject the neutrality hypothesis at conventional significance levels.

Table 3 reports our estimation results under the restriction of long-run monetary neutrality. Based on these results, the net cumulative output effects of a monetary policy shock can be calculated in a straightforward way. The adjustment path of output following a temporary 1-percentage-point increase of the monetary policy shock measure is generated by a dynamic forecast of output. The percentage deviations of output from its long-run level are aggregated to give the cumulated output effect of the monetary shock. As can be seen from figure 4, the output and the cumulated output response are similar to the prediction of the forward-looking price equation (figure 2). In particular, the negative response of output to a temporary increase in the interest rate is not compensated for by positive deviations from the baseline in later periods. Most importantly, the central prediction of the backward-looking model—namely, that the initially negative response would be fully compensated for by

Table 3. Impact of Monetary Policy Shock on Real GDP (with restriction)

Change in GDP			Monetary Policy Shock		
Lag	Coefficient	Standard Error	Lag	Coefficient	Standard Error
1	0.0627	0.0637	1	0.0038	0.0018
2	−0.0128	0.0633	2	0.0026	0.0018
3	0.1072	0.0628	3	−0.0038	0.0018
4	0.0484	0.0630	4	−0.0012	0.0018
5	0.0284	0.0629	5	−0.0039	0.0018
6	−0.0054	0.0628	6	−0.0001	0.0018
7	0.0179	0.0627	7	−0.0008	0.0019
8	0.0075	0.0626	8	−0.0029	0.0019
9	0.0396	0.0622	9	−0.0021	0.0019
10	−0.0426	0.0609	10	−0.0047	0.0018
11	0.0709	0.0593	11	−0.0025	0.0019
12	0.2867	0.0602	12	−0.0035	0.0019
13	0.0227	0.0608	13	−0.0021	0.0019
14	−0.1964	0.0604	14	−0.0007	0.0018
15	−0.1510	0.0610	15	−0.0003	0.0019
16	−0.1282	0.0623	16	0.0019	0.0018
17	0.0777	0.0635	17	−0.0009	0.0018
18	0.0853	0.0632	18	−0.0024	0.0018
19	0.0557	0.0632	19	−0.0023	0.0019
20	0.0805	0.0629	20	−0.0007	0.0019
21	−0.0604	0.0631	21	−0.0011	0.0019
22	−0.0171	0.0630	22	−0.0032	0.0018
23	−0.0675	0.0630	23	0.0015	0.0019
24	0.0863	0.0631	24	0.0000	0.0019
			25	−0.0001	0.0019
			26	−0.0000	0.0019
			27	−0.0007	0.0019
			28	0.0038	0.0019
			29	0.0013	0.0019
			30	0.0035	0.0019
			31	0.0018	0.0019
			32	0.0009	0.0018
			33	0.0014	0.0018
			34	0.0047	0.0018
			35	0.0011	0.0018
			36	0.0024	0.0018

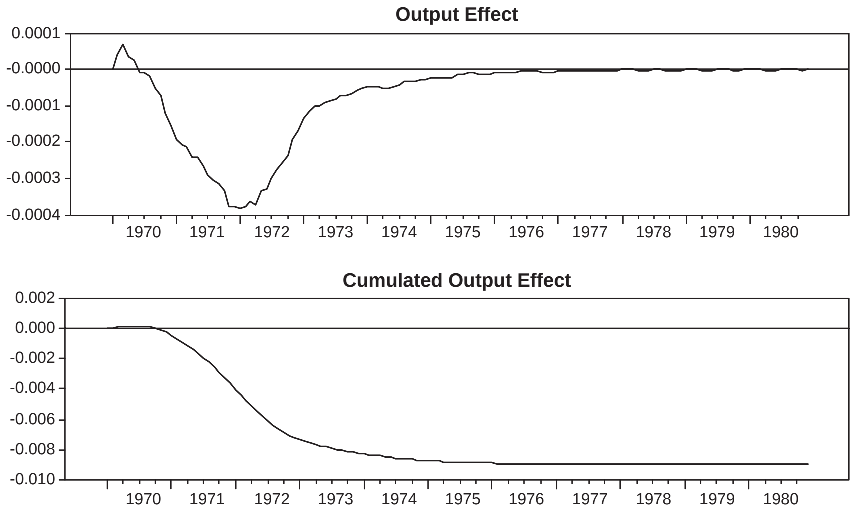
Notes: R2 = 0.81; D.W. = 2.01; s.e.e. = 0.009; N = 324. The sample period is 1970:M1–1996:M12. Coefficients and standard errors for the constant term and the monthly dummies are not reported.

Table 4. Impact of Monetary Policy Shock on Real GDP (without restriction)

Change in GDP			Monetary Policy Shock		
Lag	**Coefficient**	**Standard Error**	**Lag**	**Coefficient**	**Standard Error**
1	0.0650	0.0636	1	0.0040	0.0018
2	−0.0103	0.0632	2	0.0028	0.0018
3	0.1105	0.0626	3	−0.0037	0.0018
4	0.0519	0.0628	4	−0.0010	0.0018
5	0.0329	0.0626	5	−0.0038	0.0018
6	−0.0011	0.0626	6	0.0001	0.0018
7	0.0236	0.0622	7	−0.0007	0.0018
8	0.0138	0.0621	8	−0.0028	0.0019
9	0.0460	0.0616	9	−0.0020	0.0018
10	−0.0382	0.0606	10	−0.0045	0.0018
11	0.0754	0.0590	11	−0.0024	0.0018
12	0.2905	0.0600	12	−0.0034	0.0018
13	0.0264	0.0606	13	−0.0019	0.0019
14	−0.1927	0.0602	14	−0.0004	0.0018
15	−0.1484	0.0609	15	−0.0001	0.0018
16	−0.1268	0.0622	16	0.0021	0.0018
17	0.0807	0.0633	17	−0.0006	0.0018
18	0.0883	0.0630	18	−0.0022	0.0018
19	0.0578	0.0631	19	−0.0020	0.0018
20	0.0830	0.0628	20	−0.0004	0.0018
21	−0.0577	0.0630	21	−0.0008	0.0018
22	−0.0144	0.0628	22	−0.0029	0.0018
23	−0.0645	0.0628	23	0.0019	0.0018
24	0.0887	0.0629	24	0.0003	0.0018
			25	0.0003	0.0018
			26	0.0004	0.0018
			27	−0.0003	0.0018
			28	0.0042	0.0018
			29	0.0017	0.0018
			30	0.0039	0.0018
			31	0.0021	0.0018
			32	0.0013	0.0018
			33	0.0018	0.0018
			34	0.0050	0.0018
			35	0.0014	0.0018
			36	0.0027	0.0018

Notes: R2 = 0.81; D.W. = 2.02; s.e.e. = 0.009; N = 324. The sample period is 1970:M1–1996:M12. Coefficients and standard errors for the constant term and the monthly dummies are not reported.

Figure 4. Output and Cumulated Output Effects of Temporary Negative Interest Rate Shock (under long-run monetary neutrality)



a positive output response such that the cumulated output effect precisely sums to zero—does not seem to hold.

The estimated impulse response shows, however, one important difference from the standard New Keynesian Phillips curve, which concerns the peak of the output effect. In the Romer and Romer model, the output effect peaks in the second to third period, while the New Keynesian Phillips curve predicts the peak already in the first period. However, neither the hybrid nor the traditional Phillips curve is superior with respect to the timing of the output peak.[3] As discussed above, the peak response of output can be shifted by introducing a lagged response of demand, i.e., habit persistence in the IS curve. As was also shown above, adding demand persistence does not fundamentally alter the predictions of the two hypotheses concerning the magnitude of the cumulative output effect.

Since it is our interest to arrive at an estimate of the cumulative output effect and to test whether it is significantly different from

[3]The hybrid Phillips curve does better in terms of inflation persistence.

Table 5. Distribution of Cumulated Output Effects (YCUM)

1^{st} Percentile	5^{th} Percentile	10^{th} Percentile	Median	90^{th} Percentile	95^{th} Percentile	99^{th} Percentile
−.1726	−.0443	−.0270	−.0265	−.0265	−.0038	−.0030

Note: Sample mean = −.0252, standard error of mean = .0081, standard error = .2580, skewness = −29.6, kurtosis = 910.3, observations = 1,000.

zero, it is important to get an idea of the precision in which the cumulated output effect can be estimated. This is not a straightforward exercise, since the cumulative output effect will generally be a complicated non-linear function of the estimated coefficients. Therefore, we base our calculation of confidence bounds on Monte Carlo simulations. For that purpose we generate 1,000 random samples of coefficients from the vector of estimated Π, using the estimated variance-covariance matrix Ω of Π. For each draw of parameters, we calculate the cumulated output effect over a period of 500 months by simulating the output effect of an 1-percentage-point increase in the "Romer interest rate shock variable" over this period and summing up the output response over time. This time horizon seems sufficient, since the long-run neutrality restriction implies that output will eventually return to its baseline level. As can be seen from figure 2, the effect of an increase in money on output is completed well within 100 months. Table 5 shows the characteristics of the distribution of YCUM (evaluated after 500 quarters) of the Monte Carlo exercise. The sample mean of the distribution is significantly negative (sample mean plus standard error of sample mean), thus suggesting that the cumulative output effect of a monetary policy shock is clearly negative. Because the empirical distribution shows some negative skewness with the mean being closer to zero than the median, we also show the percentiles of the distribution (area to the left of the distribution as percent of the total area). This shows that even at the 99^{th} percentile, the values for YCUM are negative.

The results show that the probability that YCUM takes a positive value is less than 1 percent. Thus, we can reject the hypothesis of a zero cumulative output effect at the 1 percent level.

Table 6. Distribution of Backward-Looking Share Parameter in Hybrid Phillips-Curve Model

1st Percentile	5th Percentile	10th Percentile	Median	90th Percentile	95th Percentile	99th Percentile
.0004	.0026	.0048	.0543	.2460	.3338	.5257

Note: Sample mean = .0972, standard error of mean = .0038, standard error = .1211, skewness = 2.72, kurtosis = 11.14, observations = 1,000.

Similarly, we can test the forward-looking Phillips curve by testing whether the area of output effects $\sum_{s=0}^{500} Max(Y_s, 0)$ is significantly negative. For this we construct the following test statistic:

$$X = \frac{\sum_{s=0}^{500} \max(Y_s, 0)}{\sum_{s=0}^{500} \min(Y_s, 0)}.$$

Notice that X varies between 0 and -1. The backward Phillips curve is one extreme case, with exactly the same size for cumulated positive and negative output effects, implying a value of -1. The forward-looking model implies zero positive output deviations, thus implying a value of $x = 0$.

As shown above, different share parameters imply different relative sizes of these positive and negative deviations. Therefore, we can also use the Monte Carlo experiment to shed some light on the location of the share parameter (of a hybrid Phillips curve) by calculating the integrals of positive and negative effects.

Table 6 shows the distribution of X. Both the median and the mean of the empirical distribution of X are very small, suggesting a ratio between positive and negative deviations of below .1 in absolute value. Notice, however, that the mean of X is significantly different from zero; thus we also reject the forward-looking model at standard levels of significance. Comparing the estimated value of X with the values of X generated by alternative hybrid specifications (see tables 1 and 2) suggests that our estimate is consistent with a hybrid Phillips curve with a share of forward-looking price setters above 50 percent, an estimate often found in the empirical literature. It is interesting to note that the hybrid Phillips curve nevertheless makes predictions about the output effects of monetary policy which are close to the pure forward-looking model.

4. Conclusions

This paper looks into the monetary policy implications of backward-looking and forward-looking price equations. So far, no consensus has been reached among the macroeconomics profession regarding which of the two price adjustment schemes should be chosen on empirical grounds. We identify a crucial difference between the two hypotheses. The backward-looking Phillips curve predicts a strict intertemporal trade-off in the case of monetary shocks: a negative short-run response of output is followed by a period where output is above the baseline and the cumulative effect is exactly zero. In contrast, the forward-looking model allows the cumulative output effect of temporary monetary shocks to be non-zero. These contrasting predictions of the two hypotheses are tested in this paper. We find that the empirical evidence on the cumulated output effects of monetary shocks is in fact more consistent with the forward-looking model. Both qualitatively and quantitatively, the output response to a temporary interest rate innovation seems close to what is predicted by the New Keynesian Phillips curve.

References

Brayton, F., A. Levin, R. Tryon, and J. C. Williams. 1997. "The Evolution of Macro Models at the Federal Reserve Board." *Carnegie-Rochester Conference Series on Public Policy* 47 (1): 43–81.

Calvo, G. A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12 (3): 383–98.

Castelnuovo, E. 2009. "Testing the Structural Interpretation of the Price Puzzle with a Cost Channel Model." Bank of Finland Research Discussion Paper No. 20.

Chadha, B., P. R. Masson, and G. Meredith. 1992. "Models of Inflation and the Costs of Disinflation." *IMF Staff Papers* 39 (2): 395–431.

Christiano, L., M. Eichenbaum, and C. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1): 1–45.

Clarida, R., J. Galí, and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37 (4): 1661–1707.

Fuhrer, J. C. 1997. The (Un)Importance of Forward-Looking Behavior in Price Specifications." *Journal of Money, Credit, and Banking* 29 (3): 338–50.

———. 2000. "Habit Formation in Consumption and Its Implications for Monetary Policy." *American Economic Review* 90 (3): 367–90.

Fuhrer, J. C., and G. R. Moore. 1995. "Inflation Persistence." *Quarterly Journal of Economics* 110 (1): 127–59.

Galí, J., and M. Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44 (2): 195–222.

Galí, J., M. Gertler, and J. D. López-Salido. 2001. "European Inflation Dynamics." *European Economic Review* 45 (7): 1237–70.

McAdam, P., and A. Willman. 2004. "Supply, Factor Shares and Inflation Persistence: Re-Examining Euro-Area New-Keynesian Phillips Curves." *Oxford Bulletin of Economics and Statistics* 66 (s1): 637–70.

———. 2010. "Arrow-Calvo Price Staggering." *The Manchester School* 78 (6): 558–81.

McCallum, B. T. 1997. "Crucial Issues Concerning Central Bank Independence." *Journal of Monetary Economics* 39 (1): 99–112.

Roberts, J. M. 1997. "Is Inflation Sticky?" *Journal of Monetary Economics* 39 (2): 173–96.

———. 2001. "How Well Does the New Keynesian Sticky-Price Model Fit the Data?" FEDS Discussion Paper No. 2001-13, Board of Governors of the Federal Reserve System.

Romer, C., and D. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* 94 (4): 1055–84.

Rotemberg, J. J. 1982. "Sticky Prices in the United States." *Journal of Political Economy* 90 (6): 1187–1211.

Rudd, J., and K. Whelan. 2005. "New Tests of the New-Keynesian Phillips Curve." *Journal of Monetary Economics* 52 (6): 1167–81.

———. 2006. "Can Rational Expectations Sticky-Price Models Explain Inflation Dynamics?" *American Economic Review* 96 (1): 303–20.

Rudebusch, G. D., and L. E. O. Svensson. 1999. "Policy Rules for Inflation Targeting." In *Monetary Policy Rules*, ed. J. Taylor, 203–46. Chicago: University of Chicago Press.

Sims, A. C., and T. Zha. 2006. "Were There Regime Shifts in U.S. Monetary Policy?" *American Economic Review* 96 (1): 54–81.

Taylor, J. B. 1979. "Staggered Wage Setting in a Macro Model." *American Economic Review* 69 (2): 108–113.

Estimating Inflation Expectations with a Limited Number of Inflation-Indexed Bonds*

Richard Finlay and Sebastian Wende
Reserve Bank of Australia

We develop a novel technique to estimate inflation expectations and inflation risk premia when only a limited number of inflation-indexed bonds are available. The method involves pricing coupon-bearing inflation-indexed bonds directly in terms of an affine term structure model, and avoids the usual requirement of estimating zero-coupon real yield curves. We estimate the model using a non-linear Kalman filter and apply it to Australia. The results suggest that long-term inflation expectations in Australia are well anchored within the Reserve Bank of Australia's inflation target range of 2 to 3 percent, and that inflation expectations are less volatile than inflation risk premia.

JEL Codes: E31, E43, G12.

1. Introduction

Reliable and accurate estimates of inflation expectations are important to central banks, given the role of these expectations in influencing inflation and economic activity. Inflation expectations may also indicate over what horizon individuals believe that a central bank will achieve its inflation target, if at all.

A common measure of inflation expectations based on financial market data is the break-even inflation yield, referred to simply as the inflation yield. The inflation yield is given by the difference in

*The authors thank Rudolph van der Merwe for help with the central difference Kalman filter, as well as Adam Cagliarini, Jonathan Kearns, Christopher Kent, Frank Smets, Ian Wilson, and an anonymous referee for useful comments and suggestions. Responsibility for any remaining errors rests with the authors. The views expressed in this paper are those of the authors and are not necessarily those of the Reserve Bank of Australia. E-mail: FinlayR@rba.gov.au.

yields of nominal and inflation-indexed zero-coupon bonds of equal maturity. That is,

$$y^i_{t,\tau} = y^n_{t,\tau} - y^r_{t,\tau},$$

where $y^i_{t,\tau}$ is the inflation yield between time t and $t + \tau$, $y^n_{t,\tau}$ is the nominal yield, and $y^r_{t,\tau}$ is the real yield.[1] But the inflation yield may not give an accurate reading of inflation expectations. Inflation expectations are an important determinant of the inflation yield but are not the only determinant; the inflation yield is also affected by inflation risk premia, which is the extra compensation required by investors who are exposed to the risk that inflation will be higher than expected (we assume that other factors that may affect the inflation yield, such as liquidity premia, are absorbed into risk premia in our model). By treating inflation as a random process, we are able to model expected inflation and the cost of the uncertainty associated with inflation separately.

Inflation expectations and inflation risk premia have been estimated for the United Kingdom and the United States using models similar to the one used in this paper. Beechey (2008) and Joyce, Lildholdt, and Sorensen (2010) find that inflation risk premia decreased in the United Kingdom, first after the Bank of England adopted an inflation target and then again after it was granted independence. Using U.S. Treasury Inflation-Protected Securities (TIPS) data, Durham (2006) estimates expected inflation and inflation risk premia, although he finds that inflation risk premia are not significantly correlated with measures of the uncertainty of future inflation or monetary policy. Also using TIPS data, D'Amico, Kim, and Wei (2008) find inconsistent results due to the decreasing liquidity premia in the United States, although their estimates are improved by including survey forecasts and using a sample over which the liquidity premia are constant.

In this paper we estimate a time series for inflation expectations at various horizons, taking into account inflation risk premia, using a latent factor affine term structure model which is widely

[1] To fix terminology, all yields referred to in this paper are gross, continuously compounded zero-coupon yields. So, for example, the nominal yield is given by $y^n_{t,\tau} = -\log(P^n_{t,\tau})$, where $P^n_{t,\tau}$ is the price at time t of a zero-coupon nominal bond paying one dollar at time $t + \tau$.

used in the literature. Compared with the United Kingdom and the United States, there are a very limited number of inflation-indexed bonds on issue in Australia. This complicates the estimation but also highlights the usefulness of our approach. In particular, the limited number of inflation-indexed bonds means that we cannot reliably estimate a zero-coupon real yield curve and so cannot estimate the model in the standard way. Instead we develop a novel technique that allows us to estimate the model using the *price* of coupon-bearing inflation-indexed bonds instead of zero-coupon real yields. The estimation of inflation expectations and risk premia for Australia, as well as the technique we employ to do so, is the chief contribution of this paper to the literature.

To better identify model parameters, we also incorporate inflation forecasts from Consensus Economics in the estimation. Inflation forecasts provide shorter-maturity information (for example, forecasts exist for inflation next quarter) as well as information on inflation expectations that is separate from risk premia. Theoretically the model is able to estimate inflation expectations and inflation risk premia purely from the nominal and inflation-indexed bond data; inflation risk premia compensate investors for exposure to variation in inflation, which should be captured by the observed variation in prices of bonds at various maturities. This is, however, a lot of information to extract from a limited amount of data. Adding forecast data helps to better anchor the model estimates of inflation expectations and so improves model fit.

Inflation expectations as estimated in this paper have a number of advantages over using the inflation yield to measure expectations. For example, five-year-ahead inflation expectations as estimated in this paper (i) account for risk premia and (ii) are expectations of the inflation rate *in five years time*. In contrast, the five-year inflation yield ignores risk premia and gives an *average* of inflation rates *over the next five years*.[2] The techniques used in the paper are potentially

[2] In addition, due to the lack of zero-coupon real yields in Australia's case, yields-to-maturity of coupon-bearing nominal and inflation-indexed bonds have historically been used when calculating the inflation yield. This restricts the horizon of inflation yields that can be estimated to the maturities of the existing inflation-indexed bonds, and is not a like-for-like comparison due to the differing coupon streams of inflation-indexed and nominal bonds.

useful for other countries with a limited number of inflation-indexed bonds on issue.

In section 2 we outline the model. Section 3 describes the data, estimation of the model parameters and latent factors, and how these are used to extract our estimates of inflation expectations. Results are presented in section 4 and conclusions are drawn in section 5.

2. Model

2.1 Affine Term Structure Model

Following Beechey (2008), we assume that the inflation yield can be expressed in terms of an inflation stochastic discount factor (SDF). The inflation SDF is a theoretical concept, which for the purpose of asset pricing incorporates all information about income and consumption uncertainty in our model. Appendix 1 provides a brief overview of the inflation, nominal, and real SDFs.

We assume that the inflation yield can be expressed in terms of an inflation SDF, M_t^i, according to

$$y_{t,\tau}^i = -\log\left(\mathbb{E}_t\left(\frac{M_{t+\tau}^i}{M_t^i}\right)\right).$$

We further assume that the evolution of the inflation SDF can be approximated by a diffusion equation,

$$\frac{\mathrm{d}M_t^i}{M_t^i} = -\pi_t^i \mathrm{d}t - \boldsymbol{\lambda}_t^{i\prime}\, \mathrm{d}\mathbf{B}_t. \tag{1}$$

According to this model, $\mathbb{E}_t(\mathrm{d}M_t^i/M_t^i) = -\pi_t^i\mathrm{d}t$, so that the instantaneous inflation rate is given by π_t^i. The inflation SDF also depends on the term $\boldsymbol{\lambda}_t^{i\prime}\mathrm{d}\mathbf{B}_t$. Here $\mathbf{B}_t$ is a Brownian motion process and $\boldsymbol{\lambda}_t^i$ relates to the market price of this risk. $\boldsymbol{\lambda}_t^i$ determines the risk premium, and this setup allows us to separately identify inflation expectations and inflation risk premia. This approach to bond pricing is standard in the literature and has been very successful in capturing the dynamics of nominal bond prices (see Kim and Orphanides 2005, for example).

We model both the instantaneous inflation rate and the market price of inflation risk as affine functions of three latent factors. The instantaneous inflation rate is given by

$$\pi_t^i = \rho_0 + \boldsymbol{\rho}'\mathbf{x}_t, \tag{2}$$

where $\mathbf{x}_t = [x_t^1, x_t^2, x_t^3]'$ are our three latent factors.[3] Since the latent factors are unobserved, we normalize $\boldsymbol{\rho}$ to be a vector of ones, $\mathbf{1}$, so that the inflation rate is the sum of the latent factors and a constant, ρ_0. We assume that the price of inflation risk has the form

$$\boldsymbol{\lambda}_t^i = \boldsymbol{\lambda}_0 + \Lambda\mathbf{x}_t, \tag{3}$$

where $\boldsymbol{\lambda}_0$ is a vector and Λ is a matrix of free parameters.

The evolution of the latent factors $\mathbf{x}_t$ is given by an Ornstein-Uhlenbeck process (a continuous-time mean-reverting stochastic process),

$$\mathrm{d}\mathbf{x}_t = K(\boldsymbol{\mu} - \mathbf{x}_t)\mathrm{d}t + \Sigma\ \mathrm{d}\mathbf{B}_t, \tag{4}$$

where $K(\boldsymbol{\mu} - \mathbf{x}_t)$ is the drift component, K is a lower triangular matrix, $\mathbf{B}_t$ is the same Brownian motion used in equation (1), and Σ is a diagonal scaling matrix. In this instance we set $\boldsymbol{\mu}$ to zero so that $\mathbf{x}_t$ is a zero-mean process, which implies that the average instantaneous inflation rate is ρ_0.

Equations (1)–(4) can be used to show how the latent factors affect the inflation yield (see appendix 2 for details). In particular, one can show that

$$y_{t,\tau}^i = \alpha_\tau^* + \boldsymbol{\beta}_\tau^{*\prime}\mathbf{x}_t, \tag{5}$$

where α_τ^* and $\boldsymbol{\beta}_\tau^*$ are functions of the underlying model parameters. In the standard estimation procedure, when a zero-coupon inflation yield curve exists, this function is used to estimate the values of $\mathbf{x}_t$.

[3] Note that one can specify models in which macroeconomic series take the place of latent factors—as done, for example, in Hördahl (2008). Such models have the advantage of simpler interpretation but, as argued in Kim and Wright (2005), tend to be less robust to model misspecification and generally result in a worse fit of the data.

2.2 Pricing Inflation-Indexed Bonds in the Latent Factor Model

We now derive the price of an inflation-indexed bond as a function of the model parameters, the latent factors, and nominal zero-coupon bond yields, denoted $H1(\mathbf{x}_t)$. This function will later be used to estimate the model as described in section 3.2.

As is the case with any bond, the price of an inflation-indexed bond is the present value of its stream of coupons and its par value. In an inflation-indexed bond, the coupons are indexed to inflation so that the real value of the coupons and principal is preserved. In Australia, inflation-indexed bonds are indexed with a lag of between 4½ and 5½ months, depending on the particular bond in question. If we denote the lag by Δ and the historically observed increase in the price level between $t - \Delta$ and t by $I_{t,\Delta}$, then at time t the implicit nominal value of the coupon paid at time $t + \tau_s$ is given by the real (at time $t - \Delta$) value of that coupon, C_s, adjusted for the historical inflation that occurred between $t - \Delta$ and t, $I_{t,\Delta}$, and further adjusted by the current market-implied change in the price level between periods t and $t + \tau_s - \Delta$ using the inflation yield. So the implied nominal coupon paid becomes $C_s I_{t,\Delta} \exp(y^i_{t,\tau_s-\Delta})$. The present value of this nominal coupon is then calculated using the nominal discount factor between t and $t + \tau_s$, $\exp(-y^n_{t,\tau_s})$. So if an inflation-indexed bond pays a total of m coupons, where the par value is included in the set of coupons, then the price at time t of this bond is given by

$$P^r_t = \sum_{s=1}^{m} (C_s I_{t,\Delta} e^{y^i_{t,\tau_s-\Delta}}) e^{-y^n_{t,\tau_s}} = \sum_{s=1}^{m} C_s I_{t,\Delta} e^{y^i_{t,\tau_s-\Delta} - y^n_{t,\tau_s}}.$$

We noted earlier that the inflation yield is given by $y^i_{t,\tau} = \alpha^*_\tau + \boldsymbol{\beta}^{*\prime}_\tau \mathbf{x}_t$, so the bond price can be written as

$$P^r_t = \sum_{s=i}^{m} C_s I_{t,\Delta} e^{-y^n_{t,\tau_s} + \alpha^*_{\tau_s-\Delta} + \boldsymbol{\beta}^{*}_{\tau_s-\Delta}{}' \mathbf{x}_t} = H1(\mathbf{x}_t). \qquad (6)$$

Note that $\exp(-y^n_{t,\tau_s})$ can be estimated directly from nominal bond yields (see section 3.1). So the price of a coupon-bearing inflation-indexed bond can be expressed as a function of the latent factors $\mathbf{x}_t$

as well as the model parameters, nominal zero-coupon bond yields, and historical inflation. We define $H1(\mathbf{x}_t)$ as the non-linear function that transforms our latent factors into bond prices.

2.3 Inflation Forecasts in the Latent Factor Model

In the model, inflation expectations are a function of the latent factors, denoted $H2(\mathbf{x}_t)$. Inflation expectations are not equal to expected inflation yields since yields incorporate risk premia, whereas forecasts do not. Inflation expectations as reported by Consensus Economics are expectations at time t of how the CPI will increase between time s in the future and time $s+\tau$ and are therefore given by

$$\mathbb{E}_t\left(\exp\left(\int_s^{s+\tau}\pi_u^i \mathrm{d}u\right)\right) = H2(\mathbf{x}_t),$$

where π_t^i is the instantaneous inflation rate at time t. In appendix 2 we show that one can express $H2(\mathbf{x}_t)$ as

$$H2(\mathbf{x}) = \exp\Big(-\bar{\alpha}_\tau - \bar{\boldsymbol{\beta}}_\tau'(e^{-K(s-t)}\mathbf{x}_t + (I - e^{-K(s-t)})\boldsymbol{\mu}) + \frac{1}{2}\bar{\boldsymbol{\beta}}_\tau'\Omega_{s-t}\bar{\boldsymbol{\beta}}_\tau\Big). \tag{7}$$

The parameters $\bar{\alpha}_\tau$ and $\bar{\boldsymbol{\beta}}_\tau$ (and Ω_{s-t}) are defined in appendix 2, and are similar to α_τ^* and $\boldsymbol{\beta}_\tau^*$ from equation (5).

3. Data and Model Implementation

3.1 Data

Four types of data are used: nominal zero-coupon bond yields derived from nominal Australian Commonwealth Government bonds, Australian Commonwealth Government inflation-indexed bond prices, inflation forecasts from Consensus Economics, and historical inflation.

Nominal zero-coupon bond yields are estimated using the approach of Finlay and Chambers (2009). These nominal yields correspond to y_{t,τ_s}^n and are used in computing our function $H1(\mathbf{x}_t)$

from equation (6). Note that the Australian nominal yield curve has maximum maturity of roughly twelve years. We extrapolate nominal yields beyond this by assuming that the nominal and real yield curves have the same slope. This allows us to utilize the prices of all inflation-indexed bonds, which have maturities of up to twenty-four years (in practice, the slope of the real yield curve beyond twelve years is very flat, so that if we instead hold the nominal yield curve constant beyond twelve years, we obtain virtually identical results).

We calculate the real prices of inflation-indexed bonds using yield data.[4] Our sample runs from July 1992 to December 2010, with the available data sampled at monthly intervals up to June 1994 and weekly intervals thereafter; bonds with less than one year remaining to maturity are excluded. By comparing these computed inflation-indexed bond prices, which form the P_t^r in equation (6), with our function $H1(\mathbf{x}_t)$, we are able to estimate the latent factors. We assume that the standard deviation of the bond price measurement error is 4 basis points. This is motivated by market liaison which suggests that, excluding periods of market volatility, the bid-ask spread has stayed relatively constant over the period considered, at around 8 basis points. Some descriptive statistics for nominal and inflation-indexed bonds are given in table 1.

Note that inflation-indexed bonds are relatively illiquid, especially in comparison to nominal bonds.[5] Therefore, inflation-indexed bond yields potentially incorporate liquidity premia, which could bias our results. As discussed, we use inflation forecasts as a measure of inflation expectations. These forecasts serve to tie down inflation expectations, and as such we implicitly assume that liquidity premia are included in our measure of risk premia. We also assume that the existence of liquidity premia causes a level shift in estimated risk premia but does not greatly bias the estimated *changes* in risk premia.[6]

[4] Available from table F16 at www.rba.gov.au/statistics/tables/index.html.

[5] Average yearly turnover between 2003–04 and 2007–08 was roughly $340 billion for nominal government bonds and $15 billion for inflation-indexed bonds, which equates to a turnover ratio of around 7 for nominal bonds and 2½ for inflation-indexed bonds (see Australian Financial Markets Association 2008).

[6] Inflation swaps are now more liquid than inflation-indexed bonds and may provide alternative data for use in estimating inflation expectations at some point in the future. Currently, however, there is not a sufficiently long time series of inflation swap data to use for this purpose.

Table 1. Descriptive Statistics of Bond Price Data

Statistic		Time Period 1992–1995	1996–2000	2001–2005	2006–2010
Number of Bonds:	Nominal	12–19	12–19	8–12	8–14
	Inflation Indexed	3–5	4–5	3–4	2–4
Maximum Tenor:	Nominal	11–13	11–13	11–13	11–14
	Inflation Indexed	13–21	19–24	15–20	11–20
Average Outstanding:	Nominal	49.5	70.2	50.1	69.5
	Inflation Indexed	2.1	5.0	6.5	7.1

Note: Tenor in years; outstandings in billions; only bonds with at least one year to maturity are included.

The inflation forecasts are taken from Consensus Economics. We use three types of forecast:

(i) monthly forecasts of the percentage change in CPI over the current and the next calendar year

(ii) quarterly forecasts of the year-on-year percentage change in the CPI for seven or eight quarters in the future

(iii) biannual forecasts of the year-on-year percentage change in the CPI for each of the next five years, as well as from five years in the future to ten years in the future

We use the function $H2(\mathbf{x}_t)$ to relate these inflation forecasts to the latent factors, and use the past forecasting performance of the inflation forecasts relative to realized inflation to calibrate the standard deviation of the measurement errors.

Historical inflation enters the model in the form of $I_{t,\Delta}$ from section 2.2, but otherwise is not used in estimation. This is because the fundamental variable being modeled is the *current instantaneous* inflation rate. Given the inflation law of motion (implicitly defined by equations (2)–(4)), inflation expectations and inflation-indexed bond prices are affected by current inflation and so can inform our estimation. By contrast, the published inflation rate is always “old news" from the perspective of our model and so has nothing direct to say about current instantaneous inflation.[7]

3.2 The Kalman Filter and Maximum-Likelihood Estimation

We use the Kalman filter to estimate the three latent factors, using data on bond prices and inflation forecasts. The Kalman filter can estimate the state of a dynamic system from noisy observations. It does this by using information about how the state evolves over time, as summarized by the state equation, and relating the state to noisy observations using the measurement equation. In our case the latent factors constitute the state of the system and our bond prices and forecast data constitute the noisy observations. From the latent factors we are able to make inferences about inflation expectations and inflation risk premia.

The standard Kalman filter was developed for a linear system. Although our state equation (given by equation (14)) is linear, our measurement equations, using $H1(\mathbf{x}_t)$ and $H2(\mathbf{x}_t)$ as derived in sections 2.2 and 2.3, are not. This is because we work with coupon-bearing bond prices instead of zero-coupon yields. We overcome this problem by using a central difference Kalman filter, which is a type of non-linear Kalman filter.[8]

The approximate log-likelihood is evaluated using the forecast errors of the Kalman filter. If we denote the Kalman filter's forecast of the data at time t by $\hat{\mathbf{y}}_t(\zeta, \mathbf{x}_t(\zeta, \mathbf{y}_{t-1}))$—which depends on the parameters (ζ) and the latent factors ($\mathbf{x}_t(\zeta, \mathbf{y}_{t-1})$), which in turn depend on the parameters and the data observed up to time $t-1$ ($\mathbf{y}_{t-1}$)—then the approximate log-likelihood is given by

$$\mathcal{L}(\zeta) = -\sum_{t=1}^{T} \left(\log |P_{\mathbf{y}_t}| + (\mathbf{y}_t - \hat{\mathbf{y}}_t) P_{\mathbf{y}_t}^{-1} (\mathbf{y}_t - \hat{\mathbf{y}}_t)' \right).$$

[7] Note that our model is set in continuous time; data are sampled discretely, but all quantities—for example, the inflation law of motion as well as inflation yields and expectations—evolve continuously. π_t^i from equation (2) is the current instantaneous inflation rate, not a one-month or one-quarter rate.

[8] See appendix 3 for more detail on the central difference Kalman filter.

Here the estimated covariance matrix of the forecast data is denoted by $P_{\mathbf{y}_t}$.[9] In the model the parameters are given by $\zeta = (K, \boldsymbol{\lambda}_0, \Lambda, \rho_0, \Sigma)$.

We numerically optimize the log-likelihood function to obtain parameter estimates. From the parameter estimates we use the Kalman filter to obtain estimates of the latent factors.

3.3 Calculation of Model Estimates

For a given set of model parameters and latent factors, we can calculate inflation forward rates, expected future inflation rates, and inflation risk premia.

In appendix 2 we show that the expected future inflation rate at time t for time $t + \tau$ can be expressed as

$$\mathbb{E}_t\left(\pi^i_{t+\tau}\right) = \rho_0 + \mathbf{1}' \cdot e^{-K\tau}\mathbf{x}_t.$$

The inflation forward rate at time t for time $t + \tau$, $f^i_{t,\tau}$, is the rate of inflation at time $t + \tau$ implied by market prices of nominal and inflation-indexed bonds trading at time t. It is related to the inflation yield via $y^i_{t,\tau} = \int_t^{t+\tau} f^i_{t,s}\,\mathrm{d}s$.[10] As bond prices incorporate inflation risk, so does the inflation forward rate. In our model the inflation forward rate is given by

$$\begin{aligned} f^i_{t,\tau} &= \rho_0 + \mathbf{1}' \cdot (e^{-K^*\tau}\mathbf{x}_t + (I - e^{-K^*\tau})\boldsymbol{\mu}^*) \\ &\quad - \frac{1}{2}(\mathbf{1}'(I - e^{-K^*\tau})K^{*-1}\Sigma)(\mathbf{1}'(I - e^{-K^*\tau})K^{*-1}\Sigma)'. \end{aligned}$$

See appendix 2 for details on the above and definitions of K^* and $\boldsymbol{\mu}^*$.

The inflation risk premium is given by the difference between the inflation forward rate, which incorporates risk aversion, and

[9] In actual estimation we exclude the first six months of data from the likelihood calculation to allow "burn-in" time for the Kalman filter.

[10] Note that at time t the inflation *forward* rate at time $s > t$, $f^i_{t,s}$, is known, as it is determined by known inflation yields. The inflation rate, π^i_s, that will prevail at s is unknown, however, and in our model is a random variable. π^i_s is related to the known inflation yield by $\exp(-y^i_{t,\tau}) = \mathbb{E}_t(\exp(-\int_t^{t+\tau} \pi^{i*}_s\,\mathrm{d}s))$ so that $y^i_{t,\tau} = -\log(\mathbb{E}_t(\exp(-\int_t^{t+\tau} \pi^{i*}_s\,\mathrm{d}s)))$, where π^{i*}_s is the so-called risk-neutral version of π^i_s (see appendix 2 for details).

the expected future inflation rate, which is free of risk aversion. The inflation risk premium at time t for time $t + \tau$ is given by $f^i_{t,\tau} - \mathbb{E}_t(\pi^i_{t+\tau})$.

4. Results

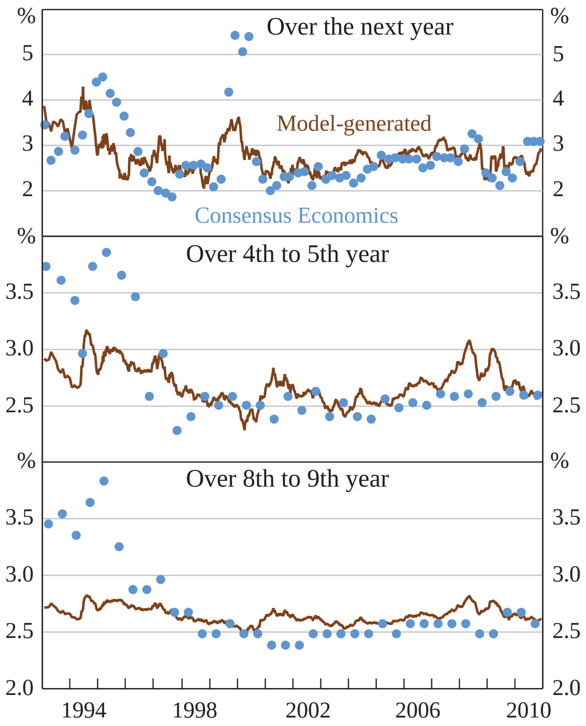
4.1 Model Parameters and Fit to Data

We estimate the model over the period July 31, 1992 to December 15, 2010 using a number of different specifications. First we estimate both two- and three-factor versions of our model. Using a likelihood-ratio test, we reject the hypothesis that there is no improvement of model fit between the two-factor model and three-factor model and so use the three-factor model. (Three factors are usually considered sufficient in the literature, with for example the overwhelming majority of variation in yields captured by the first three principal components.)

We also consider three-factor models with and without forecast data. Both models are able to fit the inflation yield data well, with a mean absolute error between ten-year inflation yields as estimated from the models and ten-year break-even inflation calculated directly from bond prices of around 5 basis points.[11] The model without forecast data gives unrealistic estimates of inflation expectations and inflation risk premia, however: ten-year-ahead inflation expectations are implausibly volatile and can be as high as 8 percent and as low as −1 percent, which is not consistent with economists' forecasts. These findings are consistent with those of Kim and Orphanides (2005), where the use of forecast data is advocated as a means of separating expectations from risk premia. Note, however, that estimates from the model with forecast data are not solely determined by the forecasts; the model estimates of expected future inflation only roughly match the forecast data and on occasion deviate significantly from them, as seen in figure 1.

[11] The divergence between model yields and those measured directly from bond data is mainly due to the different types of yields not being directly comparable—model estimates are zero-coupon yields that take into account indexation lag, while the direct measure is estimated from coupon-bearing bonds which reflect a certain amount of historical inflation.

Figure 1. Forecast Change in CPI



Source: Consensus Economics; authors' calculations.

Our preferred model is thus the three-factor model estimated using forecast data. Likelihood-ratio tests indicate that two parameters of that model (Λ_{11} and Λ_{21}) are statistically insignificant and so they are excluded. Our final preferred model has twenty freely estimated parameters, which are given in table 2. We note that the estimate of ρ_0, the steady-state inflation rate in our model, is 2.6 percent, which is within the inflation target range. The persistence of inflation is essentially determined by the diagonal entries of the K matrix, which drives the inflation law of motion as defined by equations (2)–(4). The first diagonal entry of K is 0.19, which in a single-factor model would imply a half-life of the first latent factor (being the time taken for the latent factor, and so inflation, to revert halfway back to its mean value after experiencing a shock) of around

Table 2. Parameter Estimates for Final Model (Model Estimated 1992–2010)

Parameter	Index Number *(i)*		
	1	**2**	**3**
ρ_0	2.64 (0.26)	—	—
$(K)_{1i}$	0.19 (0.02)	0	0
$(K)_{2i}$	−2.88 (0.05)	1.75 (0.05)	0
$(K)_{3i}$	1.11 (0.05)	1.74 (0.05)	0.80 (0.01)
$(\Sigma)_{ii}$	0.11 (0.02)	1.51 (0.10)	0.96 (0.02)
$\lambda_{0,i}$	0.12 (0.01)	0.10 (0.01)	−0.01 (0.00)
$(\Lambda)_{1i}$	0	55.44 (0.32)	15.31 (0.06)
$(\Lambda)_{2i}$	0	−107.80 (0.26)	−8.91 (0.06)
$(\Lambda)_{3i}$	−12.38 (0.08)	−144.22 (0.45)	−73.07 (0.20)

Note: ρ_0 and $(\Sigma)_{ii}$ are given in percentage points; standard errors are shown in parentheses.

3½ years. The half-lives of the other two latent factors would be five and ten months.

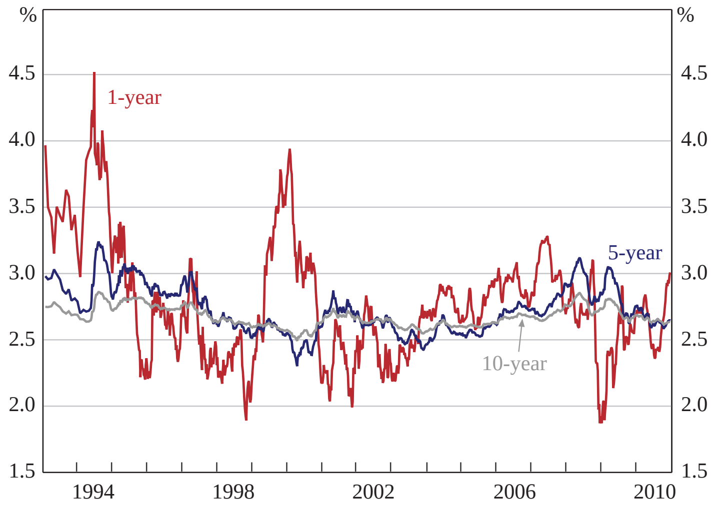
4.2 Qualitative Discussion of Results

4.2.1 Inflation Expectations

Our estimated expected future inflation rates at horizons of one, five, and ten years are shown in figure 2. Two points stand out immediately: one-year-ahead inflation expectations are much more volatile than five- and ten-year-ahead expectations and, as may be expected, are strongly influenced by current inflation (not shown); longer-term inflation expectations appear to be well anchored within the 2 to 3 percent target range.

We see that there is a general decline in inflation expectations from the beginning of the sample until around 1999, the year before the introduction of the Goods and Services Tax (GST). The estimates suggest that the introduction of the GST on July 1, 2000 resulted in a large one-off increase in short-term inflation expectations. This is reflected in the run-up in one-year-ahead inflation expectations over calendar year 1999, although the peak in

Figure 2. Expected Inflation Rates



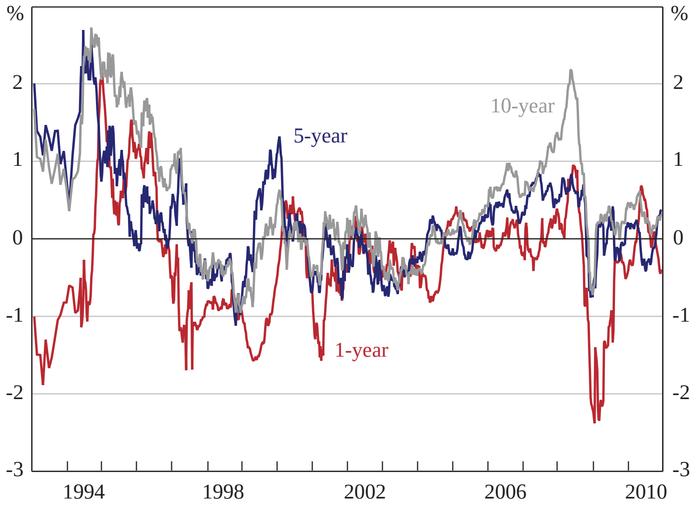
the estimated expectations is below the actual peak in year-end CPI growth of 6.1 percent.[12] Of particular interest, however, is the non-responsiveness of five- and ten-year-ahead expectations, which should be the case if the inflation target is seen as credible.

Long-term expectations increased somewhat between mid-2000 and mid-2001, perhaps prompted by easier monetary conditions globally as well as relatively high inflation in Australia. Interestingly, there appears to have been a sustained general rise in inflation expectations between 2004 and 2008 at all horizons. Again this was a time of rising domestic inflation, strong world growth, a boom in the terms of trade, and rising asset prices.

In late 2008 the inflation outlook changed and short-term inflation expectations fell dramatically, likely in response to forecasts of very weak global demand caused by the financial crisis. Longer-term expectations also fell before rising over the early part of 2009 as authorities responded to the crisis. The subsequent moderation of longer-term expectations, as well as the relative stabilization of short-term expectations over 2010 suggests that financial market

[12] The legislation introducing the GST was passed through Parliament in June 1999.

Figure 3. Inflation Risk Premia



participants considered the economic outlook and Australian authorities' response to the crisis sufficient to maintain inflation within the target range.

The latest data, corresponding to December 2010, shows one-year-ahead inflation expectations exceeding 3 percent, close to the Reserve Bank forecast for inflation of 2.75 percent over the year to December 2011 given in the November 2010 *Statement of Monetary Policy*. Longer-term model-implied inflation expectations as of December 2010 are for inflation close to the middle of the 2 to 3 percent inflation target range.

4.2.2 Inflation Risk Premia

Although more volatile than our long-term inflation expectation estimates, long-term inflation risk premia broadly followed the same pattern—declining over the first third of the sample, gradually increasing between 2004 and 2008 before falling sharply with the onset of the global financial crisis, and then rising again as markets reassessed the likelihood of a severe downturn in Australia (figure 3). The main qualitative point of difference between the two series is in their reaction to the GST. As discussed earlier, the estimates of long-term inflation expectations remained well anchored during the GST period, whereas (as we can see from figure 3) the estimates of long-term risk premia rose sharply. As the terminology suggests, inflation expectations represent investors' central forecast for inflation, while risk premia can be thought of as representing second-order information—essentially how uncertain investors are about their central forecasts and how much they dislike this uncertainty. So while longer-dated expectations of inflation did not change around the introduction of the GST, the rise in risk premia indicates a more variable and uncertain inflation outlook.

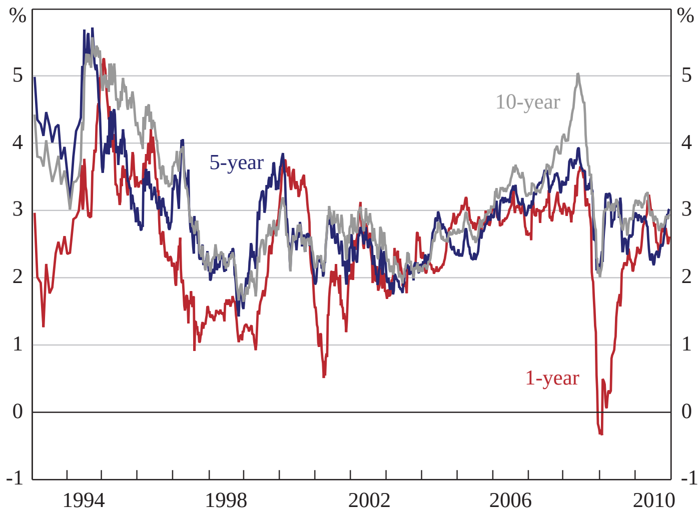
Although our estimates show periods of negative inflation risk premia, indicating that investors were happy to be exposed to inflation risk, this is probably not the case in reality. In our model, inflation risk premia are given by forward rates of inflation (as implied by the inflation yield curve), less inflation expectations. The inflation yield curve is given as the difference between nominal and real yields. Hence if real yields contain a liquidity premium, they will be higher, shifting the inflation yield curve down and reducing the estimated inflation risk premia to below their true level. The inflation-indexed bond market is known to be relatively illiquid in comparison with the nominal bond market, and this provides a plausible explanation for our negative estimates.

If the illiquidity in the inflation-indexed bond market is constant through time, then the level of our estimated risk premium will be biased but *changes* in the risk premium should be accurately estimated. Market liaison suggests that an assumption of relatively constant liquidity, at least during normal times, is not an unreasonable one; as noted earlier, for example, bid-ask spreads have stayed relatively constant over most of the period under consideration. The trough in inflation risk premia around late 2008 and early 2009 may be one exception to this, however, with the liquidity premia for inflation-indexed bonds relative to nominal bonds possibly increasing (in line with increases in liquidity premia for most assets relative to highly rated and highly liquid government securities at this time).

4.2.3 Inflation Forward Rates

The inflation forward rate reflects the relative prices of traded nominal and inflation-indexed bonds and is given by the sum of inflation

Figure 4. Inflation Forward Rates



expectations and inflation risk premia. As estimates of longer-term inflation expectations are relatively stable, movements in the five- and ten-year inflation forward rates tend to be driven by changes in estimated risk premia. The inflation forward rate, as shown in figure 4, generally falls during the first third of the sample, rises around the time of the GST, and rises between 2004 and 2008 before falling sharply with the onset of the financial crisis and then rising again.[13]

One notable feature of figure 4 is the negative inflation forward rates recorded in late 2008. This phenomenon is essentially due to very low break-even inflation rates embodied in the bond price data (two-year-ahead nominal less real yields were only around 90 basis points at this time), together with high realized inflation over 2008; as break-even inflation rates reflect around five months of historical inflation, a low two-year break-even inflation rate and high historical inflation necessarily implies a very low or even negative inflation

[13]Note that studies using U.S. and UK data essentially start with the inflation forward rate, which they decompose into inflation expectations and inflation risk premia. Due to a lack of data, we cannot do this, and instead we estimate inflation forward rates as part of our model.

forward rate in the near future. The low break-even inflation rates in turn are due to the yields on inflation-indexed bonds rising relative to the yields on nominal bonds.

5. Discussion and Conclusion

The model just described is designed to give policymakers accurate and timely information on market-implied inflation expectations. It has a number of advantages over existing sources for such data, which primarily constitute either break-even inflation derived from bond prices or inflation forecasts sourced from market economists.

As argued, break-even inflation as derived directly from bond prices has a number of drawbacks as a measure of inflation expectations: such a measure gives average inflation over the tenor of the bond, not inflation at a certain date in the future; government bonds in Australia are coupon bearing, which means that yields of similar-maturity nominal and inflation-indexed bonds are not strictly comparable; there are very few inflation-indexed bonds on issue in Australia, which means that break-even inflation can only be calculated at a limited number of tenors; inflation-indexed bonds are indexed with a lag, which means that their yields reflect historical inflation, not just future expected inflation; and finally, bond yields incorporate risk premia so that the level of, and even changes in, break-even inflation need not give an accurate read on inflation expectations. Our model addresses each of these issues: we model inflation-indexed bonds as consisting of a stream of payments where the value of each payment is determined by nominal interest rates, historical inflation, future inflation expectations, and inflation risk premia. This means we are able to produce estimates of expected future inflation at any time and for any tenor which are free of risk premia and are not affected by historical inflation.

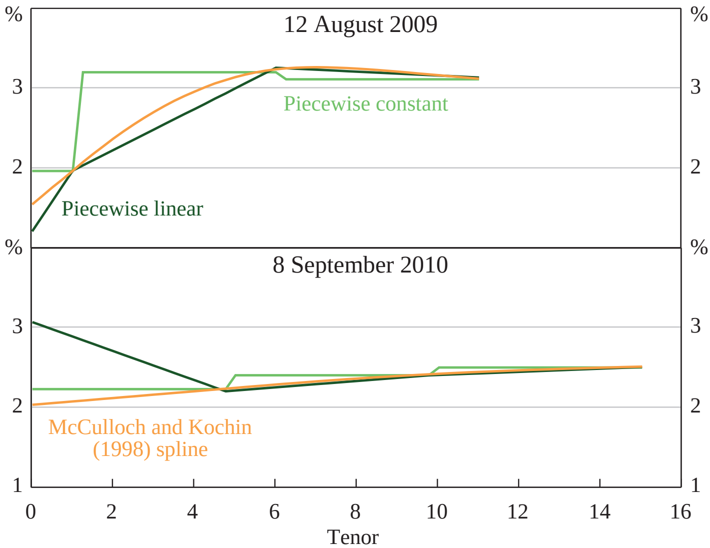
Model-derived inflation expectations also have a number of advantages over expectations from market economists: unlike survey-based expectations, they are again available at any time and for any tenor, and they reflect the agglomerated knowledge of all market participants, not just the views of a small number of economists. By contrast, the main drawback of our model is its complexity—break-even inflation and inflation forecasts have their faults but are transparent and simple to measure, whereas our model, while addressing a number of faults, is by comparison complex and difficult to estimate.

Standard affine term structure models, which take as inputs zero-coupon yield curves and give as outputs expectations and risk premia, have existed in the literature for some time. Our main contribution to this literature, apart from the estimation of inflation expectations and inflation risk premia for Australia, is our reformulation of the model in terms of coupon-bearing bond prices instead of zero-coupon yields. In practice, zero-coupon yields are not directly available but must be estimated, so by fitting the affine term structure model directly to prices, we avoid inserting a second arbitrary yield-curve model between the data and our final model. When many bond prices are available, this is only a small advantage, as accurate zero-coupon yields can be recovered from the well-specified coupon-bearing yield curve. When only a small number of bond prices are available, our method provides a major advantage: one can fit a zero-coupon yield curve to only two or three far-spaced coupon-bearing yields, and indeed McCulloch and Kochin (1998) provide a procedure for doing this, but there are limitless such curves that can be fitted with no a priori correct criteria to choose between them.

The inability to pin down the yield curve is highlighted in figure 5, which shows three yield curves—one piecewise constant, one piecewise linear and starting from the current six-month annualized inflation rate, and one following the method of McCulloch and Kochin (1998)—all fitted to inflation-indexed bond yields on two different dates. All curves fit the bond data perfectly, as would any number of other curves, so there is nothing in the underlying data to motivate a particular choice, yet different curves can differ by as much as 1 percentage point. Our technique provides a method for removing this intermediate curve-fitting step and estimating directly with the underlying data instead of the output of an arbitrary yield-curve model. The fact that we price bonds directly in terms of the underlying inflation process also allows for direct modeling of the lag involved in inflation indexation and the impact that historically observed inflation has on current yields, a second major advantage.

In sum, the affine term structure model used in this paper addresses a number of problems inherent in alternative approaches to measuring inflation expectations, and produces plausible measures of inflation expectations over the inflation targeting era. Given the complexity of the model and the limited number of inflation-indexed

Figure 5. Zero-Coupon Real Yield Curves



bonds on issue, some caution should be applied in interpreting the results. A key finding of the model is that long-term inflation expectations appear to have been well anchored within the inflation target over most of the sample. Conversely, one-year-ahead inflation expectations appear to be closely tied to CPI inflation and are more variable than longer-term expectations. Given the relative stability of our estimates of long-term inflation expectations, changes in five- and ten-year inflation forward rates, and so in break-even inflation rates, are by implication driven by changes in inflation risk premia. As such, our measure has some benefits over break-even inflation rates in measuring inflation expectations.

Appendix 1. Yields and Stochastic Discount Factors

The results of this paper revolve around the idea that inflation expectations are an important determinant of the inflation yield. In this section we make clear the relationships between real, nominal, and inflation yields; inflation expectations; and inflation risk premia. We also link these quantities to standard asset pricing models as discussed, for example, in Cochrane (2005).

Real and Nominal Yields and SDFs

Let M_t^r be the real SDF or pricing kernel, defined such that

$$P_{t,\tau} = \mathbb{E}_t\left(\frac{M_{t+\tau}^r}{M_t^r}x_{t+\tau}\right) \tag{8}$$

holds for any asset, where $P_{t,\tau}$ is the price of the asset at time t which has (a possibly random) payoff $x_{t+\tau}$ occurring at time $t+\tau$. A zero-coupon inflation-indexed bond maturing at time $t+\tau$, $P_{t,\tau}^r$, is an asset that pays one real dollar, or equivalently one unit of consumption, for certain. That is, it is an asset with payoff $x_{t+\tau} \equiv 1$. If we define the real yield by $y_{t,\tau}^r = -\log(P_{t,\tau}^r)$, we can use equation (8) with $x_{t+\tau} = 1$ to write

$$y_{t,\tau}^r = -\log\left(P_{t,\tau}^r\right) = -\log\left(\mathbb{E}_t\left(\frac{M_{t+\tau}^r}{M_t^r}\right)\right). \tag{9}$$

This defines the relationship between real yields and the continuous-time real SDF.

A zero-coupon nominal bond maturing at time $t+\tau$ is an asset that pays one nominal dollar for certain. If we define Q_t to be the price index, then the payoff of this bond is given by $x_{t+\tau} = Q_t/Q_{t+\tau}$ units of consumption. Taking $x_{t+\tau} = Q_t/Q_{t+\tau}$ in equation (8), we can relate the nominal yield $y_{t,\tau}^n$ to the nominal bond price $P_{t,\tau}^n$ and the continuous-time real SDF by

$$y_{t,\tau}^n = -\log\left(P_{t,\tau}^n\right) = -\log\left(\mathbb{E}_t\left(\frac{M_{t+\tau}^r}{M_t^r}\frac{Q_t}{Q_{t+\tau}}\right)\right).$$

Motivated by this result, we define the continuous-time *nominal* SDF by $M_{t+\tau}^n = M_{t+\tau}^r/Q_{t+\tau}$, so that

$$y_{t,\tau}^n = -\log\left(P_{t,\tau}^n\right) = -\log\left(\mathbb{E}_t\left(\frac{M_{t+\tau}^n}{M_t^n}\right)\right). \tag{10}$$

Inflation Yields and the Inflation SDF

The inflation yield is defined to be the difference in yields between zero-coupon nominal and inflation-indexed bonds of the same maturity,

$$y_{t,\tau}^i = y_{t,\tau}^n - y_{t,\tau}^r. \tag{11}$$

As in Beechey (2008), we define the continuous-time *inflation* SDF, $M^i_{t+\tau}$, such that the pricing equation for inflation yields holds—that is, such that

$$y^i_{t,\tau} = -\log\left(\mathbb{E}_t\left(\frac{M^i_{t+\tau}}{M^i_t}\right)\right). \tag{12}$$

All formulations of $M^i_{t+\tau}$ which ensure that equations (9), (10), and (11) are consistent with equation (12) are equivalent from the perspective of our model. One such formulation is to define the inflation SDF as

$$M^i_{t+\tau} = \frac{M^n_{t+\tau}}{\mathbb{E}_t(M^r_{t+\tau})}. \tag{13}$$

We can then obtain equation (12) by substituting equations (9) and (10) into equation (11) and using the definition of the inflation SDF given in equation (13). In this case we have

$$\begin{aligned} y^i_{t,\tau} = y^n_{t,\tau} - y^r_{t,\tau} &= -\log\left(\mathbb{E}_t\left(\frac{M^n_{t+\tau}}{M^n_t}\right)\right) + \log\left(\mathbb{E}_t\left(\frac{M^r_{t+\tau}}{M^r_t}\right)\right) \\ &= -\log\left(\frac{M^r_t}{M^n_t}\mathbb{E}_t\left(\frac{M^n_{t+\tau}}{\mathbb{E}_t\left(M^r_{t+\tau}\right)}\right)\right) = -\log\left(\mathbb{E}_t\left(\frac{M^i_{t+\tau}}{M^i_t}\right)\right), \end{aligned}$$

as desired. If one assumed that $M^r_{t+\tau}$ and $Q_{t+\tau}$ were uncorrelated, a simpler formulation would be to take $M^i_{t+\tau} = 1/Q_{t+\tau}$. Since $M^n_{t+\tau} = M^r_{t+\tau}/Q_{t+\tau}$, in this case we would have $\mathbb{E}_t(M^n_{t+\tau}/M^n_t) = \mathbb{E}_t(M^r_{t+\tau}/M^r_t)\mathbb{E}_t(Q_t/Q_{t+\tau})$, so that $y^n_{t,\tau} = -\log(\mathbb{E}_t(M^r_{t+\tau}/M^r_t)) - \log(\mathbb{E}_t(Q_t/Q_{t+\tau}))$ and $y^i_{t,\tau} = y^n_{t,\tau} - y^r_{t,\tau} = -\log(\mathbb{E}_t(Q_t/Q_{t+\tau})) = -\log(\mathbb{E}_t(M^i_{t+\tau}/M^i_t))$, as desired.

Interpretation of Other SDFs in Our Model

We model M^i_t directly as $\mathrm{d}M^i_t/M^i_t = -\pi^i_t\mathrm{d}t - \boldsymbol{\lambda}^{i\prime}_t\mathrm{d}\mathbf{B}_t$, where we take π^i_t as the instantaneous inflation rate and $\boldsymbol{\lambda}^i_t$ as the market price of inflation risk. Although very flexible, this setup means that in our model the relationship between different stochastic discount factors in the economy is not fixed.

In models such as ours there are essentially three quantities of interest, any two of which determine the other: the real SDF, the nominal SDF, and the inflation SDF. As we make assumptions about only one of these quantities, we do not tie down the model completely. Note that we *could* make an additional assumption to tie down the model. Such an assumption would not affect the model-implied inflation yields or inflation forecasts, however, which are the only data our model sees and so, in the context of our model, would be arbitrary.

This situation of model ambiguity is not confined to models of inflation compensation such as ours. The extensive literature which fits affine term structure models to nominal yields contains a similar kind of ambiguity. Such models typically take the nominal SDF as driven by $\mathrm{d}M_t^n/M_t^n = -r_t^n\mathrm{d}t - \boldsymbol{\lambda}_t^{n\prime}\mathrm{d}\mathbf{B}_t$, where once again the real SDF and inflation process are not explicitly modeled, so that, similar to our case, the model is not completely tied down.

Inflation Expectations and the Inflation Risk Premium

Finally, we link our inflation yield to inflation expectations and the inflation risk premium. The inflation risk premium arises because people who hold nominal bonds are exposed to inflation, which is uncertain, and so demand compensation for bearing this risk. If we set $m_{t,\tau} = \log(M_{t+\tau}^r/M_t^r)$ and $q_{t,\tau} = \log(Q_{t+\tau}/Q_t)$, which are both assumed normal, and use the identity $\mathbb{E}_t(\exp(X)) = \exp(\mathbb{E}_t(X) + \frac{1}{2}\mathbb{V}_t(X))$ where X is normally distributed and $\mathbb{V}(\cdot)$ is variance, we can work from equation (11) to derive

$$y_{t,\tau}^i = \mathbb{E}_t(q_{t,\tau}) - \frac{1}{2}\mathbb{V}_t(q_{t,\tau}) + \mathbb{C}\mathrm{ov}_t(m_{t,\tau}, q_{t,\tau}).$$

The first term above is the expectations component of the inflation yield, while the last two terms constitute the inflation risk premium (incorporating a “Jensen’s” or “convexity” term).

Appendix 2. The Mathematics of Our Model

We first give some general results regarding affine term structure models, then relate these results to our specific model and its interpretation.

Some Results Regarding Affine Term Structure Models

Start with the latent factor process,

$$\mathrm{d}\mathbf{x}_t = K(\boldsymbol{\mu} - \mathbf{x}_t)\mathrm{d}t + \Sigma\, \mathrm{d}\mathbf{B}_t.$$

Given $\mathbf{x}_t$, we have, for $s > t$ (see, for example, Duffie 2001, p. 342),

$$\begin{aligned}\mathbf{x}_s &= e^{-K(s-t)}\left(\mathbf{x}_t + \int_t^s e^{K(u-t)}K\boldsymbol{\mu}\,\mathrm{d}u + \int_t^s e^{K(u-t)}\Sigma\,\mathrm{d}B_u\right)\\ &\overset{D}{=} e^{-K(s-t)}\mathbf{x}_t + (I - e^{-K(s-t)})\boldsymbol{\mu} + \boldsymbol{\epsilon}_{t,s},\end{aligned} \tag{14}$$

where '$\overset{D}{=}$' denotes equality in distribution and $\boldsymbol{\epsilon}_{t,s} \sim N(\mathbf{0}, \Omega_{s-t})$ with

$$\begin{aligned}\Omega_{s-t} &= e^{-K(s-t)}\left(\int_t^s e^{K(u-t)}\Sigma\Sigma' e^{K'(u-t)}\,\mathrm{d}u\right)e^{-K'(s-t)}\\ &= \int_0^{s-t} e^{-Ku}\Sigma\Sigma' e^{-K'u}\,\mathrm{d}u.\end{aligned}$$

Further, if we define

$$\pi_t = \rho_0 + \boldsymbol{\rho}'\mathbf{x}_t,$$

then since $\int_t^{t+\tau}\pi_s\,\mathrm{d}s$ is normally distributed,

$$\begin{aligned}&\mathbb{E}_t\left(\exp\left(-\int_t^{t+\tau}\pi_s\,\mathrm{d}s\right)\right)\\ &\quad= \exp\left(-\mathbb{E}_t\left(\int_t^{t+\tau}\pi_s\,\mathrm{d}s\right) + \frac{1}{2}\mathbb{V}_t\left(\int_t^{t+\tau}\pi_s\,\mathrm{d}s\right)\right)\end{aligned}$$

with

$$\begin{aligned}\int_t^{t+\tau}\pi_s\,\mathrm{d}s &= \int_t^{t+\tau}\rho_0 + \boldsymbol{\rho}'\mathbf{x}_s\,\mathrm{d}s\\ &= \int_t^{t+\tau}\rho_0 + \boldsymbol{\rho}'\Big(e^{-K(s-t)}\mathbf{x}_t + (I - e^{-K(s-t)})\boldsymbol{\mu}\\ &\qquad + e^{-K(s-t)}\int_t^s e^{K(u-t)}\Sigma\mathrm{d}B_u\Big)\,\mathrm{d}s\end{aligned}$$

$$= \int_t^{t+\tau} \rho_0 + \boldsymbol{\rho}'(e^{-K(s-t)}\mathbf{x}_t + (I - e^{-K(s-t)})\boldsymbol{\mu})\mathrm{d}s$$
$$+ \int_t^{t+\tau} \boldsymbol{\rho}' \left(\int_u^{t+\tau} e^{-K(s-t)}\mathrm{d}s \right) e^{K(u-t)}\Sigma\, \mathrm{d}B_u, \quad (15)$$

where we have used a stochastic version of Fubini's theorem to change the order of integration (see, for example, Da Prato and Zabczyk 1993, p. 109). Evaluating the inner integral of line (15), using Itô's isometry (see, for example, Steele 2001, p. 82) and making the change of variable $s = t + \tau - u$, we have

$$\mathbb{E}_t \left(\int_t^{t+\tau} \pi_s \, \mathrm{d}s \right) = \int_0^{\tau} \rho_0 + \boldsymbol{\rho}'(e^{-Ks}\mathbf{x}_t + (I - e^{-Ks})\boldsymbol{\mu})\mathrm{d}s$$
$$\mathbb{V}_t \left(\int_t^{t+\tau} \pi_s \, \mathrm{d}s \right) = \int_0^{\tau} (\boldsymbol{\rho}'(I - e^{-Ks})K^{-1}\Sigma)^2 \, \mathrm{d}s,$$

where for $\mathbf{x}$, a vector we define $\mathbf{x}^2 = \mathbf{x}'^2$ as the vector dot-product $\mathbf{x}'\mathbf{x}$. Hence

$$\mathbb{E}_t \left(\exp\left(- \int_t^{t+\tau} \pi_s \, \mathrm{d}s \right) \right) = \exp\left(- \int_0^{\tau} \boldsymbol{\rho}' e^{-Ks}\mathbf{x}_t \, \mathrm{d}s \right.$$
$$\left. - \int_0^{\tau} \rho_0 + \boldsymbol{\rho}'(I - e^{-Ks})\boldsymbol{\mu} - \frac{1}{2}(\boldsymbol{\rho}'(I - e^{-Ks})K^{-1}\Sigma)^2 \mathrm{d}s \right).$$

Now for $M'_{1,\tau} = (I - e^{-K\tau})K^{-1}$ we have

$$\int_0^{\tau} \boldsymbol{\rho}' e^{-Ks}\mathbf{x}_t \mathrm{d}s = \boldsymbol{\rho}'(I - e^{-Kt})K^{-1}\mathbf{x}_t = \boldsymbol{\rho}' M'_{1,\tau}\mathbf{x}_t,$$

while

$$\int_0^{\tau} \boldsymbol{\rho}'(I - e^{-Ks})\boldsymbol{\mu} \, \mathrm{d}s = \boldsymbol{\rho}'(\tau I + e^{-K\tau}K^{-1} - K^{-1})\boldsymbol{\mu}$$
$$= \boldsymbol{\rho}'(\tau I - M'_{1,\tau})\boldsymbol{\mu}$$

and

$$\begin{aligned}
&\int_0^\tau -\frac{1}{2}(\boldsymbol{\rho}'(I-e^{-Ks})K^{-1}\Sigma)^2\mathrm{d}s \\
&\quad = -\frac{1}{2}\boldsymbol{\rho}'K^{-1}\left(\int_0^\tau (I-e^{-Ks})\Sigma\Sigma'(I-e^{-K's})\mathrm{d}s\right)K^{-1\prime}\boldsymbol{\rho} \\
&\quad = -\frac{1}{2}\boldsymbol{\rho}'K^{-1}(\tau\Sigma\Sigma' - \Sigma\Sigma' M_{1,\tau} - M_{1,\tau}'\Sigma\Sigma' + M_{2,\tau})K^{-1\prime}\boldsymbol{\rho},
\end{aligned}$$

where from Kim and Orphanides (2005), for example,

$$\begin{aligned}
M_{2,\tau} &= \int_0^\tau e^{-Ks}\Sigma\Sigma' e^{-K's}\,\mathrm{d}s \\
&= -vec^{-1}(((K\otimes I)+(I\otimes K))^{-1}vec(e^{-K\tau}\Sigma\Sigma' e^{-K'\tau} - \Sigma\Sigma')).
\end{aligned}$$

Putting this together, we have

$$\mathbb{E}_t\left(\exp\left(-\int_t^{t+\tau}\pi_s\,\mathrm{d}s\right)\right) = \exp(-\alpha_\tau - \boldsymbol{\beta}_\tau'\mathbf{x}_t) \tag{16}$$

with

$$\begin{aligned}
\alpha_\tau &= \tau\rho_0 + \boldsymbol{\rho}'(\tau I - M_{1,\tau}')\boldsymbol{\mu} \\
&\quad - \frac{1}{2}\boldsymbol{\rho}'K^{-1}(\tau\Sigma\Sigma' - \Sigma\Sigma' M_{1,\tau} - M_{1,\tau}'\Sigma\Sigma' + M_{2,\tau})K^{-1\prime}\boldsymbol{\rho} \qquad (17) \\
\boldsymbol{\beta}_\tau &= M_{1,\tau}\boldsymbol{\rho}. \qquad (18)
\end{aligned}$$

Equivalent formulas are available in Kim and Orphanides (2005).

Bond Price Formula

If we model the SDF according to

$$\begin{aligned}
\mathrm{d}M_t/M_t &= -\pi_t\mathrm{d}t - \boldsymbol{\lambda}_t'\,\mathrm{d}\mathbf{B}_t \\
\pi_t &= \rho_0 + \boldsymbol{\rho}'\mathbf{x}_t,\ \ \boldsymbol{\lambda}_t = \boldsymbol{\lambda}_0 + \Lambda\mathbf{x}_t \\
\mathrm{d}\mathbf{x}_t &= K(\boldsymbol{\mu} - \mathbf{x}_t)\mathrm{d}t + \Sigma\,\mathrm{d}\mathbf{B}_t, \qquad (19)
\end{aligned}$$

then the price of a zero-coupon bond at t paying one dollar at $t+\tau$ is given by (see, for example, Cochrane 2005)

$$\mathbb{E}_t\left(\frac{M_{t+\tau}}{M_t}\right) = \mathbb{E}_t\left(\exp\left(-\int_t^{t+\tau}\pi_t + \frac{1}{2}\boldsymbol{\lambda}_t'\boldsymbol{\lambda}_t\,\mathrm{d}t - \int_t^{t+\tau}\boldsymbol{\lambda}_t'\mathrm{d}\mathbf{B}_t\right)\right)$$
$$= \mathbb{E}_t\left(\exp\left(-\int_t^{t+\tau}\pi_s^*\,\mathrm{d}s\right)\right), \tag{20}$$

where π_s^* is like π_s in equation (19) above but with

$$\mathrm{d}\mathbf{x}_t = K^*(\boldsymbol{\mu}^* - \mathbf{x}_t)\mathrm{d}t + \Sigma\,\mathrm{d}\mathbf{B}_t,$$

where $K^* = (K + \Sigma\Lambda)$ and $\boldsymbol{\mu}^* = {K^*}^{-1}(K\boldsymbol{\mu} - \Sigma\boldsymbol{\lambda}_0)$. (Here π_s^* is the "risk-neutral" version of π_s.) Hence we can price bonds via equation (16) using K^* and $\boldsymbol{\mu}^*$ in place of K and $\boldsymbol{\mu}$ in equations (17) and (18). We can write equation (20) as

$$\exp(-\alpha_\tau^* - \boldsymbol{\beta}_\tau^{*\prime}\mathbf{x}_t) = \mathbb{E}_t\left(\exp\left(-\int_t^{t+\tau}\pi_s^*\,\mathrm{d}s\right)\right).$$

In terms of the inflation yield from equation (12), this can be written as $y_{t,\tau}^i = \alpha_\tau^* + \boldsymbol{\beta}_\tau^{*\prime}\mathbf{x}_t$.

Inflation Forecast Formula

Inflation expectations are reported in terms of percentage growth in the consumer price index, *not* average inflation (the two differ by a Jensen's inequality term). As such, expectations at time t of how the CPI will grow between time $s > t$ and time $s + \tau$ in the future correspond to a term of the form

$$\mathbb{E}_t\left(\exp\left(\int_s^{s+\tau}\pi_u\,\mathrm{d}u\right)\right) = \mathbb{E}_t\left(\mathbb{E}_s\left(\exp\left(-\int_s^{s+\tau}-\pi_u\,\mathrm{d}u\right)\right)\right)$$
$$= \mathbb{E}_t(\exp(-\bar{\alpha}_\tau - \bar{\boldsymbol{\beta}}_\tau'\mathbf{x}_s))$$
$$= \exp\left(-\bar{\alpha}_\tau - \bar{\boldsymbol{\beta}}_\tau'(e^{-K(s-t)}\mathbf{x}_t + (I - e^{-K(s-t)})\boldsymbol{\mu}) + \frac{1}{2}\bar{\boldsymbol{\beta}}_\tau'\Omega_{s-t}\bar{\boldsymbol{\beta}}_\tau\right),$$

where the last line follows since $\mathbf{x}_s|\mathbf{x}_t \sim N(e^{-K(s-t)}\mathbf{x}_t + (I - e^{-K(s-t)})\boldsymbol{\mu}, \Omega_{s-t})$. Here $\bar{\alpha}_\tau$ and $\bar{\boldsymbol{\beta}}_\tau$ are equivalent to α_τ and $\boldsymbol{\beta}_\tau$ from equations (17) and (18) but with the market price or risk $\boldsymbol{\lambda}_t$ set to zero and using $-\rho_0$ and $-\boldsymbol{\rho}$ in place of ρ_0 and $\boldsymbol{\rho}$. So if the CPI is expected to grow by 3 percent between s and $s+\tau$, for example, we would have

$$\tau \log(1+3\%) = -\bar{a}_\tau - \bar{\boldsymbol{\beta}}_\tau'(e^{-K(s-t)}\mathbf{x}_t + (I - e^{-K(s-t)})\boldsymbol{\mu}) + \frac{1}{2}\bar{\boldsymbol{\beta}}_\tau'\Omega_{s-t}\bar{\boldsymbol{\beta}}_\tau.$$

Appendix 3. Central Difference Kalman Filter

The central difference Kalman filter is a type of sigma-point filter. Sigma-point filters deal with non-linearities in the following manner:

- First, a set of points *around* the forecast of the state is generated. The distribution of these points depends on the variance of the forecast of the state.
- The measurement equations (functions $H1(\mathbf{x}_t)$ and $H2(\mathbf{x}_t)$) are used to calculate a set of forecast observation points. This set of points is used to estimate a mean and variance of the data forecasts.
- The mean and variance of the data forecasts are then used to update the estimates of the state and its variance.

The algorithm we use is that of an additive-noise central difference Kalman filter, the details of which are given below. For more details on sigma-point Kalman filters, see van der Merwe (2004).

- *Step 1.* Initialize the state vector and its covariance matrix to their unconditional expected values,

$$\hat{\mathbf{x}}_0 = [0,0,0]'$$
$$P_{\mathbf{x}_0} = \Omega_\infty.$$

- *Step 2.* Loop over $k = 1:n$, where n is the length of our data set.
- *Step 2a.* Time-update equations:

$$\hat{\mathbf{x}}_k^- = e^{-\mathbf{K}d_k}\hat{\mathbf{x}}_{k-1}$$
$$\mathbf{P}_{\mathbf{x}_k}^- = e^{-\mathbf{K}d_k}\mathbf{P}_{\mathbf{x}_{k-1}}e^{-\mathbf{K}'d_k} + \Omega_{d_k},$$

where d_k is the time in years between data point k and data point $k-1$.

- *Step 2b.* Create the sigma points,

$$X_k^0 = \hat{\mathbf{x}}_k^-$$
$$X_k^i = \hat{\mathbf{x}}_k^- + \left(h\sqrt{\mathbf{P}_{\mathbf{x}_k}^-}\right)_i \qquad i = 1,\ldots,L$$
$$X_k^i = \hat{\mathbf{x}}_k^- - \left(h\sqrt{\mathbf{P}_{\mathbf{x}_k}^-}\right)_i \qquad i = L+1,\ldots,2L,$$

where $(\sqrt{P_{\mathbf{x}_k}^-})_i$ is the ith column of the matrix square root of $P_{\mathbf{x}_k}^-$, L is the number of latent factors, and h is the central difference step size, which is set to $\sqrt{3}$.

- *Step 2c.* Propagate the sigma points through the pricing functions $H1(\cdot)$ and $H2(\cdot)$. Let m_k be the number of observed inflation-indexed bond prices in period k. Let n_k be the number of observed inflation forecasts in period k. For each observed price $j = 1,\ldots,m_k$ we propagate each sigma point X_k^i, $i = 0,\ldots,2L$ through the pricing function for bond j in period k, $H1_{k,j}(\cdot)$. For each observed forecast $j = m_k+1,\ldots,m_k+n_k$ we propagate each sigma point X_k^i, $i = 0,\ldots,2L$ through the pricing function for forecast j in period k, $H2_{k,j}(\cdot)$. Denote the output by φ_k, which is a matrix of dimension $n_k + m_k$ by $2L+1$ with elements

$$(\varphi_k)_{j,i} = \begin{cases} H1_{k,j}(X_i) & i = 0,\ldots,2L, \quad j = 1,\ldots,m_k \\ H2_{k,j}(X_i) & i = 0,\ldots,2L, \quad j = m_k{+}1,\ldots,m_k{+}n_k. \end{cases}$$

Denote the ith column of φ_k by $\boldsymbol{\varphi}_k^i$.

- *Step 2d.* Observation-update equations. For weightings of

$$w_0^{(m)} = \frac{h^2 - L}{h^2} \quad w_i^{(m)} = \frac{1}{2h^2} \qquad \forall i \geq 1$$
$$w_i^{(c_1)} = \frac{1}{4h^2} \quad w_i^{(c_2)} = \frac{h^2-1}{4h^4} \qquad \forall i \geq 1,$$

the estimate of the price vector is given by a weighted average of the $\boldsymbol{\varphi}_k^i$s,

$$\hat{\mathbf{y}}_k = \sum_{i=0}^{2L} w_i^{(m)} \boldsymbol{\varphi}_k^i,$$

and the estimated covariance matrix of $\hat{\mathbf{y}}_k$ is given by

$$P_{\mathbf{y}_k} = \sum_{i=1}^{L} \left[w_i^{(c_1)} \left(\boldsymbol{\varphi}_k^i - \boldsymbol{\varphi}_k^{L+i} \right)^{[2]} + w_i^{(c_2)} \left(\boldsymbol{\varphi}_k^i + \boldsymbol{\varphi}_k^{L+i} - 2\boldsymbol{\varphi}_k^0 \right)^{[2]} \right] + R_k,$$

where R_k is the covariance matrix of the noise present in the observed prices. Here $(\cdot)^{[2]}$ denotes the vector outer product.

Next the estimate of the covariance between the state estimate and the price estimate is given by

$$P_{\mathbf{x}_k \mathbf{y}_k} = \sqrt{w_1^{(c1)} P_{\mathbf{x}_k}^-} \left[\boldsymbol{\varphi}_k^{1:L} - \boldsymbol{\varphi}_k^{L+1:2L} \right]^T.$$

- *Step 2e.* Calculate the Kalman gain matrix G_k,

$$G_k = P_{\mathbf{x}_k \mathbf{y}_k} P_{\mathbf{y}_k}^{-1}.$$

- *Step 2f.* Update the state estimates,

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + G_k(\mathbf{y}_k - \hat{\mathbf{y}}_k)$$
$$P_{\mathbf{x}_k} = P_{\mathbf{x}_k}^- - G_k P_{\mathbf{y}_k} G_k^T,$$

where $\mathbf{y}_k$ is the vector of observed prices.

References

Australian Financial Markets Association. 2008. *Australian Financial Markets Report.*

Beechey, M. 2008. "Lowering the Anchor: How the Bank of England's Inflation-Targeting Policies Have Shaped Inflation Expectations and Perceptions of Inflation Risk." FEDS Working Paper No. 2008-44, Board of Governors of the Federal Reserve System.

Cochrane, J. 2005. *Asset Pricing.* Princeton, NJ: Princeton University Press.

D'Amico, S., D. Kim, and M. Wei. 2008. "Tips from TIPS: The Informational Content of Treasury Inflation-Protected Security Prices." FEDS Working Paper No. 2008-30, Board of Governors of the Federal Reserve System.

Da Prato, G., and J. Zabczyk. 1993. *Stochastic Equations in Infinite Dimensions.* Number 45 in *Encyclopaedia of Mathematics and Its Applications.* Cambridge, UK: Cambridge University Press.

Duffie, D. 2001. *Dynamic Asset Pricing Theory.* Princeton, NJ: Princeton University Press.

Durham, J. 2006. "An Estimate of the Inflation Risk Premium Using a Three-Factor Affine Term Structure Model." FEDS Working Paper No. 2006-42, Board of Governors of the Federal Reserve System.

Finlay, R., and M. Chambers. 2009. "A Term Structure Decomposition of the Australian Yield Curve." *Economic Record* 85 (271): 383–400.

Hördahl, P. 2008. "The Inflation Risk Premium in the Term Structure of Interest Rates." *BIS Quarterly Review* (September): 23–38.

Joyce, M., P. Lildholdt, and S. Sorensen. 2010. "Extracting Inflation Expectations and Inflation Risk Premia from the Term Structure: A Joint Model of the UK Nominal and Real Yield Curves." *Journal of Banking and Finance* 34 (2): 281–94.

Kim, D., and A. Orphanides. 2005. "Term Structure Estimation with Survey Data on Interest Rate Forecasts." FEDS Working Paper No. 2005-48, Board of Governors of the Federal Reserve System.

Kim, D., and J. Wright. 2005. "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates." FEDS Working Paper No. 2005-33, Board of Governors of the Federal Reserve System.

McCulloch, J., and L. Kochin. 1998. "The Inflation Premium Implicit in the U.S. Real and Nominal Term Structures of Interest Rates." Ohio State University Working Paper No. 98-12.

Steele, J. 2001. *Stochastic Calculus and Financial Applications.* New York: Springer.

van der Merwe, R. 2004. *Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models.* Dissertation, Oregon Health and Science University.

The Geographical Composition of National External Balance Sheets: 1980–2005*

Chris Kubelec[a] and Filipa Sá[b]
[a]Royal Bank of Scotland
[b]Trinity College, University of Cambridge

This paper constructs a data set on the stocks of bilateral external assets and liabilities for eighteen countries in the period from 1980 to 2005. It distinguishes between four asset classes: foreign direct investment, portfolio equity, debt, and foreign exchange reserves. Network methods are used to explore the key facts that emerge from the data. We find that there has been a remarkable increase in interconnectivity over the past two decades and that this has been centered around a small number of countries. In a simulation exercise we show that shocks to one of the central countries generate much larger losses to the network than shocks to the periphery.

JEL Code: F3.

1. Introduction

Financial globalization was one of the most striking phenomena of the last two decades. But until recently very little was known about the size and composition of countries' external financial assets and liabilities. This gap was partly narrowed by the work of Lane and

*An earlier version of this paper was published as Bank of England Working Paper No. 384 (Copyright © 2010 Bank of England). Suggestions from an anonymous referee substantially improved the paper. We wish to thank Philip Lane, Vasileios Madouros, and Peter Sinclair for helpful discussions. We also wish to thank Gian Maria Milesi-Ferretti and Andreas Baumann for help with data. Participants at the Bank of England seminar and at the BIS CGFS workshop on the use of BIS international financial statistics provided useful comments. Amy Varty and Richard Edghill provided excellent research assistance. The data set constructed in this paper is available at www.econ.cam.ac.uk/teach/filipasa/publications.htm. Corresponding author (Sá) address: Trinity College, Trinity Street, CB2 1TQ, Cambridge, UK. Tel.: +44 1223337731. E-mail: fgs22@cam.ac.uk.

Milesi-Ferretti (2001, 2007), which provides estimates of the total external financial assets and liabilities of 145 countries from 1970 to 2007. These data cover different asset classes—foreign direct investment (FDI), portfolio equity, debt, financial derivatives, and foreign exchange reserves—and show that there has been a marked increase in the ratio of foreign assets and liabilities to GDP, particularly since the mid-1990s. This increase has been especially pronounced among industrial countries, where financial integration has exceeded trade integration. However, very little is known about the geographical composition of external assets and liabilities.

The key contribution of this paper is to go beyond *total* external assets and liabilities by constructing a data set of stocks of *bilateral* assets and liabilities. Our study is pioneer in providing a comprehensive picture of bilateral external assets and liabilities across countries. Existing data sets suffer from several gaps along both a cross-sectional and time-series dimension. This paper fills these gaps and constructs a complete data set of bilateral external positions for a group of eighteen countries, covering the period from 1980 to 2005.

Another contribution of our study is to provide a global perspective across asset classes. Existing studies on bilateral financial flows and stocks focus on a single asset class. This paper looks at four different asset classes: FDI, portfolio equity, debt, and foreign exchange reserves. For FDI, equity, and debt we collect data from a variety of sources and fill gaps using gravity models, which are the workhorse models for trade in goods. They explain trade flows between countries i and j by their sizes (GDP) and a variety of variables capturing the geographical and historical proximity between the two countries (distance, common language, common border, colonial links, etc.). These models have more recently been applied to bilateral financial stocks and flows. For reserves we adopt a different procedure and start by constructing the currency composition, which is then translated into the geographical composition.

Martin and Rey (2004) develop a theoretical framework that delivers an equilibrium relation between bilateral asset flows, the size of the home and host countries, and transportation and information costs. Their model provides a theoretical foundation for gravity models applied to trade in assets. Okawa and van Wincoop (2010) add information asymmetries to a static portfolio choice model. Similarly to Martin and Rey, their model delivers an equation where bilateral asset holdings are driven by the size of the source and host countries and the information asymmetry between them. Because the information asymmetry cannot be directly observed, it is captured empirically by variables such as the distance between the two countries, whether they share a common border or a common language, etc.

Empirically these models have been applied to different asset classes. Stein and Daude (2007) focus on the determinants of FDI stocks in OECD countries in the late 1990s and find that differences in time zones have a negative and significant effect on the location of FDI. Portes and Rey (2005) use a gravity model to explain bilateral cross-border equity flows between fourteen economies in the period from 1989 to 1996. They find that the model performs at least as well as when applied to trade in goods and there is a significant and negative effect of distance on equity transactions. Lane and Milesi-Ferretti (2008) use a gravity model to explain stocks of bilateral portfolio equity in 2001. They find that bilateral equity holdings are strongly correlated with bilateral trade in goods and services and are also positively associated with measures of proximity. Rose and Spiegel (2004) use a gravity equation to explain bilateral debt flows and also find that bilateral trade appears to have a positive and significant effect on bilateral lending.

Consistent with previous studies, we find gravity models to have very good explanatory power when applied to bilateral financial stocks. Standard gravity variables have a significant effect on financial stocks: countries that are less distant or share a common border or a common language have stronger financial linkages across all three asset classes. We also confirm the findings in Stein and Daude (2007) on the negative effect of time difference on FDI stocks and find that this is true for equity and debt holdings as well.

The idea that variables such as distance and cultural affinities may explain a large proportion of cross-border asset flows and stocks is perhaps surprising. Unlike goods, financial assets are not subject to transportation costs. Also, if investors wish to diversify their portfolios, they may choose to invest in more distant countries, where the business cycle has a low or negative correlation with their own country's business cycle. The fact that gravity variables perform at least as well in explaining financial positions as they do in explaining trade suggests that financial markets are not frictionless but are segmented by information asymmetries and familiarity effects.

After describing the data construction in detail, we apply a number of tools from network analysis to examine the key stylized facts that emerge from the data. The international financial system can be seen as a network, where nodes represent countries and links represent bilateral financial assets. We observe that there has been a remarkable increase in interconnectivity over the past two decades: financial links have become larger and countries have become more open. We also find that the global financial network is centered around a small number of nodes, which have many and large links.

The global trade network also shows an increase in interconnectivity over time. However, while the financial network is centered around the United States and the United Kingdom, the trade network shows strong intracontinental links and is arranged in three clusters: a European cluster (centered on Germany), an Asian cluster (centered on China), and an American cluster (centered on the United States).

The configuration of the international financial network has important implications for the stability of the international financial system. We discuss how the combination of high interconnectivity and a small number of hubs makes for a "robust yet fragile" system, where a disturbance to one of the central countries would be transmitted rapidly and widely. To illustrate how shocks propagate through the network, we perform a simulation exercise where asset values in the shock country drop by 10 percent. This exercise shows that the largest losses to the network occur after a shock to the United States: the value of all other countries' assets as a percentage of their GDP falls by 7 percent. The countries that suffer the largest losses following a shock to the United States are the other "hubs" in the network: Hong Kong, Singapore, and the United Kingdom. By contrast, shocks to peripheral countries have a much smaller effect: shocks to Argentina, India, Portugal, Mexico, and Brazil all generate losses of less than 0.4 percent of the combined GDP of all countries except the shock country.

The rest of the paper is organized as follows. Section 2 describes data sources on bilateral financial assets and liabilities and the

Table 1. Country Coverage

Developed Countries	Emerging Markets
Australia	Argentina
Canada	Brazil
France	Mexico
Germany	China
Italy	Hong Kong
Japan	India
Portugal	Korea
Spain	Singapore
United Kingdom	
United States	

techniques used to fill in gaps in those sources. Section 3 uses network methods to show the key stylized facts that emerge from the data and compares the international financial and trade networks. Section 4 discusses the implications of the configuration of the international financial network for the stability of the international financial system. Section 5 concludes.

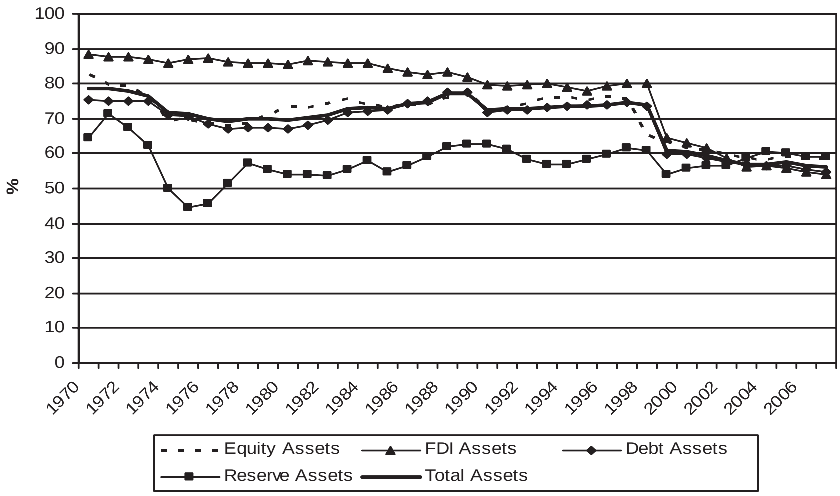
2. Data Construction

2.1 Country Selection and Treatment of Financial Centers

The data are constructed at annual frequency and include eighteen countries, listed in table 1. The sample was selected to include both emerging and developed economies located in different continents. To measure the proportion of total external assets that is accounted for by our sample, we use the data compiled by Lane and Milesi-Ferretti and compute the share of total external assets in their sample of 145 countries that is accounted for by the 18 countries in our sample.

Figure 1 shows how this share has changed over time for different asset classes. Until the late 1990s, the share of the world's total external assets excluding reserves accounted for by our sample was between 70 percent and 90 percent. This fraction dropped to around 60 percent in the 2000s. The five countries outside our sample whose

Figure 1. Percentage of World's Total Assets Accounted for by the Eighteen Countries in Our Sample



Source: Lane and Milesi-Ferretti (2001, 2007) data set.

shares of the world's total external assets excluding reserves most increased from 1997 to 2007 were Ireland and Luxembourg (which are financial centers), Norway, the Netherlands, and Austria. These five countries accounted for about 9 percent of the world's total external assets excluding reserves in 1997 and 14 percent in 2007. Given that the share accounted for by the eighteen countries in our sample dropped by about 20 percentage points in this period, this implies that the gains in share were distributed over a large number of countries. Our sample captures between 50 percent and 60 percent of the world's total reserves.

Some of the countries in the sample—the United Kingdom, the United States, Singapore, and Hong Kong—are important financial centers and are both final destinations and intermediaries of foreign investment. Balance-of-payments statistics are constructed on the basis of the residence principle. For example, if a German resident invests in a Chinese company and directs the investment via a financial institution located in the United Kingdom, balance-of-payments data would register the transaction as an asset of Germany in the

United Kingdom and an asset of the United Kingdom in China, even though the United Kingdom has only acted as an intermediary.[1]

Most available data sets on bilateral financial links follow the residence principle. A notable exception is the Bank for International Settlements (BIS) consolidated banking statistics, which contain information on cross-border assets held by banks and are based on the nationality of the reporting bank, netting out intragroup positions. The BIS also collects data based on residence (locational banking statistics).[2] Which data are preferable depends on the question being addressed. Data based on residence are useful to detect broad trends in cross-border links from a geographical perspective, while data based on nationality may be preferable for analyzing the international transmission of shocks through the banking system. However, this depends on whether foreign subsidiaries and branches fund themselves locally or in their country of nationality. For example, suppose that Santander in the United Kingdom (part of a Spanish group) borrows from households in the United Kingdom to lend to China. Consolidated data would treat this as an investment of Spain in China. This may be appropriate to study the effect of a shock in China on Santander as a group. However, it would not be appropriate to study the implications of a shock in the United Kingdom for cross-border capital flows. For this question locational data would be preferable. Since no type of data is clearly preferable in all circumstances and residence-based data are more widely available, we follow the balance-of-payments methodology and construct the data set based on the residence principle.

2.2 General Approach for FDI, Equity, and Debt

The construction of data for FDI, equity, and debt follows a six-step procedure:

- *Step 1.* Collect data on bilateral assets from a variety of sources.

[1]Felettigh and Monti (2008) show that there can be significant differences between bilateral links based on the residence principle and ultimate exposures.

[2]The BIS consolidated banking statistics are described in detail in McGuire and Wooldridge (2005). For a useful discussion of the differences between consolidated and locational baking statistics, see McGuire and Tarashev (2008).

- *Step 2.* Compute geographical weights.

  By dividing assets of country i in country j (A_{ijt}) by total external assets of country i (A_{it}), obtain the percentage of assets of country i which are held in country j (w_{ijt}):

$$w_{ijt} = \frac{A_{ijt}}{A_{it}}.$$

  Weights do not necessarily add up to 1, since the eighteen countries in the sample do not account for a country's total external assets.

- *Step 3.* Estimate gravity models for geographical weights.

  Missing data are estimated separately for each asset class using the following gravity model:

$$\log\left(\frac{w_{ijt}}{1 - w_{ijt}}\right) = \phi_i + \phi_j + \phi_t + \alpha X_{ij} + \beta Z_{ijt} + \varepsilon_{ijt}. \quad (1)$$

  w_{ijt} is the proportion of assets of country i held in country j in year t. We estimate the model on weights rather than stocks of foreign assets because stocks would be non-stationary, implying that the usual distributions for OLS estimates would be invalid. The dependent variable is the logit of weights. This is a standard transformation to deal with proportions data, transforming (1) into a linear model which can be estimated by OLS.[3]

  ϕ_i and ϕ_j are dummy variables for each source and host country and ϕ_t are time dummies. Host-country fixed effects control for characteristics that make countries attractive to foreign investment. Source-country fixed effects control for characteristics that make countries more diversified, investing a smaller share in a larger number of countries. X_{ij} is a set of bilateral variables which are standard in trade gravity

[3]Taking logs eliminates observations for which weights are zero. Given the small proportion of zeros in the data (less than 10 percent), eliminating them should not have much influence on the results. Also, eliminating zeros may be less problematic than estimating a model that fits over both zero and non-zero observations. This is because the determinants of whether a country has any financial linkages with another country may be different from the determinants of the size of the exposures, given that countries are linked.

models and measure the geographical and historical proximity between economies: common border, common language, colonial links, distance, and time difference. The colony dummy is asymmetric and is equal to 1 if country i is a former colonizer of country j. This variable is asymmetric to reflect the fact that while former colonizers may have preferential status when they invest in former colonies, former colonies may not have preferential status when investing in former colonizers. Z_{ijt} is a set of time-varying regressors.

- *Step 4.* Combine actual with estimated weights.

  After estimating gravity models for geographical weights, we use the estimated coefficients to obtain out-of-sample predictions of weights for those years and country pairs for which data are missing. We then combine actual weights with those predicted values to obtain a data set on asset weights with no missing observations ($\tilde{w}_{ijt}$).
- *Step 5.* Transform geographical weights into stocks of foreign assets.

  To transform geographical weights into stocks of foreign assets, we multiply the weights obtained in step 4 by total external assets of country i reported in the Lane and Milesi-Ferretti (2007) data set:

$$\widetilde{A}_{ijt} = \tilde{w}_{ijt} \times A_{it,LMF}.$$

  This step ensures that bilateral stocks of foreign assets incorporate some adjustment for valuation effects arising from exchange rate movements and changes in asset prices. Lane and Milesi-Ferretti introduce this adjustment in their data; therefore it will also be incorporated into our estimates of bilateral stocks. This is potentially important, since valuation effects have been shown to be sizable (see Gourinchas and Rey 2007b).[4]

[4] A more accurate method to adjust for valuation effects would be to do it directly on bilateral stocks, taking into account changes in bilateral exchange rates and in stock market valuations in the host country. By taking the adjustment from Lane and Milesi-Ferretti, we are applying the adjustment on total external assets to bilateral assets, rather than making it specific to each country pair.

- *Step 6.* Construct liabilities from assets.

  The data set is constructed taking the assets perspective. This last step uses the fact that assets and liabilities should be symmetric and constructs liabilities from assets:

$$Liabilities_{ijt} = Assets_{jit}.$$

  Liabilities of country i with country j at year t equal assets of country j in country i at year t.

2.3 FDI

2.3.1 Data

The main source of data on FDI assets is the OECD International Direct Investment by Country data set, which contains FDI data at book value reported by OECD members starting in 1981. There are many missing values in the data. To the extent possible, missing observations are filled with data from the United Nations Conference on Trade and Development (UNCTAD). But even after combining the data sets, there are significant gaps in the data. Table 2 lists the percentage of missing data for each source country. Coverage is better for developed economies, but there is a large fraction of missing data for Mexico, Argentina, and India. Overall, approximately 44 percent of the data on bilateral FDI are missing and need to be estimated.

Because the OECD and UNCTAD report data on both assets and liabilities, it would in principle be possible to combine the two and reduce the extent to which bilateral positions need to be estimated. We do not follow this approach because different methods are used to report FDI assets and liabilities. Liabilities are reported following the ultimate beneficial owner (UBO) principle, according to which the source of inward FDI is allocated to the country of ultimate ownership. The equivalent principle on the assets side would be the country of ultimate destination (CUD) principle, according to which outward FDI would be allocated to the country of final destination. However, while the UBO principle is widely adopted in the production of FDI statistics, the CUD principle is not the norm; i.e., liabilities are reported following the ultimate ownership principle and assets are reported following the residence principle adopted

Table 2. Proportion of Missing Data

Source Country	FDI	Equity	Debt
Argentina	84%	63%	76%
Australia	40%	68%	62%
Brazil	67%	68%	78%
Canada	3%	63%	0%
China	76%	89%	94%
France	19%	63%	0%
Germany	0%	67%	0%
Hong Kong	77%	72%	79%
India	84%	84%	76%
Italy	26%	63%	0%
Japan	15%	63%	0%
Korea	15%	68%	78%
Mexico	86%	85%	86%
Portugal	52%	65%	62%
Singapore	54%	64%	77%
Spain	76%	64%	11%
United Kingdom	16%	64%	0%
United States	6%	63%	0%
Full Sample	**44%**	**69%**	**43%**

Notes: Proportions are computed after filling in missing values using the index of stock market liberalization. For equity, the CPIS only reports data for 1997 and the period from 2001 to 2005. Data for all other years are missing. For debt, data for Argentina, China, Hong Kong, Korea, and Singapore are from the IMF CPIS only. Therefore, data are missing for all years except 1997 and 2001 to 2005.

in the balance-of-payments statistics. Since we choose to follow the balance-of-payments methodology, we focus only on assets and make no use of data on liabilities.

2.3.2 Estimation

FDI asset weights are estimated using model (1). The gravity variables are obtained from the Distances Database compiled by the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII). The set of time-varying regressors includes GDP per capita in countries i and j and the degree of openness of country j to inward

FDI. GDP per capita captures the degree of development and is obtained from the World Bank's *World Development Indicators*. It is measured at constant prices and is PPP adjusted. The degree of openness of country j to inward FDI is a time-varying index constructed from the tables in Kaminsky and Schmukler (2003), which report the chronology of stock market liberalization and classify countries into three degrees of liberalization over time:

(i) *No liberalization*: Foreign investors are not allowed to hold domestic equity and cannot repatriate capital, dividends, and interest until five years after the initial investment.

(ii) *Partial liberalization*: The country is open to foreign investment, but with some restrictions.

(iii) *Full liberalization*: Foreign investors are allowed to hold domestic equity and to repatriate capital, dividends, and interest without restrictions.

We transform this classification into a numerical variable which takes the value 0 if country j is not liberalized in year t, 1 if it is partially liberalized, and 2 if it is fully liberalized.[5] As well as being used as a control in regression (1), this index is used to fill in some of the missing data prior to estimation. Table 3 illustrates how this is done, using as an example FDI assets of the United Kingdom in China. We know the stock of assets of the United Kingdom in China in 1991, while China was still closed to FDI. Because there would have been no inward flows to China from 1980 to 1990, the stock of assets in that period should equal the stock in 1991, adjusted for valuation effects due to changes in exchange rates and asset prices. To adjust for valuation effects, we assume that FDI assets of the United Kingdom in China in that period grow at the same rate as

[5]Some countries in our sample are not studied by Kaminsky and Schmukler (2003). For those countries, we use information on the timing of stock market liberalization from other studies and code it according to the criteria used by Kaminsky and Schmukler. For China we use information in OECD (2000), Prasad and Wei (2005), and Bekaert, Harvey, and Lundblad (2007). For India we use Ahluwalia (2002) and Reserve Bank of India (2006).

Table 3. Using the Liberalization Index on Inward FDI to Fill in Missing Data

Year	FDI Assets of United Kingdom in China	Liberalization Index on Inward FDI FDI in China
1980	8	0
1981	8	0
1982	10	0
1983	13	0
1984	19	0
1985	30	0
1986	44	0
1987	60	0
1988	77	0
1989	100	0
1990	124	0
1991	150	0
1992	157	1
1993	271	1
1994	184	1
1995	270	1
1996	778	1
1997	776	1
1998	566	1
1999	2027	1
2000	2246	1
2001	3055	1
2002	5177	1
2003	3229	1
2004	3645	1
2005	5364	1

Sources: OECD and UNCTAD; values in millions of U.S. dollars.
Note: Highlighted values are filled in using the liberalization index.

total Chinese FDI liabilities. We take the value in 1991 as the starting point and build stocks backwards using the growth rate of total Chinese liabilities.

Turning to the estimation results, column 1 of table 4 reports results of a model where FDI asset weights are only explained by

Table 4. Estimation Results for FDI Weights

	(1) Host-Country FE	(2) Host- and Source-Country FE	(3) Gravity Variables	(4) Model for Prediction
Border			0.394***	0.340***
			(0.119)	(0.113)
Language			1.585***	1.598***
			(0.095)	(0.094)
Colony			0.507***	0.481***
			(0.092)	(0.096)
Log(Distance)			−0.681***	−0.681***
			(0.043)	(0.040)
Time Difference			−0.054***	−0.054***
			(0.010)	(0.009)
Log($GDPpc_{it}$)				0.750***
				(0.295)
Log($GDPpc_{jt}$)				1.817***
				(0.137)
Index Liberalization FDI_{jt}				0.379***
				(0.054)
N	3810	3810	3810	3810
R^2	0.41	0.50	0.68	0.71
Marginal R^2 of Gravity Variables			0.36	
Marginal R^2 of Time-Varying Variables				0.04

Notes: Robust standard errors are in parentheses. * denotes significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level. Regression (4) includes time dummies. The marginal R^2 of the gravity variables indicates the percentage improvement in the R^2 from including these variables, over and above the model with only host- and source-country fixed effects. The marginal R^2 of time-varying variables indicates the percentage improvement in the R^2 from the time-varying variables (including time dummies) over and above the model with fixed effects and the gravity variables.

host-country fixed effects. The predictive power is relatively good, with an R^2 of 41 percent. Column 2 adds source-country fixed effects, with an improvement in the R^2 to 50 percent. Including the standard gravity variables further increases the R^2 to 68 percent, which is high and consistent with the results in other empirical studies. The standard gravity variables are significant and have the expected signs: FDI weights are larger for countries that share a common border or a common language and have colonial links. Distance and time difference have a significant negative effect on FDI weights.

Time-varying controls are included in column 4. Countries with larger GDP per capita receive larger shares of FDI investment. This illustrates the paradox discussed in Lucas (1990)—that capital tends to flow to rich countries even though the marginal product of capital is larger in poor countries—and is consistent with the findings in Papaioannou (2009). Countries whose markets are more liberalized to FDI also receive larger investment shares. However, the improvement in the R^2 from including these time-varying controls is marginal.

We also experimented with additional controls. Previous studies have found a significant effect of bilateral trade on bilateral asset holdings. There are at least two reasons why this may be the case. First, bilateral trade may capture an additional familiarity effect, over and above the gravity variables. Second, countries may use financial investment to hedge against shocks in countries with which they trade. We extended the model to include trade weights, measured as the ratio of trade (exports plus imports) between countries i and j over total trade of country i, using data from the International Monetary Fund (IMF) Direction of Trade Statistics (DOTS). Trade weights were found to have a positive but insignificant effect in explaining FDI weights and were not included in the model used for prediction.[6]

Another variable we experimented with was the volatility in bilateral exchange rates measured as the standard deviation in the rate of change of monthly bilateral exchange rates on a three-year rolling window. Exchange rates were obtained from the IMF International Financial Statistics (IFS). Bilateral financial positions may be smaller when the bilateral exchange rate is more volatile because there is more uncertainty about the returns. This variable turned out to have an insignificant effect on FDI asset weights and was

[6]Only variables with a p-value lower than 0.25 were kept in the model used for prediction.

excluded from the model used for prediction. The insignificant effect of bilateral exchange rates is consistent with the findings in Portes and Rey (2005) and Lane and Milesi-Ferretti (2008).

2.4 Equity

2.4.1 Data

Data on portfolio equity assets are collected from the IMF Coordinated Portfolio Investment Survey (CPIS), which covers all countries in our sample except China. The time coverage, though, is quite limited: a pilot survey was conducted in 1997 and a regular annual survey was introduced in 2001 for an extended group of participating countries. Table 2 lists the proportion of missing data by source country. Given limited time coverage of the CPIS, over 60 percent of data are missing for all countries and need to be estimated. For China this proportion is higher since it does not participate in the CPIS.

As for FDI, we only use data on assets and make no use of data on liabilities. This is because, while countries that participate in the CPIS are required to report assets, liabilities are reported on a voluntary basis. For the few countries in our sample that report liabilities, there is a big discrepancy between liabilities and assets reported by creditors. Because of this discrepancy, we use only reported assets.

2.4.2 Estimation

Table 5 shows the results of estimating model (1) on equity weights. Host-country fixed effects explain 46 percent of the variation in equity weights. Introducing source-country fixed effects increases the R^2 to 55 percent. The coefficients on the gravity variables are significant and have the expected signs except for colonial links, which is negative. This suggests that investors may prefer to invest in countries with a similar degree of development as their home country regardless of historical colonial links. The inclusion of these variables leads to a significant improvement in the R^2, which rises to 71 percent.

Table 5. Estimation Results for Equity Weights

	(1) Host-Country FE	(2) Host- and Source-Country FE	(3) Gravity Variables	(4) Model for Prediction
Border			0.820***	0.820***
			(0.185)	(0.187)
Language			1.729***	1.736***
			(0.143)	(0.141)
Colony			−0.792***	−0.805***
			(0.203)	(0.192)
Log(Distance)			−0.453***	−0.433***
			(0.074)	(0.072)
Time Difference			−0.107***	−0.110***
			(0.017)	(0.017)
Log($GDPpc_{jt}$)				4.063***
				(0.769)
Exchange Rate Volatility				−0.003**
				(0.001)
Index Liberalization Equity_{jt}				2.452***
				(0.603)
N	1341	1341	1341	1341
R^2	0.46	0.55	0.71	0.72
Marginal R^2 of Gravity Variables			0.29	
Marginal R^2 of Time-Varying Variables				0.01

Notes: Robust standard errors are in parentheses. * denotes significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level. Regression (4) includes time dummies. The marginal R^2 of the gravity variables indicates the percentage improvement in the R^2 from including these variables, over and above the model with only host- and source-country fixed effects. The marginal R^2 of time-varying variables indicates the percentage improvement in the R^2 from the time-varying variables (including time dummies) over and above the model with fixed effects and the gravity variables.

The set of time-varying controls includes GDP per capita in country j, bilateral exchange rate volatility, and the degree of openness of country j to inward equity investment.[7] The results suggest that investors invest more in countries that are more open to inward equity investment and have a larger GDP per capita. They also invest more when the volatility of the bilateral exchange rate is smaller. However, these time-varying variables do not have a large explanatory power and lead to a very small improvement in the R^2.

As for FDI, the index of openness to inward equity investment is used to estimate missing data. However, while for FDI it was possible to take a data point when the host country was still closed and build the data backwards using the growth rate of its total liabilities (as illustrated in table 3), for equity the data start when all countries were already open. Since it is not possible to build the data backwards in the same way as for FDI, we simply impose zero bilateral weights for the period when the host country was closed to inward equity investment.[8]

We also experimented with other control variables. To capture stock market returns and correlations in returns, we included averages, standard deviations, and the correlation coefficient of daily stock market indices in the host and source countries. These variables were insignificant and were not included in the final regression. GDP per capita in country i, stock market capitalization in country j, and trade weights were also insignificant.

[7]The degree of openness to inward equity investment was constructed in the same way as for FDI. In fact, FDI can be seen as a type of portfolio equity investment where the degree of ownership exceeds 10 percent of the firm's equity. Countries may liberalize their stock markets to foreign portfolio equity investment and remain closed to FDI by introducing a ceiling on the percentage of total equity that can be owned by foreign residents. The only country in our sample where the index of liberalization is different for equity and FDI is Korea, where foreign portfolio equity investment was partially liberalized in 1991, while foreign FDI investment remained restricted. Both types of investment were fully liberalized in 1998.

[8]The only exception to this rule is equity investment of Hong Kong in China. China was closed to inward equity investment until 1992. However, given the strong political and administrative links between the two countries, we do not impose zeros for Hong Kong's equity investment in China pre-1992.

2.5 Debt

2.5.1 Data

Data on portfolio debt assets are collected from the IMF CPIS and the BIS locational banking statistics. The BIS data set has the advantage of having a much longer time coverage, going back to 1977 for most advanced countries. However, it has the limitation of only reporting debt assets held by banks, while the CPIS has much broader coverage. The data sets also differ in the assets covered: while the CPIS only covers portfolio debt, the BIS also covers loans and deposits.

To test whether it is sensible to combine data from the BIS and the CPIS, we compute the correlation coefficient between the asset weights generated by the two data sources. The correlation coefficient is large (80 percent), suggesting that it is reasonable to combine them. By default, we use asset weights computed from the BIS data and complete it with weights computed from the CPIS data whenever possible. After combining the two data sets, approximately 43 percent of the data are missing (table 2). The gaps are especially pronounced for China, which is not covered by either data set, and for countries not covered by the BIS locational banking statistics, for which we only have data after the CPIS was introduced in 1997.

As for the other asset classes, we make no use of data on liabilities. For CPIS data we face the same problems as with equity: very few countries report liabilities and, where they do, there is a large difference between liabilities and assets reported by creditors. For BIS data we cannot use liabilities to build assets by symmetry because the data are not symmetric: banks in country i report assets held against *banks and non-banks* in country j, while *banks* in country j report liabilities against both banks and non-banks in country i.

2.5.2 Estimation

Table 6 reports the results of estimating model (1) on debt weights. The model with only host-country fixed effects explains 49 percent of the variation in debt weights. Adding source-country fixed effects increases the R^2 to 57 percent, and adding standard gravity variables further improves the R^2 to 69 percent. Common border was excluded

Table 6. Estimation Results for Debt Weights

	(1)	**(2)**	**(3)**	**(4)**
	Host-Country FE	**Host- and Source-Country FE**	**Gravity Variables**	**Model for Prediction**
Language			1.081***	1.001***
			(0.077)	(0.081)
Colony			−0.261***	−0.170**
			(0.078)	(0.082)
Log(Distance)			−0.423***	−0.367***
			(0.042)	(0.044)
Time Difference			−0.119***	−0.114***
			(0.010)	(0.010)
Log($GDPpc_{jt}$)				0.892***
				(0.120)
Trade Weights_{ijt}				1.160***
				(0.449)
Exchange Rate Volatility_{ijt}				−0.003***
				(0.001)
N	4187	4187	4187	4187
R^2	0.49	0.57	0.69	0.70
Marginal R^2 of Gravity Variables			0.21	
Marginal R^2 of Time-Varying Variables				0.01

Notes: Robust standard errors are in parentheses. * denotes significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level. Regression (4) includes time dummies. The marginal R^2 of the gravity variables indicates the percentage improvement in the R^2 from including these variables, over and above the model with only host- and source-country fixed effects. The marginal R^2 of time-varying variables indicates the percentage improvement in the R^2 from the time-varying variables (including time dummies) over and above the model with fixed effects and the gravity variables.

from the set of gravity variables because it had no significant effect on debt weights. The colony dummy has a negative sign, as in the model for equity. This suggests that for types of investment which imply a larger degree of commitment, such as FDI, former colonizers tend to invest in former colonies. However, for equity and debt investment they seem to prefer countries with a similar degree of development regardless of colonial links.

As for equity, the results suggest that investors tend to invest larger shares in more-developed countries—the Lucas paradox—and in countries with lower exchange rate volatility with respect to the currency of the source country. In contrast with the result for FDI and equity, bilateral trade weights have a significant and positive effect on debt weights. This is consistent with the findings in Rose and Spiegel (2004), who show that borrowers fear that default on their debt may lead to a reduction in international trade; therefore, creditors systematically lend more to countries with whom they have closer trade links.[9]

We experimented with additional controls and estimated the model including bond market capitalization and measures of bond returns. These variables turned out insignificant and were not included in the model used for prediction.

2.6 Reserves

The construction of the reserves data follows a different approach. While for FDI, equity, and debt investors choose *where* to invest, for reserves they choose *in which currency* to invest. We follow a two-step procedure to obtain the geographical composition of reserves. First, we obtain the currency composition. Then we translate it into the geographical composition: if country i holds an amount X of reserves in U.S. dollars, we take X as being the amount of reserve assets that country i holds in the United States. For simplification we focus on the four main reserve currencies: the U.S. dollar, the euro, the pound, and the yen. These capture the bulk of countries' foreign exchange reserves. Also for simplification we treat reserves of country i denominated in euros as being assets of country i in

[9]Unlike for FDI and equity, the set of time-varying controls does not include the degree of liberalization of the host country to inward debt investment. This is because we were unable to construct an index which captures restrictions only to *inward* investment. A time-series index for capital account restrictions is available in Kaminsky and Schmukler (2003). This captures restrictions to borrowing abroad by banks and corporations (which could be interpreted as restrictions to debt capital *inflows*) as well as exchange rates and other restrictions to capital *outflows*. Because it confounds restrictions to inward and outward investment, we decided not to use it.

Germany. For the period before the introduction of the euro we use the deutsche mark.[10]

Data on the currency composition of reserves are confidential and not readily available. The BIS Multilateral Surveillance Statistics contain data on the currency composition of reserves for countries in the G-10 since 1994. This gives us data for six countries in our sample: France, Germany, Italy, Japan, the United Kingdom, and the United States. Given the remarkable stability of currency weights over time, we assume that weights stay constant from 1980 to 1994. For the remaining countries the IMF collects data in the COFER (Currency Composition of Official Foreign Exchange Reserves) data set. Although the numbers are only released as aggregates across industrialized and developing countries, disaggregated data have been used in some previous studies. We follow the approach in Lane and Shambaugh (2007) and use the results reported in those studies to obtain estimates of the currency composition of reserves for countries that are not part of the G-10.

The studies we use are Dooley, Lizondo, and Mathieson (1989) and Eichengreen and Mathieson (2000), who adopt the following specification to explain the currency composition of reserves:

$$\begin{aligned} share_{ict} = c + \alpha_1 dollar\ peg_{ict} + \alpha_2 other\ peg_{ict} \\ + \beta share\ trade_{ijt} + \gamma share\ debt\ payments_{ict} + \varepsilon_{ict}. \end{aligned} \tag{2}$$

The dependent variable is the share of foreign exchange reserves held by country i in currency c at time t, obtained from COFER. The regression includes dummy variables equal to 1 if country i pegs to the U.S. dollar or to another currency, the share of trade between country i and country j at time t (where country j is the country that issues currency c), and the share of debt service payments of country i in currency c at time t. The share of trade is calculated as the sum of exports and imports between countries i and j divided by total exports plus imports plus debt service payments of country i. The share of debt payments in currency c is calculated as

[10] A more precise way of dealing with euro reserves would be to allocate them according to the relative GDP of each country in the euro area. Here we take a shortcut and allocate all euro reserves to Germany.

service payments of country i on debt denominated in currency c divided by total exports plus imports plus debt service payments of country i.

Eichengreen and Mathieson (2000) report the results of estimating this model for a sample of eighty-four emerging and transition economies for the period 1979–96. We collect data for the right-hand-side variables and multiply by the estimated coefficients reported in their paper to obtain estimates of the currency composition of reserves.[11]

Data on exchange rate regimes are obtained from Levy-Yeyati and Sturzenegger (2005). They report an index which classifies exchange rate regimes in three categories: floating, intermediate, and fixed. We transform this index into a binary variable, which takes the value 0 if the country has a floating regime and 1 if the country has an intermediate regime or a peg. We construct one indicator for U.S. dollar pegs and another for other currency pegs. Data on trade are collected from the IMF Direction of Trade Statistics. Debt service payments are obtained by multiplying the six-month euro currency deposit rates, obtained from Datastream, by the amount of debt outstanding, obtained from the World Bank's *Global Development Finance.*

This approach gives us estimates of the currency composition of reserves which seem sensible when compared with the reserve shares that countries occasionally report in announcements and media interviews. For example, China is reported to hold roughly 70 percent of its reserves in dollars, 20 percent in euros, and 10 percent in other currencies. Our estimation gives 79 percent in dollars and 21 percent in euros.

3. A Look at the Data

The international financial system can be seen as a network, where nodes represent countries and links represent bilateral financial assets. Our data set provides information on the links and allows us to study how the global financial network has changed over time.

[11] We use the coefficients reported in table 3 of Eichengreen and Mathieson (2000).

In this section we use network methods to give a flavor of the data set and show the key stylized facts that emerge from it.

3.1 Financial Network

Figure 2 looks at the evolution of the global financial network. In each year t links are given by the ratio of bilateral assets (including all asset classes) to GDP of the source country:

$$link_{ijt} = \frac{Assets_{ijt}}{GDP_{it}}.$$

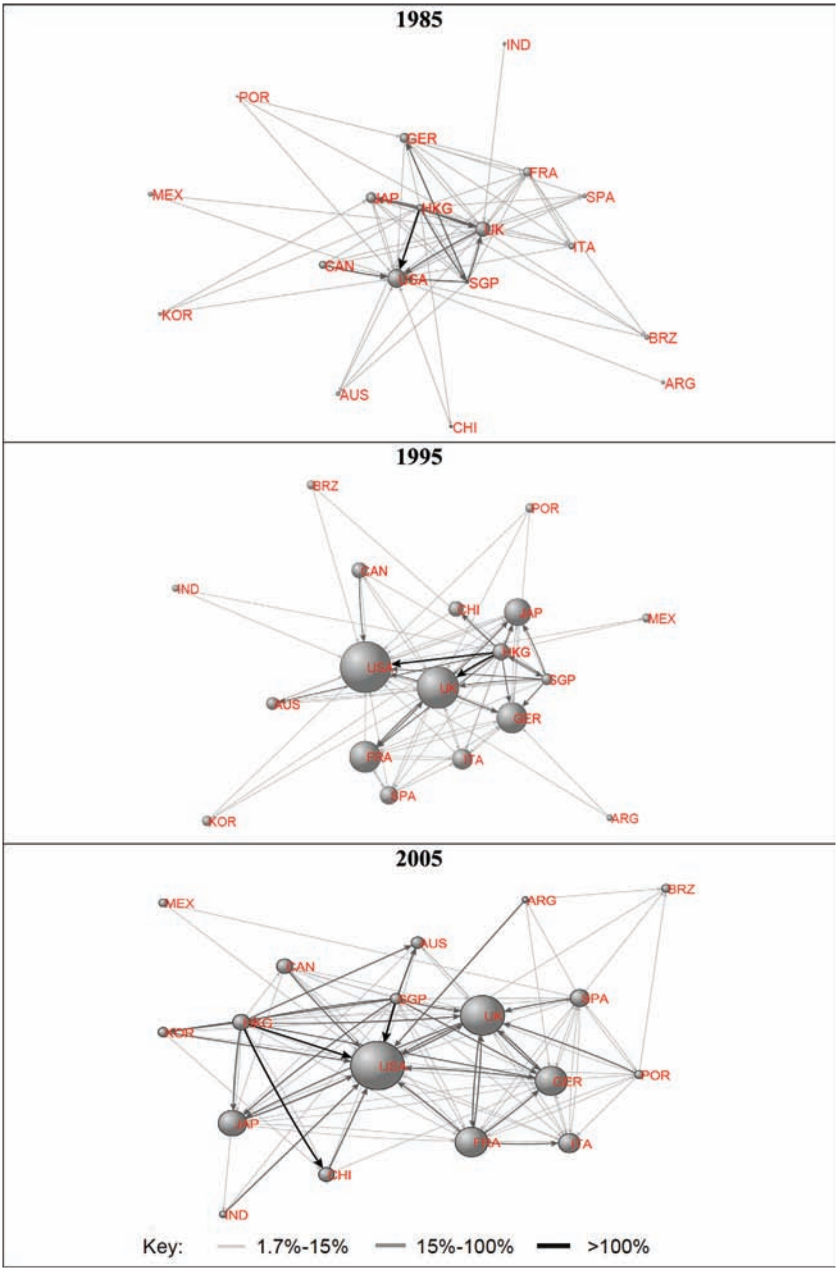
The network is *directed:* an arrow pointing from county i to j represents the value of country i's assets in country j scaled by country i's GDP. It is also *weighted* because links represent the strength of the connections between nodes and not simply whether a connection exists or not. To simplify the diagrams, we impose a cutoff and represent only the strongest links (where the ratio defined above is higher than 1.7 percent). This cutoff is chosen in such a way that every node is linked to at least one other node in every year. The thickness of the lines indicates the size of the links, and the size of the nodes is proportional to the country's financial openness, measured by the sum of its total external assets and liabilities. Pairs of countries with stronger links are placed closer to each other.[12]

Table 7 provides some summary statistics. Skewness is a measure of the asymmetry of a distribution. A positive value indicates that there are many country pairs with small links and few country pairs with large links. Kurtosis is a measure of the "peakedness" of a distribution. A large value for kurtosis indicates that the distribution has "fat tails."[13] Average path length is the average of the shortest distance between all pairs of nodes in the network. Clustering measures the probability that, given that country i is directly linked to countries j and k, country j is also directly linked to country k. Small

[12]This is achieved using the Kamada-Kawai algorithm, which positions nodes in the space so that their geometric distance reflects the strength of the links between them. The network charts were produced using Pajek (a program for analysis and visualization of large networks).

[13]A normal distribution has skewness equal to 0 and kurtosis equal to 3.

Figure 2. International Financial Network



Notes: Links are given by the ratio of bilateral assets to GDP of the source country. The size of the nodes is proportional to the country's financial openness, measured by the sum of its total external assets and liabilities.

Table 7. Summary Statistics on the International Financial Network

	1985	1995	2005
Skewness	7.63	7.16	5.22
Kurtosis	69.60	61.55	35.27
Average Path Length	2.07	2.14	1.93
Clustering Coefficient	0.55	0.63	0.60

values for average path length and large values for the clustering coefficient indicate that the network is highly interconnected.[14]

A few findings emerge:

- *The interconnectivity of the global financial network has increased significantly over the past two decades.* This can be seen from the increase in the size of the nodes and the increase in number and size of the links. It is also confirmed by the large values of the clustering coefficient and the reduction in average path length over time. In 2005 there are less than two degrees of separation on average between any two nodes.
- *The distribution of financial links exhibits a long tail.* Measures of skewness and kurtosis show the asymmetry compared with the normal distribution. A small number of countries ("hubs") have large links to other countries, but most links are small.

To study which countries are the main sources and destinations of international investment, table 8 reports measures of network centrality for each node (country), following the approach of von Peter (2007).

The key findings that emerge from the centrality measures are as follows:

[14]Detailed definitions of these statistics are presented in the appendix. Average path length and clustering depend on the cutoff chosen for the links. Imposing a cutoff enables us to apply these statistics (which were developed for unweighted networks) to our network. Because the global financial network is complete—i.e., all pairs of nodes are linked even if the size of financial assets and liabilities is very small—these statistics would be meaningless if we had not imposed a cutoff.

Table 8. Measures of Network Centrality—Finance, 2005

	In-Degree	Out-Degree	Closeness	Betweenness	Intermediation	Prestige
United States	100.00 (1)	35.29 (7)	1.00 (1)	24.67 (1)	49.89 (1)	7.41 (1)
Germany	82.35 (2)	35.29 (8)	0.85 (2)	11.18 (4)	9.28 (3)	2.68 (3)
Hong Kong	23.53 (9)	76.47 (1)	0.81 (3)	7.34 (6)	2.35 (7)	1.30 (11)
Singapore	23.53 (10)	76.47 (2)	0.81 (4)	6.70 (7)	1.16 (9)	1.22 (14)
United Kingdom	64.71 (3)	70.59 (3)	0.77 (5)	21.82 (2)	15.46 (2)	4.33 (2)
Spain	41.18 (6)	52.94 (5)	0.74 (6)	16.46 (3)	5.60 (4)	1.72 (6)
France	58.82 (4)	52.94 (4)	0.71 (7)	9.21 (5)	5.26 (5)	2.31 (4)
Italy	41.18 (7)	29.41 (10)	0.65 (8)	0.00	2.35 (8)	1.70 (7)
Japan	47.06 (5)	35.29 (9)	0.65 (9)	4.90 (8)	2.57 (6)	2.03 (5)
Canada	29.41 (8)	29.41 (11)	0.63 (10)	0.00 (13)	1.14 (11)	1.59 (8)
Portugal	17.65 (12)	41.18 (6)	0.63 (11)	1.18 (9)	0.68 (14)	1.17 (16)
Australia	23.53 (11)	23.53 (12)	0.61 (12)	0.00	1.15 (10)	1.42 (9)
Korea	17.65 (13)	17.65 (13)	0.61 (13)	0.90 (10)	0.61 (15)	1.22 (13)
China	17.65 (14)	17.65 (14)	0.59 (14)	0.79 (11)	0.89 (13)	1.32 (10)
Argentina	5.88 (17)	17.65 (15)	0.57 (15)	0.00	0.22 (16)	1.07 (17)
Brazil	17.65 (15)	5.88 (16)	0.57 (16)	0.00	1.10 (12)	1.23 (12)
India	11.76 (16)	5.88 (17)	0.55 (17)	0.00	0.11 (18)	1.07 (18)
Mexico	5.88 (18)	5.88 (18)	0.53 (18)	0.00	0.18 (17)	1.19 (15)

Notes: Numbers in parentheses indicate the ranking. In-degree, out-degree, betweenness, and intermediation are expressed in percent.

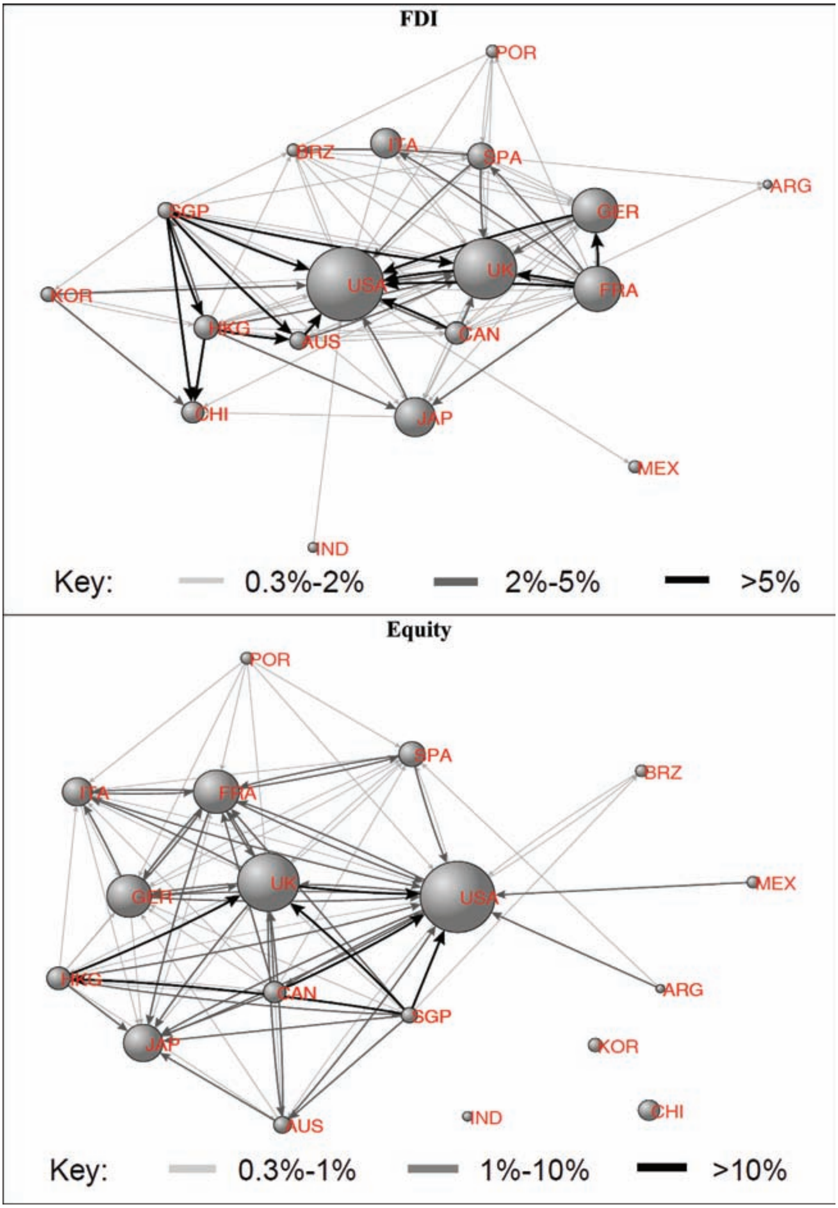
- *The United States, the United Kingdom, and Germany are the main recipients of foreign investment.* This can be seen by the number of arrows pointing to these nodes and by the high value of *in-degree centrality*, which measures the number of links that arrive at a node divided by the maximum number of links.
- *Financial centers—Hong Kong, Singapore, and the United Kingdom—are the main originators of foreign investment*, as can be seen by the number of arrows pointing out and the high value of *out-degree centrality*, which measures the number of links that depart from a node divided by the maximum number of links.
- *The countries which are located closer to other nodes in the network are the United States, Germany, Hong Kong, Singapore, and the United Kingdom. Closeness centrality* is the inverse of the average distance between countries, where distance is measured by the number of links on the shortest path. A country which is directly connected to all other countries, such as the United States, has a closeness score equal to 1.
- *The United States and the United Kingdom are the main countries connecting other nodes.* This is captured by *betweenness centrality*, which measures the frequency with which a country lies on the shortest path between two other countries, and *intermediation centrality*, which captures the intensity of links by incorporating portfolio shares.
- *The United States and the United Kingdom also score highest in terms of prestige centrality. Prestige centrality* (or eigenvector centrality) reflects the importance of the counterparties. A country with high prestige is one that is linked to others that have high prestige themselves.

3.2 Financial Network—Asset Composition

To analyze differences across asset classes, figure 3 represents the networks with links given by the ratio of assets to GDP of the source country for each asset class in 2005.[15] Centrality measures have also been calculated for each of these networks.

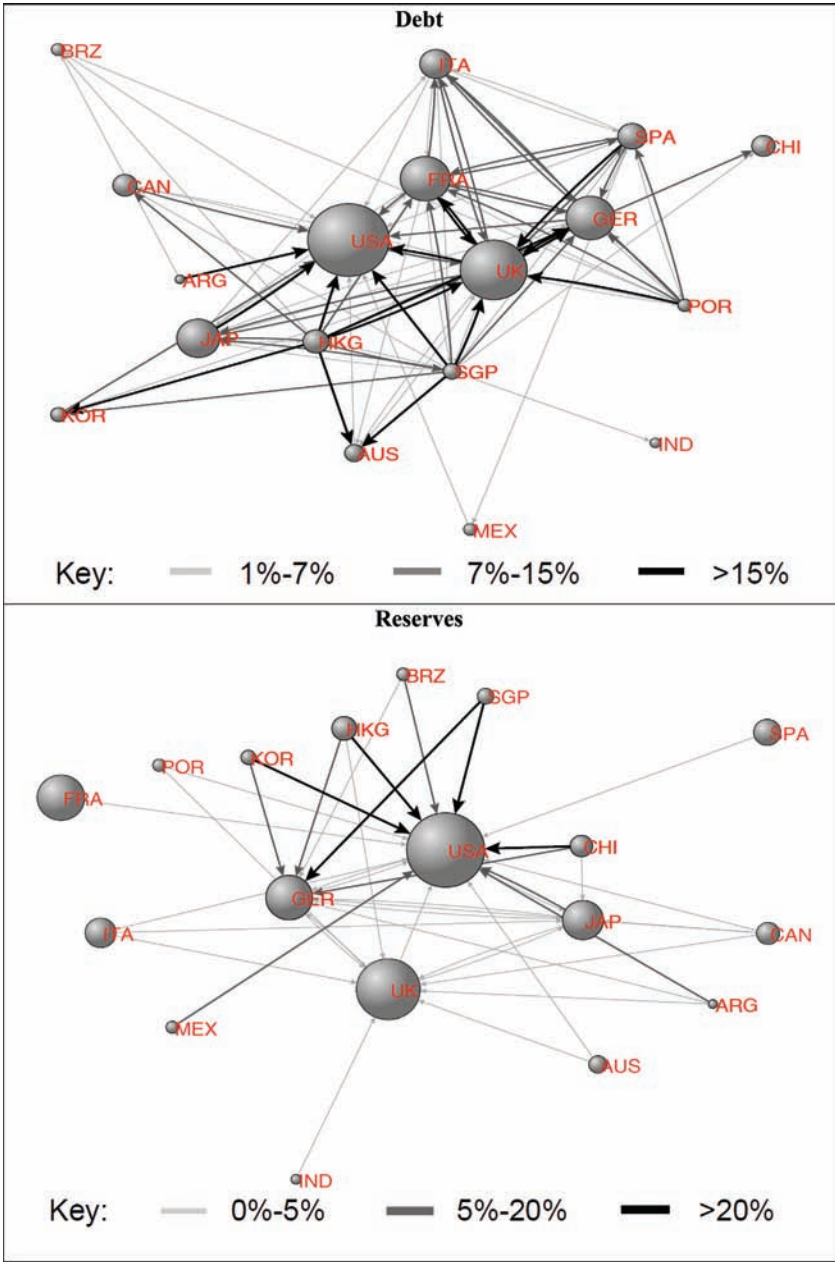
[15] The cutoff for deletion of the smallest links is 0.3 percent for FDI and equity and 1 percent for debt. No cutoff is imposed for reserves.

Figure 3. International Financial Network by Asset Class, 2005



(continued)

Figure 3. (Continued)



Notes: Links are given by the ratio of bilateral assets to GDP of the source country for each asset class. The size of the nodes is proportional to the country's financial openness, measured by the sum of its total external assets and liabilities.

Table 9. Summary Statistics on the International Trade Network

	1985	1995	2005
Skewness	6.55	6.73	5.67
Kurtosis	53.03	55.93	42.90
Average Path Length	1.56	1.67	1.89
Clustering Coefficient	0.36	0.39	0.48

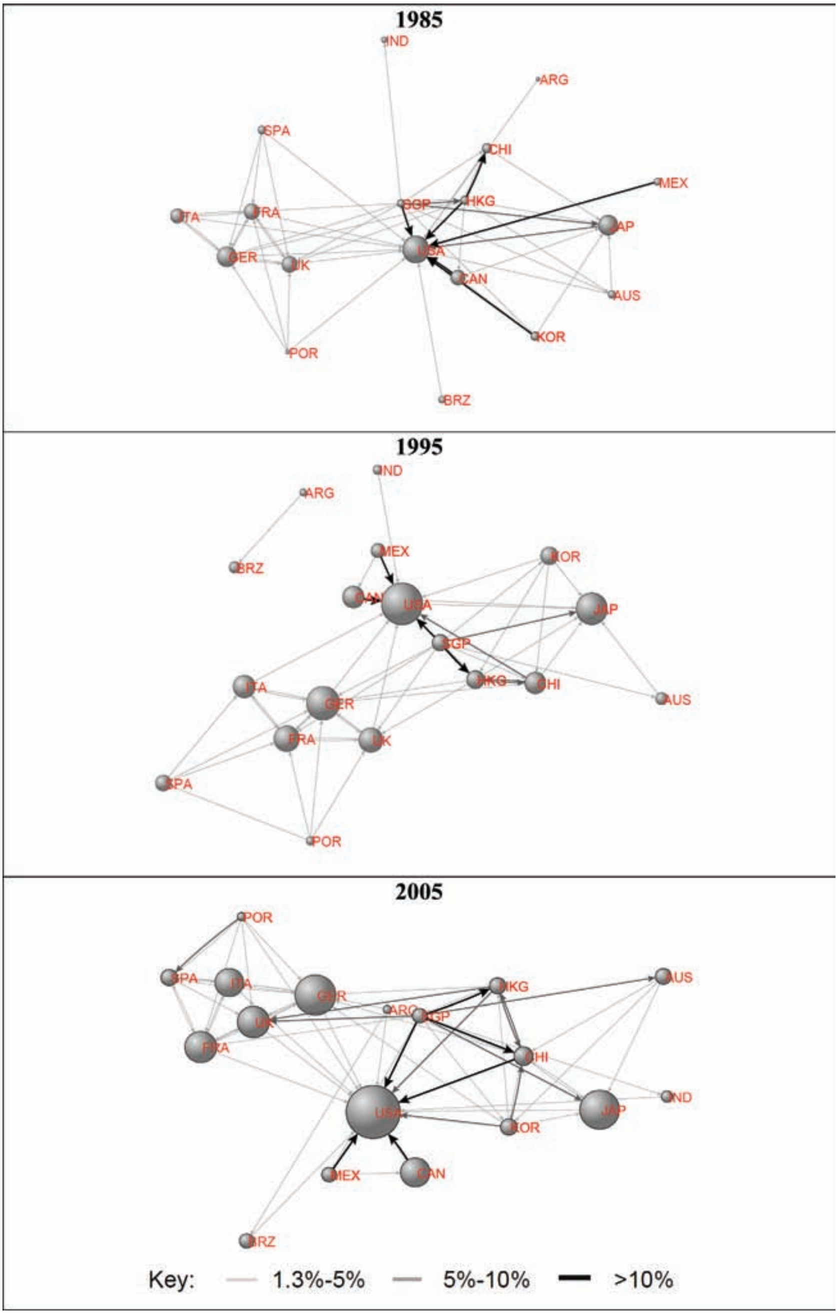
The United States and the United Kingdom emerge as the main recipients of foreign investment for FDI, equity, and debt and have the highest scores for in-degree centrality. Singapore and Hong Kong score low as recipients of foreign investment, but they score high as originators. There are some interesting differences across asset classes. The equity network is less dense than for other asset classes, with some countries (China, Korea, and India) being unconnected. The United States is ranked first in out-degree centrality for FDI and second for equity investment, but is ranked only tenth as originator of debt investment. This is consistent with the finding that U.S. foreign assets tend to be composed mostly of equity and FDI, while its foreign liabilities tend to be composed mostly of bonds (Gourinchas and Rey 2007a). For reserves, the network is less dense because we only measure reserve holdings in four currencies: dollars, euros, pounds, and yens. Among these currencies, the dollar is clearly dominant. The large arrow linking China to the United States reflects the large value of China's foreign exchange reserves: over US $0.8 trillion in 2005, most of which was held in dollar assets. This value has increased steadily to over US $2.3 trillion in 2009.

3.3 Comparison with the Trade Network

Table 9 reports summary statistics and figure 4 represents the directed trade network. Links are given by the ratio of bilateral exports to GDP of the source country:

$$link_{ijt} = \frac{Exports_{ijt}}{GDP_{it}}.$$

Figure 4. International Trade Network



Notes: Links are given by the ratio of bilateral exports to GDP of the source country. The size of the nodes is proportional to the country's trade openness, measured by the sum of its total exports and imports.

An arrow pointing from i to j is proportional to the value of country i's exports to country j, divided by the GDP of country i.[16] The thickness of the lines is proportional to the size of the links, and the size of the nodes is proportional to the country's trade openness, measured by the sum of total exports and imports. Data on bilateral trade are from the IMF Direction of Trade Statistics (DOTS).

A few findings emerge:

- Just as for the global financial network, *the interconnectivity of the global trade network increased over the last two decades.* This can be seen from the increase in the size of the nodes and the increase in the size and number of links.
- *The distribution of trade links also exhibits a long tail* with a small number of countries having large links. This asymmetry in the distribution can be seen from the large values of skewness and kurtosis.

To identify which countries play a central role in the trade network, measures of centrality in 2005 are given in table 10. These measures highlight some additional facts:

- *In all years the trade network exhibits strong intracontinental links with three clusters:* an American cluster (United States, Canada, and Mexico), an Asian cluster (Singapore, Hong Kong, China, Korea, and Japan), and a European cluster (United Kingdom, Germany, France, Spain, Italy, and Portugal). This pattern contrasts with the one found for financial links, where the United Kingdom and the United States were clearly at the center of the network linking to almost all other nodes.
- *Germany, China, and France are important trade centers and score highly both as exporters and as importers. The United States is the main importer but scores low as an exporter. The opposite is true for Singapore, which is the main exporter but scores low as an importer.*

[16]Links for which this ratio is below 1.3 percent are not shown in the figure. This cutoff is set so that every node is linked to at least one other node.

Table 10. Measures of Network Centrality—Trade, 2005

	In-Degree	Out-Degree	Closeness	Betweenness	Intermediation	Prestige
United States	88.24 (1)	5.88 (15)	0.89 (1)	5.33 (5)	28.55 (1)	5.42 (1)
Singapore	5.88 (12)	64.71 (1)	0.74 (2)	6.74 (4)	2.60 (12)	1.42 (13)
Germany	52.94 (2)	29.41 (2)	0.71 (3)	10.99 (2)	10.03 (2)	2.64 (3)
China	35.29 (3)	29.41 (3)	0.65 (4)	11.64 (1)	9.93 (3)	2.67 (2)
France	35.29 (4)	29.41 (4)	0.63 (5)	1.50 (8)	6.73 (5)	2.25 (5)
Korea	17.65 (8)	29.41 (6)	0.63 (6)	2.08 (7)	3.00 (11)	1.59 (12)
United Kingdom	35.29 (5)	17.65 (10)	0.63 (7)	0.27 (9)	7.29 (4)	2.25 (4)
Hong Kong	17.65 (9)	29.41 (7)	0.61 (8)	7.90 (3)	3.52 (9)	1.69 (11)
Japan	23.53 (6)	11.76 (13)	0.57 (9)	0.25 (10)	5.66 (7)	2.10 (7)
Portugal	5.88 (13)	29.41 (8)	0.57 (10)	0.15 (12)	0.91 (18)	1.23 (17)
Italy	17.65 (10)	23.53 (9)	0.55 (11)	0.15 (13)	4.36 (8)	1.82 (10)
Argentina	0.00 (18)	17.65 (12)	0.53 (12)	0.00	0.97 (17)	1.13 (18)
Mexico	5.88 (14)	11.76 (14)	0.53 (13)	0.18 (11)	1.50 (16)	1.85 (9)
India	5.88 (15)	5.88 (17)	0.52 (14)	0.00	1.58 (14)	1.25 (16)
Brazil	5.88 (16)	5.88 (18)	0.50 (15)	0.00	3.47 (10)	1.35 (14)
Canada	11.76 (11)	5.88 (16)	0.50 (16)	0.00	1.54 (15)	2.12 (6)
Australia	5.88 (17)	17.65 (11)	0.49 (17)	0.00	2.16 (13)	1.35 (15)
Spain	23.53 (7)	29.41 (5)	0.47 (18)	4.29 (6)	6.21 (6)	1.90 (8)

Notes: Numbers in parentheses indicate the ranking. In-degree, out-degree, betweenness, and intermediation are expressed in percent.

- *Germany appears to be the center of the European cluster and China appears to be the center of the Asian cluster.* These countries play an important role connecting other nodes, as can be seen by their high scores for betweenness and intermediation centrality.
- *The United Kingdom occupies a much less central position in the trade network than in the financial network.*

4. Implications for the Stability of the International Financial System

Higher interconnectivity entails a fundamental trade-off. On the one hand, it enhances risk sharing by allowing countries to better diversify idiosyncratic risks. On the other hand, it increases the risk of contagion. If a shock hits a highly interconnected country, its creditors will suffer losses because the profitability of their investment falls. This could generate a cascade of losses through the system.

The international financial network has long tails, with some countries having multiple and large links. A long-tailed distribution of links is a property of “scale-free” networks, whose robustness has been studied, for example, by Albert, Jeong, and Barabási (2000). Their study shows that these networks are robust to random shocks: since the majority of the nodes have only a few small links, there is a higher probability that a random shock will hit a less-connected node. However, they are very vulnerable to targeted attacks hitting the most-connected nodes.

Low average path length and a high clustering coefficient are characteristics of the so-called “small-world” networks described, for example, in Watts and Strogatz (1998). In contrast to “scale-free” networks, these networks do not exhibit much variability in the number of links of each node. This suggests that they are not particularly vulnerable to targeted attacks. However, because they are characterized by a high degree of interconnectivity, once an attack occurs, it will tend to spread more widely.

The global financial network exhibits characteristics of both “scale-free” and “small-world” networks. Because the network has a small number of nodes with multiple and large links and is highly interconnected, it is susceptible to targeted attacks affecting the key financial hubs. Disturbances to those hubs would spread rapidly and generate large losses throughout the network.[17]

To study how shocks would spread through the network, we take the configuration of links in 2005 and simulate the losses to the network following a 10 percent reduction in the value of all countries' assets in country i (the shock country). For example, if the shock originates in the United States, we start by reducing the value of other countries' assets in the United States by 10 percent. This gives the losses in the first round. In the second round, countries that hold assets in U.S. creditor countries will suffer as well because the profitability of those assets is reduced due to the losses that these countries suffer on their exposures to the United States.

To illustrate, suppose that Spain has assets in the United States and the United Kingdom and that the United Kingdom also has assets in the United States. In the first round, both Spain and the United Kingdom lose 10 percent of the value of their assets in the United States. In the second round, Spain loses on its assets in the United Kingdom due to the losses that the United Kingdom incurred because of its exposure to the United States. The first round of contagion reflects losses on direct exposures to the shock country. Later rounds of contagion reflect losses on indirect exposures. The appendix discusses the simulation in detail.

Table 11 reports losses for all countries in the network following a reduction of 10 percent in their asset values in the United States. Losses are reported both in value and in percentage of the GDP of the impact country. The countries that suffer the largest losses in percentage of GDP following a shock to the United States are the other "hubs" in the network: Hong Kong, Singapore, and the United Kingdom.[18] Figure 5 plots total losses to the network in each round following a U.S. shock. Losses are measured as a fraction of the sum of the GDP of all countries except the United States. The impact of the shock is largest in the first round and decays exponentially in subsequent rounds. This suggests that the largest losses occur due to direct exposures.

[17]These properties of the global financial network and its consequences for stability are discussed in Haldane (2009).

[18]Argentina also suffers significant losses in the first round because of its large exposure to the United States in 2005, which is apparent in figure 2.

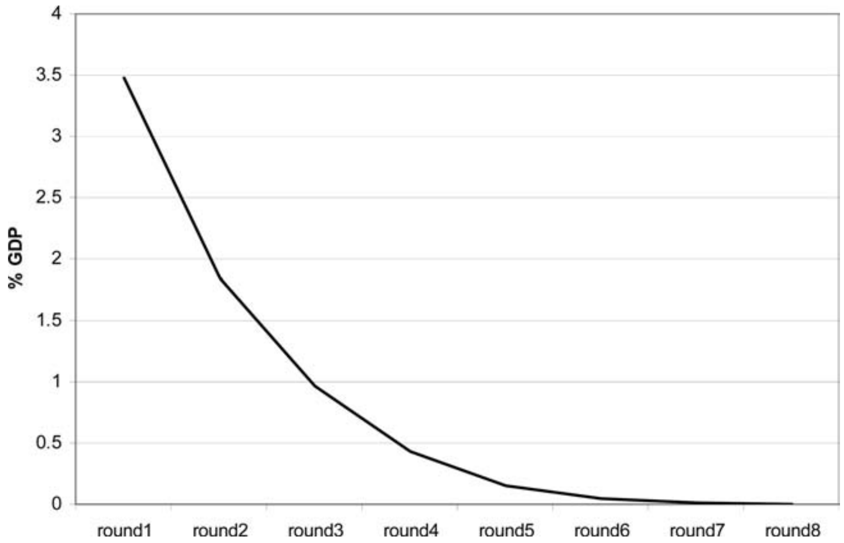
Table 11. Losses Following a Shock to the United States

Impact on	Value ($US millions)		% of GDP	
	Round 1	Total	Round 1	Total
Hong Kong	25624.96	83484.78	14.41	46.95
Singapore	15405.77	38968.57	13.20	33.39
United Kingdom	192073.50	343368.15	8.61	15.38
Argentina	13179.33	14515.94	7.19	7.92
Canada	51272.34	72980.79	4.53	6.45
Japan	193610.96	288804.83	4.26	6.35
Korea	24747.97	31970.43	3.13	4.04
France	65220.41	218109.49	3.05	10.21
Australia	21593.70	33628.65	2.93	4.56
China	65181.36	82130.22	2.90	3.66
Germany	65987.52	189934.63	2.37	6.82
India	12408.78	14513.42	1.54	1.80
Mexico	11416.76	11981.94	1.49	1.56
Spain	9219.23	64210.37	0.82	5.70
Portugal	1507.00	10168.07	0.81	5.49
Brazil	6431.20	8611.10	0.73	0.98
Italy	10775.55	60657.80	0.61	3.43
Total Impact	**785656.33**	**1568039.16**	**3.47**	**6.93**

Notes: The table reports losses following a shock that reduces the value of all countries' assets in the United States by 10 percent. Losses in round 1 come from direct exposures to the United States. Losses in later rounds are due to indirect exposures, via other countries. The total represents the sum of all losses up to round 8. The last two columns report losses in percentage of the GDP of the impact country.

The transmission of the shock would be very different it had originated in a peripheral country. Table 12 reports losses following a 10 percent reduction in asset values in Brazil. The countries most affected in the first round are the ones with the strongest links to Brazil—in particular, Spain, Portugal, and Argentina. As further rounds of losses take place, the shock spreads to other countries, especially to the most interlinked ones. When looking at total losses after eight rounds, the countries that suffer the most are the financial centers of Hong Kong, Singapore, and the United Kingdom. Still, these losses are very small compared with the ones that would occur following a shock to the United States.

Figure 5. Total Losses in Each Round Following a Shock to the United States



To understand which countries can generate the largest losses to the system, table 13 reports total losses following shocks to each country in the network. A shock to the United States is by far the most harmful, generating losses equal to 7 percent of the combined GDP of all countries except the United States. Shocks to the United Kingdom, Germany, France, and Japan also generate considerable losses. The shocks that generate the smallest losses are the ones to peripheral countries: shocks to Argentina, India, Portugal, Mexico, and Brazil generate losses of less than 0.4 percent of the combined GDP of all countries except the shock country.

This simulation looks at countries' total external assets regardless of their composition. It implicitly assumes that losses are propagated in the same way for FDI, equity, debt, and reserve assets. It should be noted, however, that different types of assets have different risk-sharing abilities. In particular, contingent asset classes (such as FDI and portfolio equity) offer better opportunities for risk sharing than non-contingent assets classes (such as bonds or bank loans). In addition, FDI is more long term in nature and would be less vulnerable to shocks than shorter-term instruments.

Table 12. Losses Following a Shock to Brazil

	Value ($US millions)		% of GDP	
Impact on	**Round 1**	**Total**	**Round 1**	**Total**
Spain	3584.90	6742.75	0.32	0.60
Portugal	543.52	1162.40	0.29	0.63
Argentina	483.53	1073.31	0.26	0.59
Hong Kong	278.54	4172.56	0.16	2.35
United States	14503.47	24279.07	0.12	0.20
France	2259.17	14104.14	0.11	0.66
Canada	1088.34	4148.63	0.10	0.37
United Kingdom	2067.86	18298.52	0.09	0.82
Singapore	103.97	1975.13	0.09	1.69
Germany	1768.44	12014.89	0.06	0.43
Italy	1096.79	4611.42	0.06	0.26
Japan	1258.00	14152.77	0.03	0.31
Australia	140.89	1652.49	0.02	0.22
Korea	69.82	1470.70	0.01	0.19
Mexico	10.78	513.45	0.00	0.07
India	9.42	633.08	0.00	0.08
China	22.86	3683.81	0.00	0.16
Total Impact	**29290.31**	**114689.13**	**0.09**	**0.34**

Notes: The table reports losses following a shock that reduces the value of all countries' assets in Brazil by 10 percent. Losses in round 1 come from direct exposures to Brazil. Losses in later rounds are due to indirect exposures, via other countries. The total represents the sum of all losses up to round 8. The last two columns report losses in percentage of the GDP of the impact country.

Hence, we would expect countries whose external assets are dominated by short-term debt instruments to suffer more losses following a shock than countries whose external assets are composed mostly of FDI. This is confirmed in the study of Milesi-Ferretti and Tille (2010), which finds that the retrenchment in international capital flows that occurred during the current financial crisis was more pronounced in countries with large net external liabilities in the form of debt. While we do not explore this issue, our data set can be used to study the role of different asset classes in the transmission of shocks.

Table 13. Total Losses Following Shocks to Each Country in the Network

Shock to	Value ($US millions)	% GDP
United States	1568039.16	6.93
United Kingdom	1268860.43	3.87
Germany	710879.23	2.21
France	633305.64	1.93
Japan	526333.97	1.73
Italy	411378.91	1.24
Spain	386418.14	1.14
Canada	335041.01	0.99
Australia	222762.98	0.65
Hong Kong	147660.26	0.42
China	136022.47	0.41
Singapore	136726.26	0.39
Korea	120970.69	0.35
Brazil	114689.13	0.34
Mexico	110800.01	0.32
Portugal	103376.24	0.30
India	42934.96	0.13
Argentina	36274.34	0.10

Notes: The table reports losses following a shock that reduces the value of all countries' assets in the shock country by 10 percent. Losses are an aggregate for all rounds up to round 8. When reported as a percentage of GDP, losses are measured as a share of the sum of the GDP of all countries in the network except the shock country.

5. Conclusions

This paper contributes to the study of financial globalization by constructing a data set on bilateral financial links for a group of eighteen countries from 1980 to 2005. We collect data from several sources and fill gaps using gravity models. Network tools are used to identify the key stylized facts that emerge from the data and to study the propagation of shocks to different countries in the network. We find a remarkable increase in interconnectivity over the past two decades, with an increase in the number and size of financial links. In addition, the distribution of financial links has a long tail, with a small number of countries having large and numerous links.

The network exhibits some "small-world" properties, with a small number of degrees of separation between nodes and a high clustering coefficient. The combination of high interconnectivity, long-tails, and "small-world" properties makes for a robust yet fragile system, where disturbances to one of the central hubs would be transmitted widely and rapidly.

The trade network also reveals an increase in interconnectivity over time. However, unlike the financial network, where the United States and the United Kingdom are at the center and intracontinental links are not particularly strong, the trade network exhibits much stronger links within continents. In particular, there is a European cluster, centered around Germany; an Asian cluster, centered around China; and an American cluster, centered around the United States. The United Kingdom plays a much less central role in the trade network than in the financial network.

Appendix

Statistical Definitions

Skewness is a measure of the asymmetry of a distribution and is defined as $\frac{E(X-\mu)^3}{(E(X-\mu)^2)^{3/2}}$. A normal distribution is symmetric and has skewness equal to 0. A positive value for skewness indicates that the distribution has a long tail on the right; i.e., there are many observations with small values of X and few observations with large values of X.

Kurtosis is a measure of the "peakedness" of a distribution and is defined as $\frac{E(X-\mu)^4}{(E(X-\mu)^2)^2}$. A normal distribution has a kurtosis equal to 3. A large value for kurtosis indicates that the distribution has "fat tails."

Network Definitions

The network can be expressed in matrix form, where the typical element A_{ij} records the value of financial assets held by country i in country j. The matrix has dimension equal to the number of countries, n, and can be read in two directions: rows of A represent assets of country i in country j, and columns of A represent liabilities of j in i. All diagonal elements are zero. Off-diagonal elements are zero for country pairs whose links are below the cutoff, defined in such a way that each country is linked to at least one other country (either as a creditor or as a debtor). The network is directed and weighted, hence A is not symmetric and its entries reflect the size of financial assets.

Two perspectives can be taken when analyzing weighted networks. One perspective looks at whether a link exists or not, regardless of the value of the link; i.e., it looks at the indicator $N_{ij} = 1$ if $A_{ij} > 0$, and 0 otherwise. Another perspective takes into account the size of the links A_{ij}.

Average path length is the average of the shortest paths between all pairs of nodes in the network. For example, if node i is directly linked to node k, the shortest path between the two nodes has length 1. Average path length is the average of this measure for all pairs of nodes.

Clustering measures the probability that, given that node i is directly linked to nodes j and k, node j is also directly linked to k. The clustering coefficient is given by $\frac{\sum_{i,j\neq i,k\neq j,k\neq i} N_{ij}N_{ik}N_{jk}}{\sum_{i,j\neq i,k\neq j,k\neq i} N_{ij}N_{ik}}$.

Measures of Network Centrality

The definitions of the centrality measures used in the paper follow closely the box in von Peter (2007). The centrality measures apply to each node and describe how that node relates to the network, taking different perspectives. Degree, closeness, and betweenness centrality are based on whether a link exists or not, regardless of the value of the link; i.e., they are based on the indicator N_{ij}. Intermediation and prestige centrality take into account the size of the links and rely on the portfolio shares $P_{ij} = A_{ij}/\sum_k A_{ik}$ for all i.

In-degree is the number of links that point to a node and is given by the sum $\sum_j N_{ji}$. *Out-degree* is the number of links departing from a node and is given by the sum $\sum_j N_{ij}$. The measures of in-degree and out-degree centrality reported in the tables scale these sums by the total possible number of links, $n - 1$.

Closeness is the inverse of the average distance from node i to all other nodes. The distance between i and j, δ_{ij}, equals the length of the shortest path. The average distance from i to all other nodes is given by $\sum_j \delta_{ij}/(n - 1)$. Closeness is the inverse of this measure.

Betweenness focuses on the nodes that the shortest path goes through. Let g_{jk} denote the number of shortest paths between j and k, and let $g_{jk}(i)$ denote the number of such paths that go through node i. The probability that node i is on the shortest path from j to k is given by $g_{jk}(i)/g_{jk}$. Betweenness of node i is the sum of these probabilities over all nodes excluding i, divided by the maximum that the sum can attain: $(\sum_{j\neq i}\sum_{k\neq i} g_{jk}(i)/g_{jk})/(n-1)(n-2)$.

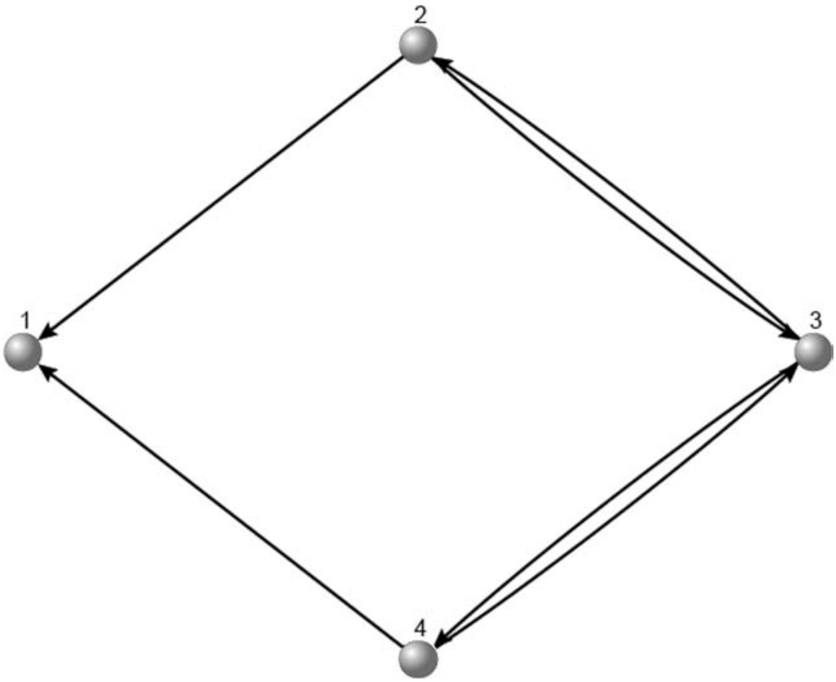
Intermediation extends the betweenness measure, taking into account the value of the links. The probability that a dollar sent by i reaches j in two steps is given by $\sum_k P_{ik}P_{kj}$. The probability that a dollar sent by i reaches j through k is given by $P_{ik}P_{kj}/\sum_k P_{ik}P_{kj}$. The intermediation measure for node k is obtained by summing these probabilities for all pairs (i,j), divided by the total number of pairs $n(n-1)$.

Prestige (or eigenvector centrality) considers the identity of the counterparties. The prestige of country i (v_i) is obtained by taking the prestige of its creditors, weighted by their portfolio shares with i, i.e., $v_i = \sum_j P_{ji}v_j$. This defines a linear system $v = P'v$, where P is the matrix of portfolio shares. The solution to this system is the eigenvector associated with the unit eigenvalue. Following von Peter (2007), we solve the alternative system $v = \frac{1}{2}P'v + e \Rightarrow v = (I - \frac{1}{2}P')^{-1}e$, where e is the unit vector. This avoids countries with a zero score contributing nothing to the centrality of others.

Simulation Exercise

To illustrate how losses to the network are calculated following a reduction of 10 percent in asset values in a given country, consider the hypothetical network with four countries shown in figure 6. An arrow from country 2 to country 1 indicates that country 2 holds assets in country 1. Suppose there is a shock that reduces asset values in country 1 by 10 percent. In the first round, countries 2 and 4 suffer losses because they have a direct exposure to country 1. Total losses in the first round equal $L^{round1} = 0.1\times P_{21}\times A_2 + 0.1\times P_{41}\times A_4$, where P_{21} and P_{41} are the portfolio shares of countries 2 and 4 in country 1, and A_2 and A_4 are the values of these countries' total assets before the shock. In the second round, country 3 also suffers losses because it holds assets in countries 2 and 4. The profitability

Figure 6. Hypothetical Network



of these assets is reduced due to the losses that countries 2 and 4 suffer on their exposure to country 1. Losses in the second round equal $L^{round2} = 0.1 \times (P_{32} \times P_{21} + P_{34} \times P_{41}) \times A_3$. The product of portfolio shares $P_{32} \times P_{21}$ can be interpreted as the exposure of country 3 to shock country 1 via country 2. There is a third and final round of contagion, in which countries 2 and 4 lose on their assets in country 3 following the losses that country 3 suffered in the second round. Losses in the third round equal $L^{round3} = 0.1 \times P_{23} \times P_{34} \times P_{41} \times A_2 + 0.1 \times P_{43} \times P_{32} \times P_{21} \times A_4$.

Note that we rule out feedback loops. For example, in the third round, we only consider the losses that country 2 suffers because of the second-round losses of country 3 on its exposures to country 4. We do not consider the second-round losses that country 3 has on its exposures to country 2; i.e., we do not consider $0.1 \times P_{23} \times P_{32} \times P_{21} \times A_2$. This is a simplifying assumption to ensure that the simulation converges to an equilibrium where no further losses to the network occur. If we had allowed for feedback loops, there would always be further rounds of contagion and the simulation would not converge.

In the actual network considered in this paper, there are eighteen countries which are all linked to each other. Ruling out feedback loops, there would be seventeen rounds of contagion. For computational tractability, we only consider losses up to round 8. Given that losses decrease exponentially—as shown in figure 5—this is a good approximation to total losses.

References

Ahluwalia, M. 2002. "Economic Reforms in India Since 1991: Has Gradualism Worked?" *Journal of Economic Perspectives* 16 (3): 67–88.

Albert, R., H. Jeong, and A. Barabási. 2000. "Error and Attack Tolerance of Complex Networks." *Nature* 406 (6794): 378–82.

Bekaert, G., C. Harvey, and C. Lundblad. 2007. "Financial Openness and the Chinese Growth Experience." Available at http://ssrn.com/abstract=991640.

Dooley, M., S. Lizondo, and D. Mathieson. 1989. "The Currency Composition of Foreign Exchange Reserves." *IMF Staff Papers* 36 (2): 385–434.

Eichengreen, B., and D. Mathieson. 2000. "The Currency Composition of Foreign Exchange Reserves: Retrospect and Prospect." IMF Working Paper No. 00/131.

Felettigh, A., and P. Monti. 2008. "How to Interpret the CPIS Data on the Distribution of Foreign Portfolio Assets in the Presence of Sizeable Cross-Border Positions in Mutual Funds. Evidence from Italy and the Main Euro-Area Countries." Banca d'Italia Questioni di Economia e Finanza Occasional Paper No. 16 (August).

Gourinchas, P. O., and H. Rey. 2007a. "From World Banker to World Venture Capitalist: U.S. External Adjustment and the Exorbitant Privilege." In *G7 Current Account Imbalances: Sustainability and Adjustment*, ed. R. Clarida. Chicago: University of Chicago Press.

———. 2007b. "International Financial Adjustment." *Journal of Political Economy* 115 (4): 665–703.

Haldane, A. 2009. "Rethinking the Financial Network." Speech delivered at the Financial Student Association, Amsterdam. Available at http://www.bankofengland.co.uk/publications/speeches/2009/speech386.pdf.

Kaminsky, G., and S. Schmukler. 2003. "Short-Run Pain, Long-Run Gain: The Effects of Financial liberalization." NBER Working Paper No. 9787.

Lane, P., and G. M. Milesi-Ferretti. 2001. "The External Wealth of Nations: Measures of Foreign Assets and Liabilities for Industrial and Developing Countries." *Journal of International Economics* 55 (2): 263–94.

———. 2007. "The External Wealth of Nations Mark II: Revised and Extended Estimates of Foreign Assets and Liabilities, 1970-2004." *Journal of International Economics* 73 (2): 223–50.

———. 2008. "International Investment Patterns." *Review of Economics and Statistics* 90 (3): 538–49.

Lane, P., and J. Shambaugh. 2007. "Financial Exchange Rates and International Currency Exposures." NBER Working Paper No. 13433.

Levy-Yeyati, E., and F. Sturzenegger. 2005. "Classifying Exchange Rate Regimes: Deeds vs. Words." *European Economic Review* 49 (6): 1603–35.

Lucas, R. 1990. "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review: Papers and Proceedings* 80 (2): 92–96.

Martin, P., and H. Rey. 2004. "Financial Super-Markets: Size Matters for Asset Trade." *Journal of International Economics* 64 (2): 335–61.

McGuire, P., and N. Tarashev. 2008. "Global Monitoring with the BIS International Banking Statistics." BIS Working Paper No. 244.

McGuire, P., and P. Wooldridge. 2005. "The BIS Consolidated Banking Statistics: Structure, Uses and Recent Enhancements." *BIS Quarterly Review* (September): 73–86.

Milesi-Ferretti, G. M., and C. Tille. 2010. "The Great Retrenchment: International Capital Flows During the Global Financial Crisis." Available at http://www.cepr.org/meets/wkcn/9/979/papers/Milesi-Ferretti_Tille.pdf.

OECD (Organisation for Economic Cooperation and Development). 2000. "Main Determinants and Impacts of Foreign Direct Investment on China's Economy." OECD Directorate for Financial, Fiscal and Enterprise Affairs, Working Paper on International Investment No. 2000/4.

Okawa, Y., and E. van Wincoop. 2010. "Gravity in International Finance." Hong Kong Institute for Monetary Research Working Paper No. 07/2010 (April).

Papaioannou, E. 2009. "What Drives International Financial Flows? Politics, Institutions and Other Determinants." *Journal of Development Economics* 88 (2): 269–81.

Portes, R., and H. Rey. 2005. "The Determinants of Cross-Border Equity Flows." *Journal of International Economics* 65 (2): 269–96.

Prasad, E., and S. Wei. 2005. "The Chinese Approach to Capital Inflows: Patterns and Possible Explanations." IMF Working Paper No. 05/79.

Reserve Bank of India. 2006. "Report of the Committee on Fuller Capital Account Convertibility." Chapter 3. Available at http://www.rbi.org.in/scripts/PublicationReportDetails.aspx?fromdate=09/1/06&SecId=21&SubSecId=.

Rose, A., and M. Spiegel. 2004. "A Gravity Model of Sovereign Lending: Trade, Default, and Credit." *IMF Staff Papers* 51 (s1): 64–74.

Stein, E., and C. Daude. 2007. "Longitude Matters: Time Zones and the Location of Foreign Direct Investment." *Journal of International Economics* 71 (1): 96–112.

von Peter, G. 2007. "International Banking Centres: A Network Perspective." *BIS Quarterly Review* (December): 33–45.

Watts, D., and S. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393 (6684): 409–410.

Diagnosing the Financial System: Financial Conditions and Financial Stress*

Scott Brave[a] and R. Andrew Butters[b]
[a]Federal Reserve Bank of Chicago
[b]Northwestern University

We approach the task of monitoring financial stability within a framework that balances the costs and benefits of identifying future crisis-like conditions based on past U.S. financial crises. Our results indicate that the National Financial Conditions Index (NFCI) produced by the Federal Reserve Bank of Chicago is a highly predictive and robust indicator of financial stress at leading horizons of up to one year, with measures of leverage playing a crucial role in signaling financial imbalances. At longer forecast horizons, we propose an alternative sub-index of the NFCI that captures the relationship between non-financial leverage, financial stress, and economic activity.

JEL Codes: G01, G17, C43.

1. Introduction

Monitoring financial stability is not unlike a medical practitioner using a person's body temperature, blood pressure, and other vital signs to make a diagnosis of their health. Identifying the magnitude of financial stress at any given point in time can depend critically on

*This article is a revision of a previously published working paper (Scott Brave and R. Andrew Butters, 2010, "Gathering Insights on the Forest from the Trees: A New Metric for Financial Conditions," Federal Reserve Bank of Chicago Working Paper No. 2010-07). The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Chicago or the Federal Reserve System. The authors thank the editor and two anonymous referees as well as Hesna Genay, Spencer Krane, Alejandro Justiniano, and Federal Reserve Bank of Chicago, 2011 Midwest Finance Association, and 2011 Computing in Economics and Finance seminar participants for their helpful comments and suggestions, and Max Lichtenstein for excellent research assistance.
Author e-mails: sbrave@frbchi.org and r-butters@kellogg.northwestern.edu.

the indicators examined and the reference points used. For instance, (i) what is a "normal" level of financial stress and subsequently a level that would warrant concern? (ii) are the risks associated with both extremely low values and high values the same? and (iii) how well does any one measure predict the true underlying state of the financial system?

What proves to be a "normal" level in medicine often tends to be a range rather than a particular value. For example, it is common to simply select some number of standard deviations from the population mean, or, equivalently, "build the range" by including everything that falls outside a desired percentage of the population. A possible source of bias in this kind of reasoning, however, might result if a priori we believe that some members of that population are in fact "sick" and their vital signs are skewed as a result. Including these members when calculating averages or percentiles will reduce the power of this metric to distinguish between states of the world, i.e., a "healthy state" and a "sick state."

In the context of financial stability, the "sick state" conforms with the notion of a financial crisis, where a number of financial indicators deviate substantially, and perhaps in a particular way, from their historical averages. If we envision every observation in time of an indicator as many different patients, some "sick" and some "healthy," and its value as their "vital sign," we can begin to build the intuition behind a statistical approach to monitoring financial stability. In fact, we can apply existing methods used in medical statistics, such as receiver operating characteristic (ROC) curve analysis, to approach this problem.

The ROC curve methodology is most commonly associated with the evaluation of the outcome of a medical test, which has at its base a Bayesian calculation of the following sort: Given a known incidence of a disease in a population, how likely is it that a positive test result is reflective of a true occurrence in sample? We take a similar approach, using known incidences of U.S. financial crises to inform our analysis of potential future stress on financial conditions. In this way, the past can be used as a guide to detecting future financial imbalances by evaluating the health of the financial system against relevant benchmarks from past financial crises.

The ROC methodology is rather flexible in that it can be used both to characterize the historical predictive ability of an indicator of financial conditions and to devise a rule based on that relationship with which to judge its future realizations. However, it leaves unanswered the question of how to appropriately weight potentially conflicting signals from a number of "vital signs." This is perhaps trivial when financial markets are operating smoothly, but it becomes more of an issue when markets become segmented. In this case, extracting a signal from a large number of indicators of financial conditions that reflects the systemic importance of each is likely to provide a more robust diagnosis.

This is the approach we take in applying the ROC methodology to a dynamic factor constructed from an unbalanced panel of 100 mixed-frequency indicators of financial activity. The resulting weekly index is the National Financial Conditions Index (NFCI) made publicly available by the Federal Reserve Bank of Chicago. Drawing on the work of Berge and Jordà (2011) in applying the ROC methodology to business-cycle dating, we then describe a statistical framework that balances the costs and benefits of identifying future crisis-like conditions based on the level of the index during past U.S. financial crises and explore the NFCI's robustness as both a contemporaneous and leading indicator of financial stress.

Our results indicate that the NFCI is 95 percent accurate in identifying historical crises contemporaneously, with a decline to 80 percent accuracy at a lead of up to one year. Furthermore, breaking down the index into subcomponents reflecting the themes of risk, credit, and leverage can enhance the nature of the signal provided by the NFCI as to the severity of the crisis, with leverage playing a crucial role in signaling financial imbalances. At horizons beyond one year, we show that a particular combination of household and non-financial business leverage measures in the NFCI proves to be a consistent leading indicator of financial stress and its impact on economic activity.

The rest of the paper proceeds as follows. In the next section, we describe the ROC methodology and its application to the NFCI and the risk, credit, and leverage sub-indexes. The following section then documents the robustness of the NFCI as a measure of financial stability. Finally, we compare its ability to capture financial stress relative to several alternatives. The concluding section summarizes our findings and discusses their potential consequences for policies aimed at promoting financial stability.

2. Monitoring Financial Stability

In this section, we outline a method of evaluating the state of the financial system based on a single composite indicator of financial conditions, the Federal Reserve Bank of Chicago's National Financial Conditions Index, or NFCI. The salient properties of the NFCI are first summarized. We then proceed to describe the ROC curve methodology that underlies our investigation of U.S. financial crises. Finally, we summarize the relevant features of the NFCI as a contemporaneous and leading indicator of financial stress.

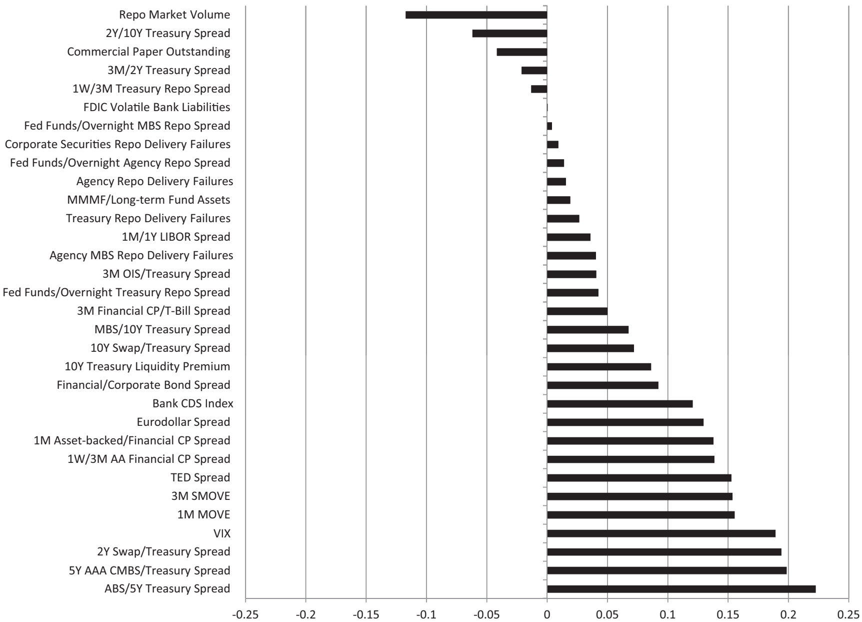
2.1 Features of the NFCI

Similar to other measures of financial conditions in the literature, such as the indexes of Illing and Liu (2006), Nelson and Perli (2007), Hakkio and Keeton (2009), Hatzius et al. (2010), and Matheson (2011), the NFCI is essentially a weighted average of a number of financial indicators where the weight given to each reflects the indicator's ability to explain the total variation among them. Indexes of this kind have the advantage of descriptively capturing the interconnectedness of their components. This attribute is desirable in the context of monitoring financial stability, as it means that more weight in the index is placed on financial indicators which have historically been systemically important.

Unlike the other financial conditions indexes in the literature, however, the NFCI uses a very flexible estimation procedure for these weights that builds off the work of Stock and Watson (2002), Doz, Giannone, and Reichlin (2006), and Aruoba, Diebold, and Scotti (2009) and is described in detail in the appendix. By allowing for the inclusion of financial indicators of varying reported frequencies that start and end at different times, it can produce a high-frequency index of financial conditions with minimal restrictions. Furthermore, the estimation strategy makes use of both cross-sectional and time-series information, exploiting the historical cross-correlations and dynamic properties of the indicators to determine their weight in the index. In this fashion, the NFCI is able to consistently extract the common signal in 100 indicators of financial activity on a weekly basis since 1973.

Volatility and credit risk measures tend to receive positive weights in the NFCI, while measures of credit and leverage tend

Figure 1. Ranking of NFCI Indicators by Factor Loadings: Risk Indicators



to be negatively correlated with the level of the index. This pattern of increasing volatility and risk premia and declining liquidity and leverage is consistent with tightening financial conditions and provides the basis for the NFCI's interpretation. Brave and Butters (2011) provide a detailed examination of the individual indicators in the NFCI. We summarize their findings in figures 1–3, which depict the weights, or factor loadings, for all 100 indicators classified into three types: risk, credit, and leverage.[1]

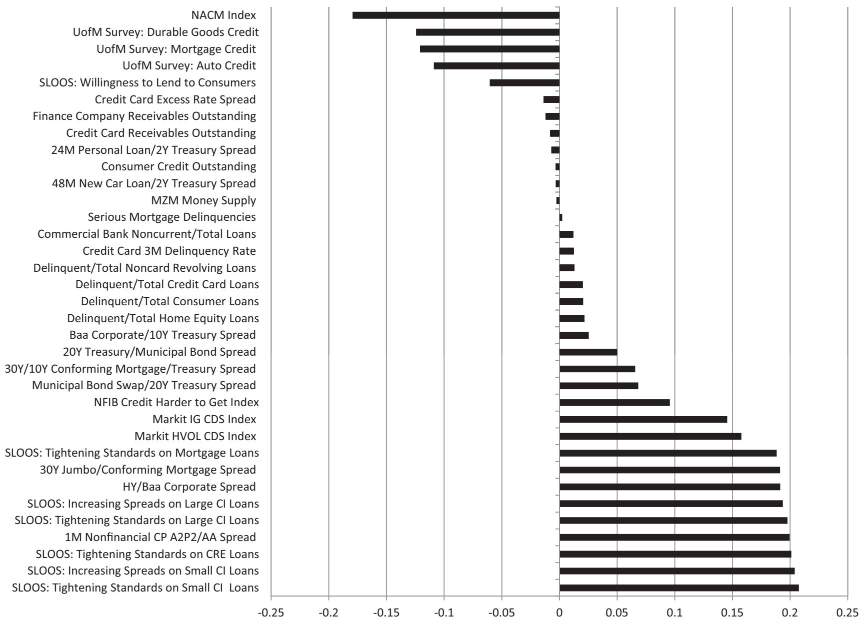
2.2 Receiver Operating Characteristics (ROC) Curve Analysis

We begin by constructing prior information on the incidence of U.S. financial crises. Ideally, we would like to have a professional

---

[1]Additional information on the NFCI, including data sources and complete data descriptions, can be found at www.chicagofed.org/nfci.

Figure 2. Ranking of NFCI Indicators by Factor Loadings: Credit Indicators

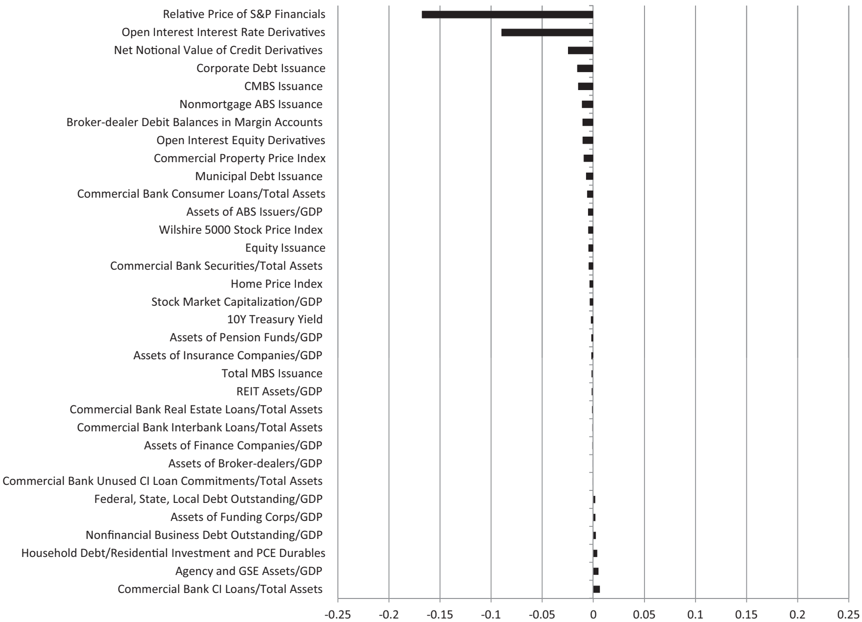


consensus much like the business-cycle dates that the National Bureau of Economic Research produces for U.S. recessions and expansions. Unfortunately, what we have to resort to are the historical accounts of events in U.S. financial history for which there is not always consensus. With this information in hand, we then follow the approach used by Berge and Jordà (2011) in evaluating indicators of the business cycle. By placing relative weights on the utility from correctly predicting crisis and non-crisis states and the disutility from making a false positive versus false negative evaluation, we develop a threshold decision rule for the transition into and out of crisis-like financial conditions.

2.2.1 Financial Crisis Dating

Table 1 provides a list of five financial crises in U.S. financial history over the last forty years, along with some of the major events that occurred during each. To arrive at these crisis episodes, we conducted

Figure 3. Ranking of NFCI Indicators by Factor Loadings: Leverage Indicators



a survey of the literature on U.S. financial crises.[2] Each decade contains at least one crisis episode, with the shortest episode extending for only two years and the longest seven years in length. The earlier episodes tend to be more concentrated in the commercial banking system, while the latter begin to take on a larger capital markets element. In determining these episodes, we paid close attention to similarities in the historical accounts on the triggers of the crisis and the policy responses that followed.

Laeven and Valencia (2008, 2010), Reinhart and Rogoff (2009), and López-Salido and Nelson (2010) go through similar exercises in producing crisis dates, albeit at a lower frequency.[3] While their start and end dates often differ significantly from those in table 1,

[2]Some examples include FDIC (1984, 1997), Schreft (1990), Spero (1999), Laeven and Valencia (2008), and Reinhart and Rogoff (2009).

[3]For the most recent crisis, the alternative dating conventions mostly conform with table 1.

Table 1. U.S. Financial Crises Since 1973

International Banking Crisis (1973–75)	1973w2	Jan. 11, 1973	DJIA peaks above 1000 and then begins to sharply decline.
	1973w19	May 11–13, 1973	Bilderberg meeting to discuss developing pressures in "petrodollars."
	1973w42	Oct. 17–18, 1974	U.S. National Bank of San Diego declared insolvent, first billion-dollar bank failure; Arab oil embargo begins.
	1974w11	Mar. 17, 1974	Oil embargo is lifted, but money-center banks and REITS continue to experience problems.
	1974w19	May 10, 1974	Regulatory agencies step in with financial assistance for Franklin National Bank.
	1974w41	Oct. 9, 1974	Franklin National Bank collapses and is acquired by European American Bank.
	1974w49	Dec. 6, 1974	DJIA bottoms out at 45% decline.
	1975w4	Jan. 19, 1975	Assisted merger of Security National Bank of New York with Chemical National Bank by regulatory agencies.
	1975w21	May 23, 1975	Regulatory agencies assist Bank of the Commonwealth to keep it afloat.
Dollar, Banking, and LDC Crises (1977–84)	1977w40	Oct. 5, 1977	Dollar peaks and then begins steep decline against major foreign currencies.
	1978w44	Nov. 1, 1978	Carter administration announcement of a dollar defense program.
	1978w51	Dec. 17, 1978	OPEC decides to keep its U.S. dollar reserves but increase oil prices in 1979.
	1980w11	Mar. 14, 1980	Carter administration announcement of imposition of credit controls.
	1980w13	Mar. 26, 1980	Regulatory agencies step in with financial assistance for First Pennsylvania National Bank.

(continued)

Table 1. (Continued)

	1980w27	July 3, 1980	Federal Reserve announces phase-out of credit controls.
	1981w43	Nov. 4, 1981	FDIC assists merger of Greenwich Savings Bank, first in a series of mutual savings bank assisted mergers.
	1982w26	July 5, 1982	Penn Square Bank fails.
	1982w31	Aug. 12, 1982	Mexico defaults on its debt, beginning of LDC crisis.
	1982w44	Nov. 6, 1982	Mexico and IMF reach accord on loan plan.
	1984w19	May 9, 1984	Run on Continental Illinois begins, bank borrows \$3.6 billion through discount window.
	1984w27	July 1, 1984	Regulators develop plan to take over Continental's bad loans.
	1984w39	Sept. 26, 1984	Resolution of Continental completed.
S&L Crisis, Black Monday, and LBO/Junk Bond Collapse (1987–91)	1986w53	Jan. 1, 1987	Federal Savings and Loan Insurance Corporation becomes insolvent.
	1987w41	Oct. 19, 1987	"Black Monday": DJIA −22.6%, S&P500 −20.4%.
	1989w32	Aug. 9, 1989	Financial Institutions Reform Recovery and Enforcement Act (FIRREA) signed into law.
	1989w41	Oct. 13, 1989	"Friday the 13th Mini Crash" helps to trigger junk bond market collapse.
	1990w3	Jan. 15, 1990	Campeau Corporation files for bankruptcy after junk bond defaults.
	1990w7	Feb. 13, 1990	Drexel Burnham Lambert files for bankruptcy, beginning of end to LBO boom.
	1990w31	Aug. 2, 1990	Iraqi invasion of Kuwait, DJIA declines 18% in three months.
	1990w52	Dec. 31, 1990	First phase of implementation of Basel I regulatory capital and leverage ratio requirements.
	1991w9	Feb. 28, 1991	Gulf War ends.

(continued)

Table 1. (Continued)

Asian Crisis, Russian Default and LTCM, Y2K, NASDAQ Bubble, 9/11, and Enron (1997–2002)	1997w43	Oct. 27, 1997	"Mini crash" brought on by Asian financial crisis.
	1998w32	Aug. 17, 1998	Russia defaults on its debt.
	1998w37	Sept. 23, 1998	Collapse of LTCM (Federal Reserve steps in with support).
	1999w29	July 27, 1999	IMF approves stand-by credit for Russian Federation.
	1999w39	Oct. 1, 1999	Fed establishes Century Date Change Special Liquidity Facility.
	1999w52	Jan. 1, 2000	Y2K passes.
	2000w35	Sept. 1, 2000	NASDAQ peaks above 4000, then begins to sharply decline.
	2001w37	Sept. 11, 2001	Terrorist attack on the World Trade Center.
	2001w50	Dec. 13, 2001	SEC enforcement action against Enron.
	2002w28	July 15, 2002	Arthur Anderson indicted.
	2002w31	July 30, 2002	Sarbanes-Oxley Act passed.
Subprime Mortgage Crisis and Aftermath (2007–Current)	2007w31	July 31, 2007	Bear Stearns liquidates two hedge funds investing in MBS.
	2008w11	Mar. 14, 2008	Bear Stearns sold to JP Morgan Chase with NY Fed support.
	2008w28	July 12–15, 2008	Fannie Mae and Freddie Mac receive government assistance, IndyMac fails.
	2008w37	Sept. 14–16, 2008	Lehman Brothers files for bankruptcy, AIG receives govt. support, Reserve Fund "breaks the buck."
	2008w39	Sept. 26, 2008	Washington Mutual Bank failure, largest failure in terms of assets to date.
	2008w40	Oct. 3, 2008	Emergency Economic Stabilization Act passed (TARP).
	2008w47	Nov. 23, 2008	Citigroup requires government assistance.
	2009w2	Jan. 16, 2009	Bank of America requires government assistance.
	2010w18	May 9, 2010	EU, ECB, and IMF announce $1 trillion aid package after Greek debt crisis.
	2011w32	Aug. 5, 2011	S&P downgrade of U.S. credit rating.

their crisis episodes are largely subsets of those that we consider, with the exception of the 1997–2002 period, which none of the others deems as evidence of crisis-like financial conditions. Our episodes share with those of López-Salido and Nelson (2010) the tumultuous events in financial markets that accompanied the 1974–75, 1981–82, and 2007–09 U.S. recessions, as well as the savings and loan crisis of the mid-1980s and early 1990s. The latter two crises are the singular focus of Laeven and Valencia (2008, 2010) and Reinhart and Rogoff (2009), as each is concerned primarily with periods characterized by substantial bank failures. We focus on the dating convention of Laeven and Valencia (2010) for the U.S. crises, as it differs the most from table 1 and López-Salido and Nelson (2010).

2.2.2 The Area Under the ROC Curve

The ROC method requires that we categorize each point in time as falling within a crisis or non-crisis period. Given the dating conventions in table 1, consider the derivation in Berge and Jordà (2011):

$$TP(c) = P[I_t \geq c || C_t = 1] \tag{1}$$

$$FP(c) = P[I_t \geq c || C_t = 0], \tag{2}$$

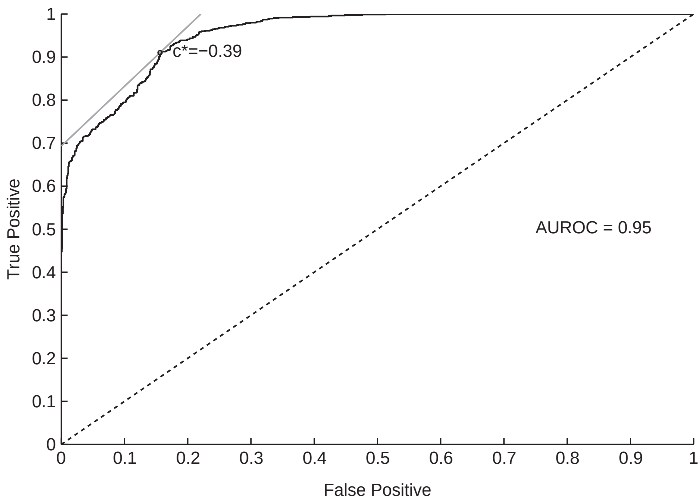
with $C_t \in \{0, 1\}$ indicating the non-crisis and crisis states of financial conditions, respectively. $TP(c)$ is typically referred to as the true positive, sensitivity, or recall rate, and $FP(c)$ is known as the false positive or 1-specificity rate for an indicator, I_t, observed with value c. The relationship between the two is described by the ROC curve. With the Cartesian convention, this curve is given by

$$\{ROC(r), r\}_{r=0}^{1}, \tag{3}$$

where $ROC(r) = TP(c)$ and $r = FP(c)$.

Figure 4 depicts this curve estimated non-parametrically for the weekly National Financial Conditions Index produced by the Federal Reserve Bank of Chicago. The closer to one or zero the area under the curve (AUROC)—depending on whether the index is positively or negatively correlated with the crisis episodes in table 1, respectively—the more predictive it is of crisis-like financial conditions, where statistical significance is judged relative to the area

Figure 4. NFCI ROC Curve with Utility Line



under the ray from the origin extending at a 45-degree angle.[4] The NFCI is highly predictive of the crisis episodes in table 1, with an AUROC of roughly 0.95 that is statistically significant at standard levels of confidence. Repeating this procedure with a lead of one and two years between the values of the NFCI and the crisis episodes produces AUROCs that remain statistically significant at 0.80 and 0.62, respectively.

Of the three crisis dating conventions, our baseline and the López-Salido and Nelson (2010) dates produce the most similar AUROC results for the NFCI, particularly at the one-year-ahead horizon. What is mostly responsible for the lower contemporaneous AUROC values of 0.76 and 0.62 when using the López-Salido and Nelson (2010) dates or those of Laeven and Valencia (2010) is the 1997–2002 period not being considered by either as a crisis and the omission of the early '70s and '80s by the latter. Still, for both alternative crisis dating conventions, the AUROC values for the NFCI remain statistically significant, and they remain so even at the three-year-ahead horizon as opposed to our baseline dates, which tend to

[4]The procedure for evaluating statistical significance is described in DeLong, DeLong, and Clarke-Pearson (1988).

already be more liberal in their definition of the beginning and ends of crises.

For the sake of comparison, consider table 2, which depicts the AUROC value for each of the indicators in the NFCI contemporaneously and at leads of up to three years.[5] Values in darkened and boldface fonts denote statistical significance at standard confidence levels based on the crisis episodes in table 1. The table categorizes each indicator into one of three types: risk, credit, or leverage. Risk indicators capture volatility and funding risk in the financial sector, while credit indicators are composed of measures of credit conditions, and leverage indicators consist of debt and equity measures.

In general, risk and credit measures tend to have very significant contemporaneous and one-year-ahead AUROC values. However, the AUROCs for risk measures at longer forecast horizons tend to be more significant than for credit measures, although a few of the latter remain significant predictors out to three years ahead. In contrast, there are considerably fewer significant leverage indicators at a contemporaneous forecast horizon, although several do appear to have a leading feature from two to three years ahead. Only a handful of indicators display a sense of being a superior contemporaneous or leading indicator to the NFCI, and many of these are neither observed weekly nor do they span all five crisis episodes in table 1, providing our motivation for the NFCI as the relevant measure for monitoring financial stability.

It is also possible to construct sub-indexes of the NFCI that reflect the different classes of indicators in table 2. Table 3 displays the AUROC for the NFCI and its three sub-indexes using our baseline and alternative crisis dates contemporaneously and up to a lead of three years. The predictive ability of the overall index is superior to any of the three sub-indexes from a year ahead to the contemporaneous horizon, although of the three sub-indexes, risk measures dominate. Two years and beyond, measures of leverage increasingly account for the ability of the NFCI to predict crisis conditions. In fact, relative to the Laeven and Valencia (2010) dates, the leverage

[5]One indicator in the NFCI, the net notional value of credit derivatives, did not have a sufficient history for the AUROC calculations.

Table 2. AUROC for Financial Indicators at Various Horizons

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Baa Corporate/10Y Treasury Spread	Credit	W	5	**0.60**	0.50	**0.45**	**0.38**
1M Non-Financial CP A2P2/AA Spread	Credit	W	2	**0.82**	**0.64**	0.53	**0.40**
20Y Treasury/Municipal Bond Spread	Credit	W	5	**0.58**	**0.55**	**0.54**	**0.56**
24M Personal Loan/2Y Treasury Spread	Credit	Q	4	**0.38**	**0.35**	**0.29**	**0.32**
30Y Jumbo/Conforming Mortgage Spread	Credit	W	2	**0.90**	**0.88**	**0.70**	**0.57**
30Y/10Y Conforming Mort./Treasury Spread	Credit	W	2	**0.68**	**0.67**	**0.56**	**0.44**
48M New Car Loan/2Y Treasury Spread	Credit	Q	4	0.44	0.44	**0.33**	**0.34**
Commercial Bank Non-Current/Total Loans	Credit	Q	4	**0.80**	**0.76**	0.61	0.50
Consumer Credit Outstanding	Credit	M	5	0.45	0.50	0.54	**0.60**
Credit Card 3M Delinquency Rate	Credit	M	2	0.50	0.55	0.56	0.57
Credit Card Excess Rate Spread	Credit	M	2	**0.63**	0.53	0.49	0.52
Credit Card Receivables Outstanding	Credit	M	2	**0.42**	0.46	0.50	0.52
Delinquent/Total Consumer Loans	Credit	M	2	0.55	**0.64**	0.60	0.51
Delinquent/Total Credit Card Loans	Credit	M	2	0.50	0.56	0.54	0.50
Delinquent/Total Home Equity Loans	Credit	M	2	0.50	0.61	**0.70**	**0.68**
Delinquent/Total Non-Card Revolving Loans	Credit	M	2	0.51	**0.63**	0.60	0.50

(continued)

Table 2. (Continued)

Indicator	**Type**	**Freq.**	**# of Crises**	**Current**	**−1 Year**	**−2 Years**	**−3 Years**
Finance Company Receivables Outstanding	Credit	M	3	0.45	0.51	0.54	0.55
HY/Baa Corporate Spread	Credit	W	2	**0.79**	**0.56**	**0.39**	**0.17**
Markit HVOL CDS Index	Credit	W	1	**1.00**	**0.80**	**0.68**	NaN
Markit IG CDS Index	Credit	W	1	**0.99**	**0.77**	**0.60**	NaN
Municipal Bond Swap/20Y Treasury Spread	Credit	W	3	**0.63**	**0.77**	**0.79**	**0.78**
MZM Money Supply	Credit	M	5	0.48	**0.56**	0.54	0.52
NACM Index	Credit	M	2	**0.14**	**0.16**	**0.30**	0.49
NFIB Credit Harder to Get Index	Credit	M	5	**0.65**	0.53	**0.37**	**0.40**
Serious Mortgage Delinquencies	Credit	Q	5	0.56	**0.62**	0.56	0.54
SLOOS: Increasing Spreads on Large CI Loans	Credit	Q	3	**0.83**	0.71	0.54	0.28
SLOOS: Increasing Spreads on Small CI Loans	Credit	Q	3	**0.81**	0.65	0.47	**0.23**
SLOOS: Tightening Standards on CRE Loans	Credit	Q	3	**0.81**	0.71	0.60	0.39
SLOOS: Tightening Standards on Large CI Loans	Credit	Q	3	**0.83**	**0.71**	0.56	0.34
SLOOS: Tightening Standards on Mortgage Loans	Credit	Q	3	0.69	0.68	0.59	0.53
SLOOS: Tightening Standards on Small CI Loans	Credit	Q	3	**0.84**	**0.74**	0.59	0.37

(continued)

Table 2. (Continued)

Indicator	**Type**	**Freq.**	**# of Crises**	**Current**	**−1 Year**	**−2 Years**	**−3 Years**
SLOOS: Willingness to Lend to Consumers	Credit	Q	5	**0.31**	**0.34**	**0.38**	0.50
UofM Survey: Auto Credit	Credit	M	4	**0.24**	**0.30**	**0.33**	**0.41**
UofM Survey: Durable Goods Credit	Credit	M	4	**0.17**	**0.27**	**0.36**	0.47
UofM Survey: Mortgage Credit	Credit	M	4	**0.22**	**0.29**	**0.34**	**0.40**
10Y Treasury Yield	Leverage	W	5	**0.54**	0.52	0.52	0.51
Agency and GSE Assets/GDP	Leverage	Q	5	0.60	0.57	0.54	0.47
Assets of ABS Issuers/GDP	Leverage	Q	4	**0.29**	**0.30**	0.39	0.60
Assets of Broker-Dealers/GDP	Leverage	Q	5	**0.37**	0.45	0.48	0.52
Assets of Finance Companies/GDP	Leverage	Q	5	0.49	0.59	0.56	0.50
Assets of Funding Corps/GDP	Leverage	Q	5	0.54	**0.62**	**0.63**	0.52
Assets of Insurance Companies/GDP	Leverage	Q	5	0.42	0.51	0.50	0.52
Assets of Pension Funds/GDP	Leverage	Q	5	0.43	0.48	0.48	0.56
Broker-Dealer Debit Balances in Margin Accounts	Leverage	M	5	**0.37**	0.49	0.54	**0.59**
CMBS Issuance	Leverage	M	3	**0.42**	0.44	0.44	0.49
Commercial Bank CI Loans/Total Assets	Leverage	M	5	**0.56**	**0.59**	**0.65**	**0.66**
Commercial Bank Consumer Loans/Total Assets	Leverage	M	5	**0.43**	**0.42**	**0.43**	0.47
Commercial Bank Interbank Loans/Total Assets	Leverage	M	5	0.50	0.51	0.52	0.51
Commercial Bank Real Estate Loans/Total Assets	Leverage	M	5	0.49	0.49	0.49	0.47

(continued)

Table 2. (Continued)

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Commercial Bank Securities/Total Assets	Leverage	M	5	0.47	**0.44**	**0.39**	**0.39**
Commercial Bank Unused CI Commitments/Total Assets	Leverage	Q	4	0.44	0.40	0.39	0.40
Commercial Property Price Index	Leverage	Q	4	0.37	0.48	0.56	0.60
Corporate Debt Issuance	Leverage	M	3	**0.39**	**0.41**	0.44	0.51
Equity Issuance	Leverage	M	3	0.44	0.46	0.54	0.52
Federal, State, Local Debt Outstanding/GDP	Leverage	Q	5	0.45	**0.38**	0.41	0.46
Home Price Index	Leverage	M	4	0.51	**0.56**	**0.56**	**0.57**
Household Debt/Residential Investment and PCE Durables	Leverage	Q	5	0.59	0.57	0.55	0.52
Municipal Debt Issuance	Leverage	M	1	0.40	0.48	0.43	0.47
Non-Financial Business Debt Outstanding/GDP	Leverage	Q	5	**0.62**	**0.70**	**0.71**	**0.67**
Non-Mortgage ABS Issuance	Leverage	M	2	0.41	0.44	0.45	0.48
Open Interest Equity Derivatives	Leverage	W	2	0.46	**0.43**	**0.42**	0.47
Open Interest Interest Rate Derivatives	Leverage	W	2	**0.35**	**0.32**	**0.39**	0.46
REIT Assets/GDP	Leverage	Q	5	0.43	0.44	0.40	0.46
Relative Price of S&P Financials	Leverage	W	3	**0.27**	**0.33**	**0.31**	**0.36**

(continued)

Table 2. (Continued)

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Stock Market Capitalization/GDP	Leverage	Q	5	0.40	0.50	0.53	0.59
Total MBS Issuance	Leverage	M	2	0.49	0.40	**0.35**	**0.29**
Wilshire 5000 Stock Price Index	Leverage	M	5	0.47	0.51	0.54	**0.58**
MBS/10Y Treasury Spread	Risk	M	4	**0.75**	**0.79**	**0.69**	0.56
10Y Swap/Treasury Spread	Risk	W	3	**0.75**	**0.69**	**0.57**	**0.41**
10Y Treasury Liquidity Premium	Risk	W	3	**0.61**	**0.58**	0.47	**0.34**
1M Asset-Backed/Financial CP Spread	Risk	W	2	**0.91**	**0.81**	**0.73**	**0.61**
1M MOVE	Risk	W	3	**0.62**	0.49	**0.46**	**0.36**
1M/1Y LIBOR Spread	Risk	W	3	0.51	**0.43**	0.49	**0.43**
1W/3M AA Financial CP Spread	Risk	W	2	0.51	0.46	**0.65**	**0.65**
1W/3M Treasury Repo Spread	Risk	W	2	**0.39**	**0.45**	**0.55**	**0.57**
2Y Swap/Treasury Spread	Risk	W	3	**0.85**	**0.81**	**0.67**	0.48
2Y/10Y Treasury Spread	Risk	W	5	**0.34**	**0.28**	**0.33**	**0.37**
3M Financial CP/T-Bill Spread	Risk	W	5	**0.65**	**0.69**	**0.70**	**0.58**
3M OIS/Treasury Spread	Risk	W	1	**0.31**	0.48	**0.73**	**0.94**
3M SMOVE	Risk	W	2	**0.46**	**0.36**	**0.31**	**0.20**
3M/2Y Treasury Spread	Risk	W	5	**0.38**	**0.34**	**0.46**	**0.46**
5Y AAA CMBS/Treasury Spread	Risk	W	2	**0.93**	**0.71**	0.51	**0.28**
ABS/5Y Treasury Spread	Risk	M	3	**0.88**	**0.88**	**0.82**	**0.65**
Agency MBS Repo Delivery Failures	Risk	W	2	**0.55**	**0.55**	0.51	**0.44**
Agency Repo Delivery Failures	Risk	W	2	**0.54**	**0.55**	0.49	**0.44**
Bank CDS Index	Risk	W	1	**1.00**	**0.80**	0.57	NaN

(continued)

Table 2. (Continued)

Indicator	**Type**	**Freq.**	**# of Crises**	**Current**	**−1 Year**	**−2 Years**	**−3 Years**
Commercial Paper Outstanding	Risk	W	2	**0.45**	0.54	**0.60**	**0.59**
Corporate Securities Repo Delivery Failures	Risk	W	2	0.48	0.54	0.48	0.45
Eurodollar Spread	Risk	W	5	**0.82**	**0.79**	**0.71**	**0.62**
FDIC Volatile Bank Liabilities	Risk	Q	4	0.45	0.54	0.62	0.63
Fed. Funds/Overnight Agency Repo Spread	Risk	W	2	0.51	0.53	0.51	**0.45**
Fed. Funds/Overnight MBS Repo Spread	Risk	W	2	0.52	0.53	0.48	**0.45**
Fed. Funds/Overnight Treasury Repo Spread	Risk	W	2	0.51	**0.55**	**0.54**	0.51
Financial/Corporate Bond Spread	Risk	M	4	**0.81**	**0.84**	**0.79**	**0.67**
MMMF/Long-Term Fund Assets	Risk	M	5	**0.74**	**0.66**	**0.63**	0.55
Repo Market Volume	Risk	W	2	**0.42**	**0.45**	**0.39**	**0.41**
TED Spread	Risk	W	4	**0.75**	**0.79**	**0.75**	**0.62**
Treasury Repo Delivery Failures	Risk	W	2	0.51	0.53	0.47	0.50
VIX	Risk	W	3	**0.91**	**0.79**	**0.62**	**0.39**

Notes: The table displays the area under the ROC curve relative to the five crisis periods in table 1 contemporaneously and at leads of up to three years for each financial indicator in the NFCI. Areas significant from 50 percent are denoted by bold (95 percent significance level) and darkened fonts (90 percent significance level).

Table 3. AUROC at Contemporaneous and Leading Horizons

Indicator	Baseline				López-Salido and Nelson (2010)				Laeven and Valencia (2010)			
	0 Yr	**−1 Yr**	**−2 Yr**	**−3 Yr**	**0 Yr**	**−1 Yr**	**−2 Yr**	**−3 Yr**	**0 Yr**	**−1 Yr**	**−2 Yr**	**−3 Yr**
NFCI	**0.95**	**0.80**	**0.62**	0.48	**0.76**	**0.79**	**0.73**	**0.59**	**0.62**	**0.56**	**0.45**	**0.32**
Risk	**0.93**	0.80	**0.64**	0.50	**0.73**	**0.77**	**0.74**	**0.60**	**0.58**	**0.53**	**0.45**	**0.32**
Credit	**0.90**	**0.73**	**0.55**	**0.45**	**0.72**	**0.73**	**0.62**	**0.55**	**0.63**	**0.55**	**0.43**	**0.39**
Leverage	**0.78**	**0.72**	**0.66**	0.53	**0.69**	**0.79**	**0.79**	**0.71**	**0.77**	**0.74**	**0.68**	**0.55**

Notes: The table displays the area under the ROC curve relative to the crisis periods denoted in the column headings contemporaneously and at leads of up to three years for the NFCI and its three sub-indexes. Areas significant from 50 percent are denoted by bold (95 percent significance level) and darkened fonts (90 percent significance level).

sub-index is the dominant signal from the contemporaneous to the two-year-ahead horizon. We explore this result further below.

2.2.3 Crisis Thresholds

The utility function depicted in figure 4 is expressed as in Baker and Kramer (2007),

$$\begin{aligned} U = U_{11}ROC(r)\pi + U_{01}(1 - ROC(r))\pi + U_{10}r(1 - \pi) \\ + U_{00}(1 - r)(1 - \pi), \end{aligned} \tag{4}$$

where U_{ij} is the utility (or disutility) associated with the prediction i, given that the true state is j, $i, j \in \{0, 1\}$ and π is the unconditional probability of observing a crisis episode during the sample period. Utility maximization implies that the optimal threshold value c^* is given by the solution to

$$\frac{\partial ROC}{\partial r} = \frac{U_{00} - U_{10}}{U_{11} - U_{01}} \frac{1 - \pi}{\pi}, \tag{5}$$

that is, the point where the slope of the ROC curve equals the expected marginal rate of substitution between the net utility of accurate crisis and non-crisis episode prediction.

Essentially, one is weighing the costs of a type I versus type II error relative to the benefits of correctly predicting the true state when attempting to separate a mixture distribution into its unique components. This intuitively amounts to deciding on the emphasis (in utility terms) one wants to put on correctly identifying either state. An example of assigning equal weight to correctly identifying both crisis and non-crisis episodes would be assigning $U_{00} = U_{11} = 1$, $U_{01} = U_{10} = -1$. In contrast, placing all the emphasis on correctly identifying financial crises, and subsequently no emphasis on the likely error of identifying the other state as a crisis, the utilities could be assigned this way: $U_{00} = 0$, $U_{11} = 1$, $U_{01} = -1$, $U_{10} = -\epsilon$, where ϵ needs to be small but non-zero in order to prevent the utility function from being degenerate. Finally, a threshold rule that puts more emphasis on identifying non-crisis periods corresponds with a utility function like $U_{00} = 1$, $U_{11} = 0$, $U_{01} = -\epsilon$, $U_{10} = -1$.

These alternative approaches are consistent with three parameterizations of the level sets of the utility function. Graphically, each rule attempts to find the unique intersection of the linear utility function with the convex ROC curve. A rule placing a very steep penalty on missing early on an occurrence of a financial crisis thus looks to intersect the upward-sloping portion of the ROC curve. A rule that places a relatively larger penalty on missing an occurrence of a non-crisis period does the opposite and instead intersects the flatter portion of the ROC curve. The equal weight, or "unbiased," rule falls somewhere in between the other two on the ROC curve.

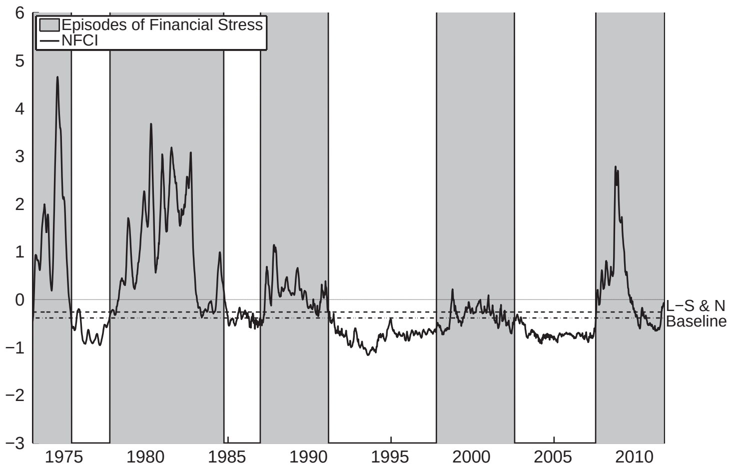
2.3 Making a Diagnosis

With equal weight on all four utilities (disutilities), the unbiased threshold rule balances the need to detect financial imbalances in advance of a crisis with the desire to avoid assigning movements in the NFCI at low levels to financial stress versus simple mean reversion. In essence, this derivation "rebases" the index based on the historical financial crises we have identified. However, as in any Bayesian analysis, it is necessary to investigate the sensitivity of the decision rule to the prior information used. For the NFCI, the unbiased threshold rule turns out to be rather robust to the dating of past financial crises.

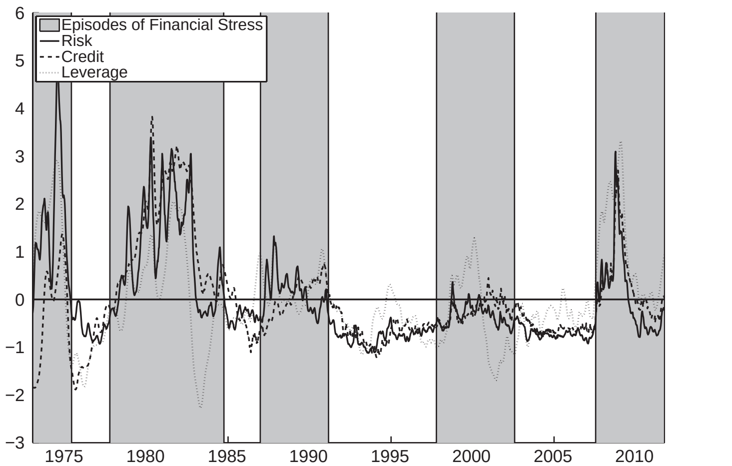
To see this, consider the top panel in figure 5, which depicts the history of the NFCI relative to its historical average and in standard deviation units, shading the crisis episodes in table 1 and plotting the unbiased contemporaneous thresholds based on those episodes and the alternative dates suggested by López-Salido and Nelson (2010). Interestingly, the resulting threshold for the index is a slightly negative number in both cases. This result is intuitive, as it is very apparent in figure 5 that the transition into and out of a crisis is often characterized by a sharp deviation from below and above the index mean, respectively.

The sub-indexes provide further evidence on the source of the initial impetus into and out of a financial crisis. The bottom panel of figure 5 plots each relative to its own mean and in its own standard deviation units. This scaling makes for easier viewing but does

Figure 5. The NFCI and Sub-Indexes



Crisis Dating Convention	Baseline	L-S & N
Contemporaneous Threshold	–0.39	–0.26



NFCI Sub-Index	Risk	Credit	Leverage
Contemporaneous Threshold	–0.39	–0.49	–0.13

obscure the relative contribution of each sub-index to the NFCI. However, it should be readily apparent from the figure that the risk measures are the dominant source of variation in the overall index, as they are nearly identical in quality to the NFCI. The most severe crises in our sample are characterized by above-average values of all three sub-indexes. Leverage is the most cyclical of the three, often leading the others into and out of a crisis. Credit instead tends to follow the more persistent risk sub-index over the course of a crisis with a slight lag.

The contemporaneous threshold values for each of the three, shown in the tables accompanying figure 5, reflect the above facts, with risk demonstrating a threshold largely in line with the NFCI, credit slightly below this, and leverage nearer its long-run average of zero. The behavior of leverage mirrors the results of Adrian and Shin (2010) on the procyclical nature of financial leverage, while the level of the threshold suggests a possible role for financial stability policy in stabilizing measures of financial leverage near their long-run averages. This remains the case even at policy-relevant horizons, where the leverage threshold is mostly unchanged and its AUROC values remain significant.

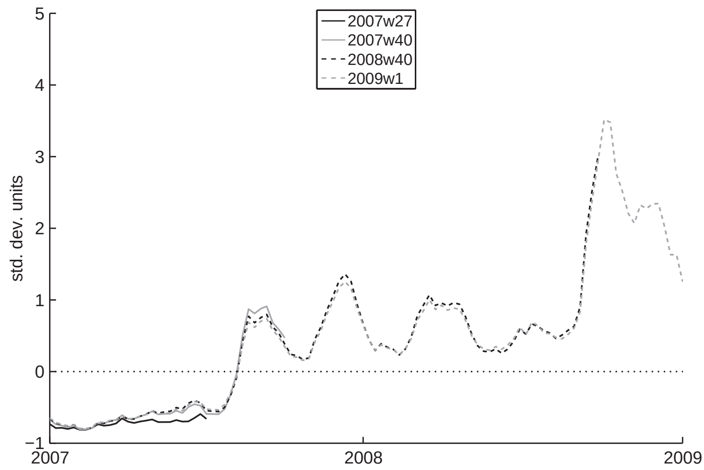
2.4 The Recent Crisis

It is helpful to consider what an unbiased threshold rule would have meant in hindsight for the most recent crisis. Such a rule based on our crisis episodes first signals the development of crisis conditions according to the NFCI in August 2007, nearly in step with our dating. Using the López-Salido and Nelson (2010) dates does little to alter this inference.[6] In fact, this is a feature common to the beginning and end of nearly all of the crises we consider. The leverage sub-index was the first to cross its threshold, as suggested by the results above. However, its leading signal was rather weak given its volatility, crossing the threshold in mid-2005 and again in early 2007, but each time falling back below it shortly thereafter.

However, using the additional information in the NFCI sub-indexes would have considerably qualified the diagnosis of the severity of the recent crisis on several occasions. For instance, the leverage sub-index began to spike first in June 2007, quickly reaching its highest point since the early 1970s. This was the case before the failure

[6] The Laeven and Valencia (2010) dates do not produce a reliable contemporaneous threshold, as they are heavily influenced by the omission of the 1973–75 period.

Figure 6. NFCI Revisions



of Lehman Brothers in the fall of 2008, which only exacerbated the increase until it exceeded its prior high from the early '70s. Furthermore, after the collapse of Lehman, credit conditions tightened more so than at any point since the early 1980s. Both credit conditions and leverage were then subsequently much slower to return to their average levels than were measures of risk, with the leverage sub-index consistently above its threshold for much of the last two years.

Such an analysis, however, omits a role to be played by the nature of the construction of the NFCI, which is potentially nontrivial for a real-time evaluation of the state of financial conditions. We can address the out-of-sample properties of the NFCI by simulating its production in the period leading up to the most recent crisis. We do so holding fixed the availability of financial indicators at the end of each quarter and using revised data to focus solely on the impact that estimating the scale of the index and the weight assigned to each financial indicator has over this period. Our simulation runs from the second quarter of 2007 through the fourth quarter of 2008, thus capturing the run-up and height of the recent crisis.

Figure 6 is reassuring in the fact that despite very large movements in many of the financial indicators, revisions to the NFCI are small over this period. This suggests very little loss in efficiency in the ROC framework from our method of index construction. The effect of the crisis is initially characterized by a one-time jump in the level of the index of about 0.2 standard deviations between the second and third quarters of 2007. Revisions afterwards are smaller but do occur between the third and fourth quarters of 2007 and 2008. Furthermore, incorporating both data revisions and staggered data availability, the publicly available history of the NFCI since April 2011 seldom demonstrates revisions larger than 0.1 standard deviations in size.

Keep in mind that this analysis still remains subject to the Lucas critique, as it holds fixed both the reaction of financial markets to past policy and policy to past financial market events. It is not intended to be a substitute for a fully specified dynamic model of the interaction between the two. At the very least, it provides a historical basis for judging the current state of the financial system and provides a sense in which measures of leverage may signal the development of financial imbalances. In the conclusion, we discuss potential ways in which our results could be used to inform a policy aimed at financial stability.

3. Robustness of the NFCI

There are several properties of the NFCI that interact closely with the ROC methodology and warrant closer inspection. For instance, our motivation in using the NFCI as a measure of financial stability is closely tied to its systemic interpretation of movements in a number of financial indicators. However, if the nature of those interactions is subject to exogenous breaks over time, the inference provided by the NFCI may in fact be biased. On the other hand, by focusing only on a subsample of its history where these breaks are not likely to have occured, some information is likely to be lost from the omission of past crises.

In what follows, we examine the robustness of the NFCI as a measure of financial stability. We focus our analysis on several key assumptions underlying its construction, namely (i) the lack of structural breaks in the correlation properties of the underlying financial indicators, (ii) the ability of these indicators to span the breadth of financial activity, and (iii) a consistent relationship with economic conditions.

3.1 Structural Breaks

To test for structural breaks in financial activity, we constructed an alternative version of the NFCI using only data from the post-1984 period. Panel A of figure 7 plots both the post-1984 and full-sample NFCI relative to their sample means and scaled by their sample standard deviations.[7] For the period of time in which they overlap, most of the difference between the two indexes appears in their levels, as the variance of the post-1984 index by which it is scaled is considerably smaller than the same measure for the NFCI. This suggests that it is primarily the lower volatility of the post-1984 period that is driving what differences we do see, and is in line with broader findings on the "Great Moderation."

That said, there are noticeable differences in the patterns of the two indexes, particularly over the last decade or so. The factor loadings provide some evidence of where the differences are coming from.[8] For the most part, they reflect only small differences between the indexes. However, it is apparent that the post-1984 index does shift weight away from the more traditional banking indicators and towards more money- and capital-market-centric indicators. We conclude from this that focusing on a more recent time period demonstrates that these indicators have become more important to overall financial conditions.

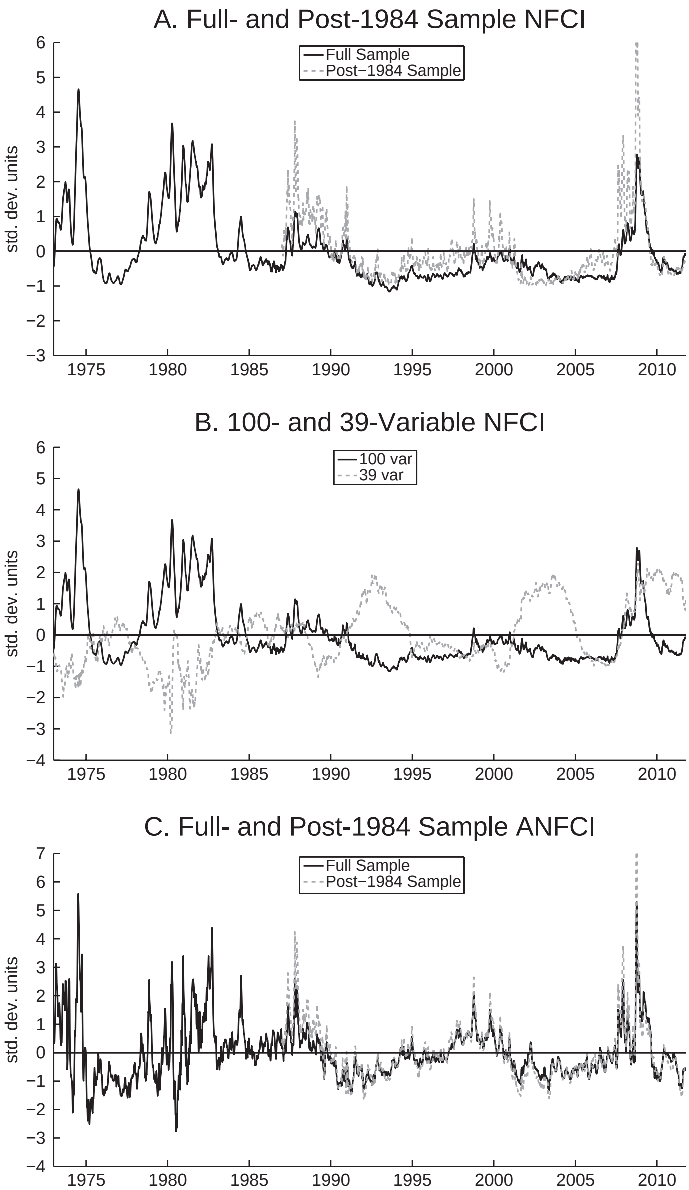
There are several reasons why this is the case. First, many of these measures reflect the growth of the "shadow" banking system over the last decade at the expense of the commercial banking system. In this sense, there does seem to be some evidence of a structural break in financial activity. Second, many of these measures also played a large part in the most recent crisis. Given how unique the recent crisis was to the period post-1984, they not surprisingly explain considerably more of the co-movement during this period.

---

[7] Just as with the full-sample index, we do not consider the first two years of estimates, so that the shorter sample index begins in 1987.

[8] The factor loadings for all of the alternative indexes considered in this section are available from the authors upon request.

Figure 7. Stability of the NFCI over Time, Indicators, and Economic Conditions



3.2 Coverage of the Financial System

An alternative to using a shorter sample period is to instead focus only on the subset of financial indicators whose history extends back over most of the sample. In this way, we can further judge whether it is possible to consistently capture financial conditions over an extended period without incorporating information from more recently developed financial markets. Panel B of figure 7 plots the NFCI computed from its thirty-nine financial indicators that extend back to 1978 against the 100-variable index over the same time period.

One can see from panel B of figure 7 that the smaller index is capturing something very different than the larger one. Except for the most recent period where both indexes demonstrate large positive values, the two are highly negatively correlated. Well-known periods of deterioration in financial conditions, such as the late 1970s and early 1980s, appear in the narrower index as very loose periods for financial conditions. In fact, many of the factor loadings are of opposite sign to those in the NFCI.

There are several explanations for the differences between the NFCI and the smaller index. First, the NFCI spans a larger cross-section of financial activity, largely due to its greater inclusion of money- and capital-market-centric measures. Second, the subset of indicators in the smaller index also contains a bias towards those more likely to be affected by the change in volatility post-1984. This can be seen in the fact that many of the same indicators in the thirty-nine-variable index also show large changes in their factor loadings in the post-1984 index.

As an example of where the smaller index seems to be diverging, consider the Treasury yield-curve indicators. Measures from both the short and long end of the curve receive large positive factor loadings in the thirty-nine-variable index, meaning that as the yield curve steepens, the index rises. In contrast, the factor loadings they receive in the NFCI are much smaller and negative, meaning that as the yield curve steepens, financial conditions tend to improve. However, even this relationship is not stable over time. In the post-1984 index, the long end of the curve receives a large positive factor loading, while the short end receives a smaller negative factor loading.

3.3 Adjusting for Economic Conditions

The results above suggest to us that at least some of the instability in the NFCI over time and indicators may be due in part to changes in the level and volatility of economic growth and inflation. To examine this hypothesis, we followed Hatzius et al. (2010) in also constructing both full- and post-1984 sample indexes where each of our 100 financial indicators was first regressed on current and lagged values of a measure of the business cycle, the three-month moving-average Chicago Fed National Activity Index (CFNAI-MA3), and inflation, three-month total PCE inflation.[9]

Panel C of figure 7 plots both the post-1984 and full-sample histories for what we call the adjusted NFCI (ANFCI). For the period of time in which they overlap, the peaks and valleys in the shorter sample index are slightly more pronounced, but the general pattern is very similar to the full-sample index. Their factor loadings confirm that the differences between the two indexes tend to be much smaller than for the NFCI and its post-1984 counterpart. This suggests that after we have accounted for the decline in the volatility of economic growth and inflation, the resulting index is much more stable.

Interestingly, during the recent crisis this was not always the case. Here, because of the large movements in several of the indicators—particularly, the money- and capital-market-centric ones—the small differences in factor loadings across indexes have a more prominent effect. Reinforcing this is the fact that the shorter sample ANFCI also shifts some weight away from the traditional banking system and towards the money and capital markets. Thus, even after accounting for changes in the volatility of economic growth and inflation, it appears that there remains some instability in the NFCI over time attributable to the growth of the shadow banking system.

[9]The number of current and lagged values was chosen for each variable using the Bayesian information criterion (BIC), with the independent variables transformed to match the frequency of observation of the financial indicator. For weekly indicators, we assumed that only lagged values enter the regression and that these values were constant over the weeks of the month.

4. Other Indicators of Financial Stress

In this section, we compare the ability of the NFCI to capture financial stress relative to other measures suggested in the literature and the reformulations of the index described above. We focus on both the contemporaneous and leading abilities of these measures. Then, we put forth alternative leading indexes derived from the NFCI that draw on the leading properties of some of the indicators observed above and compare them against the credit-to-GDP-based and bank conditions measures that have recently received attention in the literature as measures of systemic risk.

4.1 Other Indexes

Table 4 computes the AUROC for each of the alternative versions of the NFCI discussed above. The values for the NFCI are repeated at the top of the table for the sake of comparison. None of the alternative indexes proves to be a superior contemporaneous indicator to the NFCI using our baseline crisis dates. Furthermore, while the post-1984 NFCI and both the shorter and longer sample ANFCI perform slightly better at leading horizons using our baseline dates, the differences are small using the López-Salido and Nelson (2010) crisis dates. The pattern is reversed using the Laeven and Valencia (2010) crisis dates so that the NFCI is much more informative at leading horizons and less informative at the contemporaneous horizon. Interestingly, the thirty-nine-variable index is the most predictive at the contemporaneous horizon using these crisis dates.

There are several reasons for these results. First, the ANFCI, by nature of its adjustment of measures of credit by economic growth and inflation, results in a number of additional measures that exhibit similar leading characteristics to the NFCI leverage sub-index. Second, by focusing on the post-1984 period, the factor loadings for several of these measures become considerably larger. Some factor loadings even change signs relative to those in the NFCI, reflecting a different correlation pattern in the shorter sample. This is true for both risk, credit, and leverage measures, several of which demonstrate leading qualities in table 2. The Laeven and Valencia (2010) dates attenuate these concerns, as they are primarily reflective of the recent crisis where leverage measures played a large role as seen above.

Table 4. AUROC at Contemporaneous and Leading Horizons

Indicator	Baseline				López-Salido and Nelson (2010)				Laeven and Valencia (2010)			
	0 Yr	−1 Yr	−2 Yr	−3 Yr	0 Yr	−1 Yr	−2 Yr	−3 Yr	0 Yr	−1 Yr	−2 Yr	−3 Yr
NFCI	**0.95**	**0.80**	**0.62**	0.48	**0.76**	**0.79**	**0.73**	**0.59**	**0.62**	**0.56**	**0.45**	**0.32**
Post-1984 NFCI	**0.60**	**0.71**	**0.70**	**0.58**	**0.75**	**0.84**	**0.83**	**0.67**	**0.70**	**0.73**	**0.67**	0.51
39-Variable NFCI	**0.34**	**0.28**	**0.33**	**0.37**	0.51	**0.40**	**0.37**	**0.46**	**0.78**	**0.66**	**0.56**	**0.59**
ANFCI	**0.73**	**0.74**	**0.69**	**0.53**	**0.60**	**0.71**	**0.74**	**0.58**	**0.62**	**0.63**	**0.60**	**0.46**
Post-1984 ANFCI	**0.63**	**0.73**	**0.76**	**0.64**	**0.60**	**0.73**	**0.76**	**0.64**	**0.68**	**0.70**	**0.62**	0.47
STLFSI	**0.86**	**0.63**	**0.43**	**0.19**	—	—	—	—	—	—	—	—
KCFSI	**0.93**	**0.76**	0.56	**0.31**	—	—	—	—	—	—	—	—
IMF	**0.89**	**0.77**	0.58	**0.28**	—	—	—	—	—	—	—	—
Credit-to-GDP	0.59	0.54	0.48	0.40	**0.76**	**0.81**	**0.83**	**0.76**	0.75	**0.81**	**0.82**	**0.74**
Non-Financial Leverage	**0.68**	**0.73**	**0.72**	**0.63**	0.50	**0.70**	**0.83**	**0.87**	**0.58**	**0.73**	**0.86**	**0.83**
Bank Conditions	**0.69**	**0.63**	**0.55**	**0.46**	**0.66**	**0.77**	**0.79**	**0.68**	**0.65**	**0.76**	**0.77**	**0.59**
Adj. Bank Conditions	**0.64**	**0.70**	**0.73**	**0.63**	**0.58**	**0.72**	**0.77**	**0.65**	**0.62**	**0.76**	**0.78**	**0.60**

Notes: The table displays the area under the ROC curve relative to the crisis periods denoted in the column headings contemporaneously and at leads of up to three years for the NFCI and its three sub-indexes. Areas significant from 50 percent are denoted by bold (95 percent significance level) and darkened fonts (90 percent significance level).

Table 4 also computes the AUROC for several competing indexes of financial stress produced by the Federal Reserve Banks of Kansas City and St. Louis—the KCFSI and STLFSI, respectively—and the monthly variant of Hatzius et al. (2010) produced by the International Monetary Fund (IMF) and described in Matheson (2011).[10] Of the three, the IMF measure comes the closest to demonstrating the breadth of coverage of the financial system in the NFCI. In contrast, the KCFSI and STLFSI contain a smaller number of primarily what we term in the NFCI as measures of risk.

All three competing indexes have much shorter histories than the NFCI, each beginning in the 1990s, so that AUROC values in the table are omitted for the alternative crisis dating conventions given their diminished relevance. In addition, other than the STLFSI, the remaining indexes are observed at a monthly frequency. Even after adjusting for these facts when comparing the AUROC values, the NFCI remains just as predictive, if not more predictive, than any of the three competing indexes at a contemporaneous horizon.[11] While not apparent in the table, this is also true at leading horizons when judging the NFCI relative to the others solely on their shared history. All of these indexes were very low in advance of the most recent crisis, which is reflected in an AUROC value well below 0.5 at a three-year-ahead horizon.

4.2 Credit-to-GDP-Based Measures

At the bottom of table 4 are the AUROC values for the HP-filtered private credit-to-GDP measure put forth as a leading indicator of financial stress by Drehmann et al. (2010).[12] While the NFCI matches its predictive ability at short horizons, its appeal can be seen in its very high and statistically significant AUROC values at leads from two to three years using either of the alternative dating

[10] We thank Troy Matheson for kindly making available a time series of his index.

[11] The same can be said of the quarterly index of Hatzius et al. (2010).

[12] We measure private credit as the sum of household mortgage and consumer credit market debt combined with the credit market debt of non-financial businesses. For the smoothing parameter of the HP filter, we use the preferred value of lambda in Drehmann et al. (2010).

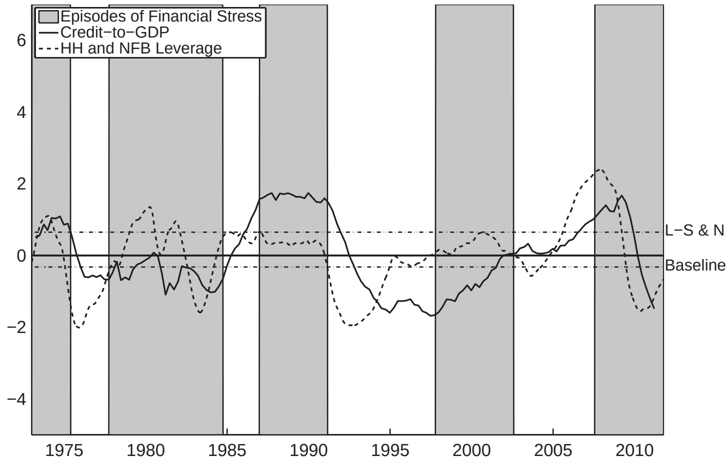
conventions we consider. Similarly, while the post-1984 NFCI performs just as well up to a lead of two years with the López-Salido and Nelson (2010) crisis dates, credit-to-GDP dwarfs it when considering the Laeven and Valencia (2010) dates.

These results are, however, very sample dependent. For instance, both indexes are superior to private credit-to-GDP at all horizons under our baseline crisis dates. To see why, consider figure 8, which plots the credit-to-GDP measure, shading the periods corresponding with our baseline crisis dates. The signal sent by this indicator is very strong for three of the five crisis episodes, i.e., the 1973–75, 1987–91, and most recent crises. It has very little to say about the 1977–84 and 1997–2002 periods we consider, hence the lower AUROC value. In contrast, the alternative crisis dates omit much of these periods, resulting in a higher AUROC value.

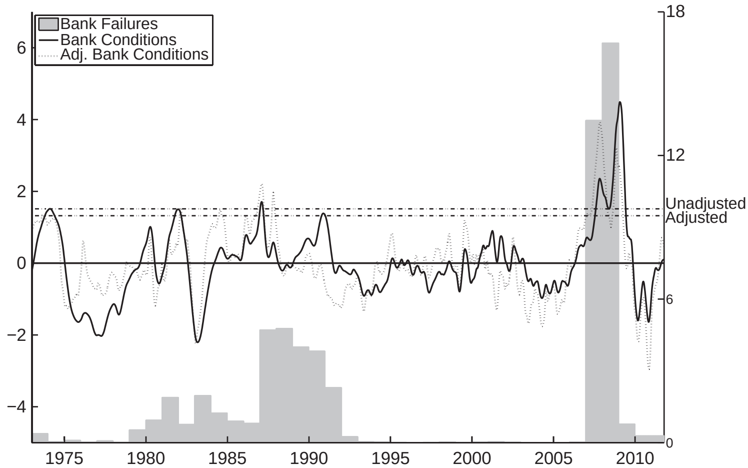
For comparison, consider the sub-index constructed from the two leverage indicators in the NFCI that most closely resemble the inputs to the private credit-to-GDP ratio of Drehmann et al. (2010), referenced as HH and NFB Leverage in the top panel of figure 8 and Non-Financial Leverage in table 4. The signal that emerges offers a picture over time that is more consistent than private credit-to-GDP. It captures all five crisis periods while still producing AUROC values at leading horizons that are on par with private credit-to-GDP using our alternative crisis dates. In addition, it leads all three of the most severe López-Salido and Nelson (2010) crises, with by far its strongest signal occurring in the run-up to the recent crisis.

There are several reasons why our measure differs so markedly from that of Drehmann et al. (2010) despite the use of similar data. The first is that it is unnecessary for us to take a stance on the nature of the trends in household and non-financial business leverage, as we focus on their growth rates. Little is lost in doing so, as it is already apparent from table 2 that the growth rate of non-financial business debt relative to GDP is a highly predictive leading indicator of financial stress. Our measure is also more flexible in that it weights household and non-financial business leverage unequally. Household leverage receives roughly 1.3 times the weight that non-financial business leverage does, as opposed to its equal treatment in the private credit-to-GDP ratio. The weighting we use for our measure based on the systemic decomposition of the NFCI turns

Figure 8. Leading Indicators of Financial Stress



Crisis Dating Convention	Baseline	L-S & N
Two-Year-Ahead Threshold	–0.32	0.65



Crisis Dating Convention	Unadj.	Adj.
Two-Year-Ahead Threshold	1.51	1.32

Note: "Bank Failures" refers to the annual percentage of failed/assisted institutional assets according to FDIC statistics relative to the total assets of the commercial banking sector taken from the Federal Reserve's H.8 statistical release.

out to be non-trivial, as there are several instances in time where the movements in the two measures are conflicting.

In addition, we measure household leverage as household mortgage and consumer debt relative to residential investment and personal consumption of durable goods rather than to overall GDP. This also turns out to be non-trivial because, expressed in this way, household leverage receives a positive weight in the NFCI. It thereby serves to amplify the signal from non-financial business leverage, whereas expressed relative to GDP it has the opposite effect. The flexibility thus provided by our measure makes it possible to produce a combined signal that is superior to that of each of the individual measures, particulary at leading horizons.

The positive weight assigned to each of these measures in the NFCI reflects the fact that rising household and non-financial business leverage are typically associated with increasingly tighter financial conditions. This is in contrast to many of the measures of financial leverage in the NFCI which instead show the opposite correlation with financial conditions. It also makes them characteristic of the financial accelerator mechanism of Bernanke, Gertler, and Gilchrist (1999). Increasingly tighter financial conditions are associated with rising risk premia and declining asset values. Hence, the net worth of households and firms is reduced at the same time that credit tightens. This leads to a period of deleveraging and ultimately to lower economic activity, making these measures relevant for both credit and business cycles.

It is also possible to produce meaningful thresholds at a leading horizon for our measure based on our baseline and the López-Salido and Nelson (2010) crisis dates. We plot these thresholds in the top panel of figure 8 for a leading horizon of two years. Instances of where this measure falls above both thresholds are characteristic of the most severe crises. It is instructive that these coincide with the three deepest and longest U.S. recessions in the post-War era. Instances in between are consistent with more moderate crises and less pronounced recessions. This suggests that policies which aim to minimize the deviation of measures of non-financial leverage from their long-run averages are likely to fall within the realm of both financial stability and the welfare costs of business cycles.

The longer leading horizon considered here is important given that implementation of any such policy would require sufficient lead time. As an example, consider again the most recent crisis. Our measure was at its historical average as recently as the end of 2004. It first crossed the upper threshold in late 2005, nearly two years before the onset of the crisis in August 2007. That said, revision error is likely to be problematic for inference, because the underlying data are subject to large revisions and lags in availability. However, our method of constructing this measure using the architecture for the NFCI should help to minimize these concerns. This is because we can project our measure forward in time based on the higher frequency availability of the NFCI, thereby disciplining its near-term movements by its correlations with the other indicators in the index. In this way, the NFCI still plays an important role in monitoring financial stability in real time.

4.3 Bank Conditions Measures

It is interesting that our measure of non-financial leverage also remains a superior leading indicator to private credit-to-GDP even when using the Laeven and Valencia (2010) crisis dates. These dates were designed to capture periods of systemic bank failure in the United States, i.e., primarily the savings and loan crisis of the late 1980s and the recent crisis. Based on this fact, it is easy to see why in figure 8 private credit-to-GDP was proposed by Drehmann et al. (2010) as a useful leading indicator of systemic risk, as its two highest peaks precede both crises. One might, however, question the magnitude of the signals with the movements in the 1980s more persistent and just as pronounced as the recent crisis. Our measure of non-financial leverage does not feature these concerns.

To assess the robustness of our measure, we also constructed sub-indexes from a broader set of NFCI indicators considered by Laeven and Valencia (2008) as indicators of systemic risk. The indicators included those in our non-financial leverage measure above as well as public debt-to-GDP, bank liabilities and conditions data, and asset prices. Laeven and Valencia (2008) also suggest using measures of economic conditions to assess systemic risk. To achieve this, we constructed sub-indexes of these measures using both the NFCI and ANFCI methodologies, the latter of which adjusts the indicators for growth in economic activity and inflation, while the former does not.

The bottom panel of figure 8 depicts both sub-indexes as well as the relative size of the assets of failed/assisted banking institutions. We include the latter to highlight the magnitudes of the two systemic banking crises. Both sub-indexes quite clearly capture the recent crisis as the worst period for bank conditions in our sample. Other spikes in bank failures are also captured in the peaks of the indexes, although few other than the peaks preceding the savings and loan crisis breach the unbiased two-year-ahead thresholds shown in the figure for both sub-indexes.

Looking again at table 4, it is clear that the difference in predictive ability at shorter forecast horizons between the two sub-indexes (referred to as Bank Conditions and Adj. Bank Conditions) is rather small, as the AUROC values are very similar across all three crisis dating conventions. At longer forecast horizons, the adjusted version is superior only in the case of our baseline crisis dates, and this is also the only instance where either of the two indexes produces a leading signal comparable to that of our measure of non-financial leverage. In this sense, there seems to be little additional value to be had from combining non-financial leverage with other indicators to assess the risk of systemic bank failure.

5. Conclusion

Our analysis suggests that the National Financial Conditions Index (NFCI) produced by the Federal Reserve Bank of Chicago is a highly robust and accurate indicator of financial stress at leading horizons of up to one year. Furthermore, breaking the index into its subcomponents of risk, credit, and leverage can enhance the nature of the signal provided by the NFCI as to the severity of the crisis, with leverage playing a crucial role in signaling financial imbalances as in Adrian and Shin (2010). Moving to a multi-factor representation of the NFCI, which makes full use of the different signals in its three types of indicators and their dynamic properties, is a non-trivial extension which we leave to future research.

At forecast horizons beyond one year, we show that a particular combination of household and non-financial business leverage measures proves to be a consistent leading indicator of financial stress and its impact on economic activity. This result is by virtue of its resemblance to the financial accelerator mechanism of Bernanke, Gertler, and Gilchrist (1999). Increases over time in these measures are associated with increasingly tighter financial conditions, which lead to periods of deleveraging and ultimately lower economic activity. By combining their signal into a single measure based on their relative weighting in the NFCI, we arrive at a combined signal that is superior to several alternative measures of systemic risk suggested in the literature.

The flexibility of the ROC methodology is its greatest asset. However, it leaves unaddressed the types of policy actions that can and should be used to address any financial imbalances it identifies, as well as the magnitude of their likely impact. At the very least, though, our analysis establishes properties that could be explored in the formulation of optimal policy within a general equilibrium framework.[13] In particular, policies which aim to minimize the deviations of financial and non-financial leverage from their long-run averages are consistent with our analysis of financial stability and may also be beneficial for minimizing the welfare costs of business cycles.

Appendix. A Dynamic Factor Model of Financial Conditions

The model for the National Financial Conditions Index is fundamentally similar to many of the dynamic factor models summarized in Stock and Watson (2011). Its state-space representation is shown below, where y_t is a vector of stationary variables that have been demeaned and standardized; ε_t and η_t are idiosyncratic error vectors with $\varepsilon_t \sim N(0, H)$ and $\eta_t \sim N(0, Q)$; α_t is a vector made up of a latent coincident factor f_t and its $L^\star - 1$ lags; and $t = 1, \ldots, \hat{T}$, where $\hat{T}$ is the longest time-series length of the collection of $\hat{N}$ financial indicators in the NFCI.

$$y_t = Z\alpha_t + \varepsilon_t \tag{6}$$

$$\alpha_{t+1} = T\alpha_t + R\eta_t,$$

$$a_1 = \mathbb{E}[\alpha_1] \text{ and } P_1 = Var[\alpha_1] \text{ given.} \tag{7}$$

[13]For instance, Brave and Genay (2011) provide some evidence that during the recent crisis, traditional and non-traditional monetary policy actions contributed to the recovery in financial conditions as measured by the NFCI and ANFCI in a joint model of Federal Reserve policy and financial conditions.

Written in this form, the latent factor is identified based on both the historical cross-correlations of the vector of variables y_t and its own historical autocorrelations embedded in the system matrix T in the state equation, (7). In contrast, the variable cross-correlations enter via the model's observation equation, (6), relating each of the $\hat{N}$ variables to the coincident latent factor via the factor loadings λ included in the system matrix Z. Identification is achieved only up to scale, as initial conditions for the mean and variance of the latent factor—a_1 and P_1, respectively—are necessary to complete the model.

This model can be estimated using the procedure outlined in Doz, Giannone, and Reichlin (2006) based on the EM algorithms of Shumway and Stoffer (1982) and Watson and Engle (1983) in order to obtain quasi-maximum-likelihood estimates of the system matrices and subsequently the latent factor. In general, it requires one pass through the Kalman filter and smoother, and then reestimation of the system matrices—Z, T, H, and Q—using ordinary least squares (OLS) estimation at each iteration.[14] The resulting sequence of log-likelihood function valuations is non-decreasing, and convergence of the algorithm is governed by its stability according to the suggested procedure in Doz, Giannone, and Reichlin (2006).

A unique characteristic of y_t in the context of the NFCI is that it contains series of varying reported frequencies and series that start and end at different times within the sample. The EM algorithm proves advantageous in this setting because it allows for a complete characterization of the data-generating process using incomplete data. However, applying this procedure to our particular data set requires some changes to the standard Kalman filter and smoother equations as described below.

Missing-Value Kalman Filter

Due to the irregular observation of the data in our framework, two extensions to the standard Kalman recursion equations need to first be made before running the EM algorithm. The first alteration involves setting up the Kalman filter to deal with missing values as

[14] A small alteration in the least-squares step is required to account for the fact that the unobserved components of the model must first be estimated.

discussed by Durbin and Koopman (2001). As one moves through time, the vector of observables for the NFCI, y_t, changes size from period to period. Consequently, it is necessary to accommodate the partially observed vector $y_t^\star$ following Durbin and Koopman (2001) by using the known matrix W_t whose rows are a subset of the rows of $I(\hat{N})$ such that $y_t^\star = W_t y_t$ to alter the system matrices at each point in time. Taking this W_t as given, the system matrices Z and H are replaced with $Z^\star = W_t Z$ and $H^\star = W_t H W_t'$, respectively. Substituting these matrices into the standard Kalman filter and smoother equations allows one to proceed as usual through the recursive equations.

Temporal Aggregation and the Harvey Accumulator

The second necessary modification involves including additional state variables that evolve deterministically to properly adjust for the varying temporal aggregation properties of the mixed-frequency data used to create the NFCI. By applying the accumulator of Harvey (1989), one can manage this data irregularity with relative ease. The goal of the accumulator is to augment the state, α_t, with a deterministically evolving indicator that is a summary of all past values of the unobserved factor aggregated in such a way as to correspond with the nature of the observed data.

More specifically, variables viewed as a “stock,” or a *snapshot* in time, will not need such aggregation of past realizations of the factor. Variables that correspond to sums or averages over the higher base frequency of the factor will need to accumulate the factor realizations over a defined period in order to properly account for the contemporaneous factor’s contribution to what is being observed. Any “stocks” that are differenced can be interpreted as sum variables and treated as such.

The NFCI includes variables that resemble both “sums” and “averages,” in addition to variables that are first-differenced at lower frequencies than the weekly (base) frequency. Combining this with monthly and quarterly frequencies of observation, the model for the NFCI requires three Harvey accumulators in the state, one each for (i) monthly averages, (ii) monthly sums, and (iii) quarterly sums.

Sum-Variables Accumulator

For both monthly and quarterly sums, we follow Aruoba, Diebold, and Scotti (2009)'s implementation of the Harvey (1989) accumulator. The accumulators for sum variables are denoted by S_t. By construction, any sum accumulator should represent the sum of all of the factor realizations that have occurred within the current period of the lower frequency of observation. Additionally, the accumulator is defined recursively so as to be included in the state-space equations (6)–(7). Analytically, the sum accumulator evolves each period according to the following equation:

$$S_{t+1} = s_t S_t + f_{t+1},$$

where s_t is a calendar-determined indicator that evolves according to

$$s_t = \begin{cases} 0 & \text{if } t \text{ is the last period (base frequency)} \\ & \quad \text{within the lower frequency} \\ 1 & \text{otherwise.} \end{cases}$$

For notational purposes, it is assumed in what follows that f_{t+1} is an AR(1) process defined by $f_{t+1} = \rho f_t + R\eta_t$. Incorporating this representation of the accumulator into the state-space model follows from a simple substitution of the contemporaneous factor as outlined by Aruoba, Diebold, and Scotti (2009).

Average-Variables Accumulator

We denote the desired accumulator for the average variables with M_t and derive it as though we are aggregating from a weekly base frequency to monthly observations.[15] By construction, this accumulator should represent the *current* average of all of the factor realizations (occurring every week) that have occurred within the current month (frequency that is being observed) and be defined

[15] All methods outlined in this section generalize fully to any particular combination of base and observation frequencies that one might encounter, with the only necessary modifications occurring in the evolution of the calendar indicator m_t or s_t.

recursively for seamless addition to the state-space equations (6)–(7). Analytically, the average accumulator evolves each period by the following equation:

$$M_{t+1} = \frac{(m_t - 1)}{m_t} M_t + \frac{1}{m_t} f_{t+1},$$

where m_t is a calendar-determined indicator that evolves:

$$m_t = \begin{cases} 1 & \text{if } t \text{ is the } \textit{last} \text{ week of the month} \\ 2 & \text{if } t \text{ is the first week of the month} \\ 3 & \text{if } t \text{ is the second week of the month} \\ \textit{etc.} & \end{cases}$$

Explicitly including the accumulator in the state requires augmenting the state and some substitution. The resulting formulation is given by

$$\begin{bmatrix} f_{t+1} \\ M_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ \frac{\rho}{m_t} & \frac{m_t - 1}{m_t} \end{bmatrix} \begin{bmatrix} f_t \\ M_t \end{bmatrix} + \begin{bmatrix} R \\ \frac{R}{m_t} \end{bmatrix} \eta_t.$$

Building the State

This section gives a more detailed explanation of how to build the system matrices of the state space given the two necessary extensions to accommodate the irregular observation of the NFCI's observables described above. The latent factor is assumed to have some finite-order dynamics $L^\star$, given by the vector of coefficients, ρ. Augmenting the state to include $L^\star - 1$ lags of f_t yields the following state equation, with $\alpha_t = [f_{t-i}]$ for $i = 0, \ldots, L^\star - 1$:

$$\alpha_{t+1} = \begin{bmatrix} \rho & 0 \\ I(L^\star - 1) & 0_{L^\star - 1 \times 1} \end{bmatrix} \alpha_t + \begin{bmatrix} 1 \\ 0_{L^\star - 1 \times 1} \end{bmatrix} \eta_t.$$

Now, in this particular representation of the state, the system matrix $\tilde{T}$ is defined by

$$\tilde{T} = \begin{bmatrix} \rho & 0 \\ I(L^\star - 1) & 0_{L^\star - 1 \times 1} \end{bmatrix}.$$

One must also augment the state (currently a $L^\star$ long vector) by the additional states needed for each of the accumulators derived in the previous section to yield the state equation (taking the ρ dynamics again as given):

$$\begin{bmatrix} \alpha_{t+1} \\ M_{t+1} \\ S_{t+1} \end{bmatrix} = \begin{bmatrix} \tilde{T} & 0 & 0 & 0 \\ \frac{\rho}{m_t} & \frac{m_t - 1}{m_t} & 0 & 0 \\ \rho & 0 & s_t & 0 \end{bmatrix} \begin{bmatrix} \alpha_t \\ M_t \\ S_t \end{bmatrix} + \begin{bmatrix} 1 \\ 0_{L^\star - 1 \times 1} \\ \frac{1}{m_t} \\ 1 \end{bmatrix} \eta_t. \tag{8}$$

It should be noted that, as written above, the $\tilde{T}$ *within* the general transition system matrix, T, here is time invariant, and subsequently the dynamics being estimated (essentially a reestimation of ρ) at each iteration of the EM algorithm are from a *time-invariant* system. However, our (accumulator-augmented) state transition system matrix (as well as the coefficient matrix on the η_t) *does vary over time* due to the different number of weeks in a given month, or quarter, and the particular evolution of the average accumulators.

Moving to the measurement equation, assume that a priori the vector of factor loadings λ is known. Then, taking the state equation (8) as given, the Z measurement system matrix is simply an $\hat{N}$ by $L^\star + 3$ matrix, where each row has the particular loading (the particular element of λ) in either the first column (if it corresponds with a weekly or stock variable) or one of the last three columns (corresponding to one of the three accumulators, i.e., monthly average, monthly sum, or quarterly sum) and zeros everywhere else. Finally, the system matrices H and Q are the standard variance-covariance matrices of ε_t and η_t, respectively, as described in Doz, Giannone, and Reichlin (2006), where H is assumed to be a diagonal matrix and Q normalizes the scale for the latent factor.

Kalman Filter and Smoother Recursive Equations

Adopting the notation of Durbin and Koopman (2001), the Kalman filter and smoother equations taking into account the state-space structure above are as follows.[16] With a_1 and P_1 given, the filter equations are

[16] For more details on the derivation of these equations, see Durbin and Koopman (2001, pp. 64–73).

$$v_t = y_t^{\star} - Z^{\star} a_t \qquad F_t = Z^{\star} P_t Z^{\star\prime} + H^{\star}$$
$$K_t = T_t P_t Z^{\star\prime} F_t^{-1} \qquad L_t = T_t - K_t Z^{\star}$$
$$a_{t+1} = T_t a_t + K_t v_t \qquad P_{t+1} = T_t P_t L_t' + R_t Q R_t',$$

where $y_t^{\star}$, $Z^{\star}$, and $H^{\star}$ are the truncated versions of the more general vector and system matrices due to any potential missing observations in the vector y_t as described above.

Likewise, the equations for the backwards smoother are given by[17]

$$r_{t-1} = Z^{\star\prime} F_t^{-1} v_t + L_t' r_t \qquad N_{t-1} = Z^{\star\prime} F_t^{-1} Z^{\star} + L_t' N_t L_t$$
$$\tilde{\alpha}_t = a_t + P_t r_{t-1} \qquad V_t = P_t - P_t N_{t-1} P_t$$
$$J_t = P_{t-1} L_{t-1}' (I - N_{t-1} P_t),$$

with $r_{\hat{T}} = 0$, and $N_{\hat{T}} = 0$.

The EM Algorithm

Given the missing observation and accumulator extensions to the Kalman filter, the system parameters can be estimated via the EM algorithm of Shumway and Stoffer (1982) and Watson and Engle (1983). Starting at some initial values for every system matrix—Z^0, H^0, T^0, Q^0, P_1^0, and a_1^0—each iteration of the algorithm consists of one pass through the Kalman filter and smoother using the system matrices estimated at the previous iteration r—Z^r, H^r, T^r, Q^r, P_1^r, and vector a_1^r. The typical critique of the algorithm, its slow convergence rate, is not problematic in this setting due to the size of the time-series and cross-section dimensions for the NFCI, which allow for consistent initial values using the iterative PCA techniques described in Stock and Watson (2002).

The initial values of the Z system matrix Z^0 include the estimated factor loadings λ, while the initial guess of the H system matrix H^0 is simply the estimated variances of the idiosyncratic errors from the same analysis. The initial guess of the T system

[17] It should be noted that the additional matrix, J_t, is being calculated so that the maximization step in the EM algorithm can take into account the uncertainty in the estimation of the state.

matrix T^0 is given by standard OLS techniques on the initial estimate of the factor F_t^0 with a chosen lag length of fifteen weeks, or roughly one quarter's worth of observations, determined based on the BIC. The initial guess of a_1^0 is set to zero and the initial guess of P_1^0 is set at some reasonable baseline value. Furthermore, we discard two years worth of estimated data to avoid any issues with this initialization.

The lack of identification that is common to these models requires that we restrict the scale of either the factor loadings or the factor. We use the normalization of Doz, Giannone, and Reichlin (2006) and restrict the variance of the state disturbances to be 1 to set the scale of the factor. By utilizing both the smoothed estimates and their covariance matrices, one can update the expectation of the conditional log-likelihood function, the (E) step. A concise version of the log-likelihood, and the one that can be computed at each iteration, is as follows:

$$\log L = -\frac{1}{2}\sum_{t=1}^{\hat{T}} \left(\log |F_t| + v_t' F_t^{-1} v_t \right), \tag{9}$$

where both v_t and F_t are given by the forward recursion equations of the Kalman filter. Then, using OLS techniques, the system matrices are reestimated as before, the (M) step. This reestimation yields new values for the system matrices—Z^{r+1}, H^{r+1}, T^{r+1}, Q^{r+1}, P_1^{r+1}, and vector a_1^{r+1}. Repeating this process yields a non-decreasing sequence of log-likelihood values. With the (E) and (M) steps completely defined, one can iterate between the two until (9) becomes stable.[18]

Constructing Sub-Indexes

To construct sub-indexes of the NFCI, post-estimation we replace the elements of λ corresponding to variables not in the sub-index with zeros. We then make an additional pass through the equations

[18] As a convergence criterion, we used $|\log L(r) - \log L(r-1)/((\log L(r) + \log L(r-1))/2)| < 10^{-6}$. With this criterion, the EM algorithm converges rather quickly, i.e., generally within 150 iterations.

of the Kalman filter and smoother with the system matrices estimated from the last (M) step of the EM algorithm for the NFCI. In this fashion, we hold fixed both the relevant factor loadings and dynamics of the latent factor from the fully estimated model. As with the NFCI, the scale of each sub-index is normalized according to the variance-covariance matrix Q. This method is equivalent to the two-step consistent estimation procedure as described in Doz, Giannone, and Reichlin (2006).

References

Adrian, T., and H. S. Shin. 2010. "Liquidity and Leverage." *Journal of Financial Intermediation* 19 (3): 418–37.

Aruoba, S. B., F. X. Diebold, and C. Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business and Economic Statistics* 27 (4): 417–27.

Baker, S. G., and B. S. Kramer. 2007. "Pierce, Youden, and Receiver Operating Characteristic Curves." *The American Statistician* 61 (4): 343–46.

Berge, T. J., and O. Jordà. 2011. "Evaluating the Classification of Economic Activity into Recessions and Expansions." *American Economic Journal: Macroeconomics* 3 (2): 246–77.

Bernanke, B. S., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*, ed. J. B. Taylor. Elsevier.

Brave, S., and R. A. Butters. 2011. "Monitoring Financial Stability: A Financial Conditions Index Approach." *Economic Perspectives* (Federal Reserve Bank of Chicago) 35 (Q1): 22–43.

Brave, S., and H. Genay. 2011. "Federal Reserve Policies and Financial Market Conditions during the Crisis." Federal Reserve Bank of Chicago Working Paper No. 2011-04.

DeLong, E., D. DeLong, and D. Clarke-Pearson. 1988. "Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach." *Biometrics* 44: 837–45.

Doz, C., D. Giannone, and L. Reichlin. 2006. "A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models." ECB Working Paper No. 674.

Drehmann, M., C. Borio, L. Gambacorta, G. Jimenez, and C. Trucharte. 2010. "Countercyclical Capital Buffers: Exploring Options." BIS Working Paper No. 317.

Durbin, J., and S. J. Koopman. 2001. *Time Series Analysis by State Space Methods.* Oxford University Press.

FDIC (Federal Deposit Insurance Corporation). 1984. *The First Fifty Years: A History of the FDIC 1933–1983.* Federal Deposit Insurance Corporation.

———. 1997. *History of the Eighties — Lessons for the Future.* Federal Deposit Insurance Corporation.

Hakkio, C. S., and W. R. Keeton. 2009. "Financial Stress: What Is It, How Can It Be Measured, and Why Does It Matter?" *Economic Review* (Federal Reserve Bank of Kansas City) (Second Quarter): 5–50.

Harvey, A. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press.

Hatzius, J., P. Hooper, F. Mishkin, K. L. Schoenholtz, and M. W. Watston. 2010. "Financial Condition Indexes: A Fresh Look After the Financial Crisis." U.S. Monetary Policy Forum Paper Series.

Illing, M., and Y. Liu. 2006. "Measuring Financial Stress in a Developed Country: An Application to Canada." *Journal of Financial Stability* 2 (3): 243–65.

Laeven, L., and F. Valencia. 2008. "Systemic Banking Crises: A New Database." IMF Working Paper No. 08/224.

———. 2010. "Resolution of Banking Crises: The Good, the Bad, and the Ugly." IMF Working Paper No. 10/146.

López-Salido, D., and E. Nelson. 2010. "Postwar Financial Crises and Economic Recoveries in the United States." Mimeo, Board of Governors of the Federal Reserve System.

Matheson, T. 2011. "Financial Conditions Indexes for the United States and Euro Area." IMF Working Paper No. 11/93.

Nelson, W. R., and R. Perli. 2007. "Selected Indicators of Financial Stability." *IFC Bulletin* (Irving Fisher Committee on Central-Bank Statistics) 23 (October): 92–105.

Reinhart, C. M., and K. S. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Crises.* Princeton, NJ: Princeton University Press.

Schreft, S. L. 1990. "Credit Controls: 1980." *Economic Review* (Federal Reserve Bank of Richmond) (Nov/Dec): 25–55.

Shumway, R. H., and D. S. Stoffer. 1982. "An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm." *Journal of Time Series Analysis* 3 (4): 253–64.

Spero, J. E. 1999. *The Failure of the Franklin National Bank: Challenge to the International Banking System.* Beard Books.

Stock, J. H., and M. W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97 (460): 1167–79.

———. 2011. "Dynamic Factor Models." In *Oxford Handbook of Economic Forecasting*, ed. M. P. Clements and D. E. Hardy, 35–60. Oxford University Press.

Watson, M. W., and R. F. Engle. 1983. "Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models." *Journal of Econometrics* 23 (3): 385–400.

A Review of Allan Meltzer's *A History of the Federal Reserve, Volume 2*[*]

Edward Nelson
Federal Reserve Board

This paper reviews Allan H. Meltzer's *A History of the Federal Reserve, Volume 2*. This two-book volume covers Federal Reserve policies from 1951 to 1986. The book represents an enormous achievement in synthesizing a great amount of archival information into a historical account grounded on economic analysis. At the same time, Meltzer's interpretation of specific eras is open to question. He does not appear to acknowledge adequately the degree to which 1950s monetary policy decisions had a solid analytical foundation. Furthermore, Meltzer's account of the shift from the 1970s inflation to the 1980s disinflation implausibly stresses a shift in policymakers' objective function. The crucial change over this period, both in the United States and other countries, is more likely to have been policymakers' improved grasp of the connections between monetary policy and inflation. The review also takes issue with Meltzer's account, in his book's epilogue, of the financial crisis from 2007 to 2009. In this epilogue, Meltzer understates the degree to which the Federal Reserve's reaction to the financial crisis was in line with the historical practice of the Federal Reserve and other central banks.

JEL Codes: E52, E58.

[*]For helpful discussions on the subject matter of this paper and comments on earlier drafts, the author is grateful to Richard Anderson, James McAndrews, and David Romer. The paper has benefited substantially from three referee reports and the comments of the co-editor (Carl Walsh). The author thanks Christine Garnier and Andrew Giffin for research assistance. The views expressed in this paper are those of the author alone and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System or its staff. Author e-mail: Edward.Nelson@frb.gov.

1. Introduction

In his review of Friedman and Schwartz (1963), Tobin (1965, p. 483) observed, "Now, thanks to the constructive action of the present Board of Governors in making available past minutes of the Board and the Open Market Committee, scholars will be able to provide a more definitive history of Federal Reserve policymaking." Volume 2 of Allan Meltzer's *A History of the Federal Reserve* (Meltzer, 2009a, 2009b) completes the most ambitious and wide-ranging attempt at an archives-based history of the type that Tobin envisaged. Many studies in the past few decades have intensively used historical FOMC minutes and transcripts as well as other items on the public record. Such material has been used in the monetary policy literature to shed light on specific eras of policy, to construct estimates of policy stance, or—something that is perhaps closest to Meltzer's aim—to bring out the objectives and thinking underlying Federal Reserve decisions.[1] Compared with these studies, Meltzer's has the advantage of length—volume 2 of the *History* totals 1,312 pages—as well as an extensive base of unreleased material, such as Federal Reserve Board staff memoranda and correspondence. Meltzer also draws on records related to monetary policy from U.S. presidential libraries and other sources.

Most of the years of Federal Reserve history covered by Friedman and Schwartz (1963)—that is, the years up to 1960—fell into the period covered in Meltzer's volume 1. Meltzer's volume 2 covers the period from 1951 to 1986. In contrast to Hetzel's (2008) recent history, Meltzer does not cover the Greenspan era. In the course of his account of 1951–86 developments, however, Meltzer has occasion to make many comments on the Greenspan era. In addition, an epilogue to the *History* discusses the financial crisis that began in 2007.

Volume 2 actually comprises *two* books, which are being retailed separately, and whose individual subtitles are "Volume 2, Book 1" and "Volume 2, Book 2." The individual books cover 1951 to 1969,

[1]See, for example, Feinman and Poole (1989), Karamouzis and Lombra (1989), Romer and Romer (1989, 2002a, 2002b, 2004a, 2004b), Hetzel (1998, 2008), Orphanides (2003), and Lindsey, Orphanides, and Rasche (2005).

and 1970 to 1986, respectively. Prospective readers who are interested solely in the 1951–69 period, and who are therefore inclined only to purchase book 1, should be warned that book 1 does not have its own index or bibliography. Instead, book 2's index and bibliography cover the whole of volume 2. This arrangement underscores the fact that the two books are intended to be regarded as a single work, and they are reviewed as such in the present paper.

It is worth bringing to the fore a theme, running throughout Meltzer's account, with which the present author will strongly take issue. Meltzer argues that the enduring change in the late 1970s in Federal Reserve policymaking was a shift in policymakers' objectives toward a greater weight on inflation stabilization relative to stabilization of real activity. It is this shift in the weights on policymaker objectives that Meltzer sees as the key to the early 1980s disinflation and subsequent Great Moderation. In Meltzer's view, this reweighting of objectives removed an inclination on the Federal Reserve's part to overreact to unemployment, an inclination that he believes lies behind the Great Inflation of the 1970s. The explanation offered by Meltzer downplays theoretical misconceptions as an explanation for poor monetary policy performance during the 1970s. Indeed, Meltzer's explanation actually *requires* considerable knowledge on Federal Open Market Committee (FOMC) members' part of the connections between inflation and monetary policy actions, because it presumes that the FOMC knowingly acquiesced to high inflation.

An alternative position, which the present author favors, is that changes in policymaker objectives are not an important component of the explanation of post-war monetary policy developments.[2] Rather, what distinguished policymakers' attitudes during the 1970s from those in other post-war decades was the fact that the 1970s Federal Reserve subscribed to a non-monetary view of inflation. In turn, a change in Federal Reserve doctrine in favor of a monetary view of inflation, rather than a change in the weight on inflation in policymakers' objective function, was what prompted the changes

[2]Constancy of policymakers' objective function also underlies the accounts of post-war Federal Reserve policy provided by Romer and Romer (2002b), Orphanides (2003), and Romer (2005).

in monetary policy behavior in the late 1970s.[3] The international dimension of the Great Inflation, which Meltzer seriously neglects, supports explanations of policy behavior that emphasize changes in doctrine over explanations that emphasize changes in the objective function. This international evidence supplements and reinforces the message from the U.S. documentary record that 1970s monetary policy was misguided primarily because policymakers in that decade succumbed to a non-monetary view of inflation.

This review will also take issue with Meltzer's history on several other matters. Consequently, in order to keep a proper perspective on these criticisms, three points deserve emphasis. First, Meltzer's book covers some thirty-five years of Federal Reserve history, and the reviewer is in essential agreement with the substance of Meltzer's account for fully half of the years covered, namely, the years 1961 to 1969 and from 1979 to 1986. Naturally, much of the review will concentrate on areas of disagreement with Meltzer: specifically, his accounts of the 1950s and the 1970s. Second, on the many episodes in the thirty-five years covered in the book, students of monetary policy will learn much from Meltzer's account, even if they disagree with certain arguments he advances. In particular, we are all in debt to Meltzer for his impressive consolidation of a vast amount of archival material into a digestible narrative account. Third, it is important to recall that Lucas (1988, p. 137) opened his contribution to a Festschrift for Meltzer by noting that "details are the way scholarship is carried out." In that spirit, evaluation of a highly scholarly effort like Meltzer's must involve devoting some space to the nuts and bolts of his study and not just to the main themes. Notwithstanding the disagreement laid out in this paper with both the details and major themes of Meltzer's account, it is undeniable that Meltzer has provided in both volumes of his *History* a goldmine of factual material and an account whose depth of scholarship makes the *History* the crowning achievement in Meltzer's fifty-year body of work on monetary issues.

[3] "The late 1970s" is used here (rather than 1979) because, as discussed in Romer and Romer (2002b) and Nelson (2005a), the monetary view of inflation was gaining ground among U.S. policymakers, and actual monetary policy was tightening considerably, over both 1978 and 1979.

This review proceeds as follows. Section 2 gives an overview of Meltzer's book and then provides some comments on the book's organization and methodology before considering Meltzer's analysis of specific eras. Section 3 disputes Meltzer's treatment of the international character of the Great Inflation. Section 4 considers Meltzer's epilogue on the financial crisis, and section 5 concludes.

2. Overview

This section describes the layout of Meltzer's volume 2 (section 2.1), then discusses some choices made for the organization of the material, and offers some comments on each of the eras that Meltzer studies (section 2.2).

2.1 A Bird's Eye View of Volume 2

Meltzer's volume 2 is laid out as follows. A forty-page chapter 1, "Introduction," sets the stage and sketches some of the themes that will emerge in subsequent chapters. Chapter 2, "A New Beginning, 1951–60," covers the first decade after the Treasury–Federal Reserve Accord reestablished effective central bank independence in the United States. This chapter considers the same decade as that which tail-ended Friedman and Schwartz's (1963) account, but provides greater detail than that in Friedman and Schwartz—as Meltzer's chapter 2 is over 200 pages—and has the benefit of greater hindsight and access to internal documents. Chapter 3 (again over 200 pages) is titled "The Early Keynesian Era: A Low Inflation Interlude, 1961–65," and deals with the shift to expansionary policies in the early 1960s. Chapter 4, "The Great Inflation: Phase I," essentially covers the remainder of the William McChesney Martin period at the Federal Reserve.

Volume 2, book 2, opens with "International Monetary Problems, 1964–71," covering the demise of the Bretton Woods arrangements. Chapter 6, "Under Controls: Camp David and Beyond," covers the first three-and-a-half years (early 1970 to mid-1973) of Arthur Burns's first term as Federal Reserve Chairman. Chapter 7, "Why Monetary Policy Failed Again in the 1970s," covers what Meltzer calls "the least successful period for post-war Federal Reserve policy" (page 843), the later part of 1973 to late 1979, which includes the remainder of Burns's tenure and the William Miller interregnum. "Disinflation" (chapter 8) takes the reader from the nomination in July 1979 of Paul Volcker to the position of Federal Reserve Chairman, to the end of 1982. Chapter 9 is "Restoring Stability, 1983–86." This is followed by a twenty-six-page reflective chapter on both volumes, "Past Problems and Future Opportunities" (chapter 10). The book closes with a fourteen-page epilogue entitled "The United States in the Global Financial Crisis of 2007–9."

2.2 Comments on the Book's Structure

Meltzer (1965) criticized Friedman and Schwartz (1963) for organizing their history without a formal theoretical framework (in the sense of a system of equations in which to frame the analytical discussion). In the course of writing his *History*, Meltzer seems to have become more sympathetic to the Friedman-Schwartz "analytical narrative" approach, as his discussion is deep in economic analysis but does not have an explicit model—for example, one patterned after the model of Brunner and Meltzer (1973)—to guide the discussion. Nevertheless, those familiar with the monetarist literature will recognize the analytical structure underlying Meltzer's discussion; many of the analytical points made by that literature are brought out by Meltzer's discussions of specific episodes and his critical discussion of the Federal Reserve's own analysis. Occasionally, however, Meltzer does not exposit the monetarist framework in the clearest fashion. For example, in referring to "the fungibility of money and credit" (page 506), Meltzer certainly does *not* mean that money and credit are interchangeable with one another—many of Meltzer's theoretical contributions emphasize precisely the opposite point! Rather, what is intended to be conveyed is that different types of money are interchangeable with one another, and that different types of credit are easily substituted for one another. Another instance in which Meltzer's discussion seems contrary to the monetarist literature is his evaluation of the relationship between money growth and variables such as nominal income and prices on the basis of the *contemporaneous* relationship between the series (see pages 196, 1168). One of the major insights of the monetarist literature (acknowledged by Christiano, Eichenbaum, and Evans 2005, p. 5, for example) is the pervasiveness of lags in the relationship between monetary policy actions and economic aggregates. By and large, however, Meltzer's book serves as a good means by which interested readers can acquaint themselves with the major aspects of monetarist analysis.

Meltzer's breakup of chapters reflects shrewd choices. The coverage of the 1950s as a single era is wise and facilitates comparison (undertaken below) with Romer and Romer's (2002a) study of 1950s Federal Reserve policy. The separate coverage given to the late 1960s inflation is to be applauded. This is preferable to treating the second half of the 1960s with the 1970s as a monolithic "Great Inflation." Meltzer's chronological breakdown is appealing because, as argued below, the reasons behind the 1970s monetary policy mistakes are different from those that underlay 1960s policy choices—although, as also contended below, the particular factors that Meltzer stresses in explaining 1970s policy failures do not seem to be the most important factors.[4] It was also wise to give international policy in the 1960s its own chapter. To a considerable extent, the United States in the 1960s was able to call the shots on monetary policy independently of its Bretton Woods obligations and of the U.S. balance-of-payments situation. It was therefore sensible to separate the issues that bear on international policy from those weighing on monetary policy decisions. Later in the book, Meltzer's division of the Volcker years into the 1979–82 and the post-1982 epochs allows the crucial years from 1983 to 1986 to receive the attention they deserve, and gets away from the preoccupation of researchers with the 1979–82 portion of the Volcker era.

Some observations follow concerning Meltzer's account of specific eras of 1951–86 monetary policy.

2.2.1 The 1950s

Meltzer's coverage of the 1950s provides an indispensable account of the first decade of post-Accord policy. The chapter adds substantially to existing factual knowledge about this period, and is

[4]The fact that the two periods of problematic policy arose from different sources of error is underscored by the fact that the bridging period of 1969–70 witnessed the pursuit of a disinflationary policy that was well informed by modern natural-rate theory. See Romer and Romer (2002b) and DiCecio and Nelson (2009) for development of this point.

also persuasive on several analytical matters. Nevertheless, there are grounds for believing that Meltzer understated the sophistication of policymaking in the 1950s. Romer and Romer (2002a, p. 127) argue that "monetary policymakers in the 1950s had figured out the essence of sensible policy," and Romer and Romer (2002b, p. 19) add that "monetary policymakers in the 1950s also had a relatively modern view of the process of disinflation." Meltzer rejects this view. Indeed, he is brusque in his dismissal of the Romer-Romer position, stating in a footnote on page 48, "I find no evidence to support their claim," with a similarly curt remark appearing on page 90. Meltzer's failure to provide a systematic comparison of his evidence for the 1950s with that of Romer and Romer is a shortcoming of his chapter on 1951–60.

A comparison of the rival accounts, taken in conjunction with the present author's research into 1950s monetary policy, suggests that, in several important respects, the evidence favors Romer and Romer over Meltzer. A few specific examples follow.

First, Meltzer (on pages 81 and 87) insists that the Federal Reserve in the 1950s did not understand the distinction between nominal and real interest rates. By contrast, Romer and Romer (2002a, p. 127) contend that "many FOMC members showed a clear understanding of the distinction between real and nominal interest rates." The one example they give, however, is a statement by Karl Bopp, Federal Reserve Bank of Philadelphia president. Meltzer may not attach weight to this example, perhaps requiring evidence that Board governors or Board senior staff understood the nominal rate/real rate distinction before he would concede that knowledge of the distinction was instilled in 1950s policymaking. As it happens, such evidence exists in a 1957 statement by Federal Reserve Chairman William McChesney Martin and in a 1959 statement by Winfield W. Riefler, senior staffer and assistant to the Chairman. Martin (1957, p. 1266) testified, "A large part of the savings of the country is mobilized in savings deposits and similar claims that call for some stated amount of dollars. If people generally come to feel that inflation is inevitable, they will not save in this form unless they are paid a much higher interest premium to compensate them for the depreciation of their saved dollars. It is for this reason that it is impossible, in a period of demand in excess of savings, to maintain lower interest rates through a policy of 'easy credit.'"[5] Likewise, in a speech in July 1959, Riefler (1959, p. 3372) noted that in "a period of price stability," it would not "be necessary to set anything aside to compensate for erosion in the purchasing power of the dollars invested in liquid assets," whereas in "a period of creeping inflation—at, say, a 3 percent annual rate," an interest rate would require a premium of 3 percent "to maintain the real purchasing power of the investment." Riefler continued, "It is this process that our monetary authorities have in mind when they take the position that the current higher levels of interest rates now prevailing are basically the result of a widespread experience with and expectation of inflation, rather than the result of restrictive monetary policies."[6] These 1950s statements by Federal Reserve Board officials clearly acknowledge the Fisher effect.

On a second point, Meltzer claims (on page 942) that Federal Reserve policymakers did not blend preemptive actions into a stabilizing monetary policy reaction function until 1994. Yet evidence that 1950s policy was stabilizing and preemptive is present in Romer and Romer's (2002a, p. 125) finding that 1950s funds rate policy responded strongly to expected future inflation and expected future economic activity. Other interpretations of the 1950s have also perceived a forward-looking dimension to monetary policy in that decade; for example, Tobin (1965, p. 485) observed that Friedman and Schwartz (1963) "praise the shifts to ease that occurred before the cyclical peaks of 1953 and 1960." It also bears mentioning that in 1958, Ralph Young, at the time the director of the Federal Reserve Board's Division of Research and Statistics, stated, "The forward-looking nature of monetary policy forces the central banking

---

[5] The existence of enlightened statements by Martin such as this renders inappropriate Meltzer's conclusion (on page 1222) that "Chairman Martin had no interest in economic or monetary explanations of events." While Martin was probably not as analytical in some respects as other Chairmen, it is surely excessive to claim that Martin had no interest in economic explanations. And while—as Meltzer so skillfully documents—monetary policy went astray for much of Martin's second decade as Chairman, the reasons for this are largely to be found in *misguided* economic analysis rather than in *lack of interest in* economic analysis.

[6] Private financial market participants over this period also referred to the Fisher relation. For example, Alan Greenspan (1959, p. 12) stated, "It is true that *expectations of price increases* tend to add an 'inflation premium' to interest rates." (Emphasis in original)

statistician to forecast, no matter how inadequate the tools for doing so" (Young 1958, p. 205).

Finally, a modern aspect of 1950s policymaking was the resolute orientation of monetary policy on inflation. Meltzer sees this orientation as only a post-1979 phenomenon, observing (on page 1091) of the refusal to ease during 1981, "Nothing like that had happened [previously] in the postwar years." In contrast, Friedman (1984, p. 26) and Romer and Romer (2002a) stress that it was monetary restraint during the second half of the 1950s, with the uncomfortable interim consequence of two recessions, that delivered the price stability of the early 1960s. The resolve of policymakers to fight inflation gives the monetary policy of the 1950s a modern flavor and seems inconsistent with Meltzer's position that post-war policymakers' objective function featured a lower weight on inflation stabilization before 1979 than after 1979.

2.2.2 The 1960s

Meltzer's account of the 1960s is a highlight of volume 2's book 1. The integration of Federal Reserve, White House, and Treasury materials, together with oral history recordings, is very well done. The use of Oval Office recordings in Meltzer's account does not wait until the Nixon years; he uses recordings of Kennedy-era conversations in chapter 3 (see especially pages 394 and 395). Aspects of Meltzer's narrative of the 1960s that will appear particularly relevant from today's perspective are the accounts of "Operation Twist" (the Federal Reserve's attempts in the early 1960s to influence the term structure through purchases of long-term government securities) and the 1966 credit crunch.[7] In addition, Meltzer highlights aspects of the emergence of overestimates of potential output during the 1960s, a feature that is central to Orphanides' (2003) account of the same period.

Just as he did in his account of the 1950s, Meltzer sees in 1960s Federal Reserve policy a failure to distinguish between nominal and

[7] An apparent omission from Meltzer's account of the 1966 credit crunch is mention of the Interest Rate Control Act of 1966. This legislation was an upshot of the financial disintermediation that occurred during the 1966 crunch. The Act extended Regulation Q-style deposit rate ceilings to thrift institutions (savings banks and savings and loan institutions). Though these powers were not given to the Federal Reserve, in practice they were used in tandem with the Federal Reserve's adjustments to Regulation Q in the late 1960s and early 1970s.

real interest rates. This is more plausible as a description of 1960s policy than of the 1950s; indeed, one 1968 speech by a senior Board staffer (Gramley 1968) leaves little doubt that high nominal interest rates in 1967 were taken as tight monetary policy. If, as argued above, policymakers had been well informed about the real/nominal rate distinction in the 1950s, how could they become ill informed by the 1960s? The answer is twofold. On the one hand, U.S. data for the 1950s featured very little autocorrelation in quarterly inflation rates: see McCallum (1994, table 1) and Erceg and Levin (2003). Under these circumstances, the assumption of constant expected inflation will tend to be a good approximation, and time-series evidence will suggest that a nominal rate/inflation connection is not present in the data. In light of the apparent absence, ex post, of a relation between nominal interest rates and inflation, policymakers in the 1960s may have reverted to a disregard of the Fisher relation—or, more precisely, may have assumed that inflation expectations were quite flat. A second dimension of policymakers' 1960s view of interest rates was that high nominal market interest rates were taken to have contractionary implications for aggregate demand because interest rate ceilings on bank deposits were seen as a potent policy tool. In actuality, as Meltzer documents, interest rate ceilings were likely ineffective for aggregate-demand-management purposes, because any banking activity discouraged by the ceilings could be taken up by unregulated financial intermediaries.

2.2.3 The 1970s

The 1970s are considered in detail in section 3 below. Highlights of Meltzer's account that deserve mention here are the details provided on the transitions from Burns to Miller and from Miller to Volcker.

The contributions of Federal Reserve Bank presidents to monetary policy in Meltzer's account are brought out mainly through references to their contributions to the FOMC's deliberations. Meltzer does note, however, the contributions that the Federal Reserve Banks of St. Louis and Minneapolis made during the 1970s as critics of prevailing FOMC orthodoxy, both via their research output and via the contributions of their presidents, especially Darryl Francis (St. Louis) and Mark Willes (Minneapolis). In light of Meltzer's attention to the insights provided by these Bank presidents, it is all the more jarring that the closing sentence of Meltzer's chapter on the period from 1973 to 1979 states, "No one suggested bold, decisive action to end inflation." While this provides a dramatic end to Meltzer's chapter on events prior to October 1979, it neglects the fact that the 1979 policy change was in the direction suggested by the Federal Reserve Bank of St. Louis. The St. Louis Bank had approved of the step-down in money growth and nominal spending growth envisaged in the Federal Reserve's successive monetary targets since 1975, but was alarmed by considerable overshoots of these targets. Viewing these misses as arising from the overly gradual adjustment by the FOMC of the federal funds rate target, the Federal Reserve Bank of St. Louis called for more willingness to tolerate funds rate fluctuations.[8] Consequently, the October 1979 change was welcomed by President Lawrence Roos of the Federal Reserve Bank of St. Louis, who said the day after the new arrangements were announced, "This is a major change of emphasis on the part of the Federal Reserve and monetary policymaking. This is what the St. Louis Fed has been preaching since the year one."[9]

2.2.4 1979 to 1982

This period has received saturation coverage in the monetary policy literature. Nevertheless, there are still plenty of new insights about the era available in Meltzer's chapter 8. One reason for this is the sheer level of detail and range of Meltzer's source material, which, in addition to voluminous Federal Reserve archival records,[10]

[8]This was of course the position taken by outside monetarists as well, including Meltzer, who, in the week prior to the 1979 regime change, pointed to the inadequacy of "policies that move the fed funds rate an eighth of a point at a time" (Meltzer 1980b).

[9]Quoted in Riesenberger (1979).

[10]One complaint, however, regarding Meltzer's use of the archives is that he does not make clear the specific contribution to our knowledge of events made by the archival material. The key problem is that Meltzer's citation procedure does not distinguish material that is special to the Federal Reserve's archives from material by Federal Reserve personnel that was published contemporaneously and has been used in previous histories of monetary policy. For example, Meltzer gives Federal Reserve Bank of New York archival boxes as his sources for 1970s speeches by Paul Volcker, even when these speeches were published in journals. Moreover, when quoting a 1984 speech by Governor Wallich, Meltzer gives "Board Records" as his source (p. 1164) instead of simply citing the version that appeared in print at the time (see Wallich 1984).

includes interviews with Volcker and members of his senior staff. Another reason is that Meltzer covers *both* the development of the Federal Reserve's "new operating procedures" in the form of nonborrowed reserves targeting—the aspect of this era focused on in the 1980s literature—and the implications of the policy change for such matters as the sacrifice ratio and policy credibility—the aspect of 1979–82 which tends to be the subject of the more recent literature.[11] Meltzer also decisively dispatches with some myths about the era that have downplayed Paul Volcker's own contribution to the disinflation that bears his name. For example, Meltzer rejects (on page 1024) the once-popular story that Volcker adopted the new procedures only after being browbeaten into doing so at a 1979 policymakers' summit in Yugoslavia. Meltzer also refutes (on pages 1081 and 1087) the claim—for which Meltzer gives no source, but which appeared in Newton (1983)—that the resolute monetary policy in 1981 was the product of President Reagan calling Volcker to account in a meeting at the White House.[12]

2.2.5 1983 to 1986

Meltzer's high-quality coverage of the Volcker years continues in his chapter 9. Particularly notable is the coverage of developments in 1983 and early 1984. Meltzer describes how the Volcker Federal Reserve, against an inauspicious background of a crisis-stricken financial system and confusion over the Federal Reserve's choice of

[11] On the other hand, Meltzer's consideration of both strands of the existing literature misses some key references: Karamouzis and Lombra (1989) is an important missing reference on the 1979–82 operating procedures, and Erceg and Levin (2003) is a major omission from Meltzer's discussion of the output costs and credibility problems associated with the Volcker disinflation.

[12] Meltzer is in error, however, on some details concerning this period. It is not accurate (according to the National Bureau of Economic Research business-cycle chronology) to describe the United States as being in recession in the spring of 1981, as Meltzer does on pages 1081–82. It is also not the case, as claimed by Meltzer on page 1035, that President Carter never publicly criticized Volcker's monetary policy during the 1980 campaign; Meltzer's own account (on page 1065) provides a counterexample to this claim. Meltzer's discussion of Volcker's record in hitting monetary targets is marred by his use of incorrect M1 growth rates: compare Meltzer's table on page 891 with the correct numbers given in Broaddus and Goodfriend (1984, p. 7).

policy instrument, managed to consolidate the disinflation and usher in the era of the Great Moderation.

The Federal Reserve reverted to an interest rate instrument from 1982 onward, and Meltzer uses this chapter and others to lay out his views on interest rate rules. These views have changed over the years. Meltzer (1980a) acknowledged that, in principle, the choice between a reserves instrument and an interest rate instrument pertained to tactics, not strategy, with monetary stability achievable under either instrument choice. But Meltzer's position was that, in practice, an interest rate instrument promotes excessive delays in interest rate adjustment and fosters a procyclical and often inflationary monetary policy. Meltzer is now critical of his former position (see pages 16 and 596). He provides the following poignant and balanced summary:

> Monetarists wanted the Federal Reserve to avoid procyclical actions by controlling money growth. . . . They recognized that if the Federal Reserve changed interest rates to control money growth, interest-rate control would be effective and countercyclical. But they did not emphasize the last point and insisted on the importance of controlling money directly. (p. 1017)

To acknowledge that an appropriately selected interest rate rule can be highly stabilizing does not, of course, rule out the possibility that monetary aggregates might have a valuable role in the formation of monetary policy decisions. Indeed, Meltzer notes that rapid money growth provided a more accurate signal of the strength of the initial stages of the post-1982 economic recovery than did real interest rates, which were historically high. Nevertheless, the preceding quotation from Meltzer does amount to an important acknowledgment on the part of a prominent monetarist that some of the monetarist thinking on policy rules required revision.[13]

The developments in Meltzer's thinking on monetary policy implementation provide an interesting parallel with the evolution of views of a Federal Reserve governor who features prominently in Meltzer's history: Henry Wallich. Wallich was not a monetarist; prior to serving as a Board member from 1974 to 1986, Wallich

---

[13]For a parallel acknowledgment by another prominent monetarist, see Schwartz (2009).

was one of the non-monetarist *Newsweek* columnists whose contributions provided a counterweight to Milton Friedman's articles for the magazine. While a *Newsweek* columnist and an academic, Wallich saw merit in the Nixon price controls (see the roundtable discussion of Friedman, Samuelson, and Wallich 1972). As a Federal Reserve governor, he came to see merit in the monetarist emphasis on the centrality of monetary policy to inflation and on money's role as an indicator, saying of the monetarists "they basically have a good idea."[14] But, in defending the short-term interest rate as a monetary policy instrument (for example, in Wallich 1982), he frequently found himself at odds with the monetarists. Over time, Wallich took on board much of the monetarist account of the effects of monetary policy, while Meltzer has come closer to Wallich's position on the choice of policy instrument. The common positions that Meltzer and Wallich ultimately reached in the wake of the events of the 1970s and early 1980s provides an example of the reconciliation that developed between monetarist and non-monetarist economists—a reconciliation that is also clear in the prominence given to monetary policy in New Keynesian economics.[15]

3. International Aspects of the Great Inflation

The principal matter on which the present review takes issue with Meltzer's volume 2 concerns his account of the Great Inflation of the 1970s. As noted above, Meltzer primarily attributes 1970s monetary policy outcomes to a low weight on inflation stabilization until 1979 in FOMC policymakers' objective function. (See, for example, Meltzer's pages 864, 1033–34, 1097, 1131, and 1217.) Meltzer gives little credence to explanations that focus on policymakers' misunderstanding of the importance of monetary policy to inflation. (See, for example, Meltzer's pages 861 and 863.) Meltzer's interpretation of the Great Inflation is hard to reconcile with the U.S. experience. Furthermore, as I detail below, his interpretation also does not come

[14] Quoted in Clark and McGinley (1984). See also Wallich (1980, 1984) and Meltzer's discussion of Wallich's views on pages 1092, 1132–34, and 1165.

[15] New Keynesian economics likewise recognizes the distinction, repeatedly stressed by Meltzer, between price-level shocks and the ongoing inflation rate; see, for example, the discussion of the New Keynesian Phillips curve in Walsh (2003, p. 517).

to grips with the Great Inflation's international character. Meltzer states the following:

> The international character of the Great Inflation is sometimes advanced as support for explanations based on errors in economic theory. The claim is that many countries made the same errors, particularly denial of the natural rate hypothesis, claiming that unemployment in the long run was independent of inflation. . . . Appealing as this argument is to economists, it fails to separate the start of inflation and its continuance. (p. 864)

This is an unsatisfactory statement. Unlike the rest of Meltzer's book, the quoted passage does not conform to good scholarly practice; referring to an argument as having been "sometimes advanced" is no substitute for citing the relevant literature and confronting the arguments in that literature. Yet Meltzer provides no bibliographical references in the paragraph containing the above-quoted passage. The omission is not just a bibliographical problem; Meltzer does not face the arguments that have been central to recent discussions of international aspects of the Great Inflation. It appears appropriate to conclude that, in a book published five years after his exchange with Romer (2005), Meltzer has not risen to the challenge posed by Romer, whose critique of Meltzer's account highlighted the international dimension of the Great Inflation.

The portion of the Great Inflation literature that covers the experience of other countries is quite small, as Cecchetti et al. (2007) note. Other than Cecchetti et al., this strand of the literature includes Nelson (2005a, 2005b) and Scrimgeour (2008). The material in those papers brings out problems with Meltzer's characterization of explanations for the Great Inflation. Contrary to the impression that Meltzer creates, the literature on international aspects of the Great Inflation does *not*, by and large, attribute international inflation to policymaker pursuit of a Phillips-curve trade-off. Indeed, such an account of policy behavior is untenable, as it does not square with the fact that policymakers in the 1970s in several major countries explicitly rejected Phillips-type trade-off ideas. For example, Prime Minister Pierre Trudeau of Canada in 1970 ruled out the existence of a long-run Phillips relation (see Nelson 2005a, p. 147). As another example, in March 1976 the UK Prime Minister, Harold Wilson, said in a television interview (Independent Television News 1976), "the old theory that if you have a lot of unemployment, you can't have inflation, has been disproved in Britain and all over the world, as some of us many years ago said it would be before very long. In 1957, I was writing articles saying that."

Explicit rejections of a long-run Phillips-curve trade-off appeared in statements by U.S. policymakers during the late 1960s and the 1970s. Meltzer correctly notes (on pages 585 and 845) that Arthur Burns rejected the notion of a long-run Phillips-curve trade-off. But it would not be correct to infer that Burns, or his counterparts in other major countries, adhered to modern inflation theories, under which monetary policy ultimately determines inflation. On the contrary, the coexistence of high inflation and high unemployment can be rationalized by unorthodox theories of inflation as well as by modern natural-rate theory.[16] In particular, pure cost-push, or non-monetary, views of inflation can lead one to the view that there is *no* connection between output gaps and inflation, because these theories take the coefficient on the output gap in the Phillips curve to be zero when the economy is below a state of full employment. Since pure cost-push views can rationalize arbitrary combinations of unemployment and inflation, it is easy to see why these views of the inflation process gained ground and persisted among policymakers in the 1970s.[17] Pure cost-push views of inflation were already prevalent in policy circles in the United Kingdom and other countries before

[16]Meltzer is therefore in error when he suggests (on page 864) that accounts of the Great Inflation that focus on policymaker adherence to erroneous theories fail to explain the durability of these theories in the face of the high inflation and rising unemployment of the 1970s.

[17]Belief in cost-push forces as one factor behind inflation had cropped up as an element of official thinking about inflation during the 1950s and 1960s; see Meltzer's chapter 2, as well as Riefler (1959) and Romer and Romer (2002b). But in those decades, U.S. policymakers still believed that the level of the output gap mattered for inflation behavior; thus, a policy that held down output could offset the effects of cost-push forces on overall inflation. Beginning in 1970, however, U.S. policymakers moved to a harder-line position—a purely non-monetary view of inflation—under which cost-push forces drove inflation even in the face of monetary restraint. This contrast between pre-1970 and post-1970 policymaker positions underscores the point that the 1960s and 1970s inflation problems of the United States arose from different sources of policy error. See DiCecio and Nelson (2009) for further discussion.

1970, but gained traction at the highest levels of U.S. policymaking only in 1970. Chairman Burns and Chairman Miller both subscribed to this non-monetary view of inflation.[18] Meltzer tells us (page 861) that there is "little reason to doubt that FOMC members recognized their role in creating inflation," but the extensive analysis of the statements of Burns and his FOMC colleagues in Hetzel (1998), Romer and Romer (2002a, 2004b), Nelson (2005b), and DiCecio and Nelson (2009) suggests the opposite conclusion.

Meltzer claims that international transmission of U.S. inflation via the Bretton Woods system can account for the international character of the 1970s inflation. An appeal to Bretton Woods does not, however, provide a valid explanation. For one thing, even under Bretton Woods, U.S. monetary policy was not the be-all and end-all in determining other member countries' monetary policies; it was possible for these countries to separate their monetary policy considerably from U.S. monetary policy. For example, Rasche (1990) for Japan, and Throop (1980) and Mankiw (1986) for the United Kingdom, find that, over the 1960s and 1970s, there was considerable divergence in the path of domestic interest rates from that prevailing in the United States. True, these countries did feature occasions on which the exchange rate was fixed so low that it triggered capital inflow on a scale beyond what could be discouraged by existing exchange controls or sterilized by domestic operations.[19] But a detailed examination of these instances establishes that the domestic monetary conditions engendered by the fixed exchange rate policy were seen as desirable from the point of view of stabilization policy. In particular, governments in fixed-rate countries in the 1970s frequently adhered to the misconception that non-monetary

[18]The evidence that Meltzer offers to the contrary (for example, on pages 846 and 861) neglects the aspect of the monetary explanation of inflation that monetary policy actions are *necessary and sufficient* for inflation control. The pure cost-push views of inflation prevailing during the 1970s held that inflation was insensitive to the level of the output gap (and hence aggregate demand policies) when the output gap was negative; they did not deny that positive output gaps, on those occasions when these did emerge, could add to inflation.

[19]This was the case, for example, in 1972 for Japan and in 1971 and 1977 for the United Kingdom. A contrasting case is provided by Germany, which in 1972–73 tightened exchange controls and was thereby in a position to commence an early disinflation (see, for example, von Hagen 1999, p. 686).

policies could take care of inflation, while at the same time believing that there was a considerable amount of economic slack that made the looser monetary conditions induced by the capital inflows desirable.[20] Furthermore, different countries left Bretton Woods in different stages. And once they moved to floating rates, some countries pursued a tighter monetary policy than the United States, while others pursued an easier policy. The different choices made by each country can be understood in terms of the theories of inflation prevalent among policymakers in each nation.

The non-monetary view of inflation was a fundamental misconception embedded in 1970s policymakers' evaluation of the structural connection between monetary policy and price stability. It would be a mistake, moreover, to contend that this misconception can be regarded as a special case of, a restatement of, or isomorphic to, the factor to which Meltzer gives prominence, namely, a low weight on inflation in policymakers' objective function. The non-equivalence between an emphasis on non-monetary views of inflation and a hypothesized low weight on inflation in the objective function is brought out immediately by the fact that Meltzer claims that this low weight prevailed *throughout* the post-war period up to 1979. Being unchanged in the 1970s from prior decades, such a postulated objective function cannot serve as a means of capturing intensified misconceptions, and resulting policy errors, on the Federal Reserve's part from 1970 onward.

It seems essential to focus on policymaker misconceptions in understanding the 1970s inflation and 1980s disinflation. Meltzer's focus on the objective function in accounting for policy developments does not seem to be an attractive alternative.[21] Contrary to Meltzer's account, the Great Inflation during the 1970s, both in

[20] As the co-editor (Carl Walsh) has observed, New Zealand in the 1970s provides an example of a country whose authorities—backed up by majority academic opinion—firmly believed that inflation was a non-monetary phenomenon and shaped their policy mix in light of this misguided belief. For further discussion, see Nelson (2005a) and Svensson (2011, pp. 1243–44).

[21] While his account places emphasis on particular policymaker misconceptions (in particular, misjudgments about long-run inflation/unemployment trade-offs) that are not those stressed here, Primiceri (2006) finds little evidence of shifts in weights in policymakers' objectives in accounting for differences in policy behavior before and after 1979.

the United States and elsewhere, is best understood as resulting from policymakers' failure to understand the central role of monetary policy in determining inflation.[22] Their non-monetary view of inflation led them to see the solution to inflation in non-monetary means such as wage-price controls and other types of incomes policy. Weak monetary policy responses to inflation in the 1970s flowed from the view that non-monetary tools were best suited to fight inflation—not from a low weight on inflation stabilization in policymakers' objective functions. The failure of non-monetary policies to cure inflation led to a debate in each country on monetary versus non-monetary views of inflation. The monetary view of inflation ultimately prevailed. Those countries for which the debate was resolved promptly in favor of the monetary view were the countries that disinflated earliest; these included Germany, Japan, and Switzerland. The nations in which non-monetary views lingered at the highest policy levels into the late 1970s—countries like the United States, the United Kingdom, and Canada—took longer to move decisively to disinflationary policies.

4. Meltzer's Epilogue

In his epilogue on post-2007 events, Meltzer is critical of the Federal Reserve's policy response. Meltzer claims (on page 1243) that his critique of post-2007 policy highlights points he has raised in the preceding *History*. In fact, elements of Meltzer's epilogue clash with the account in volume 2 of his *History*. For example, Meltzer states on page 1243, concerning post-2007 policy, "Never before had it taken responsibility as lender of last resort to the entire financial system." Yet Meltzer on page 656 referred to the "landmark" recognition in the late 1960s that "the Federal Reserve was the lender

[22] This is not to claim that important misconceptions about monetary policy's role were absent prior to the 1970s, or that 1970s misconceptions were isolated to non-monetary views of inflation. In particular, as already noted, Meltzer, like the previous literature, notes that exaggerated estimates of potential output played a part in 1960s policy errors. In addition, Meltzer argues persuasively (for example, on pages 521–22, 532, and 675) that the Federal Reserve in the 1960s placed too much weight on the scope for modest fiscal policy tightening to serve as a substitute for monetary policy tightening.

of last resort to the entire financial system," and on page 609, in discussing 1970 events, Meltzer had taken for granted "the role of lender of last resort to the financial system." Meltzer's epilogue also states (on page 1250), "Reading transcripts of Federal Open Market Committee meetings, one finds very little discussion of regulatory and supervisory credit problems." Why should that be a surprise? The FOMC is not the body primarily responsible for executing the Federal Reserve's regulatory and supervisory responsibilities.

The substance of Meltzer's criticism of the policy response to the financial crisis is that long-standing central banking principles formed in the nineteenth century (especially those that he associates with Bagehot) could have been adhered to more closely during the crisis period.[23] Meltzer should have been more explicit about how to interpret, apply, and, if necessary, modify these principles in a modern financial environment. In particular, Meltzer should have laid out explicitly the policy responses appropriate to an environment in which commercial banks issue uninsured short-term obligations on a large scale, and a vast wholesale banking market intensifies the connections between different financial institutions. When an institution encounters problems in such a system, the counterparty implications of the problems are formidable, and the time-critical nature of the policy response is greater than in previous systems. Readers of Meltzer's account will feel justified in feeling that he has not come to grips with this issue and is instead considering simpler and more compartmentalized cases. For example, a footnote on pages 1158–59 that considers the appropriate policy response to insolvency of a bank turns out to be framed on the assumption that the bank actually has some positive net worth remaining at the time of the policy intervention. Likewise, it is very hard to find references to wholesale banking in Meltzer's book; the word "wholesale" is almost always used in other contexts (such as in references to the wholesale price index). Wholesale funding markets grew in the 1970s and 1980s, yet once he gets past the mid-1970s, Meltzer's discussions of banks' managed liabilities are fleeting.

[23]In an article which was probably not available when Meltzer's volume 2 went to press, Madigan (2009) detailed some connections between Bagehot's prescriptions and recent Federal Reserve lending activity.

Meltzer states on page 1249, "By guaranteeing deposits, money market liabilities, and other instruments, the Federal Reserve prevented bank runs and further breakdown of the payments system." This is an important acknowledgment on Meltzer's part—although "the federal authorities" would have been a more accurate phrase than "the Federal Reserve."

5. Conclusion

This review of volume 2 of Allan Meltzer's *History of the Federal Reserve* has taken issue with Meltzer's interpretation of policy developments on several counts. The most important area of disagreement is with Meltzer's view that policymakers' objective function is the key to understanding the Great Inflation of the 1970s and the disinflation of the 1980s. His emphasis on shifts in the objective function seems misplaced, as it does not provide an explanation that satisfactorily accounts for the international character of the Great Inflation. The reason for the monetary policy mistakes of the 1970s, in the United States and other countries, is to be found elsewhere. Most importantly, 1970s policymakers had misguided non-monetary views on why inflation had risen and how it could be curbed, and it was the shaking off of these misconceptions that was the crucial step in shifting policymakers to disinflationary monetary actions. It bears emphasis, however, that notwithstanding the qualms expressed here about Meltzer's interpretations of particular episodes, volume 2 of his *History of the Federal Reserve* is a landmark contribution to our stock of knowledge about the development of U.S. monetary policy.

References

Broaddus, A., and M. Goodfriend. 1984. "Base Drift and the Longer Run Growth of M1: Experience from a Decade of Monetary Targeting." *Economic Review* (Federal Reserve Bank of Richmond) 70 (6): 3–14.

Brunner, K., and A. H. Meltzer. 1973. "Mr. Hicks and the 'Monetarists'." *Economica* 40 (157): 44–59.

Cecchetti, S. G., P. Hooper, B. Kasman, K. Schoenholtz, and M. W. Watson. 2007. "Understanding the Evolving Inflation Process."

Paper presented at the U.S. Monetary Policy Forum, Washington, DC, March 9.

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

Clark, L. H., and L. McGinley. 1984. "Money's Role: Monetarists Succeed in Pushing Basic Ideas but Not Their Policies." *Wall Street Journal*, December 10, pp. 1 and 16.

DiCecio, R., and E. Nelson. 2009. "The Great Inflation in the United States and the United Kingdom: Reconciling Policy Decisions and Data Outcomes." NBER Working Paper No. 14895.

Erceg, C. J., and A. T. Levin. 2003. "Imperfect Credibility and Inflation Persistence." *Journal of Monetary Economics* 50 (4): 915–44.

Feinman, J., and W. Poole. 1989. "Federal Reserve Policymaking: An Overview and Analysis of the Policy Process: A Comment." *Carnegie-Rochester Conference Series on Public Policy* 30 (1): 63–74.

Friedman, M. 1984. "Monetary Policy for the 1980s." In *To Promote Prosperity: U.S. Domestic Policy in the Mid-1980s*, ed. J. H. Moore, 23–60. Stanford, CA: Hoover Institution Press.

Friedman, M., P. Samuelson, and H. C. Wallich. 1972. "Three Views of Nixonomics and Where It Leads." *Newsweek*, January 31, 36–37.

Friedman, M. and A. J. Schwartz. 1963. *A Monetary History of the United States, 1867–1960.* Princeton, NJ: Princeton University Press.

Gramley, L. E. 1968. "The Informational Content of Interest Rates as Indicators of Monetary Policy." In *Proceedings of the 1968 Money and Banking Workshop*, 12–24. Minneapolis: Federal Reserve Bank of Minneapolis.

Greenspan, A. 1959. "Stock Prices and Capital Evaluation." *Proceedings of the Business and Economic Statistics Section* (American Statistical Association) 6 (1): 2–26.

Hetzel, R. L. 1998. "Arthur Burns and Inflation." *Economic Quarterly* (Federal Reserve Bank of Richmond) 84 (1): 21–44.

———. 2008. *The Monetary Policy of the Federal Reserve: A History.* Cambridge, UK: Cambridge University Press.

Independent Television News. 1976. "*Weekend World*, broadcast March 22, 1976." Tape of broadcast acquired from http://www.itnsource.com.

Karamouzis, N., and R. Lombra. 1989. "Federal Reserve Policymaking: An Overview and Analysis of the Policy Process." *Carnegie-Rochester Conference Series on Public Policy* 30 (1): 7–62.

Lindsey, D. E., A. Orphanides, and R. H. Rasche. 2005. "The Reform of October 1979: How It Happened and Why." *Review* (Federal Reserve Bank of St. Louis) 87 (2, part 2): 187–236.

Lucas, R. E., Jr. 1988. "Money Demand in the United States: A Quantitative Review." *Carnegie-Rochester Conference Series on Public Policy* 29 (1): 137–68.

Madigan, B. F. 2009. "Bagehot's Dictum in Practice: Formulating and Implementing Policies to Combat the Financial Crisis." In *Financial Stability and Macroeconomic Policy*, 169–89. Proceedings of the Federal Reserve Bank of Kansas City Annual Economic Symposium, Jackson Hole, Wyoming, August 20–22.

Mankiw, N. G. 1986. "The Term Structure of Interest Rates Revisited." *Brookings Papers on Economic Activity* 17 (1): 61–96.

Martin, W. M. 1957. "Statement of William McChesney Martin, Jr., Chairman, Board of Governors of the Federal Reserve System, before the Committee on Finance, U.S. Senate, August 13, 1957." In *Investigation of the Financial Condition of the United States: Hearings, Part 3: August 13, 14, 15, 16, and 19, 1957*, pp. 1257–72. Washington, DC: GPO. Typescript version of statement is available online in the Federal Reserve Bank of St. Louis's FRASER archive at http://fraser.stlouisfed.org/historicaldocs/762/download/30925/martin57_0813.pdf.

McCallum, B. T. 1994. "Identification of Inflation-Unemployment Tradeoffs in the 1970s: A Comment." *Carnegie-Rochester Conference on Public Policy* 41 (1): 231–41.

Meltzer, A. H. 1965. "Monetary Theory and Monetary History." *Swiss Journal of Economics and Statistics* 101 (4): 404–22.

———. 1980a. "Central Bank Policy: Some First Principles." *City University Annual Monetary Review* 2 (1): 27–33.

———. 1980b. "First Financial Forum, September 27, 1979: Allan Meltzer's Opening Statement." In *Issues in Monetary Policy: Three Financial Forums*, ed. M. J. Hamburger, 13–14. New York: New York University.

———. 2003. *A History of the Federal Reserve, Volume 1, 1913–1951.* Chicago: University of Chicago Press.

———. 2009a. *A History of the Federal Reserve, Volume 2, Book 1, 1951–1969.* Chicago: University of Chicago Press.

———. 2009b. *A History of the Federal Reserve. Volume 2, Book 2, 1970–1986.* Chicago: University of Chicago Press.

Nelson, E. 2005a. "Monetary Policy Neglect and the Great Inflation in Canada, Australia, and New Zealand." *International Journal of Central Banking* 1 (1): 133–79.

———. 2005b. "The Great Inflation of the Seventies: What Really Happened?" *Advances in Macroeconomics* 5 (1): Article 3, 1–50.

Newton, M. 1983. *The Fed.* New York: Times Books.

Orphanides, A. 2003. "The Quest for Prosperity Without Inflation." *Journal of Monetary Economics* 50 (3): 633–63.

Primiceri, G. 2006. "Why Inflation Rose and Fell: Policy-Makers' Beliefs and U.S. Postwar Stabilization Policy." *Quarterly Journal of Economics* 121 (3): 867–901.

Rasche, R. H. 1990. "Equilibrium Income and Interest Elasticities of the Demand for M1 in Japan." *Monetary and Economic Studies* (Bank of Japan) 8 (2): 31–57.

Riefler, W. W. 1959. "Inflation—Enemy of Growth." Paper presented at the Stanford Business Conference, Stanford, CA, July 21. Reprinted in *Employment, Growth and Price Levels, Hearings, Volume 10*, 3369–76, Joint Economic Committee, U.S. Congress. Washington, DC: GPO.

Riesenberger, B. 1979. "Actions by Fed Are Welcomed Here." *St. Louis Post-Dispatch*, October 8, 9A.

Romer, C. D. 2005. "Commentary on 'Origins of the Great Inflation' by Allan H. Meltzer." *Review* (Federal Reserve Bank of St. Louis) 87 (2, part 2): 177–86.

Romer, C. D., and D. H. Romer. 1989. "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz." *NBER Macroeconomics Annual* 4 (1): 121–84.

———. 2002a. "A Rehabilitation of Monetary Policy in the 1950's" *American Economic Review: Papers and Proceedings* 92 (2): 121–27.

———. 2002b. "The Evolution of Economic Understanding and Postwar Stabilization Policy." In *Rethinking Stabilization Policy*, 11–78. Proceedings of the Federal Reserve Bank of Kansas City Annual Economic Symposium, Jackson Hole, Wyoming, August 29–31.

———. 2004a. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* 94 (4): 1055–84.

———. 2004b. "Choosing the Federal Reserve Chair: Lessons from History." *Journal of Economic Perspectives* 18 (1): 129–62.

Schwartz, A. J. 2009. "Preface." In *The Great Contraction, 1929–1933*, ed. M. Friedman and A. J. Schwartz, ix–xii. Princeton, NJ: Princeton University Press.

Scrimgeour, D. 2008. "The Great Inflation Was Not Asymmetric: International Evidence." *Journal of Money, Credit, and Banking* 40 (4): 799–815.

Svensson, L. E. O. 2011. "Inflation Targeting." In *Handbook of Monetary Economics, Volume 3B*, ed. B. M. Friedman and M. Woodford, 1237–1302. Amsterdam: Elsevier/North Holland.

Throop, A. W. 1980. "Managed Floating and the Independence of Interest Rates." *Economic Review* (Federal Reserve Bank of San Francisco) 6 (3): 6–23.

Tobin, J. 1965. "The Monetary Interpretation of History." *American Economic Review* 55 (3): 464–85.

von Hagen, J. 1999. "Money Growth Targeting by the Bundesbank." *Journal of Monetary Economics* 43 (3): 681–701.

Wallich, H. C. 1980. "Monetary Policy During High Inflation." Remarks to the Swiss-American Society, Basel, June 10.

———. 1982. "Comments: 'Innovations in Monetary Policy: Challenge and Response' by Jürg Niehans." *Journal of Banking and Finance* 6 (1): 33–35.

———. 1984. "Recent Techniques of Monetary Policy." *Economic Review* (Federal Reserve Bank of Kansas City) 69 (5): 21–30.

Walsh, C. E. 2003. *Monetary Theory and Policy.* 2nd ed. Cambridge, MA: MIT Press.

Young, R. A. 1958. "The Statistical Foundation for Policy Formulation in the Federal Reserve System." *Proceedings of the Business and Economic Statistics Section* (American Statistical Association) 5 (1): 202–207.