FEDERAL RESERVE BANK OF ST. LOUIS
REVIEW

REVIEW

Volume 97 • Number 1

To Our Readers

In January 2015, I took over from Bill Gavin, who served for many years as editor-in-chief of the Federal Reserve Bank of St. Louis *Review*. Bill has retired from his position, and I hope to continue the tradition of high-quality *Review* articles published under Bill's capable leadership.

The *Review* will continue to publish articles written by our staff economists at the St. Louis Fed, along with occasional contributions by outside scholars, including our research fellows and academic visitors. We aim to please a broad audience, addressing economic issues at the frontier of research on economic theory and policy, with a special interest in macroeconomics, financial issues, monetary economics, banking, and monetary policy.

Our aim is to focus on the important economic ideas of our time, presented in an accessible fashion. To that end, we welcome input from you, our readers.

Stephen Williamson
Vice President and Editor-in-Chief of *Review*
Federal Reserve Bank of St. Louis

Three Scenarios for Interest Rates in the Transition to Normalcy

Diana A. Cooke and William T. Gavin

In this article, time-series models are developed to represent three alternative, potential monetary policy regimes as monetary policy returns to normal. The first regime is a return to the high and volatile inflation rate of the 1970s. The second regime, the one expected by most Federal Reserve officials and business economists, is a return to the credible low inflation policy that characterized the U.S. economy from 1983 to 2007, a period known as the Great Moderation. The third regime is one in which policymakers keep policy interest rates at or near zero for the foreseeable future; Japanese data are used to estimate this regime. These time-series models include four variables: per capita gross domestic product growth, consumer price index inflation, the policy rate, and the 10-year government bond rate. These models are used to forecast the U.S. economy from 2008 through 2013 and represent the possible outcomes for interest rates that may follow the return of monetary policy to normal. Here, "normal" depends on the policy regime that follows the liftoff of the federal funds rate target expected in mid-2015. (JEL E43, E47, E52, E58, E65)

During the fourth quarter of 2008, while in the process of rescuing a few large financial firms following the Lehman Brothers bankruptcy, the Federal Reserve added about $600 billion in excess reserves to the banking system. Throughout 2007, the banking system had operated with less than $5 billion. This action drove the interest rate on bank reserves (aka the federal funds rate or policy rate) below the Federal Open Market Committee's (FOMC) target rate of 2 percent. By the first Friday in December 2008, the effective federal funds rate had fallen to 12 basis points. On December 16, 2008, the FOMC then set the official federal funds rate target at a range of 0 to 0.25 percent, where it remains to this day. With the policy rate effectively at zero and the banking system flooded with excess reserves, the FOMC has tried to ease monetary conditions through two related policies: (i) forward guidance promises of maintaining the current policy rate even further into the future and (ii) large-scale purchases of long-term Treasury debt and agency mortgage-backed securities, which effectively exerts downward pressure on long-term interest rates. As of June 2014, this latter policy has increased the total of excess reserves to $2.6 trillion.

Within the Federal Reserve System, this situation is considered temporary and the FOMC is now debating strategies that would return both the balance sheet and the policy rate to normal. When the economy and monetary policy eventually return to normal, excess reserves would be expected to return to levels observed before the financial crisis. A useful definition of "normal" can be taken from the December 2014 FOMC long-run forecasts of real GDP growth (2.0 to 2.3 percent), inflation (2 percent), the unemployment rate (5.2 to 5.5 percent), and the policy rate (3.5 to 4.0 percent).[1] The effects of reducing excess reserves will depend on how interest rates change during the transition to normalcy. The level and volatility of interest rates will depend on the public's beliefs about future monetary policy. Carpenter et al. (2013) provide an excellent overview of the Fed's balance sheet and describe three exit strategies based on FOMC policy statements. Their projections of the Fed's balance sheet and its net income are conditioned on assumptions about future interest rates.

The "Lucas critique," a well-known econometric problem, is associated with simulating an economy under alternative policy assumptions (Lucas, 1976). The Carpenter et al. (2013) simulations are based on the implicit assumption that the U.S. economy has had one stable policy regime from about 1983 to the present and that the transition period will be an extension of this same policy regime. In this article, we discard this assumption and allow the periods with credibility, no credibility, and a zero policy rate to be separate regimes with different econometric properties.

Distinguishing between regimes is important because the major concerns surrounding the exit strategy for monetary policy arise from the interest rate implications of the transition to a separate policy regime. For example, in our judgment the "taper tantrum" of 2013 was a typical interest rate response that would naturally be associated with moving from a zero interest rate policy (ZIRP) to the credible monetary policy regime in place between 1983 and 2007, a period known as the Great Moderation. In addition, some economists and policymakers worry that the ZIRP regime will eventually lead to a loss of credibility for the Fed and a return to the high-inflation regime in the United States from about 1965 through 1979.[2]

Predicting interest rates during the transition to normalcy is complicated because it requires predicting the regime that will be in place at the end of the transition. How high interest rates are likely to rise and how likely the yield curve is to become inverted depends on people's beliefs about the policy regime. We develop three scenarios based on alternative assumptions about Federal Reserve policy.

We use a data-based scheme to identify time-series models for interest rates associated with each regime. Data from unique episodes before 2008 are used to estimate the models, which are then simulated to forecast the U.S. economy during the 2008-13 period from the point of view of 2007:Q4. The purpose of these forecasts is simply to illustrate how well the alternative regimes can explain interest rates during the past six years. Then starting from the point of view of 2013:Q4, the forecasts show interest rate expectations for the next few years.

The time-series model for each scenario generates interest rates assumed to be typical of the relevant policy regime. For example, we ask what would happen to interest rates if the Fed lost credibility for price stability, as it did in the period following the breakdown of the dollar standard agreed to at Bretton Woods after World War II. Our results show that inflation and

interest rates would become unacceptably high and volatile, as they did in the 1970s. We consider this scenario and two others: one based on policy during the Great Moderation (the U.S. economy from 1983 to 2007) and the other based on the ZIRP in Japan, where the monetary policy rate was held at or below ½ percent from 1995 through 2007.

THREE SCENARIOS FOR MONETARY POLICY

We consider three scenarios representing three different policy regimes:

- The first scenario assumes that the Fed loses credibility for its inflation objective, as it did in the 1970s, and inflation accelerates. We use U.S. data from 1965:Q1 to 1979:Q3 to estimate the *No Credibility* model.

- The second scenario assumes that the Fed has credibility and operates policy to achieve price stability (low inflation) as it did from 1987 to 2007. We use U.S. data from 1983:Q1 to 2007:Q4 to estimate the *Credibility* model.

- The third scenario assumes the Fed keeps the policy rate at or near zero permanently. The credibility for the 2 percent inflation objective is dominated by credibility for its ZIRP. We use Japanese data from 1995:Q1 to 2007:Q4 to estimate the *ZIRP* model.

The statistical relationships determining per capita output, inflation, and interest rates are assumed to depend on the monetary policy regime, which is characterized by the time-series models developed later in this article. We recognize that monetary policy is not the only reason for differences in the time-series properties of the data among our alternative sample periods. There are structural differences between the U.S. economy and the Japanese economy, as well as between the early U.S. period and the later one. For this reason, we do not emphasize results for the real economy. Our key assumption is that monetary policy, through its determination of the inflation trend, is the dominant factor driving nominal interest rates.[3]

We examine three periods corresponding to the three distinctly different monetary policy regimes. We review the historical experience to clarify how credibility matters for interest rates and inflation. Then we use our models to forecast inflation and interest rates over the financial crisis period to evaluate the range of uncertainty that may arise during a transition to normalcy.

Three Regimes (The Same Time-Series Model Estimated Over Different Episodes)

Our basic model in all three policy scenarios is a vector autoregression (VAR) including four quarterly time series: per capita gross domestic product (GDP) growth, consumer price index (CPI) inflation, a short-term policy rate, and the 10-year government bond rate. The policy rate for the United States is the overnight federal funds rate. For Japan, it is the call money rate.[4]

Our model produces a four-quarter-ahead forecast. It is written as

$$Y_{t+3} = A(L)Y_{t-1} + e_{t+3},$$

where

$$Y_t = \begin{pmatrix} GDP_t \\ CPI_t \\ RS_t \\ RL_t \end{pmatrix} \text{ and } e_t = \begin{pmatrix} e_{gdp,t} \\ e_{cpi,t} \\ e_{rs,t} \\ e_{rl,t} \end{pmatrix},$$

where GDP_t is real GDP growth minus population growth, CPI_t is the four-quarter change in the CPI, RS_t is the policy rate, and RL_t is the 10-year government bond rate. We assume that the error process is multivariate normal $e_t \sim N(0,\Sigma)$. We include the four-quarter forecast horizon rather than a one-quarter horizon because we are mainly interested in medium-term forecasts and the four-quarter specification produces better forecasts at longer horizons. Our models have identical structures, but the estimated parameters differ across the three scenarios (*No Credibility*, *Credibility*, and *ZIRP*) because the data used to estimate the models are from three episodes with very different monetary policy environments.

No Credibility Scenario: 1965-79

During the 1970s, the United States experienced a period of accelerating inflation known as the Great Inflation.[5] The top panel of Figure 1 shows that CPI inflation rose in fits and starts from just under 2 percent in 1965 to 12.2 percent in September 1979. This period was often characterized as an era of stop-go monetary policy. When inflation accelerated, the FOMC would raise the policy rate high enough to slow inflation. The relatively high policy rate would lower aggregate spending, reduce the demand for labor, and lead to a recession. The FOMC would then switch gears, lowering the policy rate sharply to stimulate spending and job growth. The stop-go nature of this policy is evident in the bottom panel of Figure 1, which shows the federal funds rate and the 10-year government bond rate from 1965 through 1979.

The relationship between the policy rate and the bond rate during the pre-1980 period displays three distinct features. First, both interest rates display rising trends and, on average, are roughly equal; the policy rate averaged just 0.6 percent less than the bond rate. Second, the policy rate was sometimes much higher than the bond rate. Third, periods with a relatively low policy rate were followed by higher inflation and inflation expectations, reflected in rising bond rates.

The lack of credibility required setting the policy rate above the bond rate to reduce inflation expectations. When the FOMC raised the policy rate too slowly, inflation expectations would rise to match the rise in the interest rate with no dampening effect on either the economy or inflation. The lack of credibility meant that to succeed in lowering inflation, the FOMC had to raise the policy rate high enough to slow the economy. This led to a belief that stabilizing inflation would likely lead to high unemployment. A corollary to this belief was that low interest rates would raise inflation, and, at the same time, lower the unemployment rate. What has not been generally recognized is that these dynamic relationships came to be part of conventional wisdom in macroeconomics during a time when the Fed had no credibility for its inflation objective.

Figure 1

United States without Fed Credibility (1965-79)

Percent Annual Rate

Percent

NOTE: Inflation is measured as the change over the previous year.

SOURCE: Data from FRED®, Federal Reserve Economic Data, Federal Reserve Bank of St. Louis: *Consumer Price Index for All Urban Consumers: All Items [CPIAUCSL]*; U.S. Bureau of Labor Statistics; http://research.stlouisfed.org/fred2/series/CPIAUCSL; accessed January 1, 2015. *10-Year Treasury Constant Maturity Rate [GS10]*; Board of Governors of the Federal Reserve System; http://research.stlouisfed.org/fred2/series/GS10; accessed January 1, 2015. *Effective Federal Funds Rate [FEDFUNDS]*; Board of Governors of the Federal Reserve System; http://research.stlouisfed.org/fred2/series/FEDFUNDS; accessed January 1, 2015.

Table 1A

No Credibility Model

	GDP_{t+3}	CPI_{t+3}	RS_{t+3}	RL_{t+3}
GDP_{t-1}	−0.16*	0.65*	0.66	0.23**
CPI_{t-1}	0.01	0.54	−0.03	0.25*
RS_{t-1}	−1.15	0.51	0.51	0.04*
RL_{t-1}	0.42	0.63	0.84	0.64
Constant	7.23	−5.74	−3.43	0.57
Adjusted R^2	0.67	0.79	0.56	0.83
SE equation	1.46	1.45	1.79	0.62
Mean dependent	2.39	6.37	6.93	7.05
SD dependent	2.54	3.15	2.70	1.52
Included observations	59			
Sample	1965:Q1–1979:Q3			

Residual correlation matrix (SE on diagonal)

	GDP	CPI	RS	RL
GDP	1.46	—	—	—
CPI	−0.03	1.45	—	—
RS	0.13	0.77	1.79	—
RL	0.24	0.55	0.78	0.62

NOTE: RS, short-run (policy) rate; RL, long-run (bond) rate; SD, standard deviation; SE, standard error. * and ** indicate significance at the 10 percent and 5 percent levels, respectively.

We use U.S. data from 1965:Q1 through 1979:Q3 for the *No Credibility* model. For this and the other models, we find that the best lag length was just one quarter based on the Schwartz Bayesian information criterion. Table 1A lists the estimates of the model and summary statistics. The standard errors in the per capita GDP growth, inflation, policy rate, and bond rate equations are 1.46, 1.45, 1.79, and 0.62 percent, respectively. These standard errors are important because they influence the inherent uncertainty in the forecasts.

The other major factor influencing the uncertainty in the forecast is the implication for the long-run trend. Table 2 presents the long-run mean forecasts for each model starting from the initial conditions in 2007 and 2013. These are dynamic forecasts under the assumption that there are no further shocks over the forecast period. The number of years to convergence depends on initial conditions and how quickly the models' equations converge to their long-run trends.[6] In the *No Credibility* model, starting from 2007 (2013) initial conditions, per capita GDP growth converges to −6.7 percent in 222 (224) years. The CPI inflation rate converges to 34.5 percent in 283 (286) years, the federal funds rate converges to 22.0 percent in 238 (240) years, and the bond rate converges to 23.7 percent in 335 (338) years. When we start with initial conditions in 2013, the long-run values are the same but the years to convergence are a bit longer because we start further from the long-run values.

Table 1B

Credibility Model

	GDP_{t+3}	CPI_{t+3}	RS_{t+3}	RL_{t+3}
GDP_{t-1}	–0.03*	0.04*	0.31*	–0.06*
CPI_{t-1}	–0.85	0.21	–0.05	–0.17
RS_{t-1}	–0.20	0.17*	0.52	0.13*
RL_{t-1}	0.63	–0.06*	0.25	0.79*
Constant	1.66	1.94	0.12	1.17
Adjusted R^2	0.32	0.16	0.68	0.76
SE equation	1.36	0.98	1.37	1.07
Mean dependent	2.18	3.15	5.27	6.66
SD dependent	1.64	1.07	2.42	2.20
Included observations	100			
Sample	1983:Q1–2007:Q4			

Residual correlation matrix (SE on diagonal)

	GDP	CPI	RS	RL
GDP	1.36	—	—	—
CPI	–0.07	0.98	—	—
RS	0.54	0.39	1.37	—
RL	0.57	0.43	0.65	1.07

NOTE: RS, short-run (policy) rate; RL, long-run (bond) rate; SD, standard deviation; SE, standard error. * and ** indicate significance at the 10 percent and 5 percent levels, respectively.

These long-run values suggest that the *No Credibility* policy regime is headed toward the type of hyperinflation that has occurred in third-world countries. Such a policy regime is politically unsustainable. Either the government changes the policy or the people change the government. In the United States, this policy regime spanned less than two decades. Political pressure from home and abroad led the Federal Reserve to abandon this regime and adopt one with a credible inflation policy (see Lindsey, Rasche, and Orphanides, 2013).

Credibility Scenario: 1983-2007

In October 1979, the Federal Reserve, under then-Chair Paul Volcker, adopted a new policy based on targeting the money supply to restore price stability. This new procedure lasted three years during which interest rates were very high and volatile. At the end of the three years, the inflation rate had fallen from double digits in 1980 to around 3 percent at the end of 1982. The Fed then switched from targeting the money supply to an indirect form of targeting the interest rate. By the time Alan Greenspan became the Fed chair in June 1987, the Fed had gained credibility for its inflation policy. The period of low inflation and credible monetary policy was accompanied by dramatic changes in the relationship between the policy rate

Table 1C

ZIRP Model

	GDP_{t+3}	CPI_{t+3}	RS_{t+3}	RL_{t+3}
GDP_{t-1}	0.01	0.22*	0.05**	0.02**
CPI_{t-1}	−1.07	0.27	0.06**	−0.05*
RS_{t-1}	−0.31	−0.28	0.02*	0.41
RL_{t-1}	0.29	0.47	0.16**	0.42
Constant	0.96	−1.05	−0.18	0.77
Adjusted R^2	0.22	0.31	0.60	0.63
SE equation	1.37	0.73	0.14	0.34
Mean dependent	1.49	0.06	0.20	1.68
SD dependent	1.55	0.87	0.22	0.56
Included observations	51			
Sample	1995:Q1–2007:Q4			

Residual correlation matrix (SE on diagonal)

	GDP	CPI	RS	RL
GDP	1.37	—	—	—
CPI	−0.11	0.73	—	—
RS	0.20	0.44	0.14	—
RL	0.42	−0.02	0.10	0.34

NOTE: RS, short-run (policy) rate; RL, long-run (bond) rate; SD, standard deviation; SE, standard error. * and ** indicate significance at the 10 percent and 5 percent levels, respectively.

and the bond rate, as shown in Figure 2. Note the contrast from the earlier period: The CPI inflation trend stabilized at about 3 percent rather quickly, but the trend in interest rates fell only gradually as inflation expectations lagged behind the actual decline in the inflation rate.

An interesting event occurred after September 2, 1992: The FOMC, worried about low job growth in a slow recovery, decided to set the policy rate at 3 percent, a rate approximately equal to the perceived trend in inflation. It was believed that such a low interest rate would cause higher inflation and, in October 1993, the bond rate began to rise from a low of 5.3 percent. The FOMC began to raise the policy rate in February 1994 but did not need to raise it above the bond rate to end this brief inflation scare.[7] The policy rate was raised to 6 percent in early 1995, but by then the bond rate had already begun to retreat from its peak at just under 8 percent in November 1994. On a 12-month moving average basis, the CPI inflation rate peaked at 2.9 percent in August 1994. During this entire episode, there were only a few instances when the policy rate was as high as the bond rate. On average, over the 1983-2007 sample period, the policy rate was 1.6 percentage points below the bond rate.

Table 1B presents the estimates of our *Credibility* model. For this period, the best lag structure in the VAR is also one quarter. The standard errors in the per capita GDP growth, infla-

Table 2

Long-Run Properties of Forecasting Models

	GDP	CPI	RS	RL
No Credibility				
Long-run value	−6.7	34.5	22.0	23.7
Years to convergence from 2007 initial conditions	222	283	238	335
Years to convergence from 2013 initial conditions	224	286	240	338
Credibility				
Long-run value	1.5	2.9	3.4	4.8
Years to convergence from 2007 initial conditions	8	12	21	15
Years to convergence from 2013 initial conditions	19	24	24	26
ZIRP				
Long-run value	1.5	−0.1	0.1	1.5
Years to convergence from 2007 initial conditions	9	5	7	6
Years to convergence from 2013 initial conditions	7	5	2	3

NOTE: Long-run values are percent annual rates. RS, short-run (policy) rate; RL, long-run (bond) rate.

tion, policy rate, and bond rate equations are 1.36, 0.98, 1.37, and 1.07 percent, respectively. Note that these standard errors are slightly smaller than those in the *No Credibility* model for the per capita GDP growth, inflation, and the policy rate equations but are actually larger for the bond rate equation. The reduction in uncertainty associated with the *Credibility* model stems largely from the much-improved properties of the long-run trends.

The middle panel of Table 2 presents the long-run mean forecasts for the *Credibility* model starting from initial conditions in 2007 and 2013. In the model starting from 2007 (2013) initial conditions, the per capita GDP growth rate converges to 1.5 percent in 8 (19) years, the infla- tion rate converges to 2.9 percent in 12 (24) years, the policy rate converges to 3.4 percent in 21 (24) years, and the bond rate converges to 4.8 percent in 15 (26) years. The key difference between the *Credibility* model and the *No Credibility* model is not in the short-run volatility but rather in the long-run trends for interest rates and inflation. Inflation and interest rates converge toward much lower values when policy is credible than when it is not.

ZIRP Scenario: 1995-2007

Our third scenario is an environment in which the policy rate is held at or near zero for an extended period. From the public's point of view, the monetary policy regime changes from having credibility for a 2 percent inflation target to also having credibility for promising to keep the policy rate near zero. A problem can arise if this promise remains in effect during the eco- nomic recovery. In this case, real returns are expected to be positive if the economy recovers.

In any equilibrium, the Fisher equation must hold: That is, the nominal interest rate must equal the real return plus the expected inflation rate. If the central bank holds the nominal

Figure 2

United States with Fed Credibility (1983-2013)

Percent Annual Rate

Percent

NOTE: Inflation is measured as the change over the previous year.

SOURCE: Data from FRED®, Federal Reserve Economic Data, Federal Reserve Bank of St. Louis: *Consumer Price Index for All Urban Consumers: All Items [CPIAUCSL]*; U.S. Bureau of Labor Statistics; http://research.stlouisfed.org/fred2/series/CPIAUCSL; accessed January 1, 2015. *10-Year Treasury Constant Maturity Rate [GS10]*; Board of Governors of the Federal Reserve System; http://research.stlouisfed.org/fred2/series/GS10; accessed January 1, 2015. *Effective Federal Funds Rate [FEDFUNDS]*; Board of Governors of the Federal Reserve System; http://research.stlouisfed.org/fred2/series/FEDFUNDS; accessed January 1, 2015.

interest rate (federal funds rate/policy rate) at zero while the economy is recovering, equilibrium dynamics will exert downward pressure on inflation. Over extended periods, a ZIRP is not consistent with a positive inflation target. The two policy objectives can persist only if real returns continue to be negative.[8]

A ZIRP can be a trap if inflation is below target, the economy is recovering, and policymakers believe that promising to hold interest rates low in the future will raise inflation. The 2 percent inflation target is not consistent with the zero interest rate (Japan's –0.1 percent inflation trend appears to be compatible with the *ZIRP* regime). In a growing economy, the *ZIRP* regime will lead to negative inflation. Policymakers will not want to raise interest rates because many believe that even small increases can have large negative effects on the real economy. For a good example of this belief applied to the Japanese experience cited here, see Ito and Mishkin (2004), who describe the hike in the policy rate from 2 basis points to 25 basis points in August 2000 as a "clear mistake." This hike occurred as many of the Japanese policymakers wanted a return to normalcy. Economic news had been positive—but not conclusive—leading to a typical hawks-versus-doves debate. Ito and Mishkin (2004, p. 146) write,

> Almost as soon as the interest rate was raised in August [2000], the Japanese economy entered into a recession. It was not known at the time, but the official date for the peak of the business cycle turned out to be October 2000. The growth rate of 2000:III turned negative, which was offset to some extent by a brief recovery in 2000:IV. But, as the economy turned into a recession, the criticism of the Bank of Japan's actions became stronger.

This sort of narrative, which is common in the financial press, has a chilling effect on any attempt to raise interest rates before the central bank is certain that the economy has reached full employment. In fact, there is no empirical evidence that such small changes in the money market rate have any measurable or sustainable effect on the real economy.[9] Moreover, every recovery is associated with uncertainty and fluctuations in news that drive observers from pillar to post. One day in the economic news there are optimistic reports about the recovery, and the next day there are worries that the economy will slide back into a recession. Such worries keep the policy rate at zero.

Since the United States has no earlier period with such a *ZIRP* regime, we use data from the Japanese economy for the 1995:Q1–2007:Q4 period to estimate the *ZIRP* model.[10] Figure 3 shows the Japanese experience with CPI inflation and interest rates. The inflation rate, although slightly negative on average, appears to fluctuate around zero. In October 1995, the Japanese bond rate was 4.6 percent but fell quickly to 2 percent and continued to drift even lower after 2000. The policy rate had been set low, at 2.25 percent, to try to stimulate the economy in 1995. However, further economic weakness led the Bank of Japan to lower the rate to ½ percent by the end of the year and to nearly zero in 1999 (the beginning of the official *ZIRP* policy). Although there have been periods when the rate was raised slightly, bad incoming news about the economy and slightly negative inflation eventually led the Japanese policymakers to lower the rate back near zero.

Table 1C shows the *ZIRP* model estimates. For the 1995:Q1–2007:Q4 period in Japan, the best lag structure is, again, just one quarter. The standard errors in the per capita GDP growth, inflation, policy rate, and bond rate equations are 1.37, 0.73, 0.14, and 0.34 percent, respectively.

Figure 3

Japan with a ZIRP (1995-2013)
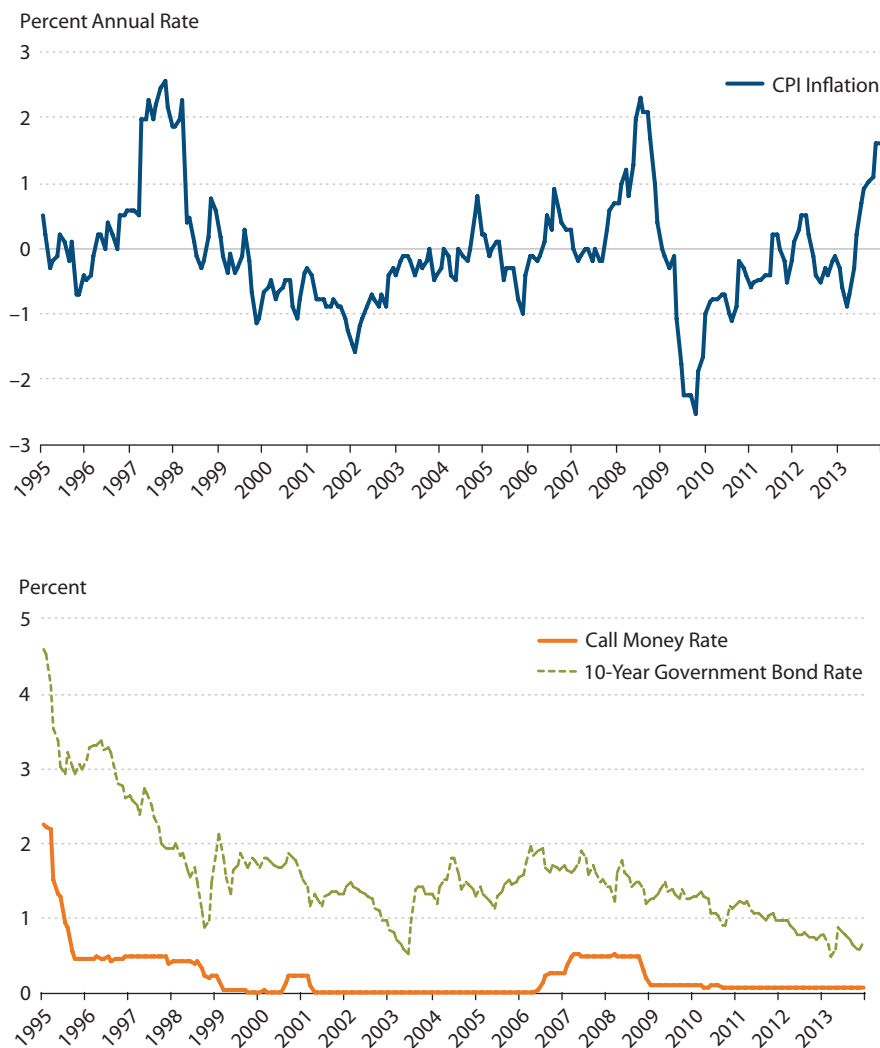
Percent Annual Rate

Percent

NOTE: Inflation is measured as the change over the previous year.

SOURCE: Data from FRED®, Federal Reserve Economic Data, Federal Reserve Bank of St. Louis: *Consumer Price Index of All Items in Japan [JPNCPIALLMINMEI]*; Organisation for the Economic Co-operation and Development; http://research.stlouisfed.org/fred2/series/JPNCPIALLMINMEI; accessed January 1, 2015. *Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for Japan [IRLTLT01JPM156N]*; Organisation for Economic Co-operation and Development; http://research.stlouisfed.org/fred2/series/IRLTLT01JPM156N; accessed January 1, 2015. *Immediate Rates: Less than 24 Hours: Call Money/Interbank Rate for Japan [IRSTCI01JPM156N]*; Organisation for Economic Co-operation and Development; http://research.stlouisfed.org/fred2/series/IRSTCI01JPM156N; accessed January 1, 2015.

The standard error in the per capita GDP growth equation is approximately the same across all three regimes although slightly higher in the *No Credibility* regime. The standard error in the inflation equation is lower when the central bank follows the *Credibility* regime and lowest in the *ZIRP* regime. The biggest differences are in the interest rate equations. The standard error in the policy rate equation falls from 1.79 to 1.37 moving from the *No Credibility* regime to the *Credibility* regime (see Tables 1A and 1B) and then to nearly zero, 0.14, in the *ZIRP* regime. The standard error in the bond rate equation actually rises from 0.62 in the *No Credibility* regime to 1.07 in the *Credibility* regime, but then falls to 0.34 in the *ZIRP* regime.

The bottom panel of Table 2 presents the long-run mean forecasts for the *ZIRP* model starting from initial conditions in 2007 and 2013. In the model starting from 2007 (2013) initial conditions, the per capita GDP growth rate converges to 1.5 percent in 9 (7) years and the inflation rate converges to –0.1 percent in 5 (5) years. The policy rate converges to 0.1 percent in 7 (2) years, and the bond rate converges to 1.5 percent in 6 (3) years.

PROJECTING INTEREST RATES IN THE POST-CRISIS ECONOMY: 2008-13

We use the long-run properties of our three times-series models to show the implications of the alternative policy regimes for per capita GDP growth, inflation, and interest rates during the period following the financial crisis. We start at the beginning of 2008 because the housing crisis was already underway. The FOMC officially adopted the *ZIRP* on December 16, 2008, when it set the policy rate target at 0 to 0.25 percent.

For each regime, we calculate dynamic stochastic forecasts using 10,000 draws of random shocks for the 2008:Q1–2013:Q4 period. We calculate the median forecast and the standard error for each quarter. In Figure 4, the median forecast is displayed as a solid red line and confidence bands of ±1 standard deviation are shown as dotted blue lines. The actual values of the predicted variables are shown as black dashed lines. Each column represents a policy regime with four rows representing per capita GDP growth, CPI inflation, the policy rate, and the bond rate.

The top row in Figure 4 shows that none of the models could predict the deep 2007-09 recession. The *No Credibility* model fails miserably for per capita GDP growth. One reason for this failure is that the trend in per capita output growth was declining throughout the *No Credibility* period and the downward trend continues into negative territory in the long run. The *ZIRP* model does the best job of predicting the downturn but still misses the negative growth in 2009. Both the *Credibility* and *ZIRP* models predict that per capita GDP growth will converge to 1.5 percent in the long run. The *ZIRP* model has the tightest confidence bands.

As shown in the second row in Figure 4, none of the models predicted the inflation decline during the recession. The *No Credibility* model predicts rising and increasingly volatile inflation. The *Credibility* model converges to a 2.9 percent long-run inflation trend and predicts too much inflation during this period. The *ZIRP* model, on the other hand, predicts that CPI inflation will converge toward zero and predicts too little inflation.

Figure 4

Out-of-Sample Forecasts for the Three Monetary Policy Regimes: 2008-13
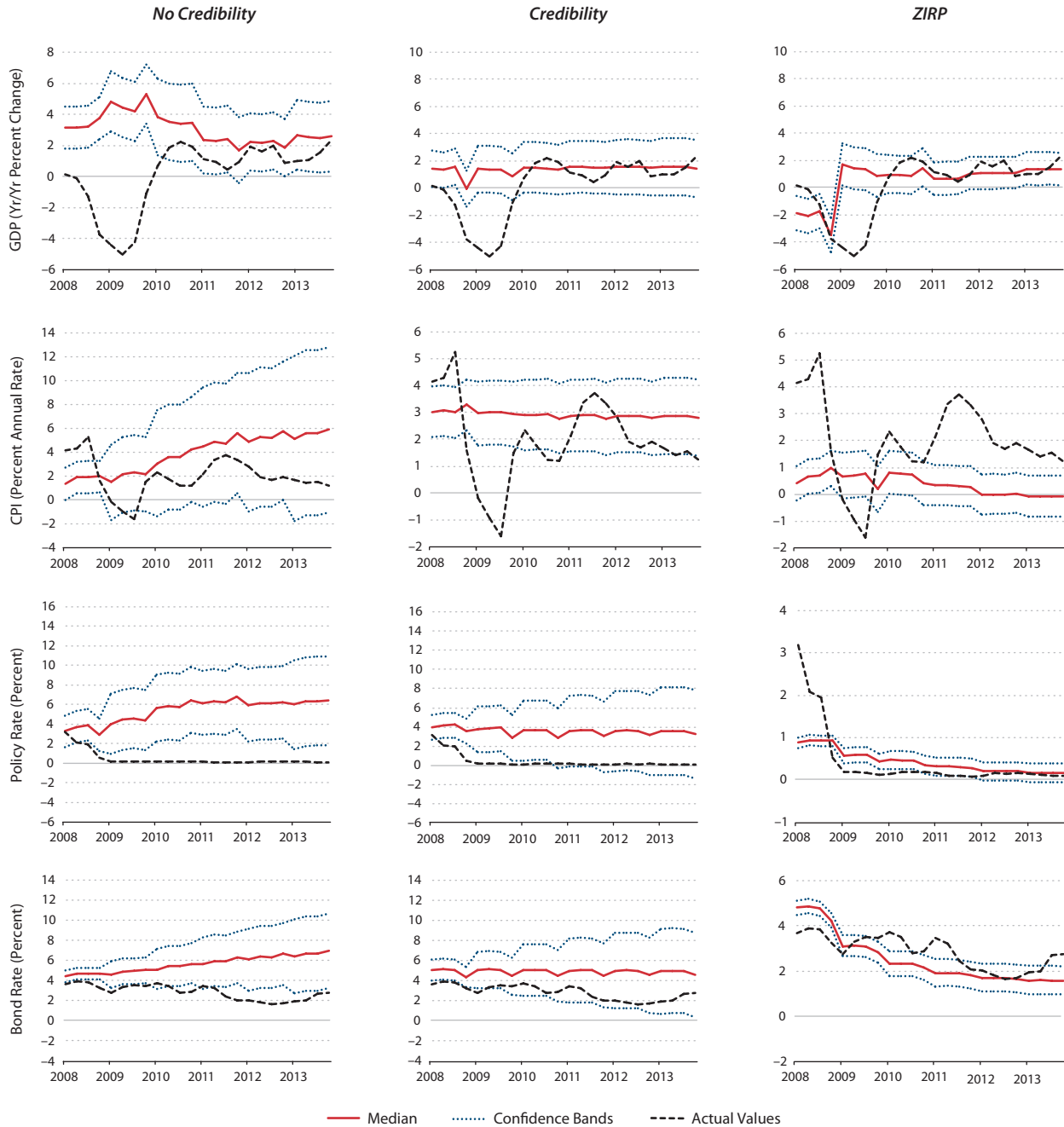
NOTE: The solid red line is the median forecast in 10,000 simulations; the dotted blue lines are the median ± 1 standard deviation; the dashed black line is the actual value.

SOURCE: Authors' calculations.

Table 3

Accuracy of Forecasts

	Median error			RMSE		
	No credibility	*Credibility*	*ZIRP*	*No credibility*	*Credibility*	*ZIRP*
GDP	−2.99	−1.28	−0.46	4.05	2.76	2.20
CPI	−1.89	−0.97	1.60	5.57	2.10	2.22
RS	−4.90	−3.15	0.01	6.11	4.66	0.47
RL	−2.77	−2.05	0.31	3.68	3.55	0.95

The third row in Figure 4 shows the forecasts for the policy rate. The *No Credibility* model predicts a high, rising, and volatile policy rate. The *Credibility* model shows the policy rate converging to a 3.4 percent trend with widening confidence bands. The *ZIRP* model, as expected, is "spot on" with a nearly perfect forecast over the past five years. All the "miss" in this model occurs in 2008, as the rate converges toward zero at the end of the year.

In the fourth row in Figure 4, the *No Credibility* model predicts a high, rising, and volatile bond rate, just as it does for inflation and the policy rate. The *Credibility* model predicts the bond rate will converge to 4.8 percent, but the confidence bands continue to widen as the forecast horizon lengthens. The actual bond rate stays below the median forecast but generally within 1 or 2 percentage points. The biggest surprise to us in this figure is the bond rate forecast from the *ZIRP* model. Here the bond rate forecast is, on average, below the actual rate, but the mean error is small relative to the other forecasts and the model correctly predicts the falling trend.

Table 3 shows the average of the quarterly median and root mean squared error (RMSE) statistics. Although the *Credibility* model appears to provide a reasonable outlook for per capita GDP growth and the best forecast for inflation, it loses dramatically to the *ZIRP* model in a comparison of interest rate forecasts. Historically, uncertainty in bond markets has been driven mainly by uncertainty about inflation expectations.[11] We expected the results from the inflation forecast to be more strongly reflected in performance of the bond rate forecast. The *ZIRP* model appears to tie down both the long and short rates. The visual evidence is shown clearly in Figure 5, which compares the Blue Chip long-range forecast of the U.S. 10-year government bond rate with the 6-year out-of-sample forecast from the *ZIRP* model. The Blue Chip long-range forecast for the bond rate is consistent with the 4.8 percent trend predicted by the *Credibility* model. The *ZIRP* model is always within 2 standard deviations of the actual rate; the long-run implication is that the bond rate will converge to a record low 1.5 percent if the Fed does not exit the *ZIRP* regime.

Forecasting Interest Rates in the Transition to Normalcy: 2014-19

The market and Fed policymakers expect to begin exiting the *ZIRP* regime in mid-2015. They plan to return to the *Credibility* regime that characterized policy from 1983 to 2007. For our purpose, we begin the simulations at the beginning of 2014. The forecasts look very similar

Figure 5

Blue Chip Consensus Forecasts Versus *ZIRP* Forecasts: 10-Year Treasury Rate (2008-13)
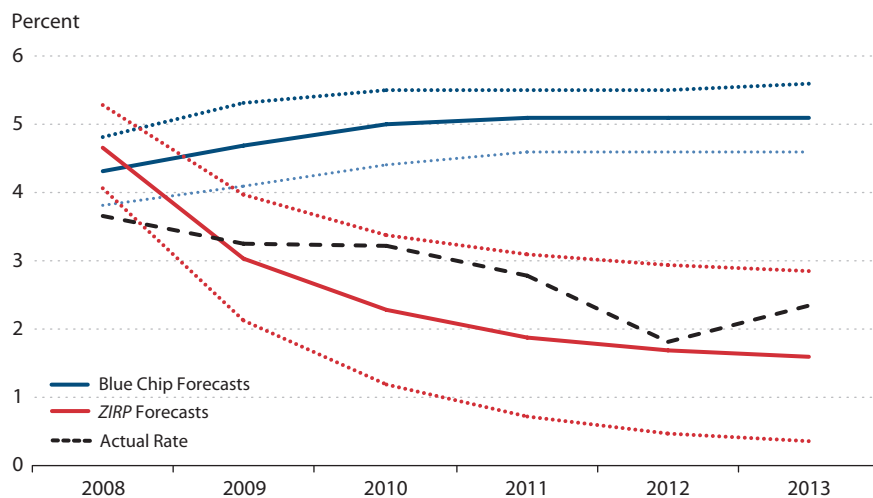
NOTE: For the Blue Chip forecasts, the solid line shows the consensus and the dotted lines show the top 10 and bottom 10 forecasts. For the *ZIRP* model, the solid line is the median forecast and the dotted lines are ±1 standard deviation. The dashed black line shows the actual data. The *ZIRP* model is estimated using Japanese data from 1995:Q1 through 2007:Q4. All forecasts are out-of-sample. Blue Chip forecasts are as reported on December 1, 2007. We thank Yi Wen for suggesting this figure.

SOURCE: Blue Chip Financial Forecasts published on December 1, 2007, and authors' calculations.

to those in Figure 4, with slightly different initial conditions. The important comparison is between the *ZIRP* and *Credibility* regimes. Nevertheless, we also report statistics for the *No Credibility* regime.

The path of interest rates during the transition to normalcy matters for many reasons. One important concern for the Fed is the effect that interest rates will have on the Fed's interest income and expenses during the transition to normalcy. Carpenter et al. (2013) provide the institutional details and simulations of the transition under alternative interest rate assumptions. Their baseline path for interest rates is based on the Blue Chip Consensus forecast reported in the December 2012 release. Under Carpenter et al.'s high interest rate scenario, they assume the paths for the policy rate and the bond rate are 1 percentage point above the Blue Chip Consensus forecast.[12] Figure 6 shows updated versions of the Blue Chip interest rate forecasts used by Carpenter et al. (2013) in their simulations. For each variable, we define the Blue Chip benchmark (BCB) forecast as the updated forecast plus 1 percentage point. In this section we ask the following questions: "How likely is it that the policy and bond rates will exceed their respective BCB forecast paths in each quarter of our six-year simulation period?" and "How often does the policy rate exceed the bond rate in each quarter of the 10,000 simulations for our three models?"

Figure 6

Blue Chip Interest Rate Forecasts (2014-19)

Federal Funds Rate

Percent

10-Year Government Bond Rate
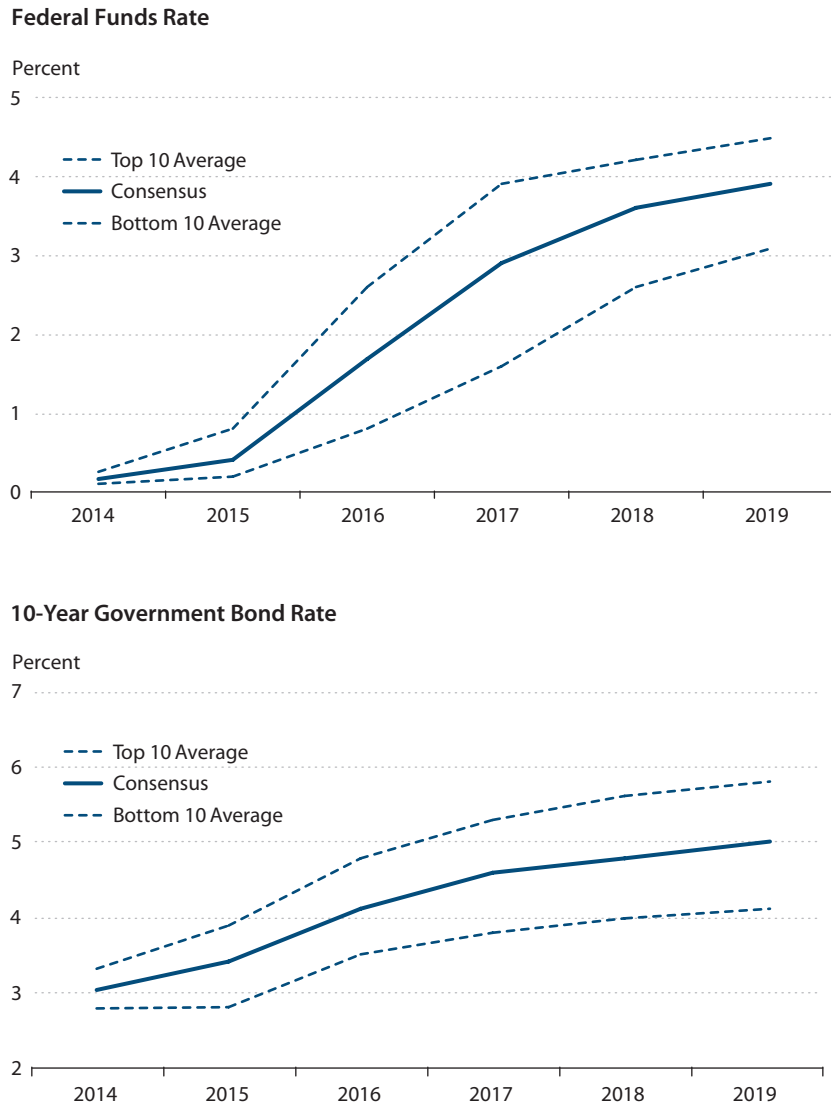
Percent

SOURCE: Blue Chip Financial Forecasts published on December 1, 2013.

Our forecasts start with actual conditions at the end of 2013. Figure 7 illustrates the answers to our questions. The top panel shows the percentage of times that the policy rates predicted by the models were above the BCB forecast for the policy rate. There is a significant likelihood that the policy rates forecasts from both the *Credibility* and *No Credibility* models will exceed the BCB forecasts in the first two years after the transition begins. For the *No Credibility* model (the blue line), the likelihood reaches a maximum of about 90 percent after three years and then declines to about 57 percent by 2019. For the *Credibility* model (the red line), with its lower trend, the likelihood averages more than 50 percent in the first two years but stabilizes around 30 percent by 2019. For the *ZIRP* model (the green line), the policy rate forecast never exceeds the BCB forecast.

The middle panel of Figure 7 shows results for the 10-year government bond rate. The likelihood of the bond rate forecast from the *No Credibility* model exceeding the BCB for the bond rate is quite low in the first year: below 10 percent. The likelihood rises gradually to 50 percent by the end of 2017. The pattern for the *Credibility* model is similar. The likelihood remains below 20 percent in 2014 and rises gradually to 30 percent by 2019. The bond rate forecasted by the *ZIRP* model never exceeds the BCB.

The bottom panel of Figure 7 plots the likelihood that the policy rate will be higher than the bond rate. In the *No Credibility* model, the yield curve is inverted (the policy rate exceeds the bond rate) much of the time in the second through the fourth years of the transition to normalcy. In the *Credibility* model, the likelihood is much lower, especially in the first year when it is at or below 10 percent. After the second year, the probability of an inverted yield curve rises to a range of 20 to 25 percent. The policy rate is almost never above the bond rate in the *ZIRP* simulations. During the last three years of the simulations, the number is positive, rising only as high as 19 out of 10,000 simulations in the final year.

Forecast Uncertainty

Our three policy regimes, *No Credibility*, *Credibility*, and *ZIRP*, were chosen to reflect the different concerns of policymakers regarding the transition to normalcy. The biggest source of uncertainty involves predicting which model will be the right one. A risk-averse decision-maker will consider the risks involved in a variety of likely outcomes.

We use VAR models to forecast inflation and interest rates. Analysis of forecasts including the 1970s and early-1980s data indicates the VAR forecasts were generally less accurate than more sophisticated forecasts that combined the forecaster's judgment with forecasts from a large econometric model. However, these "more sophisticated" forecasts typically involved periods less than two years into the future. For these short horizons, McNees (1986, 1990) finds that while VAR models perform relatively well for some variables, they do not perform well for inflation. He reports that the RMSEs were almost twice as large for the VAR forecasts as for the professional forecasters. The VAR interest rate forecasts for the 3-month Treasury bill were about 33 percent larger than the large-model alternatives.

Reifschneider and Tulip (2008) report that the RMSEs for CPI inflation forecasts for the 1986-2006 period cluster around 1 percent for horizons of 2 to 4 years. These forecasts are tethered to the official forecasts—implicit objectives—of the Fed and the government. This

Figure 7

The Likelihood of High Interest Rates: 2014-19

Policy Rate: Likelihood that Model Forecast > BCB Plus 1 Percent?
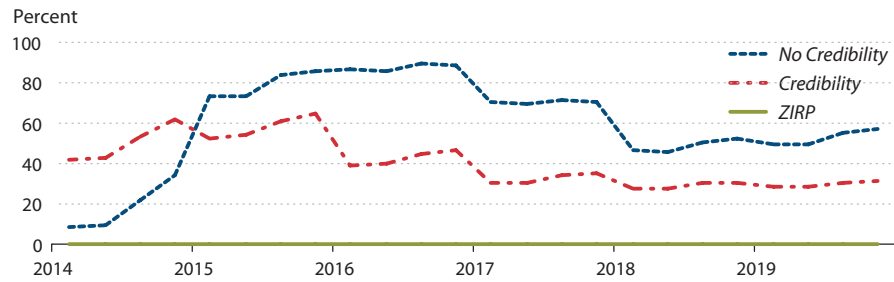
Bond Rate: Likelihood that Model Forecast > BCB Plus 1 Percent?
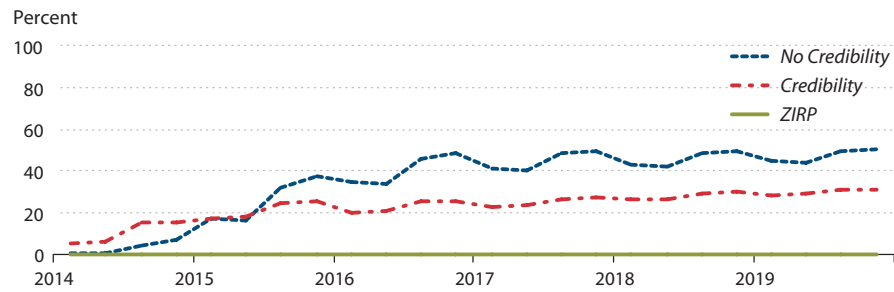
Likelihood that Policy Rate > Bond Rate?
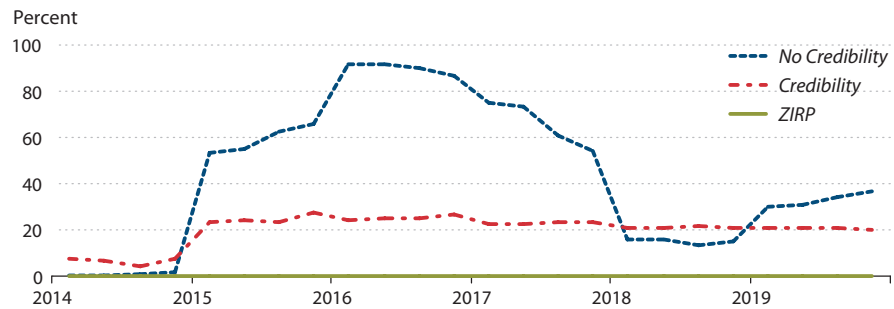
SOURCE: Authors' calculations.

level is less than our estimated uncertainty for the *Credibility* models at a 2-year horizon and much less than the uncertainty associated with the *No Credibility* models. Even in the *Credibility* regime, we do not constrain the time-series forecast to the official inflation objective. Although we know that the VAR forecasts are more disperse than typical economic forecasts, they have the advantage of simple construction and easy replication. Furthermore, we are mainly interested in characterizing the alternative regimes and comparing the relative uncertainty across them.

CONCLUSION

Our first finding is that the *ZIRP* should be treated as a separate regime with different statistical processes than observed in the United States during the Great Moderation. We assume that the *ZIRP* could be adequately modeled using Japanese data from 1995 through 2007. Our most startling result is that the *ZIRP* model, estimated using Japanese data, does the best job forecasting the U.S. data from 2008 through 2013.

A second finding is that the *ZIRP* regime leads to low and stable long-term bond rates and lower-than-expected inflation. The *ZIRP* model underpredicts inflation, which we attribute to the fact that policymakers and markets expect the FOMC to return to the *Credibility* model with a 2 percent inflation target. As noted, the 2 percent inflation target is not consistent with a *ZIRP* regime and the longer the FOMC maintains this regime, the farther the trend inflation rate will fall below the target.

A third finding is that the *No Credibility* model has terrible implications for the post-crisis period. The time-series properties of this regime strongly recommend against this model as a policy choice. The bad economic outcomes of this regime make it imperative for policymakers to take special care to avoid it. Worldwide, this policy regime has been observed in countries that lose control of their federal government budget process. When they lose the ability to curb spending or raise taxes, such governments print money to pay for spending.

Our fourth, and perhaps less obvious, finding is that any attempt to return to the *Credibility* regime will likely involve higher and more volatile interest rates, reminiscent of the volatility during the taper tantrum of May and June 2013 when then-Fed Chair Ben Bernanke announced that the Fed would gradually slow its large-scale purchases of long-term securities. Our analysis suggests that lifting off the zero lower bound will involve a period of heightened uncertainty about interest rates at both short- and long-term horizons.

We do not draw any firm conclusions from these experiments about the effects of the *ZIRP* on the real economy. In our models, the per capita GDP growth rate converges to 1.5 percent at an annual rate in both the *Credibility* and the *ZIRP* models. Our main concern is that uncertainty about which regime the economy will converge to creates a headwind that keeps the economy operating below its efficient level. A decision to adopt the *ZIRP* model should be accompanied by an explicit decision to allow inflation to run at or below zero percent, as the Japanese have done. Our analysis suggests that their recent decision to adopt a 2 percent inflation target is doomed to fail if they are not willing to raise interest rates to some normal level that is approximately equal to the sum of the inflation target and per capita real GDP growth.

The problem is that over time, if the central bank fixes the nominal interest rate and allows real factors to determine the real interest rate, then according to the Fisher equation, inflation will adjust to clear the bond market.

From the point of view of money and bond markets, the FOMC has been replicating Japan's *ZIRP* regime. The only circumstance in which future interest rates are not likely to be a problem is if the *ZIRP* policy is the new normal. In our simulations, the policy rate exceeded the bond rate about 20 to 25 percent of the time in the *Credibility* regime. In the *ZIRP* model, the yield curve was almost never inverted. If normalization is, as planned, a return to the *Credibility* model with a historically "normal"-sized balance sheet for the Fed, then one should plan for a scenario in which higher interest rates will complicate the normalization process. ∎

APPENDIX

Sources and Definitions of Data

	Inflation	Output growth	Policy rate	Bond rate
United States				
Series	Consumer Price Index for All Urban Consumers: All Items	Real Gross Domestic Product Per Capita	Effective Federal Funds Rate	10-Year Treasury Constant Maturity Rate
Series ID in FRED®	CPIAUCSL	A939RX0Q048SBEA	FEDFUNDS	GS10
Source	U.S. Bureau of Labor Statistics	U.S. Bureau of Economic Analysis	Board of Governors of the Federal Reserve System	Board of Governors of the Federal Reserve System
Units	Percent change from year ago	Percent change from year ago	Percent	Percent
Japan				
Series	Consumer Price Index of All Items in Japan	Real Gross Domestic Product in Japan Working Age Population: Aged 15-64: All Persons for Japan	Immediate Rates: Less than 24 Hours: Call Money and Interbank Rate for Japan	Long-Term Government Bond Yields: 10-year: Main (Including Benchmark for Japan)
Series ID in FRED®	JPNCPIALLQINMEI	JPNRGDPQDSNAQ LFWA64TTJPQ647S	IRSTCI01JPM156N	IRLTLT01JPQ156N
Source	Organisation for Economic Co-Operation and Development	Organisation for Economic Co-Operation and Development	Organisation for Economic Co-Operation and Development	Organisation for Economic Co-Operation and Development
Units	Percent change from year ago	Percent change from year ago	Percent	Percent

NOTE: All series are reported quarterly.

NOTES

[1] Note that the Census Bureau projects that the U.S. population will grow at an average annual rate of 0.8 percent from 2015 to 2021, implying that FOMC members project that per capita output will grow between 1.4 and 2.2 percent. See Table 1 at http://www.census.gov/population/projections/data/national/2014/summarytables.html.

[2] See, for example, the arguments by Calomiris (2012) in his comments on Campbell et al. (2012). More recently, Meltzer (2014) explains the high-inflation risk posed by the Fed's balance sheet policy.

[3] See Gavin and Kydland (1999) for evidence comparing the properties of nominal and real time-series data before and after the Volcker monetary policy reform. They show that the change in monetary policy regime had a statistically significant impact on the time-series properties of the nominal data (prices, money, and velocity), but not on the real quantities (output, consumption, investment, and hours worked).

[4] See the appendix for a more detailed description of data sources.

[5] See Nelson (2004), who explains why many economists and policymakers during this period did not consider it important for the Fed to put much emphasis on the goal of price stability.

[6] The model has converged when it is less than one-tenth of a percentage point from the long-run value.

[7] See Goodfriend (1993) for an essay on inflation scares.

[8] See Bullard (2010) for a survey of economic theories that show how an economy can become "trapped" at the zero lower bound. See Cooke and Gavin (2014, pp. 8-9) for an introductory discussion of the Fisher equation.

[9] For evidence of the contrary—showing weak empirical links among interest rates, inflation, and the real economy—see Staiger, Stock, and Watson (1997) and Stock and Watson (2003). For a good explanation about how beliefs about such relationships may drive shifts in the policy regime, see Cho, Williams, and Sargent (2002).

[10] We did not investigate the possibility that U.S. data from the 1930s may fit this definition of a *ZIRP* regime.

[11] See Gallmeyer et al. (2007).

[12] See Figure 6. This is the high interest benchmark used by Carpenter et al. (2013) when simulating alternative exit strategies. On page 26, they write "Although this shock—particularly the parallel shift—is an unlikely outcome, we present it to show the interest rate sensitivity of the portfolio."

REFERENCES

Bullard, James. "Seven Faces of 'The Peril.'" Federal Reserve Bank of St. Louis *Review*, September/October 2010, *92*(5), pp. 339-52; http://research.stlouisfed.org/publications/review/10/09/Bullard.pdf.

Calomiris, Charles W. "Comments and Discussion." *Brooking Papers on Economic Activity*, Spring 2012, pp. 55-63.

Campbell, Jeffrey R.; Evans, Charles L.; Fisher, Jonas D.M. and Justiniano, Alejandro. "Macroeconomic Effects of Federal Reserve Forward Guidance." *Brooking Papers on Economic Activity*, Spring 2012, pp. 1-54.

Carpenter, Seth B.; Ihrig, Jane E.; Klee, Elizabeth C.; Quinn, Daniel W. and Boote, Alexander H. "The Federal Reserve's Balance Sheet and Earnings: A Primer and Projections." Finance and Economics Discussion Series No. 2013-01, Federal Reserve Board, September 2013; http://www.federalreserve.gov/pubs/feds/2013/201301/201301pap.pdf.

Cho, In-Koo; Williams, Noah and Sargent, Thomas J. "Escaping Nash Inflation." *Review of Economic Studies*, January 2002, *69*(1), pp. 1-40; http://restud.oxfordjournals.org/content/69/1/1.full.pdf+html.

Cooke, Diana A. and Gavin, William T. "The Ups and Downs of Inflation and the Role of Fed Credibility." Federal Reserve Bank of St. Louis *Regional Economist*, April 2014, pp. 5-9; https://www.stlouisfed.org/publications/regional-economist/april-2014/the-ups-and-downs-of-inflation-and-the-role-of-fed-credibility.

Gallmeyer, Michael F.; Hollifield, Burton; Palomino, Francisco J. and Zin, Stanley E. "Arbitrage-Free Bond Pricing with Dynamic Macroeconomic Models." Federal Reserve Bank of St. Louis *Review*, July/August 2007, *89*(4), pp. 305-26; http://research.stlouisfed.org/publications/review/07/07/Gallmeyer.pdf.

Gavin, William T. and Kydland, Finn E. "Endogenous Money Supply and the Business Cycle." *Review of Economic Dynamics*, April 1999, *2*(2), pp. 347-69.

Goodfriend, Marvin. "Interest Rate Policy and the Inflation Scare Problem: 1979-1992." Federal Reserve Bank of Richmond *Economic Quarterly*, Winter 1993, *79*(1), pp. 1-23; https://www.richmondfed.org/publications/research/economic_quarterly/1993/winter/pdf/goodfriend.pdf.

Ito, Takatoshi and Mishkin, Frederic S. "Two Decades of Japanese Monetary Policy and the Deflation Problem." NBER Working Paper No. 10878, National Bureau of Economic Research, November 2004; http://www.nber.org/papers/w10878.pdf.

Lindsey, David E; Orphanides, Athanasios and Rasche, Robert H. "The Reform of October 1979: How It Happened and Why." Federal Reserve Bank of St. Louis *Review*, March/April 2005, *87*(2 Part 2), pp. 187-235; http://research.stlouisfed.org/publications/review/05/03/part2/Lindsey.pdf.

Lucas, Robert E. Jr. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy*, January 1976, *1*(1), pp. 19-46.

McNees, Stephen K. "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts." *Journal of Business and Economic Statistics*, January 1986, *4*(1), pp. 5-15.

McNees, Stephen K. "The Role of Judgment in Macroeconomic Forecasting Accuracy." *International Journal of Forecasting*, October 1990, *6*(3), pp. 287-99.

Meltzer, Allan H. "How the Fed Fuels the Coming Inflation." *Wall Street Journal*, May 6, 2014; http://online.wsj.com/news/articles/SB10001424052702303939404579527750249153032.

Nelson, Edward. "The Great Inflation of the Seventies: What Really Happened?" Working Paper No. 2004-001, Federal Reserve Bank of St. Louis, January 2004; http://research.stlouisfed.org/wp/2004/2004-001.pdf.

Reifschneider, David and Tulip, Peter. "Gauging the Uncertainty of the Economic Outlook From Historical Forecasting Errors." Finance and Economics Discussion Series Working Paper No. 2007-60, Federal Reserve Board, August 2007; http://www.federalreserve.gov/pubs/feds/2007/200760/200760pap.pdf. August 2008 update; http://www.petertulip.com/Reifschneider_Tulip_2008.pdf.

Staiger, Douglas; Stock, James H. and Watson, Mark W. "The NAIRU, Unemployment and Monetary Policy." *Journal of Economic Perspectives*, Winter 1997, *11*(1), pp. 33-49; http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.11.1.33.

Stock, James H. and Watson, Mark, W. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature*, September 2003, *41*(3), pp. 788-829; http://pubs.aeaweb.org/doi/pdfplus/10.1257/002205103322436197.

A Measure of Price Pressures

Laura E. Jackson, Kevin L. Kliesen, and Michael T. Owyang

The Federal Reserve devotes significant resources to forecasting key economic variables such as real gross domestic product growth, employment, and inflation. The outlook for these variables also matters a great deal to businesses and financial market participants. The authors present a factor-augmented Bayesian vector autoregressive forecasting model that significantly outperforms both a benchmark random walk model and a pure time-series model. They then use these factors in an ordered probit model to develop the probability distribution over a 12-month horizon. One distribution assesses the probability that inflation will exceed 2.5 percent over the next year; they term this probability a price pressure measure. This price pressure measure would provide policymakers and markets with a quantitative assessment of the probability that average inflation over the next 12 months will be higher than the Fed's long-term inflation target of 2 percent. (JEL C32, C35, E31)

T he Federal Reserve, like most central banks, devotes considerable economic resources to monitoring and analyzing large volumes of economic data. This effort, often termed "current analysis" by insiders, feeds directly into another, crucial aspect of central banking: forecasting key economic series such as real gross domestic product (GDP) growth, inflation, and employment. Forecasting the paths of key economic variables is an effort that flows directly from the Fed's congressionally mandated responsibility to (i) provide sufficient liquidity to achieve and maintain low inflation rates and (ii) promote maximum sustainable economic growth. This responsibility, which stems from the Federal Reserve Act and subsequent amendments, is often termed the Fed's dual mandate. Since the passage of the Dodd-Frank Wall Street Reform and Consumer Protection Act, the Fed has been handed a third monetary policy responsibility: financial stability.

In this analysis, we focus on the Fed's price stability mandate—specifically, in the context of forecasting inflation. Given its importance, Federal Reserve officials have historically been reluctant to attach an explicit definition of price stability—a rather ambiguous term that can

mean different things to different people. That reluctance changed in January 2012, when the Federal Reserve defined price stability as a numerical inflation target—2 percent—over the medium term (Board of Governors of the Federal Reserve System, 2013):

> The Federal Open Market Committee (FOMC) judges that inflation at the rate of 2 percent (as measured by the annual change in the price index for personal consumption expenditures, or PCE) is most consistent over the longer run with the Federal Reserve's mandate for price stability and maximum employment. Over time, a higher inflation rate would reduce the public's ability to make accurate longer-term economic and financial decisions. On the other hand, a lower inflation rate would be associated with an elevated probability of falling into deflation, which means prices and perhaps wages, on average, are falling—a phenomenon associated with very weak economic conditions. Having at least a small level of inflation makes it less likely that the economy will experience harmful deflation if economic conditions weaken. The FOMC implements monetary policy to help maintain an inflation rate of 2 percent over the medium term.

The Fed's inflation-targeting regime, which is similar to those of many other major central banks, thus requires the FOMC to forecast future inflation ("inflation over the medium term"). But in a large structural model such as the Board of Governors FRB/US model, the inflation process is modeled largely on the New Keynesian Phillips curve (NKPC) framework. In the NKPC model, current inflation depends on both current economic conditions—typically measured as the deviation between actual output and potential output or, equivalently, between the current unemployment rate and the natural rate of unemployment—and agents' expectations of future inflation.[1] Previous shocks matter only to the extent that they influence current conditions or expectations of future inflation. The NKPC model thus marries the Keynesian view that there is a short-run trade-off between real output (or unemployment) and inflation (by means of some "sticky price" mechanism) and the neoclassical view that, in the long run, excess money growth only leads to higher inflation (money neutrality).

We take a different approach in our analysis. First, our framework uses a pure time-series model to forecast inflation. Simple time-series models have been shown to be as accurate as larger, more complex structural models—and the resource demands on the forecaster are significantly smaller.[2] Our model is a Bayesian vector autoregressive (VAR) model augmented with a set of factors that summarize disaggregated price, employment, and interest rate data. The set of factors is derived from approximately 100 economic and financial data series, including well-known measures of inflation expectations. We find, consistent with the NKPC, that inflation expectations matter. We use standard forecast accuracy tests to test whether our dynamic factor model produces a more accurate forecast than a simple, naive forecasting model (random walk) and a benchmark time-series model that forecasts future inflation based solely on lags of previous inflation. Finally, we use our dynamic model to produce forecast probabilities. For example, policymakers usually want to know whether the probability that inflation over the next four or eight quarters will exceed the Fed's 2 percent inflation target is greater or less than the probability that it will fall short of 2 percent.[3]

In our second exercise, we consider an alternative experiment in which we forecast the probabilities that the inflation rate will be in the target zone, rise above the target zone, be positive but fall below the target zone, or fall below zero. To do this, we construct a static ordered probit model with the appropriate cutoffs for the inflation rate. The model is augmented with

the same factors used for the linear model previously described. Finally, we aggregate the various horizons' forecasted probabilities that the inflation rate will rise above the target zone to form an index that measures price pressures.

In the next section, we summarize some of the previous work on inflation forecasting. Faust and Wright (2013, henceforth FW) provide an outstanding reference for the current line of thinking; we refer readers to their paper for details but provide a helicopter view of the extant literature. The following section contains our linear forecasting exercise: Our goal is to use a large number of data series to forecast inflation at various horizons. We then describe our alternative experiment: Our goal is not to forecast the level of the inflation rate but to determine the risk that the inflation rate will exceed the Fed's inflation target.

SUMMARIZING THE EXTANT LITERATURE

The problem of forecasting future inflation has been well studied. FW provide an extensive survey of the inflation-forecasting literature, and readers seeking a comprehensive overview of this literature are encouraged to read their paper. Here, we provide a cursory summary of FW, note their key findings, and supplement FW with additional literature where appropriate.

FW compare forecasts of inflation constructed from 16 different models popular in the literature. These include, but are not limited to, VARs; the integrated moving average (1,1) [IMA(1,1)] model advocated by Stock and Watson (2007, henceforth SW); the Atkeson and Ohanian (2001, henceforth AO) random walk model; various Phillips curve models; a dynamic stochastic general equilibrium (DSGE) model; and factor models. As a robustness check, they examine whether any of these model-based forecasts are superior to three "real-time judgmental forecasts." The first two forecasts are measures of consensus among professional forecasters (e.g., the Philadelphia Fed's Survey of Professional Forecasters [SPF] or the Blue Chip survey). The third measure is the Greenbook forecasts, compiled by the Board of Governors staff; Greenbook forecasts are available to the public with a minimum five-year lag.[4]

In general, FW present four key findings from their forecasting model comparison exercise. First, judgmental forecasts are usually the most accurate across a variety of inflation measures and time horizons. Taken literally, this means that there is a forecasting equivalent of the law of large numbers at work: The average of a large group of forecasters is a close approximation to the actual (expected) value. Second, forecasts beyond one or two quarters should have some method for capturing long-run trends in inflation. This means that inflation has a long-run trend. Importantly, this long-run trend is dependent on actions by the monetary authority. Third, more shrinkage of information tends to produce better results. By shrinkage, FW mean that the best forecasts rely on a good starting point, such as a nowcast.[5] This mechanism implies that there is value in current information when forecasting future inflation. The fourth principle, which is related to the third, is that the best forecasts have "heavy handed" priors about the local mean. The third and fourth principles are deemed boundary values. In the view of FW, the best forecast thus conditions on the starting point (a nowcast) and an ending point (such as the Fed's long-run inflation target). They term this a fixed "glide path" or "swoop path."

FORECASTING INFLATION

In this section, we perform an exercise similar to that of FW but more limited in scope. Because we are not interested in reinventing the wheel, we compare only a few models, focusing instead on the effect of adding data to the model. We focus on direct forecasts, although a similar exercise could be performed for indirect forecasts.[6] Each exercise is a quasi-out-of-sample evaluation of the forecasting performance of each model. We measure performance in this section the usual way, through the sum of the mean squared deviation of the forecast from its objective.

We sidestep two important issues. First, we do not use real-time data. Rudd and Whelan (2007) show how data revisions can significantly change the value of the initial parameters from a benchmark NKPC model originally estimated by Galí and Gertler (1999). Moreover, real-time measurement can be a significant issue for price series produced from national income accounts data, such as the PCE price index or the GDP price index. Second, we do not evaluate whether the forecasts are statistically significantly different from each other. The reason for our informality about these issues is that the following exercise has been essentially performed in FW, with only slight adjustments to the data. Here, we are simply interested in whether the addition of disaggregate price and wage measures improves the forecasting performance of the factor model. Aruoba and Diebold (2010), who use a Kalman filter framework to estimate an inflation index from six indicators, provide the closest antecedent to our approach. However, they do not use their index to forecast inflation. Instead, they view their inflation index as a coincident indicator to help policymakers or forecasters better determine whether inflation movements in real time are the product of demand- or supply-side shocks.

The Models

The objective is to forecast future inflation rates, π_{t+h}, using information available at time t, Ω_t. We use three models in this section, each of which is increasingly dependent on the data. The first model, our baseline model for comparison, is the random walk forecast that AO claim works well:

$$(1) \qquad\qquad \hat{\pi}_{t+h|t} = \pi_{t-1},$$

where $\hat{\pi}_{t+h|t}$ is the forecast of π_{t+h} and π_t is the current-period inflation rate. Here, inflation is solely a function of its own previous value. The random walk forecast takes advantage of the fact that trend inflation is persistent, but the short-term movements in inflation are transitory and difficult to predict. Second, we use a simple autoregressive model with lags of inflation, $A(L)$:

$$(2) \qquad\qquad \hat{\pi}_{t+h|t} = A(L)\,\pi_t,$$

In a sense, this model nests the AO random walk specification but adds (potential) mean reversion.

In the third model, we are interested in whether additional data can help forecast inflation at longer horizons. The conventional forecasting process of monetary policymakers typically uses structural models to obtain forecasts of a few key variables such as inflation, GDP, and the unemployment rate. These structural models often rely on theoretical restrictions and conditional policy paths. However, policymakers examine many other variables when making forecasts; for example, they may use these other data to judgmentally adjust the model-based forecasts. This process is known as ad-factoring the model to produce a forecast that, to a large extent, reflects the forecasters' or policymakers' biases. Thus, information about other economic indicators should, in principle, be useful in forecasting economic variables.

Our approach follows this framework but without the large structural model. A key problem is deciding which, if any, other series to include. One problem with using more data to construct the forecast is that the informational advantage of incorporating the additional data can be outweighed by the increased parameter uncertainty. Thus, more data do not always lead to better forecasts. This is particularly true for out-of-sample forecasting where the additional data lead to overfitting the in-sample fluctuations. Many empirical studies have shown that dynamic factor models (DFMs) may provide a parsimonious way to include incoming information about a wide variety of economic activity. These models use a large dataset to extract a few common factors.[7] These factors are time-series variables such as inflation or employment growth. Many researchers have argued that DFMs can be used to improve empirical macroeconomic analysis and forecasting of key variables that inform the decisionmaking process of monetary policymakers. This DFM forecasting process has been termed a data-rich environment.[8]

Using a large amount of data in the forecasting process has been popular with forecasters and policymakers for two reasons. First, important variables are likely to be omitted in small-dimension VARs. Effectively, this means that the more variables added to the model, the fewer degrees of freedom available to the forecaster. Second, the use of factor-augmented VARs (FAVARs) is consistent with the stochastic structure of a DSGE model, which is currently in vogue among many central banks.[9] How so? Consider that at any point in time the economy is hit by numerous shocks, such as a surge in oil prices, a change in the tax environment, a collapse in asset prices, or a new technological innovation that significantly changes the production and distribution of a large swath of the nation's goods and services. These shocks affect the nation's key macroeconomic variables that matter to policymakers. DFMs, then, attempt to track the evolving equilibrium of these key variables, much as DSGE models are designed to do.[10] We construct forecasts using a FAVAR to assess the effect of various data series. A factor is a method of summarizing information in a number of different kinds of series (e.g., commodity prices, employment series). The FAVAR is essentially a standard VAR augmented with a set of factors. Although the factors are intended to summarize large sets of data and prevent (or reduce) parameter proliferation, this does not necessarily imply there will not be overfitting in-sample.

We are interested in using a large number of (standardized) predictors summarized by the $N \times 1$ period-t vector X_t. The predictive content of a large vector of indicators can be condensed into a smaller set of K factors, $\mathbf{F}_t$, where

(3)
$$X_t = \Lambda \mathbf{F}_t + \nu_t,$$

where $\mathbf{F}_t$ is a period-t ($K \times 1$) vector of factors, $K << N$, Γ is an ($N \times K$) matrix of loadings, $\nu_t \sim N(0,\Omega)$, and Ω is diagonal. The diagonality assumption implies that the observed correlation across elements of the members of X_t is produced primarily by the factors. We can impose additional assumptions on the factor loadings to identify the factors. In particular, we assume that the loadings on some variables are zero; these zero restrictions are described later with the data.

The VAR that relates inflation, the other macro variables, and the factors is

(4)
$$\begin{bmatrix} \mathbf{y}_{t+h} \\ \mathbf{F}_{t+h} \end{bmatrix} = A(L) \begin{bmatrix} \mathbf{y}_t \\ \mathbf{F}_t \end{bmatrix} + \boldsymbol{\varepsilon}_{t+h},$$

where $\mathbf{y}_t$ is a time-t vector of macroeconomic series of interest (say, unemployment and inflation), $A(L)$ is a matrix polynomial in the lag operator, and $\boldsymbol{\varepsilon}_{t+h}$ is a vector multivariate normal innovation with zero mean and covariance matrix Σ. We construct the forecast of inflation from equation (4) by computing the expectation. In principle, equation (4) could include any number of additional variables; we suppress these for ease of exposition.

Equation (3) relates the factors to the large set of data that we want to summarize, and equation (4) relates the macroeconomic variables to lags of themselves and lags of the factors. Note that the contemporaneous factors do not inform the macroeconomic variables, and vice versa, except through the contemporaneous correlation in the error terms, which are assumed to be mean zero.

Estimation, Forecasting, and Data

The AO model requires no estimation because it is a random walk forecast. The autoregressive forecast is simply a standard autoregressive model with 12 lags of the dependent variable—either the 12-month percent change (logs) in the seasonally adjusted all-items consumer price index for all urban consumers (CPI-U) or the seasonally adjusted personal consumption expenditures chain-weighted price index (PCEPI). We estimate the factor model using Bayesian methods, conditioning on the factors generated using principal components. In generating the factors, we impose zero restrictions on the factor loadings described later.

The inflation rate is the object of interest, which we use two sets of data to predict. The first set of predictive data is the year-to-year percent change in the CPI or the PCEPI; these data enter into the VAR components of the models and include lags of the headline CPI or PCEPI inflation rate. The second set of data is used to construct the factors in the FAVAR; these data are listed in the appendix.

The Factor Model Framework

Table 1 condenses the data series from the appendix into the nine sets of predictive data that form the nine factors used in the FAVAR model. These data are composed of (1) consumer

Table 1

Types of Data Used in Factor Estimation

Description	No. of individual series
1. Consumer price indexes	23
2. Producer price indexes	7
3. Commodity prices	12
4. House and commercial property prices	6
5. Labor markets	11
6. Financial	14
7. Inflation expectations	17
8. Business and consumer surveys	6
9. Foreign prices	8
Total No. of series	104

NOTE: See the appendix for individual series, data transformations, and sources.

price indexes, (2) producer price indexes, (3) commodity prices, (4) housing and commercial property prices, (5) labor market indicators, (6) financial variables, (7) inflation expectations, (8) survey data, and (9) foreign price variables. In choosing these variables, we wanted to focus first on monthly data of consumer and producer price indexes—the most obvious measure of price pressures. We also wanted to use series that measure prices in other dimensions, such as house and commercial property prices that influence rents. Similarly, we include certain commodity prices (e.g., crude oil prices) that affect the prices of goods and services consumed by consumers and producers.

From a broader standpoint, labor market variables have long been used by forecasters to help forecast inflation. According to the SPF, roughly two-thirds of survey participants incorporate some type of Phillips curve in their forecasting model.[11] As noted earlier, expectations of financial and nonfinancial market participants (e.g., consumers and firms) underpin the New Keynesian model. Thus, financial market expectations and surveys of consumers and businesses represent about a quarter of our 104 variables. Finally, Neely and Rapach (2011), Ciccarelli and Mojon (2010), and others have documented that foreign prices strongly influence the U.S. domestic inflation rate. Thus, we include several foreign prices.

We estimate a single factor from each category, assuming that the factor for category i does not load on variables in category j, equating to zero restrictions on the loadings. This approach allows for establishing a direct interpretation of the nature of each type of factor (e.g., summarizing consumer prices, producer prices, and so on). The alternative approach would be to extract a set of factors from the entire set of predictive variables. However, this makes it difficult to obtain a clear, definitive interpretation of which factor represents which source of inflationary pressures. We estimate the factors over two sample periods: February 1964–December 2013 and January 1983–December 2013. The latter period is sometimes referred to as the Great Moderation, which refers to the fact that the volatility of output, inflation, and many other

macroeconomic time-series variables was much larger before 1983 than after 1983.[12] We use a method for generating the principal components with unbalanced panels to estimate the factors. That is, the date of the first observation for all series is not the same; we then generate a separate factor for each subgroup determined earlier. This process yields an unbalanced panel of factors. Most factors begin in January 1964; the exception is the factor constructed using inflation expectations measures, the earliest of which (University of Michigan surveys of consumers) begins in January 1978. Finally, we perform two experiments. In the first experiment, we conduct out-of-sample forecast experiments using monthly revised data from the February 1964–December 2013 period. Here, we include eight different factors, excluding those related to inflation expectations. In the second experiment, we repeat the out-of-sample forecasts with data from the January 1983–December 2013 period and use a set of nine factors, now including the inflation expectations factor.

Factor Loadings

Tables 2 and 3 show the top three series in each category according to the magnitude of their loadings for the samples starting in 1964 and 1983, respectively. The loadings can provide insight because they reflect the correlation between each individual series and the factors: The greater the loading, the greater the correlations between the factor and the series in question. The factor model procedure produces an estimate of $\hat{\Lambda}\mathbf{F}_t$, with the possibility that the sign of either component will change between different runs of the estimation method. Thus, if the sign of a factor changes, the sign of the corresponding loading will change as well. For forecasting purposes, we are concerned only with the product $\hat{\Lambda}\mathbf{F}_t$ and therefore impose no restriction to maintain a consistent sign over the two subsamples. As a result, we analyze the absolute magnitude of the loadings and ignore any variation in their signs between the post-1964 and post-1983 periods.

There is little difference between the factor loadings across the samples for most factors, which suggests a stable relationship. For the first factor, consumer price indexes, the ordering of the factor loadings changes somewhat. For example, the core PCEPI (which excludes food and energy prices) is highly correlated in the full sample but less so in the post-Great Moderation sample. A few other factors change composition across samples. Foreign prices and the survey data change the ordering of the largest factor loadings, but the top three data series remain the same. The series that comprise the inflation expectations factor change composition, but this is likely due to data availability.

Forecasting Specifics

As noted earlier, we augment our FAVAR model with eight or nine factors. Eight factors are used in the full sample; the ninth factor is derived from the inflation expectations series and is included in the post-1983 period. In the first experiment, we estimate the three models—AO, AR(12), and FAVAR—through December 1989. Our first out-of-sample forecasts inflation using data available up through January 1990. We then forecast horizons from 0 (January 1990) to 12 months ahead (January 1991). Forecasts are constructed using direct methods: We regress

Table 2

Three Largest Factor Loadings Within Each Category: Full Post-February 1964 Sample

Consumer price indexes

FRB Cleveland: Median CPI, 1-month percent change	−1.21
FRB Atlanta: Sticky CPI, 1-month percent change	1.18
PCE chain-type price index, market-based excluding food and energy	−1.18

Producer prices

PPI: Final demand	1.02
PPI: Final demand goods	1.01
PPI: Final demand excluding food and energy	1.01

Commodity prices

KR-CRB Spot Commodity Price Index: All Commodities	1.38
CRB Spot Raw Industrials Price Index	1.22
U.S. retail gasoline price: Regular grade	1.20

House and commercial property prices

Case-Shiller Composite 20-City House Price Index	1.01
FHFA House Price Index, Purchase Only	1.01
CoreLogic National House Price Index (SA, Jan. 2000 = 100)	1.00

Labor markets

Civilian unemployment rate	1.37
Civilian unemployment rate gap estimate	1.37
Average hourly earnings: Private goods-producing, all employees	1.36

Financial

10-Year Treasury yield, constant maturity	1.61
5-Year Treasury yield, constant maturity	1.59
30-Year Treasury yield, constant maturity	1.55

Inflation expectations

TIPS spread, 5-year	1.001
TIPS spread, 7-year	1.001
TIPS spread, 10-year	1.001

Surveys

ISM: Nonmanufacturing Prices Paid index	−1.28
NFIB: Percent of firms planning to raise average selling prices, net	−1.17
ISM: Manufacturing Prices Paid index	−1.07

Foreign prices

Euro area harmonized overall CPI	1.17
U.S. Import Price Index, All Imports	1.15
U.S. Import Price Index, Nonpetroleum Imports	1.15

NOTE: CRB, Commodity Research Bureau; FHFA, Federal Housing Finance Agency; ISM, Institute for Supply Management; NFIB, National Federation of Independent Business; PPI, producer price index; SA, seasonally adjusted; TIPS, Treasury inflation-protected securities.

SOURCE: Authors' calculations.

Table 3

Three Largest Factor Loadings Within Each Category: Full Post-January 1983 Sample

Consumer price indexes

FRB Cleveland: 16% Trimmed mean CPI, 1-month percent change	1.35
FRB Dallas: Trimmed mean, 1-month PCE inflation rate	1.34
FRB Atlanta: Sticky CPI, 1-month percent change	1.32

Producer prices

PPI: Final demand	1.02
PPI: Final demand goods	1.01
PPI: Final demand excluding food and energy	1.01

Commodity prices

KR-CRB Spot Commodity Price Index: All Commodities	1.42
CRB Spot Raw Industrials Price Index	1.30
CRB Spot Livestock and Products Price Index	1.12

House and commercial property prices

Case-Shiller Composite 20-City House Price Index	1.23
CoreLogic National House Price Index (SA, Jan. 2000 = 100)	1.18
FHFA House Price Index: Purchase Only	1.07

Labor markets

Civilian unemployment rate	1.46
Civilian unemployment rate gap estimate	1.43
Average hourly earnings: Private goods-producing, all employees	1.29

Financial

10-Year Treasury yield, constant maturity	−1.60
5-Year Treasury yield, constant maturity	−1.58
Yield on Treasury long-term composite bond	−1.53

Inflation expectations

TIPS spread, 30-year	1.11
FRB Cleveland, 5-Year expected inflation rate	−1.10
FRB Cleveland, 7-Year expected Inflation rate	−1.10

Surveys

ISM: Nonmanufacturing Prices Paid Index	−1.40
ISM: Manufacturing Prices Paid Index	−1.26
NFIB: Percent of firms planning to raise average selling prices, net	−1.22

Foreign prices

U.S. Import Price Index, All Imports	−1.41
Euro area harmonized overall CPI	−1.42
U.S. Import Price Index: Nonpetroleum Commodities	−1.34

NOTE: CRB, Commodity Research Bureau; FHFA, Federal Housing Finance Agency; ISM, Institute for Supply Management; NFIB, National Federation of Independent Business; PPI, producer price index; SA, seasonally adjusted; TIPS, Treasury inflation-protected securities.

SOURCE: Authors' calculations.

Figure 1

CPI and PCEPI Inflation

12-Month Percent Change
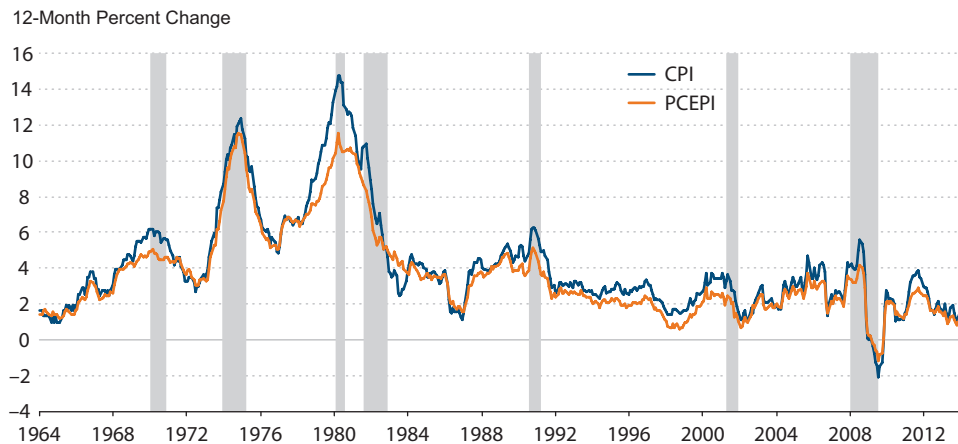
NOTE: The shaded bars indicate recessions as determined by the National Bureau of Economic Research.

SOURCE: Bureau of Economic Analysis, Bureau of Labor Statistics, National Bureau of Economic Research, and Haver Analytics.

directly the forward data on the available information. We use a recursive estimation scheme so that all past information is incorporated into the model estimates. We have three FAVAR models: 1 lag, 6 lags, and 12 lags. Because estimation of the FAVAR is computationally intensive, we reestimate the model only once per year in January, when we assume all data from the previous year are available. The forecasts are constructed monthly, which means the principal components are updated monthly, but the forecasts are constructed using that year's estimate of the model parameters.

Results

Figure 1 plots the actual inflation series we forecast. Inflation—whether measured by CPI or PCEPI—was rising, on net, from the beginning of our sample to roughly 1981 and was highly variable. Inflation fell sharply after the Volcker disinflation and has averaged around 3 percent—and generally has been much less volatile—after 1983.

Table 4 shows the root mean squared errors (RMSEs) for the two sample periods and all three models. An RMSE is a standard measure of forecast accuracy, as it penalizes a forecast if it has a higher variance and bias of its forecast errors (actual less predicted). A lower RMSE indicates a better forecast relative to another forecast. In the table, the RMSEs are benchmarked to a baseline forecast's RMSE, which we define as the random walk (AO). Thus, in Table 4 a value less than 1 indicates that the model outperforms the AO model (better forecast accuracy) and any value greater than 1 indicates that the AO has a smaller RMSE over the relevant forecast horizon.

Table 4

RMSEs of Forecasting Models

Forecast horizon

	t	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10	t+11	t+12
Full sample													
CPI													
AO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(12)	0.97	1.03	1.06	1.10	1.14	1.18	1.19	1.22	1.23	1.24	1.24	1.27	1.28
FAVAR(1)	0.98	0.99	0.97	0.95	0.94	0.94	0.93	0.92	0.91	0.89	0.91	0.96	1.01
FAVAR(6)	0.95	0.97	0.96	0.93	0.92	0.95	0.99	1.05	1.11	1.14	1.16	1.19	1.23
FAVAR(12)	1.00	1.03	1.04	1.07	1.13	1.20	1.26	1.27	1.30	1.30	1.31	1.33	1.36
PCEPI													
AO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(12)	0.95	1.00	1.03	1.06	1.10	1.12	1.15	1.17	1.19	1.20	1.21	1.24	1.26
FAVAR(1)	1.00	1.00	0.97	0.97	0.98	1.00	1.00	1.00	1.02	1.03	1.06	1.10	1.16
FAVAR(6)	0.97	1.00	1.00	1.02	1.05	1.09	1.15	1.24	1.31	1.35	1.37	1.40	1.45
FAVAR(12)	1.07	1.12	1.13	1.17	1.24	1.32	1.38	1.43	1.48	1.52	1.54	1.57	1.60
Sample starting in 1983													
CPI													
AO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(12)	0.93	0.98	1.01	1.03	1.04	1.05	1.05	1.06	1.05	1.05	1.03	1.04	1.04
FAVAR(1)	1.05	1.03	0.97	0.94	0.93	0.91	0.88	0.85	0.81	0.80	0.80	0.82	0.83
FAVAR(6)*	1.05	1.15	1.16	1.17	1.10	1.06	0.98	0.97	0.95	0.93	0.91	0.91	0.92
FAVAR(12)*	1.15	1.28	1.30	1.36	1.35	1.37	1.33	1.31	1.29	1.11	1.28	1.25	1.19
PCEPI													
AO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(12)	0.95	0.99	1.03	1.05	1.06	1.07	1.07	1.07	1.07	1.06	1.05	1.06	1.06
FAVAR(1)	1.05	1.02	0.97	0.95	0.95	0.94	0.91	0.89	0.87	0.86	0.85	0.87	0.87
FAVAR(6)*	1.10	1.12	1.08	1.08	1.04	1.03	0.99	1.00	1.00	0.98	0.96	0.95	0.97
FAVAR(12)*	1.27	1.29	1.32	1.42	1.39	1.40	1.31	1.28	1.24	1.20	1.16	1.19	1.19

NOTE: *FAVAR models with multiple lags using limited samples to produce forecasts starting in January 1995, not January 1990.

We now consider some key findings from the full sample. First, the AO model performs reasonably well across most horizons. The AO model clearly outperforms the AR(12) model—except for the contemporaneous period ($t = 0$)—for both measures of inflation.

In the full sample, the FAVAR(1) model is generally more accurate in forecasting CPI inflation than the AO model for times t to $t + 11$. The FAVAR(6) model performs much better through the first half of the forecast horizon (up through 6 months). The forecasting accuracy of the FAVAR(12) model is worse than the FAVAR(1) and FAVAR(6) models across all horizons. The longer the forecast horizon, the worse the FAVAR(12) model performs—indeed, worse than the AR(12) model. In the full sample, the FAVAR models generally do not forecast PCE inflation as well as the AO model. The random walk model tends to dominate all other models for forecasting PCE inflation in the full sample.

Table 4 also shows the forecasting performance in the post-1983 sample. In this experiment, the model is estimated with data from January 1983 through December 1994 and then out-of-sample forecasting begins in January 1995. In this experiment, we add the inflation expectations factor (for a total of nine factors). As before, the models are then reestimated a year later and out-of-sample forecasts are produced. Table 4 clearly indicates that adding the inflation expectations factor to the FAVAR model produces markedly smaller RMSEs for both inflation measures than either the AO or AR(12) models. Indeed, at 6 and 12 months ahead, the FAVAR(1) forecast for CPI inflation produces RMSEs that are 12 percent and 17 percent smaller, respectively, than the AO model. The RMSEs are a bit larger for the 6-lag FAVAR. For PCEPI inflation, the RMSEs are a bit larger than for CPI inflation, but again the FAVAR(1) and FAVAR(6) models perform measurably better than the AO or AR(12) models. The FAVAR(12) model has a much higher RMSE for both CPI and PCEPI inflation than the other models across all horizons.

Figures 2 through 5 plot actual inflation and the forecast for contemporaneous inflation for the FAVAR(1) and FAVAR(6) models for both samples. The figures also plot out-of-sample forecasts for January 2014–January 2015. As shown in Figure 2, using the full-sample estimation, the FAVAR(1) model forecasts that CPI inflation would rise to about 2.5 percent in January 2015. However, the FAVAR(6) model forecasts that inflation would remain about unchanged. Figure 3 shows the same pattern for PCEPI inflation. Using the post-1983 sample, both models predict higher inflation in 2014 and early 2015 relative to the end of 2013. Figures 4 and 5 are consistent with the view of the FOMC, which foresees inflation eventually returning to its 2 percent inflation target.

A MEASURE OF PRICE PRESSURES

In the previous section, we considered the problem of forecasting the value of inflation at some horizon. To evaluate those forecasts, we compared the point value of the forecasted distribution with the realized value. Forecasts farther from the realization yield larger penalties for the model. In some circumstances the distance from the realization is less important than, say, the direction of the change. Recently, the Federal Reserve announced a target zone for inflation. When inflation is above the target zone, the Fed has a substantially higher probability

Figure 2

Actual Versus Forecasted CPI Inflation: Full Sample
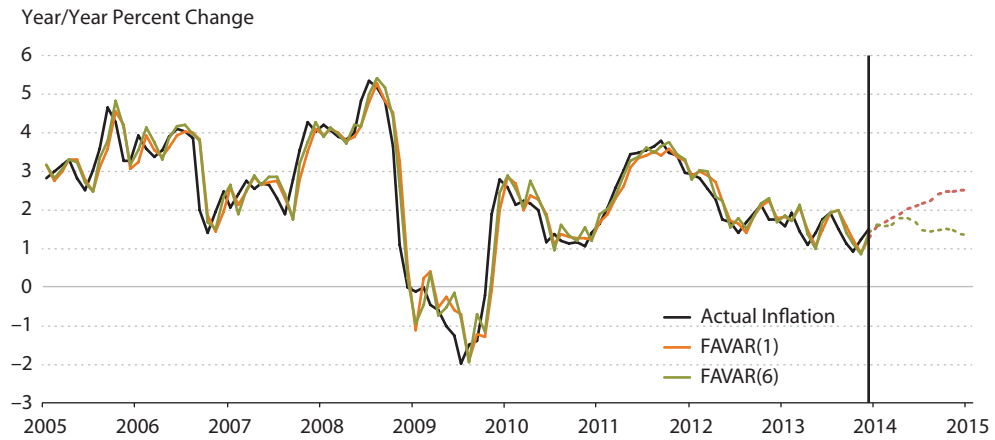
NOTE: Forecasts are plotted for January 2014 to January 2015.

SOURCE: Authors' calculations.

Figure 3

Actual Versus Forecasted PCEPI Inflation: Full Sample
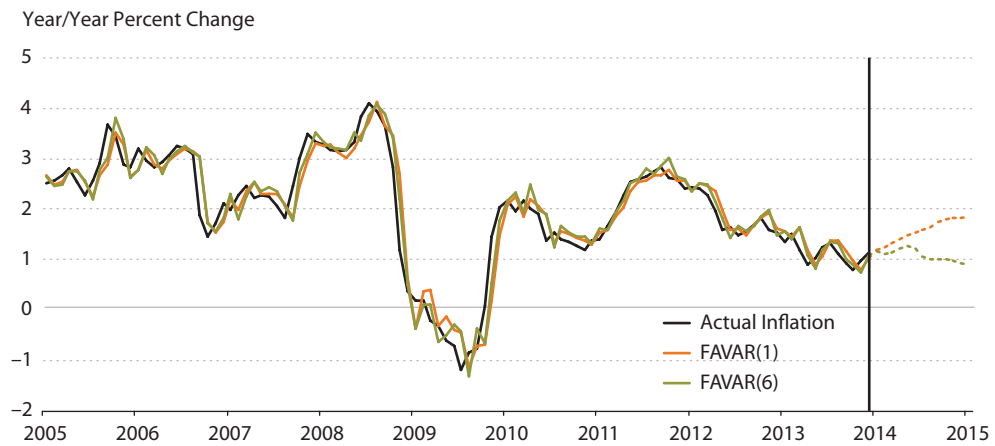
NOTE: Forecasts are plotted for January 2014 to January 2015.

SOURCE: Authors' calculations.

Figure 4

Actual Versus Forecasted CPI Inflation: Post-1983 Sample
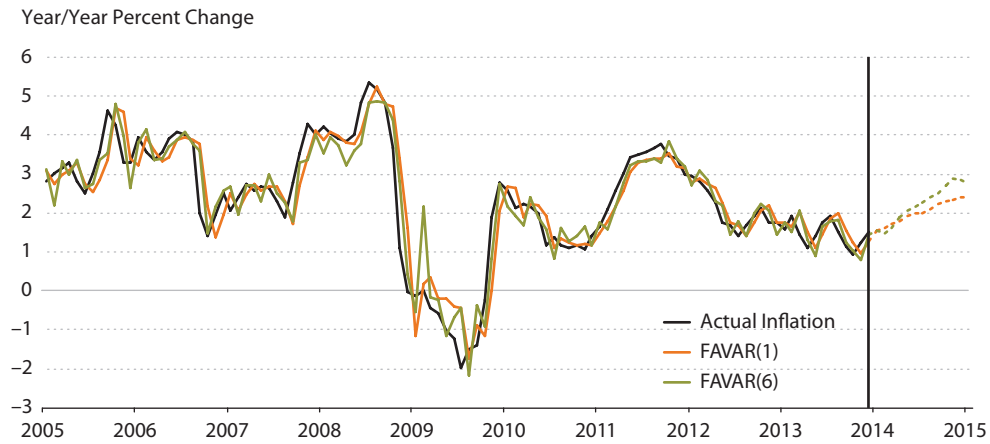
Year/Year Percent Change

NOTE: Forecasts are plotted for January 2014 to January 2015.

SOURCE: Authors' calculations.

Figure 5

Actual Versus Forecasted PCEPI Inflation: Post-1983 Sample

Year/Year Percent Change
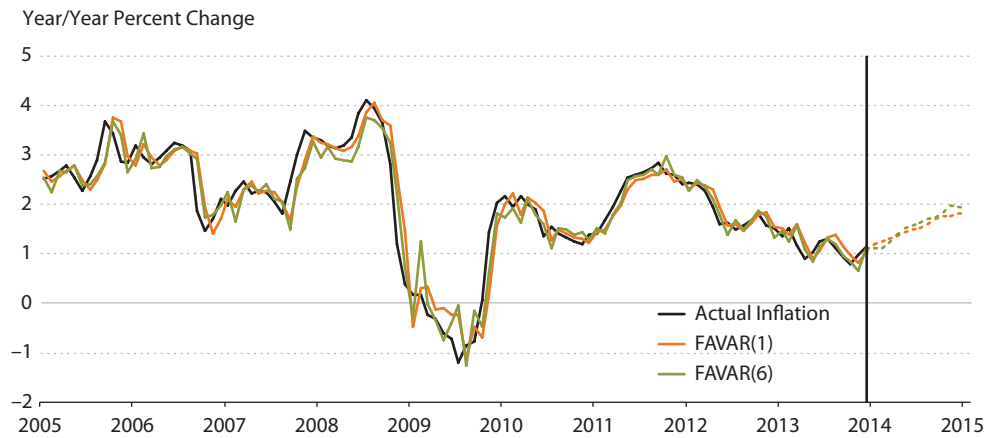
NOTE: Forecasts are plotted for January 2014 to January 2015.

SOURCE: Authors' calculations.

of increasing the federal funds rate to combat inflation; when inflation is below the target zone, the Fed has a higher probability of lowering the federal funds rate to stimulate the economy. Thus, it might be important to assess the probability that inflation will move above the target zone over some horizon. In this section, we consider a forecast of this sort. We use this forecasting model to construct an index that we call the price pressure measure (PPM), which reflects the likelihood that inflation will be above the target zone in the next year.

The Model

Our objective is to forecast the probabilities that inflation will rise above or fall below the target zone. We define the discrete variable $\Pi_t \in \{1,2,3,4\}$, where the discrete outcomes correspond to

(5)
$$
\begin{aligned}
\Pi_t &= 1 \quad \text{if} \quad\quad\; \pi_t \leq 0 \\
\Pi_t &= 2 \quad \text{if} \quad 0 < \pi_t \leq 1.5 \\
\Pi_t &= 3 \quad \text{if} \quad 1.5 < \pi_t \leq 2.5 \\
\Pi_t &= 4 \quad \text{if} \quad\quad\; 2.5 < \pi_t
\end{aligned}
$$

and π_t is the period-t inflation rate (12-month percent changes). The bounds on the right-hand side of the conditions outlined in (5) are determined by the FOMC statement about the target zone. The third condition establishes a set of bounds symmetric around the Fed's inflation target: $1.5 \leq \pi_t < 2.5$. We are interested in forecasting $\Pi_{t+h|t}$, the h-period-ahead value of the discrete variable conditional on the information at time t. Let $\pi_{t+h|t}$ be the "forecast" of inflation conditional on time-t information. Suppose that this forecast is given by

(6)
$$
\pi_{t+h|t} = G\left(\pi_{t-1}, \mathbf{F}_t\right) + \varepsilon_t;
$$

then the forecast $\Pr[\Pi_{t+h|t} = k]$, for example, can be obtained by determining the probability that $\pi_{t+h|t} > 2.5$. Because the ε_t are assumed normal, the model is the familiar ordered probit augmented with the set of factors, $\mathbf{F}_t$.

Estimation

As in the previous section, the model is estimated with the Gibbs sampler, a Bayesian method that iteratively draws each parameter from its conditional distribution. In the sampler, we treat the factor identified by principal components as a known quantity. Multiple draws from the sampler approximate the full joint density. We sample the latent forecast $\{\pi_{t+h|t}\}_{t=1}^{T}$, conditional on the factors and the model parameters, from a truncated normal, where the truncation points are given by (5). Here, T represents the h periods before the end of the estimation sample as we lose some data because of the direct forecasting scheme. Assuming a normal prior, we can draw the model parameters from the normal conjugate posterior distribution, conditional on $\{\pi_{t+h|t}\}_{t=1}^{T}$. The forecasts will be probabilities, $\Pr[\Pi_{t+h|t} = k]$, which are determined by obtaining the area under the normal cumulative distribution function between the truncation points conditional on the forecasted value $\pi_{t+h|t}$.[13]

Forming the Index

The objective of forming the PPM is to assess the likelihood that inflation will rise above the target. We computed $\{\Pr[\pi_{t+h|t} > 2.5]\}_{h=0}^{12}$ using the ordered probit model. We can compute a weighted sum of these probabilities to form our PPM:

$$PPM = \sum_{h=0}^{12} w_h \Pr\left[\pi_{t+h|t} > 2.5\right],$$

where w_h is the weight placed on horizon h and $\Sigma_h w_h = 1$. The nature of the weights depends on whether longer or shorter horizon forecasts are more valued. In this case, we opt for equal weighting.

Results

Our PPM measures the probability that the expected inflation rate (12-month percent changes) over the next 12 months—the forecast horizon—will exceed 2.5 percent. This bin (2.5 percent) exceeds the Fed's 2 percent inflation target. We calculate our PPMs from these two modes. These PPMs are plotted in Figure 6 for CPI and PCEPI inflation using the 1- and 6-lag ordered probit models.[14] We plot the smoothed series, which is a six-month moving average. Figure 6 shows that, over most of this sample period (January 1990–January 2014), the PPM for CPI inflation was greater than 0.5. By contrast, the probabilities for PCEPI inflation exceeded 0.5 by an appreciably smaller percentage over the sample period. Since the end of the recent recession, the PPMs have been significantly below 0.5 for PCEPI inflation but moderately less so for CPI inflation. In one sense, the models are picking up the fact that inflation was higher before the recession and that CPI inflation is generally higher on average than PCEPI inflation. For the January 1990–December 2007 period, CPI inflation averaged 2.9 percent and PCEPI inflation averaged 2.3 percent. However, since January 2008, CPI inflation has averaged 2 percent and PCEPI inflation has averaged 1.7 percent.

At any point in time, the PPMs plotted in Figure 6 are unweighted averages of the probability that the forecasted inflation rate will average more than 2.5 percent over the next 12 months. However, policymakers know that a standard error around the point estimate is associated with any forecast. For example, the Bank of England's fan charts contain both point estimates and error bands around these point estimates that can be thought of as probabilities. In the simplest terms, if monetary policymakers project that inflation over the following year will be 2 percent, there is some probability that inflation will be less than 2 percent and some probability that inflation will be more than 2 percent.

The ordered probit model estimated earlier provides probabilities that inflation will exceed 2.5 percent, on average, over the next 12 months. But our model also allows us to assess the probability that inflation will average something different. In this case, we structure the model to assess the probability that inflation will fall within one of four bins: less than zero (deflation); 0 percent to 1.5 percent; 1.5 percent to 2.5 percent; and more than 2.5 percent. The last bin is our PPM plotted in Figure 6; Figure 7 plots the other three probabilities. For ease of discussion, we condense the second and third bins into one, leaving two sets of

Figure 6

Price Pressure Measure: Probability That Inflation Exceeds 2.5 Percent
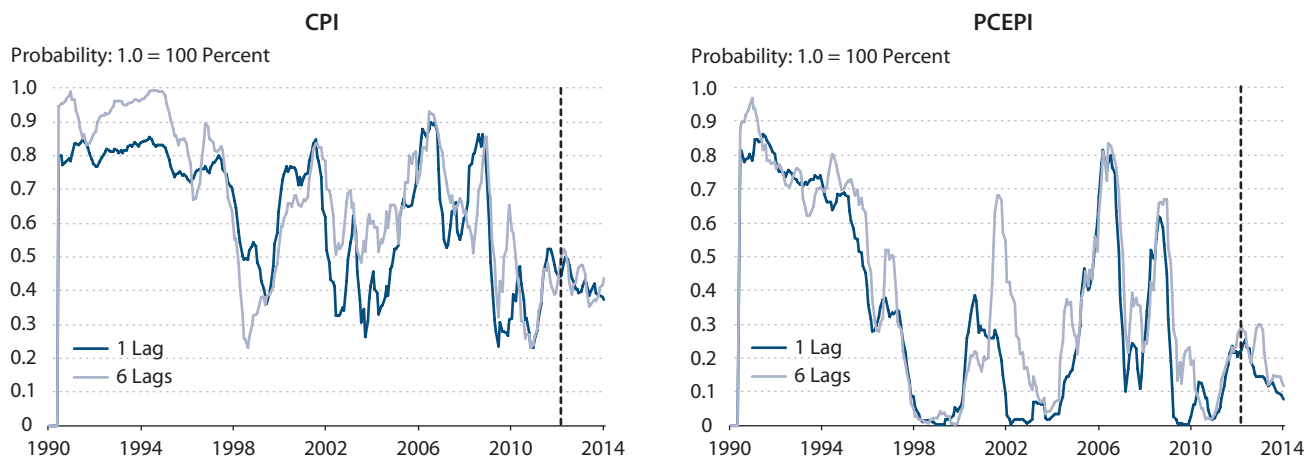
CPI

Probability: 1.0 = 100 Percent

PCEPI

Probability: 1.0 = 100 Percent

1 Lag
6 Lags

1 Lag
6 Lags

NOTE: Sample is limited to 1983 to present. The vertical dashed line indicates March 2012.

SOURCE: Authors' calculations.

probabilities: Inflation will be less than zero (deflation) over the next 12 months and inflation will average between 0 percent and 2.5 percent.

In March 2012, policymakers observed that the CPI had increased by 2.3 percent over the previous year (March 2011–March 2012). The outlook of professional forecasters, as judged by the Blue Chip Consensus (BCC), was that the CPI inflation rate (four-quarter per-cent changes) would average 2.1 percent from 2012:Q2 to 2013:Q2. However, as noted earlier, policymakers generally eschew point estimates in favor of probabilities.[15] In this case, as shown in Figure 7, it is the probability that inflation will be above or below the forecast consensus.

In March 2012, the model predicted a 45 percent probability that CPI inflation would average more than 2.5 percent from April 2012 to April 2013 (see Figure 6). This relatively high probability could have reflected the fact that crude oil prices rose by 24 percent from September 2011 to March 2012. However, the model also predicted an equal probability that inflation would average between 0 percent and 2.5 percent, with only a 10 percent probability that inflation would average less than zero over the next 12 months. (This date is noted by the vertical dashed line in the figure.)[16] But forecasters and policymakers were instead surprised because inflation fell from 2.3 percent in April 2012 to 1.1 percent in April 2013. The model performed reasonably well if one takes into account the probability of deflation: There was a greater than 50 percent probability that inflation would be less than 2.5 percent.

A year later, in March 2013, the model lowered the probability that inflation would average more than 2.5 percent over the next year (April 2013–April 2014) from 45 percent to 43 per-cent (see Figure 6). However, the model raised the probability that inflation would be between 0 percent and 2.5 percent over the following year from 45 percent to 50 percent. Likewise, the

Figure 7

Price Pressure Measure: Ordered Probit Model
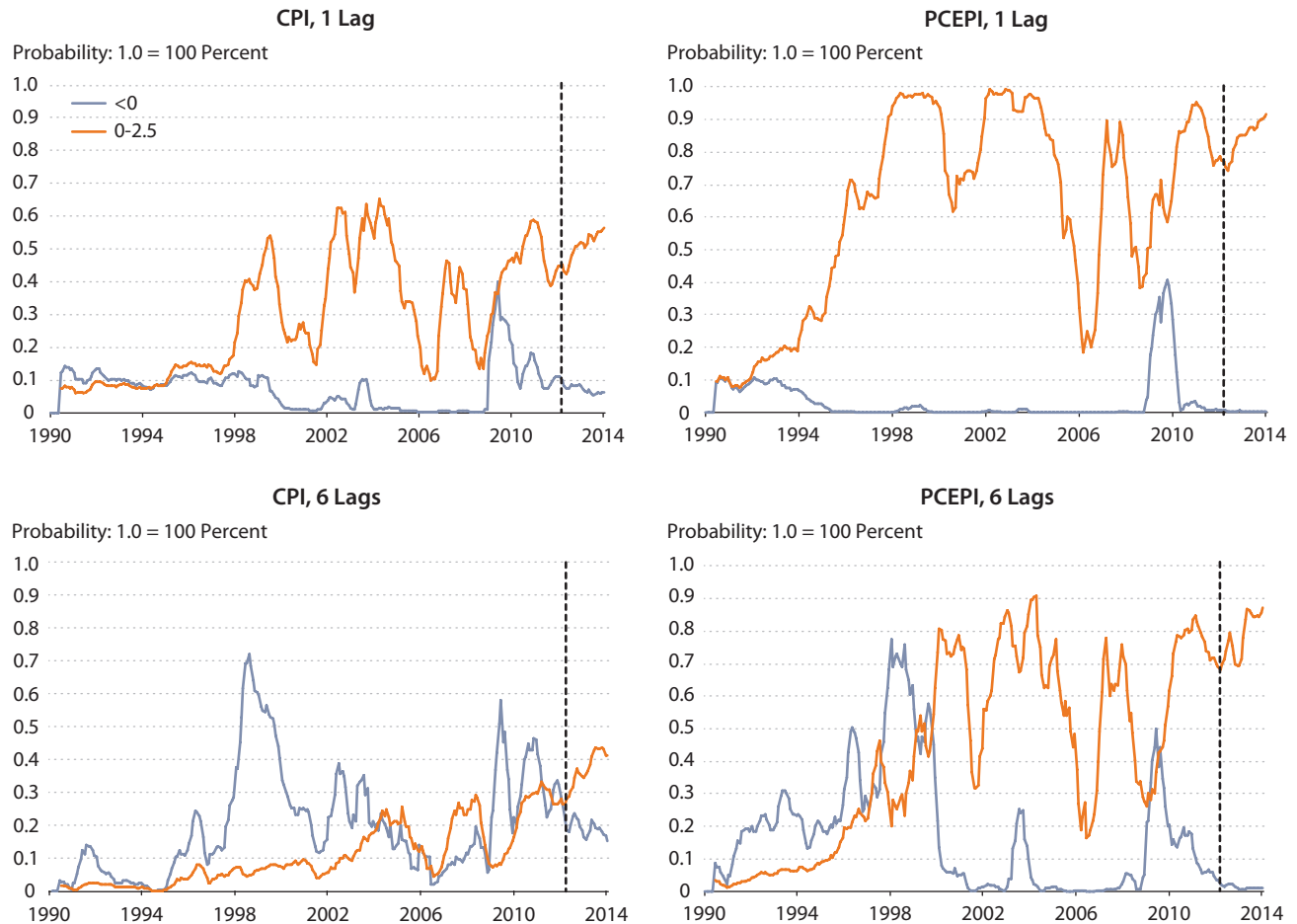
CPI, 1 Lag

Probability: 1.0 = 100 Percent

PCEPI, 1 Lag

Probability: 1.0 = 100 Percent

CPI, 6 Lags

Probability: 1.0 = 100 Percent

PCEPI, 6 Lags

Probability: 1.0 = 100 Percent

NOTE: Sample is limited to 1983 to present. The vertical dashed line indicates March 2012.

SOURCE: Authors' calculations.

model lowered the probability of deflation from 10 percent to 7 percent. By contrast, in March 2013 the BCC forecasters predicted the CPI inflation rate (four-quarter percent changes) would average 2.0 percent from 2013:Q2 to 2014:Q2—virtually unchanged from their year-ahead forecast published a year earlier.

The preceding discussion focuses on the CPI inflation rate. Although many contracts and prices are indexed to the CPI, FOMC policymakers instead prefer to target the PCEPI inflation rate. The upper-right panel in Figure 7 plots the smoothed PPMs for the PCEPI inflation rate from the probit 1-lag model. A similar story emerges here as well. In March 2012, as seen in Figure 6, the model predicted only a 22 percent probability that inflation would average more than 2.5 percent over the next 12 months (April 2012–April 2013). The model predicted a 77

percent probability that inflation would average between 0 and 2.5 percent. The probability that inflation would average less than zero (deflation) was less than 1 percent. Although the BCC does not forecast the PCEPI inflation rate, forecasts for the FOMC's preferred inflation measure are reported in the Philadelphia Fed's SPF. In its February 2012 report, the SPF predicted the PCEPI would increase from 1.7 percent (quarterly rate, annualized) in 2012:Q1 to 2 percent in 2013:Q1. Mirroring the dip in the CPI inflation rate, the PCEPI inflation rate unexpectedly slowed from 2 percent in April 2012 to 1 percent in April 2013.[17] In this case, the model performed well, perceiving a relatively high probability that inflation would remain below 2.5 percent.

The unexpected slowing in inflation, perhaps not surprisingly, affected the probability distributions a year later, in March 2013. By then, the model estimated the probability that PCEPI inflation would average between 0 and 2.5 percent over the next 12 months (April 2013–April 2014) had increased from 77 percent to 85 percent. The probability of deflation was lowered from 0.8 percent to 0.3 percent. As shown in Figure 6, the probability that inflation would average more than 2.5 percent declined from 22 percent to 14 percent. Despite this marked shift in the probability distribution, in mid-February 2013 the SPF was still projecting that PCEPI inflation would increase to 2 percent in 2014:Q1. Once again, the actual data are more consistent with our model: From April 2013 to March 2014 (the latest available data), the 12-month change in the PCEPI inflation rate increased from 1 percent to 1.2 percent. The two lower charts in Figure 7 show the PPMs using the 6-lag probit for the post-1983 sample. They show trends broadly similar to the 1-lag model.

Out-of-Sample PCEPI Inflation Forecasts

Table 4 indicates that the best model for forecasting inflation one year ahead is the FAVAR(1) for CPI inflation estimated using the post-1983 sample. Although the FAVAR(1) RMSEs for PCEPI inflation are slightly larger, this section nonetheless focuses on this measure because the FOMC targets the PCEPI inflation rate. Recall that Figure 5 plots the PCEPI inflation forecasts for January 2014–January 2015. For purposes of comparison, FOMC participants in December 2013 projected that the PCEPI inflation rate would increase by 1.5 percent in 2014 (2013:Q4–2014:Q4).[18] Thus, our preferred inflation forecasting model expected inflation to rise by slightly more than the FOMC's projection, from 1 percent in December 2013 to 1.8 percent in December 2014 and in January 2015.[19]

Figure 8 shows the probability distribution of this out-of-sample forecast from January 2014 to January 2015. The upper-left panel indicates that the model predicts a very small— roughly zero—probability of deflation over this horizon. For this forecast, we can separate the 0 percent to 2.5 percent probability distribution into two bins: 0 to 1.5 percent (upper-right panel) and 1.5 percent to 2.5 percent (lower-left panel). In the upper-right panel, the model predicts a more than 50 percent probability that PCEPI inflation would remain in the 0 to 1.5 percent range through the first five months of 2014. Thereafter, the model predicts a higher probability—averaging slightly less than 50 percent—that PCEPI inflation would rise to more than 1.5 percent but remain below 2.5 percent. The lower-right panel shows an increasing probability—over the second half of 2014—that inflation would increase by more than 2.5 percent by the end of the forecast horizon.

Figure 8

PCEPI Inflation Probabilities

Probability That PCEPI Inflation Will Be Negative (Jan. 2014–Jan. 2015)

Probability: 1.0 = 100%
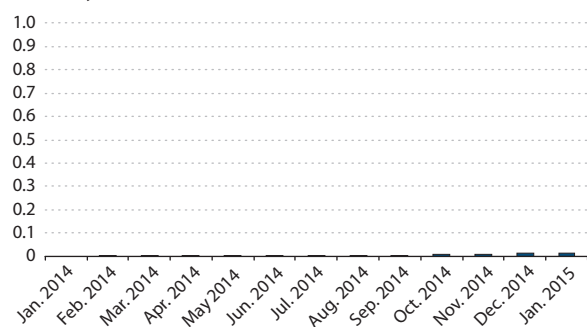
Probability That PCEPI Inflation Will Be 0% to 1.5% (Jan. 2014–Jan. 2015)
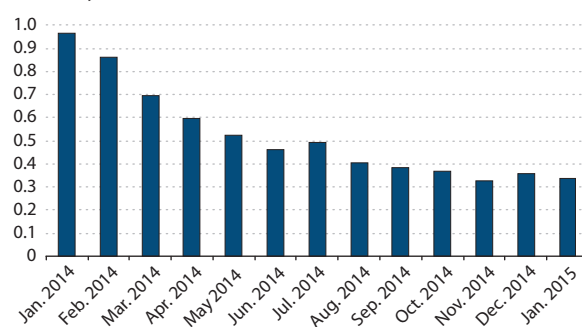
Probability: 1.0 = 100%

Probability That PCEPI Inflation Will Be 1.5% to 2.5% (Jan. 2014–Jan. 2015)
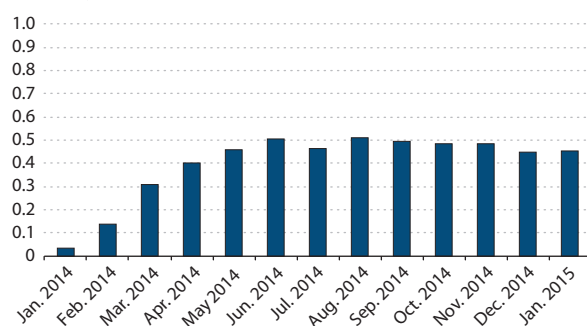
Probability: 1.0 = 100%

Probability That PCEPI Inflation Will Be Above 2.5% (Jan. 2014–Jan. 2015)

Probability: 1.0 = 100%
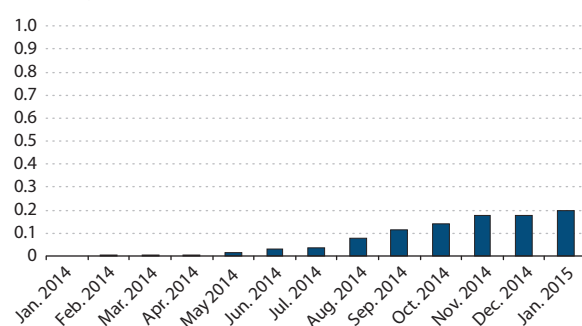
SOURCE: Authors' calculations.

CONCLUSION

The FOMC, like most major central banks, devotes significant resources to forecasting key economic variables such as real GDP growth, employment, and inflation. The outlook for these variables also matters a great deal to businesses and financial market participants. For example, when decisions are made to expend scarce resources or price financial assets, such decisions—which must be made in the present—are based on expectations of future economic conditions. In this article, we present a factor-augmented Bayesian vector autoregressive forecasting model that significantly outperforms both a benchmark random walk model and a pure time-series model. The empirical literature has shown that random walk models tend to be among the most accurate across a variety of simple time-series model specifications. A key

innovation in our article is the use of nine factors in an ordered probit model to assess the probability distribution of the model's point forecasts. We term these probabilities a price pressure measure. Our measure shows a relatively high probability that inflation in 2014 would be higher than that projected by the FOMC in its December 2013 Summary of Economic Projections.[20] ∎

APPENDIX

Data Used to Construct Factors, Their Transformation, and Their Source

	Description	Transformation	Source
Consumer price indexes			
1	PCE: Chain-type price index (SA, 2009 = 100)	DLN	BEA
2	PCE: Goods: Chain-type price index (SA, 2009 = 100)	DLN	BEA
3	PCE: Services: Chain-type price index (SA, 2009 = 100)	DLN	BEA
4	PCE: Less food and energy: Chain-type price index (SA, 2009 = 100)	DLN	BEA
5	PCE: Food and beverages purchased for off-premises consumption: Chain-type price index (SA, 2009 = 100)	DLN	BEA
6	PCE: Energy goods and services: Chain-type price index (SA, 2009 = 100)	DLN	BEA
7	Market-based PCE: Chain-type price index (SA, 2009 = 100)	DLN	BEA
8	Market-based PCE excluding food and energy: Chain-type price index (SA, 2009 = 100)	DLN	BEA
9	PCE: Imputed rental of owner-occupied nonfarm housing price index (SA, 2009 = 100)	DLN	BEA
10	CPI-U: All items (SA, 1982-84 = 100)	DLN	BLS
11	CPI-U: All items less food and energy (SA, 1982-84 = 100)	DLN	BLS
12	CPI-U: Food (SA, 1982-84 = 100)	DLN	BLS
13	CPI-U: Energy (SA, 1982-84 = 100)	DLN	BLS
14	CPI-U: Owners' Equivalent Rent of Primary Residence (SA, December 1982 = 100)	DLN	BLS
15	FRB Dallas: Trimmed mean 1-month PCE inflation, annual rate (%)	LVL	FRBDAL
16	FRB Cleveland Median CPI (SA, % change)	LVL	FRBCLE
17	FRB Cleveland 16% Trimmed mean CPI (SA, % change)	LVL	FRBCLE
18	FRB Atlanta Sticky price CPI (SA, % change)	LVL	FRBATL
19	FRB Atlanta Core sticky CPI (SA, % change)	LVL	FRBATL
20	FRB Atlanta Sticky CPI excluding shelter (SA, % change)	LVL	FRBATL
21	FRB Atlanta Core sticky CPI excluding shelter (SA, % change)	LVL	FRBATL
22	FRB Atlanta Flexible CPI (SA, % change)	LVL	FRBATL
23	FRB Atlanta Core flexible CPI (SA, % change)	LVL	FRBATL
Producer prices			
24	PPI: Final demand (SA, November 2009 = 100)	DLN	BLS
25	PPI: Final demand goods (SA, November 2009 = 100)	DLN	BLS
26	PPI: Final demand services (SA, November 2009 = 100)	DLN	BLS
27	PPI: Final demand less foods and energy (SA, April 2010 = 100)	DLN	BLS
28	PPI: Intermediate demand processed goods (SA, 1982 = 100)	DLN	BLS
29	PPI: Intermediate demand services (SA, November 2009 = 100)	DLN	BLS
30	PPI: Intermediate demand processed energy goods (SA, 1982 = 100)	DLN	BLS
Commodity prices			
31	Refiners' acquisition cost of crude oil: Composite: DOE ($/barrel)	DLN	EIA
32	Natural gas price: Henry hub, Louisiana ($/MMBTU)	DLN	WSJ
33	U.S. retail gasoline price: Regular grade (Average, cents/gallon)	DLN	EIA
34	U.S. retail diesel fuel price including taxes (Average, $/gallon)	DLN	EIA
35	Brent–WTI price spread ($/barrel)	DLV	EIA/WSJ

APPENDIX, cont'd

Description	Transformation	Source
Commodity prices, cont'd.		
36 KR-CRB Spot Commodity Price Index: All commodities (1967 = 100)	DLN	CRB
37 KR-CRB Spot Commodity Price Index: Metals (1967 = 100)	DLN	CRB
38 KR-CRB Spot Commodity Price Index: Textiles and fibers (1967 = 100)	DLN	CRB
39 KR-CRB Spot Commodity Price Index: Raw industrials (1967 = 100)	DLN	CRB
40 KR-CRB Spot Commodity Price Index: Foodstuffs (1967 = 100)	DLN	CRB
41 KR-CRB Spot Commodity Price Index: Fats and oils (1967 = 100)	DLN	CRB
42 KR-CRB Spot Commodity Price Index: Livestock and products (1967 = 100)	DLN	CRB
House and commercial property prices		
43 FHFA House Price Index: Purchase only, United States (SA, January 1991 = 100)	DLN	FHFA
44 Freddie Mac House Price Index, United States (December 2000 = 100)	DLN	FHLMC
45 S&P/Case-Shiller 20-City Composite Home Price Index (SA, January 2000 = 100)	DLN	S&P
46 CoreLogic National House Price Index (SA, January 2000 = 100)	DLN	CORE/H
47 GSA Commercial Property Price Index (NSA, August 2007 = 100)	DLN	GSA
48 Houses Under Construction: Fixed-Weighted Price Index (NSA, 2005 = 100)	DLN	CENSUS
Labor markets		
49 Average hourly earnings of production and nonsupervisory employees: Goods-producing industries (SA, $/hr)	DLN	BLS
50 Average hourly earning of production and nonsupervisory employees: Private service-providing industries (SA, $/hr)	DLN	BLS
51 Average hourly earnings: Goods-producing industries (SA, $/hr)	DLN	BLS
52 Average hourly earnings: Private service-providing industries (SA, $/hr)	DLN	BLS
53 Average weekly hours: Production and nonsupervisory employees: Overtime: Manufacturing (SA, hr)	DLN	BLS
54 Civilian unemployment rate: 16 yr + (SA, %)	DLN	BLS
55 Civilian unemployment rate, long-term unemployed (27 weeks or more)	DLN	BLS
56 Civilian unemployment rate, short-term unemployed (less than 27 weeks)	DLN	BLS
57 Civilian unemployment rate gap estimate	DLV	BLS & AC
58 Civilian long-term unemployment rate gap estimate	DLV	BLS & AC
59 Civilian short-term unemployment rate gap estimate	DLV	BLS & AC
Financial		
60 Federal funds (effective) rate (% p.a.)	DLN	FRB
61 2-Year Treasury note yield at constant maturity (% p.a.)	DLN	FRB
62 5-Year Treasury note yield at constant maturity (% p.a.)	DLN	FRB
63 10-Year Treasury note yield, constant maturity (% p.a.)	DLN	FRB
64 Long-term Treasury composite, over 10 years (% p.a.)	DLN	TREASURY
65 30-Year Treasury bond yield, constant maturity (% p.a.)	DLN	FRB
66 Yield spread, 10-yr Treasury note less 3-month Treasury bill	LVL	FRB
67 Adjusted monetary base including deposits to satisfy clearing balance contracts (SA, $ bill.)	DLN	FRBSTL
68 Money stock: M1 (SA, $ bill.)	DLN	FRB
69 Money stock: M2 (SA, $ bill.)	DLN	FRB

APPENDIX, cont'd

	Description	Transformation	Source
Financial, cont'd.			
70	Money stock: MZM (zero maturity) (SA, $ bill.)	DLN	FRB/H
71	Nominal broad trade-weighted exchange value of the US$ (January 1997 = 100)	DLN	FRB
72	St. Louis Fed Financial Stress Index (Above 0 = Above-average financial stress)	DLV	FRBSTL
73	Chicago Fed National Financial Conditions Index (+ = Tighter than average)	DLV	FRBCHI
Inflation expectations			
74	TIPS spread, 5-year	LVL	FRB
75	TIPS spread, 7-year	LVL	FRB
76	TIPS spread, 10-year	LVL	FRB
77	TIPS spread, 20-year	LVL	FRB
78	TIPS spread, 30-year	LVL	FRB
79	Cleveland Fed 1-Year expected inflation rate (%)	LVL	FRBCLE
80	Cleveland Fed 2-Year expected inflation rate (%)	LVL	FRBCLE
81	Cleveland Fed 3-Year expected inflation rate (%)	LVL	FRBCLE
82	Cleveland Fed 5-Year expected inflation rate (%)	LVL	FRBCLE
83	Cleveland Fed 7-Year expected inflation rate (%)	LVL	FRBCLE
84	Cleveland Fed 10-Year expected inflation rate (%)	LVL	FRBCLE
85	Cleveland Fed 20-Year expected inflation rate (%)	LVL	FRBCLE
86	Cleveland Fed 30-Year expected inflation rate (%)	LVL	FRBCLE
87	University of Michigan 1-year-ahead inflation expectations, median	LVL	TR/UMICH
88	University of Michigan 1-year-ahead inflation expectations, variance	LVL	TR/UMICH
89	University of Michigan 5- to 10-year-ahead inflation expectations, median	LVL	TR/UMICH
90	University of Michigan 5- to 10-year-ahead inflation expectations, variance	LVL	TR/UMICH
Surveys			
91	ISM: Manufacturing Prices Index (NSA, 50+ = Economic expansion)	DLN	ISM
92	ISM: Nonmanufacturing Prices Index (SA, 50+ = Economic expansion)	DLN	ISM
93	Philly Fed Business Outlook Survey: Future Prices Paid Diffusion Index (SA, %Balance)	DLN	FRBPHIL
94	NFIB: Percent planning to raise average selling prices, net (SA, %)	DLV	NFIB
95	NFIB: Percent planning to raise worker compensation, net (SA, %)	DLV	NFIB
96	1-Year-Ahead Expected change in unemployment rate, net response	LVL	TR/UMICH
Foreign prices			
97	U.S. Import Price Index: All imports (NSA, 2000 = 100)	DLN	BLS
98	U.S. Import Price Index: Nonpetroleum imports (NSA, 2000 = 100)	DLN	BLS
99	Euroarea 11-18: HICP: Total (SA, 2005 = 100)	DLN	EUROSTAT
100	Canada: Consumer Price Index (NSA, 2010 = 100)	DLN	OECD
101	Mexico: Consumer Price Index (NSA, 2010 = 100)	DLN	OECD
102	Japan: CPI: All items including imputed rent (NSA, 2010 = 100)	DLN	OECD
103	Developing Asia: Consumer prices (2005 = 100, NSA)	DLN	IMF
104	Western Hemisphere: Consumer prices (2005 = 100, NSA)	DLN	IMF

APPENDIX, cont'd

Description	Transformation	Source
Nomenclature: By transformation		
DLN: Change in logs		
DLV: Change in levels		
LVL: Levels		
Nomenclature: By data source		
AC: Authors' calculation		
BEA: Bureau of Economic Analysis		
BLS: Bureau of Labor Statistics		
CENSUS: U.S. Census Bureau		
CRB: Commodity Research Bureau		
CORE: CoreLogic		
EIA: U.S. Energy Information Administration		
EUROSTAT: Eurostat		
FHFA: Federal Housing Finance Agency		
FHLMC: Federal Home Loan Mortgage Corporation		
FRB: Board of Governors of the Federal Reserve System		
FRBATL: Federal Reserve Bank of Atlanta		
FRBCHI: Federal Reserve Bank of Chicago		
FRBCLE: Federal Reserve Bank of Cleveland		
FRBDAL: Federal Reserve Bank of Dallas		
FRBPHIL: Federal Reserve Bank of Philadelphia		
FRBSTL: Federal Reserve Bank of St. Louis		
GSA: Green Street Advisors		
H: Haver Analytics		
IMF: International Monetary Fund		
ISM: Institute for Supply Management		
NFIB: National Federation of Independent Business		
OECD: Organisation for Economic Co-operation and Development		
S&P: Standard & Poor's		
TR: Thomson Reuters		
TREASURY: U.S. Department of the Treasury		
UMICH: University of Michigan Survey Research Center		
WSJ: *Wall Street Journal*		

NOTE: CPI, consumer price index; CPI-U, Consumer Price Index for All Urban Consumers; DOE, U.S. Department of Energy; HICP, Harmonised Index of Consumer Prices; KR-CRB, Knight-Ridder Commodity Research Bureau; MMBTU, 1 million British thermal units; MZM, money zero maturity; NSA, not seasonally adjusted; p.a., per annum; PCE, adjusted personal consumption expenditure chain-weighted price index; PPI, producer price index; SA, seasonally adjusted; TIPS, Treasury inflation-protected securities; WTI, West Texas Intermediate.

NOTES

[1] Technically, the NKPC posits that the current-period's inflation rate depends on the next-period's inflation rate and the aggregate real marginal cost of firms in the economy. It is further assumed that aggregate real marginal cost is proportional to the difference between actual and potential output; see Rudd and Whelan (2007). Mavroeidis, Plagborg-Møller, and Stock (2014) highlight the numerous limitations of the NKPC based on the various measures of inflation expectations.

[2] In April 2014, the Board of Governors released the model code and datasets for the staff's workhorse forecasting model, FRB/US. An interested analyst with access to the software required to run FRB/US can now, in principle, generate forecasts from large, structural macroeconometric models; see "FRB/US: About" (http://www.federalreserve.gov/econresdata/frbus/us-models-about.htm).

[3] See Yellen (2014).

[4] Greenbook forecasts can be found at http://www.federalreserve.gov/monetarypolicy/fomc_historical.htm.

[5] A nowcast, sometimes called a tracking forecast, uses a variety of incoming data flows during a quarter to estimate that quarter's inflation rate; see Giannone, Reichlin, and Small (2008).

[6] A direct forecast relates the period-t data directly to the h-period-ahead data. The indirect forecast models a one-period-ahead relationship and propagates that forward, treating the shorter-horizon data as given.

[7] See Gavin and Kliesen (2008).

[8] See Stock and Watson (1999); Bernanke and Boivin (2003); Bernanke, Boivin, and Eliasz (2005); and Giannone, Reichlin, and Sala (2005). Stock and Watson have instead focused on forecasting.

[9] See Smets and Wouters (2007).

[10] See Gavin and Kliesen (2008) for a discussion on this point.

[11] See the August 16, 2013, survey report published by the Philadelphia Fed (http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/2013/survq313.cfm).

[12] See McConnell and Perez-Quiros (2000).

[13] For more information about the estimation, contact the authors.

[14] Recall that the RMSE results in Table 4 suggested that the best-performing models were the FAVAR(1) and FAVAR(6) for CPI inflation for the post-1983 period.

[15] Greenspan (2004) provides a fuller discussion in the context of a Bayesian-type model.

[16] Note that the chart plots smoothed probabilities, which are six-month moving averages. Thus, for March 2012, these are the average probabilities for the six months ending in March 2012.

[17] In quarterly terms, at an annual rate, PCEPI inflation fell from 1.3 percent in 2012:Q2 to 1 percent in 2013:Q1.

[18] The FOMC's projections are published quarterly and are termed the Summary of Economic Projections. See http://www.federalreserve.gov/monetarypolicy/fomcprojtabl20140618.htm.

[19] Converting our monthly forecasts into quarterly forecasts also reveals an expected 1.8 percent increase in the PCEPI from 2013:Q4 to 2014:Q4.

[20] See "Minutes of the Federal Open Market Committee, December 17-18, 2013" (http://www.federalreserve.gov/monetarypolicy/fomcminutes20131218.htm).

REFERENCES

Aruoba, S. Borağan and Diebold, Francis X. "Real-Time Macroeconomic Monitoring: Real Activity, Inflation, and Interactions." *American Economic Review: Papers and Proceedings*, May 2010, *100*(2), pp. 20-24.

Atkeson, Andrew and Ohanian, Lee E. "Are Phillips Curves Useful for Forecasting Inflation?" Federal Reserve Bank of Minneapolis *Quarterly Review*, Winter 2001, *25*(1), pp. 2-11; https://www.minneapolisfed.org/research/qr/qr2511.pdf.

Bernanke, Ben S. and Boivin, Jean. "Monetary Policy in a Data-Rich Environment." *Journal of Monetary Economics*, April 2003, *50*(3), pp. 525-46.

Bernanke, Ben S.; Boivin, Jean and Eliasz, Piotr. "Measuring Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics*, February 2005, *120*(1), pp. 387-422.

Board of Governors of the Federal Reserve System. "Why Does the Federal Reserve Aim for 2 Percent Inflation Over Time?" *Current FAQs*, September 26, 2013; http://www.federalreserve.gov/faqs/economy_14400.htm.

Ciccarelli, Matteo and Mojon, Benoit. "Global Inflation." *Review of Economics and Statistics*, August 2010, *92*(3), pp. 524-35.

Faust, Jon and Wright, Jonathan H. "Forecasting Inflation," in Graham Elliott and Allan Timmermann, eds., *Handbook of Economic Forecasting*. Volume 2A. Amsterdam: North Holland, 2013, pp. 2-56.

Galí, Jordi and Gertler, Mark. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics*, October 1999, *44*(2), pp. 195-222.

Gavin, William T. and Kliesen, Kevin L. "Forecasting Inflation and Output: Comparing Data-Rich Models with Simple Rules." Federal Reserve Bank of St. Louis *Review*, May/June 2008, *90*(3, Part 1), pp. 175-92; http://research.stlouisfed.org/publications/review/08/05/GavinKliesen.pdf.

Giannone, Domenico; Reichlin, Lucrezia and Sala, Luca. "Monetary Policy in Real Time," in Mark Gertler and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2004*. Cambridge, MA: MIT Press, 2005, pp. 161-200.

Giannone, Domenico; Reichlin, Lucrezia and Small, David. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics*, May 2008, *55*(4), pp. 665-76.

Greenspan, Alan. "Risk and Uncertainty in Monetary Policy." *American Economic Review*, May 2004, *94*(2), pp. 33-40.

Mavroeidis, Sophocles; Plagborg-Møller, Mikkel and Stock, James H. "Empirical Evidence on Inflation Expectations and the New Keynesian Phillips Curve." *Journal of Economic Literature*, March 2014, *52*(1), pp. 124-88.

McConnell, Margaret M. and Perez-Quiros, Gabriel. "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?" *American Economic Review*, December 2000, *90*(5), pp. 1464-76.

Neely, Christopher J. and Rapach, David E. "International Comovements in Inflation Rates and Country Characteristics." *Journal of International Money and Finance*, November 2011, *30*(7), pp. 1471-90.

Rudd, Jeremy and Whelan, Karl. "Modeling Inflation Dynamics: A Critical Review." *Journal of Money, Credit, and Banking*, February 2007, *39*(Suppl. s1), pp. 156-70.

Smets, Frank and Wouters, Rafael. "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSG Approach." *American Economic Review*, June 2007, *97*(3), pp. 586-606.

Stock, James H. and Watson, Mark W. "Forecasting Inflation." *Journal of Monetary Economics*, October 1999, *44*(2), pp. 293-335.

Stock, James H. and Watson, Mark W. "Why Has Inflation Become Harder to Forecast?" *Journal of Money, Credit, and Banking*, February 2007, *39*(Suppl. s1), pp. 3-33.

Yellen, Janet L. "Monetary Policy and the Economic Recovery." Remarks at the Economic Club of New York, April 16, 2014; http://www.federalreserve.gov/newsevents/speech/yellen20140416a.htm.

Risk Aversion at the Country Level

Néstor Gandelman and Rubén Hernández-Murillo

This article estimates the coefficient of relative risk aversion for 75 countries using data on self-reports of personal well-being from the 2006 Gallup World Poll. The analysis suggests that the coefficient of relative risk aversion varies closely around 1, which corresponds to a logarithmic utility function. The authors conclude that their results support the use of the log utility function in numerical simulations of economic models. (JEL D80, D31, I31, O57)

An individual's attitude about risk underlies economic decisions about the optimal amount of retirement or precautionary savings to set aside, investment in human capital, public or private sector employment, and entrepreneurship, among other things. In the aggregate, these micro-level decisions can influence a country's growth and development.

Although there is a vast literature on measuring risk aversion, estimates of the coefficient of relative risk aversion vary widely—from as low as 0.2 to 10 and higher. Probably the most widely accepted measures lie between 1 and 3.[1] The most common approach to measuring risk aversion is based on a consumption-based capital asset pricing model (CAPM). Hansen and Singleton (1982), using the generalized method of moments (GMM) to estimate a CAPM, report that relative risk aversion is small. Hall (1988) shows that minor changes in the specification and instruments cause the results to vary substantially. Neely, Roy, and Whiteman (2001), in turn, explain this difference, arguing that CAPM-based estimations fail to provide robust results because difficulties in predicting consumption growth and asset returns from available instruments lead to a near identification failure of the model. In this article, we follow a different approach.

We build on the methodology first outlined in Layard, Mayraz, and Nickell (2008). Using happiness data to estimate how fast the marginal utility of income declines as income increases,

they use an iterated maximum likelihood procedure that assumes a constant relative risk aversion (CRRA) utility function. Under this assumption, the elasticity of the marginal utility of income corresponds to the parameter of relative risk aversion. In Gandelman and Hernández-Murillo (2013), we also used this methodology to estimate the coefficient of relative risk aversion using pooled data from cross-sectional and panel datasets. Instead of maximum likelihood, in this article we use the GMM to perform the estimation. As with maximum likelihood, the GMM provides consistent and asymptotically normal estimates, but it does not rely on the normality assumption. Using the GMM also provides asymptotically correct standard errors for the coefficient of relative risk aversion, whereas the iterated maximum-likelihood procedure used in Layard, Mayraz, and Nickell (2008) and Gandelman and Hernández-Murillo (2013) does not easily provide a measure of the standard error for the parameter of interest.

The CRRA utility function is often used in applied theory and empirical work because of its tractability and appealing implications.[2] Assuming a CRRA form for the utility function, nevertheless, has been criticized. For example, Geweke (2001) warns about the potential limitations of assuming a CRRA utility function for traditional growth models. He argues that, under this assumption, the existence of expected utility, and hence of an operational theory of choice, depends on distributional assumptions about macroeconomic variables and about prior information that do not necessarily hold. Because many distributions are difficult to distinguish econometrically, these assumptions may lead to widely different implications for choice under uncertainty. Another potential limitation is that, in dynamic models with a CRRA per-period utility function with time-separable preferences, the coefficient of relative risk aversion is also the reciprocal of the elasticity of intertemporal substitution (EIS). Epstein and Zin (1989, 1991) address this issue with a generalization of the standard preferences in a recursive representation in which current utility is a constant elasticity function of current consumption and future utility. This more-flexible representation of utility allows for differentiation between the coefficient of relative risk aversion and the EIS and is useful for explaining problematic aspects of asset pricing behavior.[3] We acknowledge these criticisms, but we follow the happiness literature in assuming a CRRA form for the utility function because it provides a straightforward framework that can be used to estimate a measure of risk aversion summarized in a single parameter. This simple form is particularly useful when, as is our case, the only available data are cross-sectional observations on subjective well-being and income.

In estimating risk aversion, the literature has focused almost exclusively on developed countries.[4] Moreover, with the exception of Szpiro (1986) and Szpiro and Outreville (1988), to the best of our knowledge no additional study has yet applied a homogenous methodology for estimating risk aversion for a large set of both high- and low-income countries. Szpiro (1986) initially used property/liability insurance data to estimate relative risk aversion for 15 developed countries. Szpiro and Outreville (1988) augmented the analysis to 31 countries, including 11 developing countries. Gandelman and Porzecanski (2013) use a slightly different approach. They apply different assumptions about relative risk aversion to a sample of 117 developing and developed countries from the Gallup World Poll to calibrate how much happiness inequality is due to income inequality.

In this article, we fill this gap in the literature by eliciting risk-aversion measures for 75 countries, including 52 developing countries, from self-reports of personal well-being from the 2006 Gallup World Poll. This study is important for several reasons. First, applying the same methodology to different countries is useful for assessing the robustness of the estimates. Second, the study is a starting point for further research of cross-country differences in risk aversion and their correlation with multiple variables of interest. Third, dynamic stochastic general equilibrium models often rely on calibrated estimates of risk aversion for developed countries, usually without measures of the relevant parameters for developing countries. This study includes developing countries.

Our estimates show that individual country estimates of relative risk aversion vary between 0 (implying a linear utility function) and 3 (implying more risk aversion than log utility). We construct Wald tests for the null hypotheses that the coefficient of relative risk aversion equals 0, 1, or 2: 0 indicates a linear utility function in terms of income; 1 indicates a logarithmic utility function; and 2 corresponds to a value often used in the literature, which indicates a higher degree of concavity.[5] Our sample includes 23 developed countries and 52 developing countries. Detailed outcomes of the hypothesis tests for the coefficients in both developed and developing countries are presented in the Results section. In brief, we reject the null hypothesis that the coefficient of relative risk aversion equals 1 in only 2 of the 23 developed countries and only 10 of the 52 developing countries. We reject that it equals 0 or 2 for most developed countries and many developing countries. Furthermore, an analysis of the distribution of the estimates indicates that for both developed and developing countries, most of the estimates are concentrated in the vicinity of 1. We conclude that this result supports the use of the log utility function in numerical simulations.

DATA

The main variables of interest in the Gallup World Poll are self-reports of (i) satisfaction with life and (ii) household income.[6] We also use the following individual control variables: age, gender, marital status, employment status, and residence in urban areas.

The self-reports of well-being from the Gallup World Poll are answers to the following question: *Please imagine a ladder/mountain with steps numbered from zero at the bottom to ten at the top. Suppose we say that the top of the ladder/mountain represents the best possible life for you and the bottom of the ladder/mountain represents the worst possible life for you. If the top step is 10 and the bottom step is 0, on which step of the ladder/mountain do you feel you personally stand at the present time?* Henceforth, we do not distinguish well-being from happiness.

Table 1 reports summary statistics for the key variables in our estimations: the happiness scores, household income, and the control variables. The data for the 75 countries include 40,655 individual observations. The sample is split into developed countries (23) and developing countries (52) following the World Bank criterion: A country is defined as developing if its gross national income per capita is less than $12,000 U.S. dollars in 2010.[7]

Table 1 shows summary statistics of the data for all the countries and for the countries divided into the two groups. On a scale of 0 to10, the means of average reported happiness were 5.5 for the overall sample, 6.7 for developed countries, and 4.9 for developing countries.[8] In terms of the control variables, the overall sample includes individuals with an average age

Table 1

Summary Statistics

Variable	All countries (N = 75)				Developed countries (n = 23)				Developing countries (n = 52)			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
No. obs.	542	150	230	1,241	552	95	418	867	538	170	230	1,241
Happiness	5.5	1.2	3.4	7.8	6.7	0.8	5.3	7.8	4.9	0.8	3.4	7.2
Income (%)	92.3	8.6	47.1	109.8	97.8	3.5	91.0	101.7	89.9	9.1	47.1	109.8
Age (yr)	42.4	2.8	36.3	47.7	44.7	1.5	42.0	47.7	41.3	2.7	36.3	46.9
Female (%)	55.6	6.4	42.4	72.2	58.0	6.1	49.3	72.2	54.5	6.4	42.4	69.2
Married (%)	69.1	10.0	36.2	90.2	68.9	6.5	55.6	82.8	69.3	11.2	36.2	90.2
Urban (%)	44.6	19.2	5.0	87.4	42.9	15.0	24.6	75.8	45.3	20.9	5.0	87.4
Employed (%)	59.9	14.0	23.7	88.3	71.6	9.4	52.6	88.3	54.7	12.6	23.7	86.9

NOTE: Developed countries are those with gross national income per capita greater than $12,000 USD in 2010. Statistics are the country averages of the variable. Income is expressed relative to the country average. The mean does not equal 100 percent because outlier observations were trimmed. No. obs., number of observations; SD, standard deviation.

of 42.4 years, with slightly more women (55.6 percent) than men (44.4 percent), more married (69.1 percent) than single (30.9 percent), less than half (44.6 percent) living in an urban area, and over half (59.9 percent) employed. Comparing the samples for the developed and developing countries, the average age of the adults in the developed countries is older (44.7 years old vs. 41.3 years old), with a slightly larger percentage of women (58.0 percent vs. 54.5 percent) and a significantly higher percentage employed (71.6 percent vs. 54.7 percent). Both include about the same percentages of married individuals (around 69.0 percent), while slightly more individuals live in urban areas in developing countries than in developed countries (45.3 percent vs. 42.9 percent).

ESTIMATION

To perform the estimation we make several assumptions. First, we assume a CRRA form for the utility function. Second, because consumption data are not available, we assume that the utility function can be expressed in terms of income. Furthermore, because the measure of income typically available in happiness surveys (including the Gallop World Poll) is current household income, as opposed to permanent individual income, the utility function we estimate represents per-period utility instead of lifetime utility. Finally, because we are using self-reports of well-being as a proxy for utility, we make assumptions about the comparability of the responses across individuals.

Utility Function

We assume that an individual's experienced utility, u, can be explained, in addition to income, y, by a (row) vector of individual characteristics x via the function U: $u = U(y,\text{x})$. We

assume that the relation U is common to all individuals in a given country and is of the following form:

(1)
$$u = U(y, \mathrm{x}) = \alpha + \gamma g(y) + \mathrm{x}\beta,$$

where α and γ are scalars, β is a column vector of the coefficients for the controls x, and g is a CRRA utility function for the relation with income

(2)
$$g(y) = \begin{cases} \dfrac{y^{1-\rho} - 1}{1-\rho} & \text{if } \rho \neq 1, \\ \log(y) & \text{if } \rho = 1, \end{cases}$$

where ρ corresponds to the Arrow-Pratt coefficient of relative risk aversion. According to this specification, income enters the utility function as a proxy for consumption. In other words, this specification assumes that the effect of income on reported happiness corresponds to the causal effects of consumption on utility. While we follow previous studies in making this assumption, we recognize that it is not trivial and acknowledge its potential limitations.[9]

We also assume that reported happiness, h, is linked to experienced utility via a monotonically increasing function $f: h = f(u)$.[10] For simplicity, as in most of the literature, we assume that the relation f is common to all individuals. Furthermore, we assume that reported happiness scores are cardinally comparable across individuals, which implies that the relation f is linear. The cardinality assumption justifies the estimation with ordinary least squares (OLS) as in Layard, Mayraz, and Nickell (2008) and Gandelman and Hernández-Murillo (2013). Alternatively, assuming that happiness scores are ordinally comparable would justify the estimation with ordered probit or ordered logit. Ferrer-i-Carbonell and Frijters (2004) report that the results from either assumption are indistinguishable in most studies using cross-sectional datasets, and since OLS estimates are easier to interpret, this method is often preferred. The results may differ when using panel data, however, if time-invariant effects are important. Therefore, Ferrer-i-Carbonell and Frijters (2004) argue that one can practically assume that happiness scores are interpersonally comparable *both* cardinally and ordinally.

Layard, Mayraz, and Nickell (2008) studied the implications of relaxing the linearity assumption on f. They were concerned especially that the bounded happiness scale would induce compression of the responses, particularly at the top of the scale. The authors found a small degree of concavity near the top of the scale, which implies that the estimate of the coefficient of interest may be biased upward under the linearity assumption. However, the authors determined that relaxing the linearity assumption had only a small effect on their conclusions, and therefore we maintain this assumption in our exercise.

Estimation: Happiness and Utility

The estimated equation for a representative country is therefore

(3)
$$h_i = \alpha + \gamma g(y_i) + \mathrm{x}_i\beta, + v_i,$$

where $i = 1,\ldots, n$ indexes individuals, h_i is the index of reported happiness (on a 0 to 10 scale), and v_i represents an error term that is independent of experienced utility, u_i.[11]

We estimate the model with the GMM. Stacking the individual observations and letting $h = (h_1, h_2,\ldots, h_n)'$, the estimated equation is a nonlinear vector-valued function $H: R^{K+3} \rightarrow R^n$, of the parameters $= (\alpha, \gamma, \rho, \beta')'$, $h = H(\theta)$, where β is a $(K \times 1)$ vector of coefficients for the control variables x_i. Because of the CRRA assumption, we have more parameters than independent variables, so we need an appropriate set of instruments to conduct the estimation. Following Stewart (2011), we construct the set of instruments taking advantage of the nonlinearity of the specification as $Z = [J(\theta), X]$.[12] $J(\theta)$ is the $n \times (K + 3)$ Jacobian matrix of first derivatives of the function H with respect to the parameter vector θ, where each row corresponds to the vector $(1, g(y_i), \gamma m(y_i), x_i)$, where $m(y_i) = \dfrac{\partial g(y_i)}{\partial \rho}$, and X is the $n \times (K + 2)$ data matrix, where each row corresponds to the vector $(1, y_i, x_i)$. Therefore, the matrix of instruments Z simplifies to a matrix with the following characteristic row: $z_i = (1, g(y_i), \gamma m(y_i), y_i, x_i)$.[13]

RESULTS

Table 2 reports the estimates of the relative risk aversion coefficient for the 75 countries in our sample.[14] The estimates range between 0 and 3. The median and simple averages of the country estimates are 0.94 and 0.98, respectively. The average coefficient among developed countries is 0.92, while that among developing countries is 1.00. For each country we report Wald tests of the null hypotheses that the coefficient of relative risk aversion equals 0, 1, or 2. The null hypothesis that ρ equals 0 is rejected at the 10 percent level in 13 of the 23 developed countries and 34 of the 52 developing countries. In turn, the null hypothesis that ρ equals 1 is rejected at the 10 percent level in 2 developed countries and 10 developing countries. Finally, the null hypothesis that ρ equals 2 is rejected at the 10 percent level in 17 developed countries and 36 developing countries.

Figures 1 and 2 present the individual country estimates of the coefficient of relative risk aversion for developed and developing countries, respectively. The plots include 90 percent confidence intervals. The plots again indicate that, for most of the estimates in the middle of the distribution, we cannot reject that the coefficient is equal to 1. As shown in Figure 3, this conclusion is confirmed by a plot of the kernel density estimators, which indicates that most of the estimates for both the developed and developing countries are concentrated in the vicinity of 1. In addition, the distribution of the estimates for the developed countries seems to contain relatively more observations between 0 and 0.5, whereas that for the developing countries seems to contain relatively more observations around 2.

Table 2A

Relative Risk Aversion by Country (Developed Countries)

	Country	ρ	χ^2 $H_0: \rho = 0$	χ^2 $H_0: \rho = 1$	χ^2 $H_0: \rho = 2$	No. obs.
1	Australia	1.17	21.86*	0.47	10.93*	594
2	Austria	1.08	2.84*	0.02	2.03	465
3	Belgium	1.55	7.20*	0.92	0.59	533
4	Canada	0.83	7.01*	0.31	14.12*	867
5	Croatia	0.31	0.23	1.16	6.92*	489
6	Estonia	0.51	1.70	1.58	14.56*	488
7	Finland	0.57	1.22	0.70	7.73*	433
8	France	1.43	2.08	0.19	0.33	490
9	Germany	0.77	6.06*	0.53	15.35*	630
10	Greece	1.08	6.32*	0.03	4.61*	555
11	Ireland	0.35	0.27	0.91	5.87*	443
12	Japan	0.44	1.19	1.85	14.55*	550
13	Korea	0.27	0.61	4.53*	25.38*	604
14	Netherlands	0.10	0.02	1.36	6.08*	531
15	New Zealand	1.15	8.75*	0.16	4.70*	565
16	Norway	1.16	2.29	0.05	1.18	647
17	Poland	0.38	0.62	1.62	11.11*	513
18	Portugal	1.07	9.91*	0.04	7.44*	418
19	Slovenia	0.83	7.49*	0.33	15.07*	527
20	Switzerland	1.21	3.69*	0.11	1.59	528
21	Taiwan	2.45	16.88*	5.91*	0.57	566
22	United Kingdom	1.03	17.71*	0.01	15.85*	640
23	United States	1.39	18.85*	1.48	3.64*	610

NOTE: Developed countries are those with gross national income per capita greater than $12,000 USD in 2010. The chi-square statistics correspond to the likelihood ratio tests for the null hypotheses that $\rho = 0$, $\rho = 1$, or $\rho = 2$. * indicates statistical significance at the 10 percent level. No. obs., number of observations.

Table 2B

Relative Risk Aversion by Country (Developing Countries)

	Country	ρ	χ^2 $H_0: \rho = 0$	χ^2 $H_0: \rho = 1$	χ^2 $H_0: \rho = 2$	No. obs.
1	Albania	0.14	0.24	8.73*	40.90*	453
2	Argentina	1.20	4.03*	0.11	1.78	410
3	Armenia	0.57	2.12	1.21	13.38*	520
4	Azerbaijan	1.85	15.97*	3.37*	0.10	565
5	Bangladesh	1.30	11.51*	0.61	3.34*	661
6	Belarus	0.09	0.02	1.66	7.28*	528
7	Benin	0.21	0.30	4.49*	22.91*	467
8	Bolivia	0.16	0.16	4.63*	22.10*	450
9	Bosnia & Herzegovina	0.72	6.84*	1.03	21.60*	889
10	Botswana	0.94	29.44*	0.12	37.55*	453
11	Brazil	0.63	0.33	0.11	1.52	612
12	Bulgaria	1.06	14.58*	0.04	11.53*	466
13	Burundi	2.17	4.06*	1.18	0.02	451
14	Cameroon	0.82	3.41*	0.17	7.18*	504
15	Chile	1.13	20.56*	0.26	12.38*	481
16	Dominican Republic	0.32	0.83	3.68*	22.53*	332
17	Ecuador	1.39	5.87*	0.46	1.14	548
18	El Salvador	0.54	2.15	1.60	15.94*	387
19	Georgia	0.88	3.26*	0.06	5.25*	541
20	Ghana	0.63	4.40*	1.54	20.97*	379
21	Honduras	0.91	4.51*	0.05	6.56*	230
22	India	0.92	1.28	0.01	1.76	1,241
23	Indonesia	1.24	9.70*	0.36	3.70*	758
24	Kosovo	1.03	6.15*	0.01	5.46*	521
25	Kyrgyz Republic	1.81	7.54*	1.50	0.09	564
26	Lao People's Dem. Rep.	0.39	0.50	1.21	8.44*	627

NOTE: Developed countries are those with gross national income per capita greater than $12,000 USD in 2010. The chi-square statistics correspond to the likelihood ratio tests for the null hypotheses that $\rho = 0$, $\rho = 1$, or $\rho = 2$. * indicates statistical significance at the 10 percent level. No. obs., number of observations.

Table 2B, cont'd

Relative Risk Aversion by Country (Developing Countries)

	Country	ρ	χ^2 $H_0: \rho = 0$	χ^2 $H_0: \rho = 1$	χ^2 $H_0: \rho = 2$	No. obs.
27	Lithuania	1.23	18.51*	0.64	7.27*	452
28	Macedonia	1.34	15.43*	1.00	3.71*	563
29	Madagascar	0.72	2.33	0.36	7.45*	618
30	Malaysia	1.93	1.71	0.40	0.00	497
31	Mexico	0.78	1.22	0.10	3.02*	469
32	Moldova	1.19	8.58*	0.23	3.91*	545
33	Montenegro	2.10	11.38*	3.14*	0.03	322
34	Mozambique	1.11	19.22*	0.19	12.38*	486
35	Myanmar	1.01	10.72*	0.00	10.28*	749
36	Panama	0.18	0.25	4.83*	23.92*	476
37	Paraguay	0.47	0.23	0.29	2.39	480
38	Peru	1.44	6.72*	0.63	1.01	359
39	Russia	0.65	5.02*	1.46	21.69*	1,000
40	Senegal	1.89	4.65*	1.03	0.02	407
41	Serbia	0.27	0.35	2.60	14.54*	815
42	South Africa	1.29	36.15*	1.79	11.13*	458
43	Sri Lanka	0.68	4.23*	0.91	15.72*	692
44	Tajikistan	1.19	4.96*	0.12	2.33	523
45	Tanzania	1.26	7.11*	0.30	2.46	395
46	Uganda	0.67	20.24*	5.04*	80.79*	497
47	Ukraine	0.44	0.44	0.69	5.41*	564
48	Uruguay	0.90	11.74*	0.15	17.59*	485
49	Uzbekistan	2.96	14.59*	6.40*	1.54	551
50	Venezuela	2.08	11.13*	2.99*	0.01	452
51	Vietnam	1.15	18.74*	0.32	10.18*	558
52	Zimbabwe	0.04	0.00	0.93	3.88*	518

NOTE: Developed countries are those with gross national income per capita greater than $12,000 USD in 2010. The chi-square statistics correspond to the likelihood ratio tests for the null hypotheses that $\rho = 0$, $\rho = 1$, or $\rho = 2$. * indicates statistical significance at the 10 percent level. No. obs., number of observations.
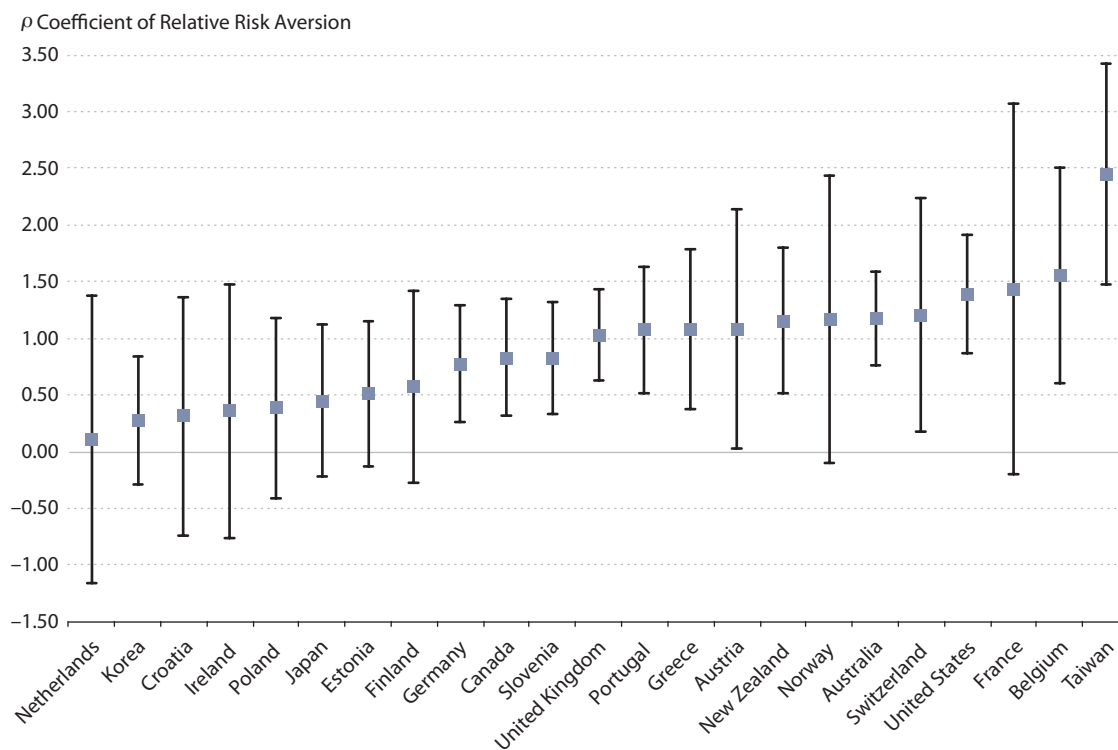
Figure 1

Relative Risk Aversion Among Developed Countries

NOTE: The squares indicate point estimates. The vertical lines represent the 90 percent confidence intervals.

CONCLUSION

The financial economics literature has made a significant effort to find adequate measures of risk aversion, but in general has focused on providing estimates for a limited set of mostly developed countries. Szpiro and Outreville (1988), for example, study 31 countries, including only 11 developing countries. Their methodology uses insurance data and primarily tests the hypothesis of constant relative risk aversion, which cannot be rejected for the majority of countries considered. In this article, we modify the methodology of Layard, Mayraz, and Nickell (2008) and Gandelman and Hernández-Murillo (2013) to estimate the coefficient of relative risk aversion using subjective well-being data for 75 countries, including 52 developing countries.

Our individual country estimates range from 0 to 3, with an average of 0.98. Wald tests for the vast majority of countries indicate that the coefficient of relative risk aversion is smaller than 2 and largely in the vicinity of 1. These estimates of relative risk aversion are smaller than those found for individual countries by Szpiro and Outreville (1988); their estimates range between 1 and 5, with an average of 2.89. Our estimates are close to the results of Layard, Mayraz, and Nickell (2008) and Gandelman and Hernández-Murillo (2013).
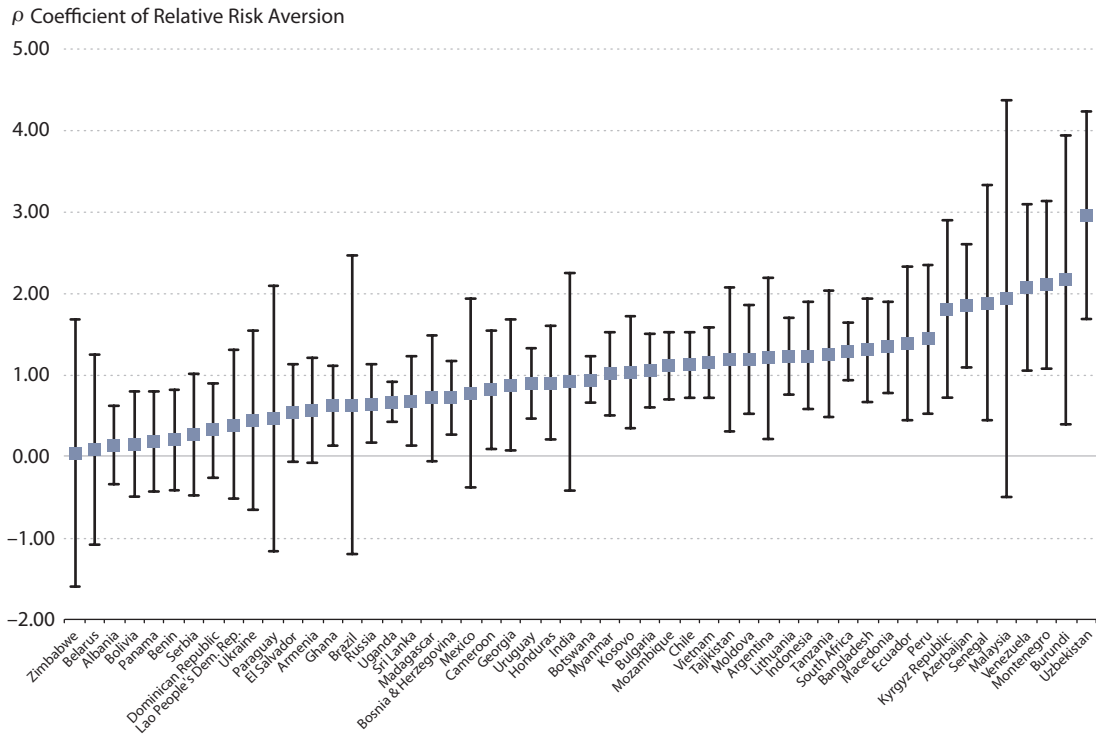
Figure 2

Relative Risk Among Developing Countries

NOTE: The squares indicate point estimates. The vertical lines represent the 90 percent confidence intervals.

Many economic models, including dynamic general stochastic equilibrium models, require estimates of key parameters, including the coefficient of relative risk aversion. Our findings support the use of the log form for the utility function in such exercises, which corresponds to a coefficient of unity for the coefficient of relative risk aversion. Our results also inform the construction of models in which it is important to allow for differing parameterizations for developed and developing countries. ∎

Figure 3

Kernel Density Estimates of the ρ **Distribution**

Density

Developed Countries
Developing Countries

ρ
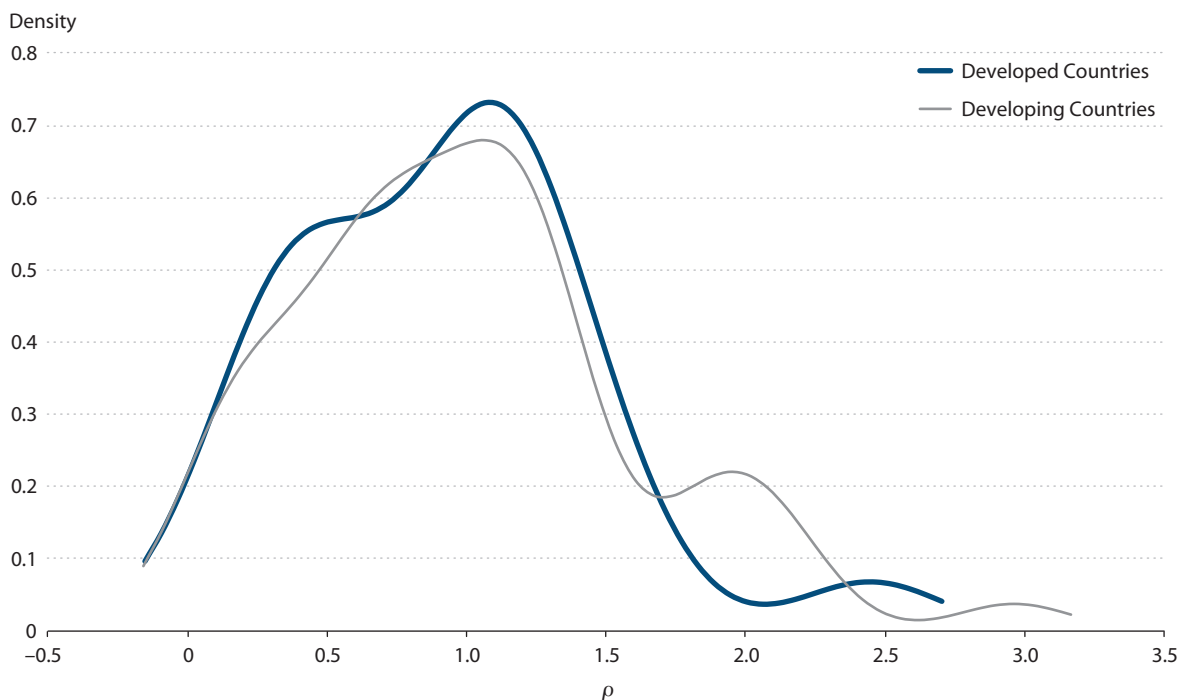
NOTES

[1] See Chetty (2006); Campo et al. (2011); Friend and Blume (1975); Gandelman and Hernández-Murillo (2013); Garcia, Lugar, and Renault (2003); Gordon and St-Amour (2004); Hansen and Singleton (1983); Kapteyn and Teppa (2011); Layard, Mayraz, and Nickell (2008); Mankiw (1985); Szpiro (1986); and Weber (1975).

[2] For example, the CRRA utility function implies stationary risk premia and interest rates even in the presence of long-run economic growth. See Mehra and Prescott (2008) for additional discussions on the implications on the equity premium.

[3] See Kocherlakota (1990) for a criticism of the Epstein-Zin approach and Kocherlakota (1996) for a more in-depth analysis and its implication for the equity premium puzzle.

[4] For an exception, see Gandelman and Hernández-Murillo (2013), who estimate measures of risk aversion for groups of countries classified by income level.

[5] The log utility function has the property that in a trade-off between present and future consumption, the income and substitution effects, in response to changes in the interest rate, exactly offset.

[6] Household income data are reported in 29 brackets. We use the midpoint of the brackets as the measure of income, and for the top bracket we use a value equal to twice the previous midpoint value. In our estimations, income is expressed as deviations from the country average. This normalization facilitates the numerical estimation and has no effect on the estimates of the risk aversion coefficient.

[7] The authors' income definitions are based on the 2010 World Bank income classification groups. The current year's groupings and more information about the classification process can be found on the World Bank's website (http://data.worldbank.org/news/2015-country-classifications).

[8] The reported income means differ from 100 percent because we trimmed outlier observations from the sample.

[9] For further discussion, see Clark, Frijters, and Shields (2008) and the references therein.

[10] For this discussion we follow loosely the notation of MacKerron (2012).

[11] The coefficients are identified up to an affine transformation of the utility function in equation (1).

[12] To be sure that our results are not affected by outliers in the income reports, we run a regression of the log of relative income on individual controls and trim observations in the bottom and top 5 percent of the distribution of residuals, as in Layard, Mayraz, and Nickell (2008).

[13] We implement the estimation with Stata version 12.0 using a wrapper function for the built-in GMM procedure for which we provide the explicit derivatives of the moment conditions. The programs are available upon request from the authors.

[14] We eliminate from the sample developed and developing countries for which the estimation procedure does not find a value for ρ.

REFERENCES

Campo, Sandra; Guerre, Emmanuel; Perrigne, Isabelle; and Vuong, Quang. "Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders." *Review of Economic Studies*, January 2011, *78*(1), pp. 112-47; http://restud.oxfordjournals.org/content/78/1/112.long.

Chetty, Raj. "A New Method of Estimating Risk Aversion." *American Economic Review*, December 2006, *96*(5), pp. 1821-34; http://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.96.5.1821.

Clark, Andrew E.; Frijters, Paul and Shields, Michael A. "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles." *Journal of Economic Literature*, March 2008, *46*(1), pp. 95-144; http://www.jstor.org/stable/27646948.

Epstein, Larry G. and Zin, Stanley E. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption Growth and Asset Returns: A Theoretical Framework." *Econometrica*, July 1989, *57*(4), pp. 937-69; http://www.jstor.org/stable/1913778.

Epstein, Larry G. and Zin, Stanley E. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption Growth and Asset Returns: An Empirical Analysis." *Journal of Political Economy*, April 1991, *99*(2), pp. 263-86; http://www.jstor.org/stable/2937681.

Ferrer-i-Carbonell, Ada and Frijters, Paul. "How Important Is Methodology for the Estimates of the Determinants of Happiness?" *Economic Journal*, July 2004, *114*(497), pp. 641-59; http://www.jstor.org/stable/3590299.

Friend, Irwin and Blume, Marshall E. "The Demand for Risky Assets." *American Economic Review*, December 1975, *65*(5), pp. 900-22; http://www.jstor.org/stable/1806628.

Gandelman, Néstor and Porzecanski, Rafael. "Happiness Inequality: How Much Is Reasonable?" *Social Indicators Research*, January 2013, *110*(1), pp. 257-69; http://link.springer.com/article/10.1007/s11205-011-9929-z.

Gandelman, Néstor and Hernández-Murillo, Rubén. "What Do Happiness and Health Satisfaction Data Tell Us About Relative Risk Aversion?" *Journal of Economic Psychology*, August 2013, *39*, pp. 301-12; http://www.sciencedirect.com/science/article/pii/S0167487013001116.

Geweke, John. "A Note on Some Limitations of CRRA Utility." *Economics Letters*, June 2001, *71*(3), pp. 341-45; http://www.sciencedirect.com/science/article/pii/S0165176501003913.

Garcia, René; Luger, Richard and Renault, Eric. "Empirical Assessment of an Intertemporal Option Pricing Model with Latent Variables." *Journal of Econometrics*, September-October 2003, *116*(1-2), pp. 49-83; http://www.sciencedirect.com/science/article/pii/S0304407603001039.

Gordon, Stephen and St-Amour, Pascal. "Asset Returns and State-Dependent Risk Preferences." *Journal of Business and Economic Statistics*, July 2004, *22*(3), pp. 241-52; http://www.jstor.org/stable/1392594.

Hall, Robert. "Intertemporal Substitution in Consumption." *Journal of Political Economy*, April 1988, *96*(2), pp. 339-57; http://www.jstor.org/stable/1833112.

Hansen, Lars Peter and Singleton, Kenneth J. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica*, September 1982, *50*(5), pp. 1269-86; http://www.jstor.org/stable/1911873.

Hansen, Lars Peter and Singleton, Kenneth J. "Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns." *Journal of Political Economy*, April 1983, *91*(2), pp. 249-65; http://www.jstor.org/stable/1832056.

Kapteyn, Arie and Teppa, Federica. "Subjective Measures of Risk Aversion, Fixed Costs, and Portfolio Choice." *Journal of Economic Psychology*, August 2011, *32*(4), pp. 564-80; http://www.sciencedirect.com/science/article/pii/S0167487011000602.

Kocherlakota, Narayana R. "Disentangling the Coefficient of Relative Risk Aversion from the Elasticity of Intertemporal Substitution: An Irrelevance Result." *Journal of Finance*, March 1990, *45*(1), pp. 175-90; http://www.jstor.org/stable/2328815.

Kocherlakota, Narayana R. "The Equity Premium: It's Still a Puzzle." *Journal of Economic Literature*, March 1996, *34*(1), pp. 42-71; http://www.jstor.org/stable/2729409.

Layard, Richard; Mayraz, Guy and Nickell, Stephen John. "The Marginal Utility of Income." *Journal of Public Economics*, August 2008, *92*(8-9), pp. 1846-57; http://www.sciencedirect.com/science/article/pii/S0047272708000248.

MacKerron, George. "Happiness Economics from 35,000 Feet." *Journal of Economic Surveys*, September 2012, *26*(4), pp. 705-35; http://onlinelibrary.wiley.com/doi/10.1111/j.1467-6419.2010.00672.x/full.

Mankiw, N. Gregory. "Consumer Durables and the Real Interest Rate." *Review of Economics and Statistics*, 1985, *67*(3), pp. 353-62.

Mehra, Rajnish; and Prescott, Edward C. "The Equity Premium: ABCs," in Mehra Rajnish, ed., *Handbook of the Equity Risk Premium*. Chap. 1. Amsterdam: Elsevier, 2008.

Neely, Christopher J.; Roy, Amlan and Whiteman, Charles H. "Risk Aversion Versus Intertemporal Substitution: A Case Study of Identification Failure in the Intertemporal Consumption Capital Asset Pricing Model." *Journal of Business & Economic Statistics*, October 2001, *19*(4), pp. 395-403; http://www.tandfonline.com/doi/abs/10.1198/07350010152596646#.VKsfi_PnaUk.

Stewart, Kenneth G. "The Optimal Construction of Instruments in Nonlinear Regression: Implications for GMM Inference." Econometrics Working Paper No. EWP1107, University of Victoria, May 2011; http://www.uvic.ca/socialsciences/economics/assets/docs/econometrics/ewp1107.pdf.

Szpiro, George G. "Relative Risk Aversion Around the World." *Economics Letters*, 1986, *20*(1), pp. 19-21; http://www.sciencedirect.com/science/article/pii/0165176586900728.

Szpiro, George G. and Outreville, Jean-François. "Relative Risk Aversion Around the World: Further Results." *Journal of Banking and Finance*, 1988, *6*(Supplement 1), pp. 127-28; http://www.sciencedirect.com/science/article/pii/0378426688900635.

Weber, Warren E. "Interest Rates, Inflation, and Consumer Expenditures." *American Economic Review*, 1975, *65*(5), pp. 843-58; http://www.jstor.org/stable/1806624.

The Welfare Cost of Business Cycles with Heterogeneous Trading Technologies

YiLi Chien

The author investigates the welfare cost of business cycles in an economy where households have heterogeneous trading technologies. In an economy with aggregate risk, the different portfolio choices induced by heterogeneous trading technologies lead to a larger consumption inequality in equilibrium, while this source of inequality vanishes in an economy without business cycles. Put simply, the heterogeneity in trading technologies amplifies the effect of aggregate output fluctuation on consumption inequality. The welfare cost of business cycles is, therefore, larger in such an economy. In the benchmark economy with a reasonably low risk aversion rate, the business cycle cost is 6.49 percent perperiod consumption for an average household when the model is calibrated to match the risk premium. (JEL C68, D61, D14, E32, G11, G12)

I n a calibrated representative agent model, Lucas (1987) shows a very insignificant welfare gain from the elimination of business cycles. His work suggests that the benefits of stabilizing the cyclical fluctuations in an economy are very limited. Hence, studying the business cycle might not be the top priority in macroeconomics. More recently, Lucas (2003) has argued that most macroeconomics models still fail to generate a sizable welfare cost associated with business cycles.

In this article, I investigate the welfare cost of business cycles in an economy where households have heterogeneous trading technologies. In contrast to most research in the incomplete market literature, the menu of assets available in this economy is quite rich. Moreover, households in this model have heterogeneous abilities to access the menu of assets available on the market. My article distinguishes between passive traders, who hold fixed portfolios of stocks and bonds, and active traders, who frequently adjust their portfolios in response to changes in investment opportunities.

The welfare cost of business cycles is defined as the average welfare difference between two economies: one with and one without aggregate output fluctuations. Given the heterogeneous agent economy, the average welfare is computed by taking the expectation not only

YiLi Chien is a senior economist at the Federal Reserve Bank of St. Louis.

over time but also across all idiosyncratic features of the population. In other words, this welfare measure acts as if people were asked ex ante which economy they would like to be born in. Hence, the measure of business cycles can be thought of as the amount of consumption compensation newborns should receive such that they are indifferent in their expected utilities between the two economies.

In the equilibrium of the calibrated economy, heterogeneous trading technologies result in a clear difference between active traders and passive traders with respect to their portfolio choices. In response to the high risk premium, households with more sophisticated trading technologies take greater aggregate risks by holding a large fraction of equities in their portfolios. They also optimally adjust their portfolios in response to changes in investment opportunities. On the other hand, households with less sophisticated trading technologies take a more cautious approach. On average, they bear less aggregate risk by holding a smaller fraction of equities in their portfolios and do not actively respond to changes in market conditions. The active traders ultimately earn a high rate of return on their portfolios, accumulating more wealth and enjoying a high level of consumption, while the passive investors earn a low return on their portfolios, thereby acquiring relatively low levels of wealth and consuming less. Hence, heterogeneous trading technologies induce more consumption inequality in this economy. Most importantly, the higher consumption inequality across the population leads to lower welfare under my welfare measure.

Clearly, the source of consumption inequality depends heavily on the risk premium level and the variation in the market price of risk; both are linked tightly to business cycle fluctuations. A reduction of aggregate output volatility helps reduce not only the size but also the time variation of the risk premium. This reduction downplays the role of portfolio choice and, hence, reduces consumption inequality and improves welfare in the economy. Without aggregate risk, the inequality in consumption caused by the heterogeneous trading technologies disappears since the composition and timing of portfolio choice no longer affect the rate of return. In short, all assets are risk free and offer exactly the same rate of return. A more sophisticated trading technology does not offer any advantage in an environment without aggregate risk.

I conjecture that heterogeneous trading technologies may contribute to the welfare cost of business cycles. In an economy with aggregate risk, the different portfolio choices across households lead to a larger consumption inequality in equilibrium, while this source of inequality vanishes in an economy without business cycles. In short, the consumption inequality caused by aggregate output fluctuation is amplified by the heterogeneous trading technologies. The welfare cost of business cycles is, therefore, larger in such an economy.

This article uses the modified macroeconomics model developed by Chien, Cole, and Lustig (2011) to evaluate the conjecture quantitatively. Their model incorporates heterogeneous trading technologies into an otherwise standard macroeconomics model. In my use of the model, the heterogeneity in trading technologies is calibrated to match the high risk premiums seen in the historical U.S. data. The welfare cost is measured by the average percentage of per-period consumption compensation received by a newly born household in an economy without business cycles, such that this household is indifferent to an environment with aggregate fluctuations. I find that the welfare gain from the elimination of business cycles is large,

with a reasonably low risk aversion coefficient. In my benchmark case where the risk aversion coefficient is 4, the business cycle costs each household in my economy 6.49 percent of per-period consumption. The welfare cost is significantly larger than that calculated by Lucas (1987).

I also compute the case in which all households are active traders and are endowed with the same sophisticated trading technologies. Given the parameter values in my benchmark calibration, the results show a low risk premium and a much smaller cost of business cycles. The importance of this computational exercise is twofold. First, it shows how an inferior investment technology among some of the investors influences the patterns of return in asset markets. If all households make no investment mistakes, the asset pricing result is dampened compared with that in my benchmark economy. Second, it demonstrates a large welfare loss resulting from poor investment strategies. This exercise shows that the welfare cost of business cycles is much smaller if no household makes investment errors. The heterogeneous trading technologies contribute significantly and mostly to the welfare cost number in my benchmark economy.

The assumption of heterogeneous trading technologies is critical to my results. The question thus arises: How realistic is the assumption of heterogeneous trading technologies? The answer can be found in empirical studies and data that have shown a high amount of heterogeneity in household portfolio choices. Different households behave as if they had access to different menus of tradable assets. In the United States, a majority of households do not invest directly in equity despite the sizable historical equity premiums. Even for those who participate in the equity market, most do not frequently adjust the composition of their portfolios, regardless of the large countercyclical variation of Sharpe ratios (SRs) in the equity market. Put simply, they miss the market timing. However, a small fraction of households hold a large share of stock and constantly change their equity position in response to the high variable-risk premiums. Therefore, these households end up richer but have more exposure to aggregate risk. Parker and Vissing-Jorgensen (2009) show that the consumption of the richest 10 percent of U.S. households is five times more exposed to aggregate risk than that of average households.

This article is closely related to the body of literature in which the distribution effects on consumption inequality might justify a large welfare cost of business cycles. Krusell and Smith (1999) propose the idea that business cycles might worsen the consumption inequality across the population while the impact on average households is insignificant. The higher cost of business cycles is due to the distributional impact of consumption among the rich and poor. Evidently, the distributional impact is missing in a representative agent economy. Krusell et al. (2009) use an incomplete market model calibrated to the wealth distribution in the United States to evaluate the welfare cost of business cycles. Using the same parameter for risk aversion as in Lucas (1987), they find the welfare cost is approximately 0.1 percent of household consumption. Although the welfare cost is one magnitude larger than that calculated by Lucas, it is still negligible in an economic sense. Storesletten, Telmer, and Yaron (2001) consider the welfare cost of business cycles in an environment with countercyclical variations in idiosyncratic shock. A more volatile idiosyncratic income risk during recessions can amplify the cost of aggregate risk in individual consumption and leads to a higher distributional impact. Although the welfare cost of business cycles is still insignificant, the cost increases rapidly as

the risk aversion coefficient increases. Krebs (2007) extends the concept of idiosyncratic labor income shock by adding a permanent job displacement risk. The risk of job displacement is assumed to be closely associated with the business cycle. He finds a sizable cost of business cycles related to the importance of displacement risks.

The central idea of this corpus of literature is to translate a low scale of aggregate risk into a large consumption inequality. I follow this concept, but the large consumption inequality in my model is caused by a novel feature: heterogeneity in trading technologies. Most articles in this literature operate under incomplete market models in which all households can trade only a very limited menu of tradable assets. However, the actual menu of assets that households can trade is quite rich. Instead of assuming a limited set of tradable assets, I introduce a heterogeneous ability to access the menu of assets, motivated by the empirical evidence of heterogeneity in portfolio choices. With heterogeneous trading technologies, households' total incomes differ not only because of their idiosyncratic risk in labor income but also because of the variations in their investment returns resulting from the heterogeneity in trading technologies. In addition, heterogeneous trading technologies affect the return of portfolio choices only in an economy with aggregate risk. Without business cycles, the cost of consumption inequality from different trading technologies disappears. Therefore, the heterogeneity in trading technologies only enlarges the consumption inequality in an economy with aggregate risk and, hence, amplifies the cost of business cycles.

Alvarez and Jermann (2004) demonstrate a close link between the cost of business cycles and risk premiums. Their work offers an alternative and intuitive way to measure the cost of business cycles by using asset pricing data. The cost of business cycles can be considered as the valuation difference between two consumption claims: One pays a constant stream of consumption and the other pays a stochastic stream of consumption. In a representative agent framework, the former claim represents a consumption stream in an economy without aggregate shocks and the latter represents a consumption stream with aggregate shocks. Hence, their work illustrates that, under a representative agent economy, the welfare cost of business cycles can be approximated by the risk premium between an aggregate stochastic consumption claim and a risk-free asset, regardless of the assumptions about utility function. Based on their observation, one can infer that if a model generates a high risk premium, then it might also imply a larger welfare cost of business cycles. My calibrated model produces a realistic result for asset pricing; however, my model's mechanism differs from that in Alvarez and Jermann (2004). The large welfare cost results mainly from the consumption inequality induced by heterogeneous trading technologies, not directly from the variations in aggregate consumption over time.

This article also relates to the fast-growing body of literature on household finance. Campbell (2006) points out that some households might make various mistakes when facing complicated financial decisions. This article evaluates the welfare cost of some of these mistakes. In the model economy, passive traders make two types of mistakes. Households that do not participate in the equity market forgo the large equity premium. Also, those equity investors who do not frequently change their portfolio choices miss the market timing. By comparing the results of two model economies, one with heterogeneous trading technologies

and the other without heterogeneous trading technologies, I demonstrate that these investment mistakes not only affect the risk premium patterns but also cause a large welfare cost. If all households consist of active traders who do not make any investment mistakes, then the risk premium is low and stable in the calibrated economy. Moreover, the welfare cost of business cycles is almost negligible and similar to the result found by Lucas (1987). This finding emphasizes the importance of the study of household finance because preventing investment mistakes can considerably improve welfare.

The next section describes the environment and trading technologies, followed by a section discussing the calibration of the model. Then the results and sensitivity analyses are presented. The final section offers the conclusion.

MODEL

The model setup closely follows that in Chien, Cole, and Lustig (2011). The novel feature of their model is the imposition of restrictions on the menu of assets that households are able to trade, which defines the trading technology a household owns. These restrictions are imposed exogenously to capture the observed portfolio behavior of most households.

I refer to households as *passive traders* if they take their portfolio composition as given and simply choose how much to save or dissave in each period. Other households constantly manage their portfolios in response to changes in the investment opportunity set. I refer to these households as *active traders* since they optimally adjust the composition of their portfolios every period. Note that the passive traders are completely rational except in their portfolio choice decisions. They fully acknowledge the rate of return on their portfolios and adjust their consumption and saving decisions accordingly. Hence, the results are clearly driven by the only additional novel assumption—heterogeneous trading technologies—in contrast to most research in the incomplete market literature.

Environment

This endowment economy consists of a continuum of heterogeneous households subject to both idiosyncratic income shocks and aggregate endowment shocks. The total measure of households is normalized to 1. The heterogeneity across households arises from two assumptions. In the planning period $t = 0$, households receive a one-time permanent shock to their trading technologies, while all other characteristics of the households are identical. Starting at period 1, these households also differ in terms of the realization of an idiosyncratic income shock at all subsequent periods. All households start with the same amount of initial wealth and face an identical stochastic process of idiosyncratic income shocks.

In the model, time is discrete, infinite, and indexed by $t = 0,1,2,\ldots$. The first period, $t = 0$, is the planning period in which financial contracting takes place. I use $z_t \in Z$ to denote the aggregate shock in period t and $\eta_t \in N$ to denote the idiosyncratic shock in period t. The variable z^t denotes the history of aggregate shocks, and similarly η^t denotes the history of idiosyncratic shocks for a household. The idiosyncratic events η are i.i.d. across households with the

mean normalized to 1. I use $\pi(z^t, \eta^t)$ to denote the unconditional probability of state (z^t, η^t) being realized. The events are first-order Markov, and I assume that

$$\pi\left(z^{t+1}, \eta^{t+1} \mid z^t, \eta^t\right) = \pi\left(z_{t+1} \mid z_t\right) \pi\left(\eta_{t+1} \mid \eta_t\right).$$

Note that the probability of idiosyncratic events does not depend on the realization of aggregate shocks. As I show later, this article does not consider the countercyclical variation of idiosyncratic risk. I introduce some additional notation: $z^{t+1} \succ z^t$ or $\eta^{t+1} \succ \eta^t$ denotes that the left node is a successor node to the right node. I use $\{z^\tau \succ z^t\}$ to denote the set of successor aggregate histories from z^t onward.

There is a single nondurable good available for consumption in each period, and its aggregate supply is denoted by $Y_t(z^t)$, which evolves according to

$$Y_t\left(z^t\right) = \exp\{z_t\} Y\left(z^{t-1}\right),$$

with $Y(z^0) = 1$. This endowment good comes in two forms. The first part is nondiversifiable income subject to idiosyncratic risk and is denoted by $\gamma Y(z^t)\eta^t$; hence γ is the share of income that is nondiversifiable. Nondiversifiable income cannot be traded in financial markets and may be considered labor income. The second part of the endowment good is diversifiable income, which is not subject to idiosyncratic shocks, and is denoted by $(1-\gamma)Y_t(z^t)$.

All households are infinitely lived, and rank stochastic consumption streams according to the following utility function:

(1)
$$U(\{c\}) = \sum_{t \geq 1, (z^t, \eta^t)}^{\infty} \beta^t \pi\left(z^t, \eta^t\right) \frac{c_t\left(z^t, \eta^t\right)^{1-\alpha}}{1-\alpha},$$

where $\alpha > 0$ denotes the coefficient of relative risk aversion and $c_t(z^t, \eta^t)$ denotes the household's consumption in state (z^t, η^t).

Assets Traded

Three assets are available in this economy: equity, bond, and contingent claims on aggregate shocks. All of these assets are claims on diversifiable income. The actual menu of assets that a household can trade depends on its trading technology. However, this is still an incomplete market economy because there is no state-contingent claim on idiosyncratic shocks.

Following Abel (1999), I simply consider equity as a leveraged claim on aggregate diversifiable income $((1-\gamma)Y_t(z^t))$. The leverage ratio is assumed to be constant over time and denoted by ψ. Let $B_t(z^t)$ denote the supply of a one-period risk-free bond in period t and $R_{t,t-1}^f(z^{t-1})$ denote the risk-free rate between periods $t-1$ and t given the aggregate history z^{t-1}. With a constant leverage ratio, the total supply of $B_t(z^t)$ has to be adjusted such that

$$B_t\left(z^t\right) = \psi\left[W_t\left(z^t\right) - B_t\left(z^t\right)\right],$$

where $W_t(z^t)$ is the price of a claim on aggregate diversifiable income. Because the aggregate diversified income can be decomposed into the interest payment to bond holders and the dividend payment to shareholders, the dividend payment, $D_t(z^t)$, is denoted by

$$D_t(z^t) = (1-\gamma)Y_t(z^t) - R_{t,t-1}^f(z^{t-1})B_{t-1}(z^{t-1}) + B_t(z^t).$$

Traders who invest a fraction $\psi/(1+\psi)$ of their wealth in bonds and the rest in equity hold the market portfolio. I denote the price of equity (a claim on dividend payment $D_t(z^t)$) by $V_t(z^t)$.

The third available asset is the aggregate state-contingent claims. I denote the price of a unit claim on the final good in aggregate state z^{t+1} acquired in aggregate state z^t by $Q_t(z_{t+1}, z^t)$.

I consider a household entering the period with a net financial wealth $\hat{a}_t(z^t, \eta^t)$. This household buys securities in financial markets (state-contingent claims $a_t(z^{t+1}, \eta^{t+1})$, risk-free bonds $b_t(z^t, \eta^t)$, and equity shares $s_t^D(z^t, \eta^t)$) and with consumption $c_t(z^t, \eta^t)$ in the goods markets subject to the following one-period budget constraint:

$$\sum_{z^{t+1} \succ z^t, \eta^{t+1} \succ \eta^t} Q_t(z_{t+1}, z^t)a_t(z^{t+1}, \eta^{t+1})\pi(\eta_{t+1}|\eta_t) + s_t^D(z^t, \eta^t)V_t(z^t) + b_t(z^t, \eta^t) + c_t(z^t, \eta^t)$$
$$\leq \hat{a}_t(z^t, \eta^t) + \gamma Y_t(z^t)\eta_t, \forall z^t, \eta^t,$$

where $\hat{a}_t(z^t, \eta^t)$, the agent's net financial wealth in state (z^t, η^t), is given by the payoffs of his or her state-contingent claim acquired last period, the payoffs from his or her equity position, and the risk-free bond payoffs:

$$\hat{a}_t(z^t, \eta^t) = a_{t-1}(z^t, \eta^t) + s_t^D(z^{t-1}, \eta^{t-1})\left[D_t(z^t) + V_t(z^t)\right] + R_{t,t-1}^f(z^{t-1})b_{t-1}(z^{t-1}).$$

Trading Technology

There are two main classes of traders: active traders and passive traders. Active traders are able to trade state-contingent claims on aggregate shocks. They change their portfolio composition of equity and bonds optimally every period in response to the variations of state-contingent prices. These active traders make no mistakes in their investment choices. In contrast, passive traders cannot trade state-contingent claims, and their portfolio choice is limited by an exogenously assigned and fixed target ϖ for the equity share. I refer to these traders as passive precisely because of their inelastic response to the changes in investment opportunities. These passive traders potentially make two kinds of investment mistakes. First, they miss the market timing if the volatility of the market price of risk is not constant in the equilibrium. Second, for those passive traders who hold small or zero fractions of equity in their portfolios, they relinquish the risk premiums. The welfare cost of their mistakes may be large in the equilibrium, exhibiting a large risk premium and a volatile SR in equity.

In addition, households face exogenous limits on their net asset positions, or solvency constraints,

(2) $$\hat{a}_t\left(z^t,\eta^t\right) \geq 0.$$

Equation (2) reflects the fact that traders cannot borrow against their future nondiversifiable income.

Measurability Restrictions

The portfolio restrictions implied by the different trading technologies can be translated into restrictions on the evolution of net wealth. These restrictions on net wealth are called measurability constraints. Measurability constraints allow us to derive an aggregation pricing kernel and to avoid searching for all the equilibrium prices that clear all markets (see Chien, Cole, and Lustig, 2011, for a detailed discussion). Below, I list the measurability restrictions for each type of trader.

Active Trader. Since idiosyncratic shocks are not spanned for the active traders, their net wealth needs to satisfy

$$\hat{a}_t\left(z^t,\left[\eta_t,\eta^{t-1}\right]\right) = \hat{a}_t\left(z^t,\left[\tilde{\eta}_t,\eta^{t-1}\right]\right),$$

for all t and η_t, $\tilde{\eta}_t \in N$. These constraints guarantee that the net asset positions are the same across all realizations of idiosyncratic shocks in each period since the active traders are not allowed to trade state-contingent claims on idiosyncratic shocks.

Passive Trader. Passive traders who hold a fixed fraction ϖ in levered equity and $1 - \varpi$ in noncontingent bonds in their portfolio earn a portfolio return:

$$R_t^p\left(\varpi,z^t\right) = \varpi R_{t,t-1}^d\left(z^t\right) + (1-\varpi) R_{t,t-1}^f\left(z^{t-1}\right),$$

where $R_{t,t-1}^d(z^t)$ denotes the equity return between periods t and $t-1$ given the realization of history state z^t. Hence, their net financial wealth satisfies this measurability restriction:

$$\frac{\hat{a}_t\left(\left[z_t,z^{t-1}\right],\left[\eta_t,\eta^{t-1}\right]\right)}{R_t^p\left(\varpi,\left[z_t,z^{t-1}\right]\right)} = \frac{\hat{a}_t\left(\left[\tilde{z}_t,z^{t-1}\right],\left[\tilde{\eta}_t,\eta^{t-1}\right]\right)}{R_t^p\left(\varpi,\left[\tilde{z}_t,z^{t-1}\right]\right)},$$

for all t, z_t, $\tilde{z}_t \in Z$, and η_t, $\tilde{\eta}_t \in N$. This restriction is straightforward to understand: The net asset holding at the beginning of period t, $\hat{a}_t$, divided by the portfolio return between periods t and $t-1$, R_t^p, state by state, should all be equal to the net wealth at the end of period $t-1$. Two fixed portfolio choices are worth mentioning here. First, if $\varpi = 1/(1 + \psi)$, then this trader holds the market in each period and earns the return on a claim on aggregate diversifiable income. Second, some passive traders do not participate in the equity market and hold only risk-free assets. I call them nonparticipants, who can be thought of as having a zero equity target share, $\varpi = 0$.

Competitive Equilibrium

The competitive equilibrium for this economy is defined in a standard manner. It consists of a list of bond, equity, and state-contingent claims holdings; a consumption allocation; and

a list of bond, equity, and state-contingent prices such that (i) given these prices, a trader's asset and consumption choices maximize his or her expected utility subject to the budget constraints, the solvency constraints, and the measurability constraints, and (ii) all asset markets clear.

CALIBRATION

This section discusses the calibration of the parameters, the endowment processes, and the composition of trader pools. The next section uses a calibrated version of the model to evaluate the welfare effect of eliminating business cycles. To compute the equilibrium of this economy, I follow the algorithm described by Chien, Cole, and Lustig (2011), who use truncated aggregate histories as state variables. I track the lagged aggregate histories for up to seven periods.

Preferences and Endowments

Lucas (2003) suggested that a reasonable risk aversion coefficient should lie between 1 and 4. My benchmark calibration set the coefficient of relative risk aversion α to 4. To check the robustness of my results with respect to the choice of risk aversion rate, I conduct a sensitivity analysis as detailed in the "Sensitivity Analysis" subsection. The model is calibrated to annual data. The time discount factor β is set to 0.95. Following Chien, Cole, and Lustig (2012), the fraction of nondiversifiable output is set to 90 percent, which is also close to the value in Mendoza, Quadrini, and Ríos-Rull (2009): 88.75 percent.

The process of aggregate output is calibrated to match the aggregate consumption growth moments from Alvarez and Jermann (2001) and Mehra and Prescott (1985). The average consumption growth rate is 1.83 percent and the standard deviation (SD) is 3.15 percent. The autocorrelation of consumption growth is –0.14. Expansions are more frequent than recessions: Seventy-three percent of realizations are states of high aggregate consumption growth. I calibrate the labor income process as in Storesletten, Telmer, and Yaron (2004, 2007), except that I eliminate the countercyclical variation of labor income risk. The variance of labor income risk is constant in this model. An invariant labor income risk setup highlights the role of the new feature (heterogeneous trading technologies) considered in this article. The main driving force of my result is the heterogeneity in trading technology, not the countercyclical variation of labor income risk. The Markov process for $\log\eta$ has an SD of 0.71, and the autocorrelation is 0.89. I use a four-state discretization for both aggregate and idiosyncratic risk. The elements of the process for $\log\eta$ are 0.38 and 1.61 for low and high shocks, respectively.

The equity in my model is simply a leveraged claim on diversifiable income. In the Financial Accounts of the United States (formerly the Flow of Funds Accounts tables), the ratio of corporate debt to net worth is around 0.65, suggesting a leverage parameter ψ of 2. However, Cecchetti, Lam, and Mark (1990) report that the SR of the growth rate of dividends is at least 3.6 times that of aggregate consumption, suggesting that the appropriate leverage level is over 3. Following Abel (1999) and Bansal and Yaron (2004), I choose to set the leverage parameter ψ to 3.

Composition of Trader Pools and Equity Target Share

I set the fraction of nonparticipants at 50 percent based on the fact that 51.1 percent of households reported owning stocks directly or indirectly in a recent Survey of Consumer Finances. To match the large equity premium (7.53 percent) measured in postwar U.S. data, a relatively small fraction of active traders needs to bear the large amount of residual aggregate risks created by nonparticipants. Hence, I set the share of active traders at 10 percent and the share of passive equity traders at 40 percent.

Among those households that hold equity, I am not able to distinguish between active traders and passive equity traders in the data. It is difficult to calibrate the target equity share of passive equity traders, since I do not know who they are. However, empirical studies have shown that rich households tend to be more sophisticated traders. Therefore, I consider the richest 10 percent of households to be active traders and the poorest 50 percent of households to be nonparticipants. The target equity share of passive equity traders is therefore calibrated to match the average fraction of equity among those households that possess a percentile of wealth between 50 percent and 90 percent (the middle wealthy). According to the data from the Survey of Consumer Finances, the average equity share among these middle wealthy households is 24.2 percent. I, therefore, set the equity target share of the passive equity traders at 24 percent. This calibration also reflects the observation that the rich tend to hold a higher fraction of equity than the poor.

QUANTITATIVE RESULTS

I consider two cases in the quantitative exercise. The first case is the benchmark economy, where the parameters are calibrated as described earlier. The second case I consider is another economy with no heterogeneity in trading technologies. All households are able to access all assets available on the market with no restrictions. Table 1 reports moments of asset prices in both of the economies considered. These results are generated by simulating data from a model with 12,000 agents for 10,000 periods. Panels A and B report results for the benchmark economy and the economy with no heterogeneity in trading technologies, respectively.

Asset Prices

The "Asset pricing" section of Table 1 shows the maximum unconditional SR, or market price of risk $\left(\frac{\sigma(m)}{E(m)}\right)$, the SD of the maximum SR $\left(Std\left(\frac{\sigma_t(m)}{E_t(m)}\right)\right)$, the equity risk premium $E\left(R_{t+1,t}^D - R_{t+1,t}^f\right)$, the SD of excess returns $\sigma\left(R_{t+1,t}^D - R_{t+1,t}^f\right)$, the SR on equity, the mean risk-free rate $E\left(R_{t+1,t}^f\right)$, and the SD of the risk-free rate $\sigma\left(R_{t+1,t}^f\right)$.

Benchmark Economy. In the benchmark economy, the maximum SR is 0.37 and the SD of the maximum SR is 4.04 percent. The equity premium is 7.54 percent and the SR on equity is 0.37. The average risk-free rate is 1.91 percent and its volatility is 2.27 percent. Clearly, the benchmark economy generates several key features of asset pricing observed in the data,

Table 1

Results of Benchmark and NHT Economy

	Panel A: Benchmark economy	Panel B: NHT economy
Active traders (%)	10	100
Passive equity traders (%)	40	0
Nonparticipants (%)	50	0
Asset pricing		
Market price of risk: $\dfrac{\sigma(m)}{E(m)}$	0.3739	0.1528
SD of market price of risk: $Std\left[\dfrac{\sigma_t(m)}{E_t(m)}\right]$ (%)	4.0440	1.0106
Equity risk premium: $E\left(R^D_{t+1,t}-R^f_{t+1,t}\right)$ (%)	7.5368	3.0077
SD of equity premium: $\sigma\left(R^D_{t+1,t}-R^f_{t+1,t}\right)$ (%)	20.3867	19.7216
SR	0.3697	0.1525
Risk-free rate: $E\left(R^f_{t+1,t}\right)$ (%)	1.9141	3.0900
SD of risk-free rate: $\sigma\left(R^f_{t+1,t}\right)$ (%)	2.2729	2.2539
Approximation		
R^2	>0.9995	>0.9999
Welfare cost		
Welfare cost of business cycle (%)	6.49	1.45

NOTE: NHT refers to an economy with no heterogeneity in trading technologies. Based on Storesletten, Telmer, and Yaron's (2007) calibration of idiosyncratic shocks without countercyclical variation risk and Alvarez and Jermann's (2001) calibration of aggregate consumption growth shocks. Parameters: $\alpha = 4$, $\beta = 0.95$, and the collateralized share of income is 10 percent. The results are generated by simulating an economy with 12,000 agents and 10,000 periods.

such as high equity premiums, a low and stable risk-free interest rate, and a relatively volatile SR.[1]

The large fraction of nonparticipant traders is critical for the results with high risk premiums. Those households that hold only risk-free assets do not take on any aggregate risk since their portfolio return is independent of the realization of aggregate shocks. Additionally, passive equity traders take on only a limited amount of aggregate risk because of their relatively low and constant target equity share. Therefore, a large amount of aggregate risk has to be absorbed by a small fraction of active traders. In equilibrium, a high risk premium is necessary so that active traders are willing to bear these extra aggregate risks. The key mechanism is to concentrate the aggregate risk in a small fraction of the population.

No Heterogeneous Trading Economy. In an economy where all households are active traders, the asset pricing results are dampened. Compared with the benchmark case, the maxi-

mum SR is only 0.15 and the SD of the maximum SR decreases to 1.01 percent. The equity premium decreases to 3.01 percent and the SR on equity is only 0.15. The average risk-free rate increases to 3.09 percent and its volatility remains roughly the same, 2.25 percent. The heterogeneity in trading technologies considerably affects the patterns of asset pricing results. The reason for the low equity premium is clear: The aggregate risk is equally borne by all households, and there is no concentration of risk in a small fraction of households as in the benchmark economy.

Approximation. In general, the prices of state-contingent claims depend on the entire aggregate history, which is intractable in computation. Following Chien, Cole, and Lustig (2011), I use truncated aggregate histories as state variables to forecast state-contingent prices. To show the accuracy of my approximation, I report the implied R-squared value from a linear regression of the actual realization of state-contingent prices on the predicted state-contingent prices, which are based on the truncated aggregate histories. This measure of precision is close to that of Krusell and Smith (1998). As shown in Table 1, the R-squared value for this regression is higher than 0.9995 in the benchmark case and higher than 0.9999 in the case without heterogeneous trading technologies. This result shows that the approximation is accurate and comparable to others reported in the literature for models with heterogeneous agents and incomplete markets.

Welfare Costs of Business Cycles

The welfare cost of eliminating business cycles is defined as the average welfare difference between two economies: one with aggregate shocks and the other without aggregate shocks. Given the fact that households are heterogeneous in terms of their wealth, income shocks, and trading technologies in the long-run equilibrium, the average welfare of one economy is computed by taking the expectation across all idiosyncratic features of the population. In addition, I measure the average welfare gap between the two economies by calculating the percentage of per-period consumption. Therefore, the welfare cost is defined as the expected percentage of consumption compensation for a household in an economy without business cycles, so that this household is indifferent to joining the benchmark economy. The welfare cost is reported at the bottom of Table 1.

Benchmark Economy. In the benchmark economy, the welfare cost of business cycles is 6.49 percent. This number means that the average household in the benchmark economy is willing to relinquish up to 6.49 percent of its per-period consumption to be in the other economy without aggregate uncertainty, all else being equal. The welfare cost is much larger than most of the findings in the body of literature. This result demonstrates that heterogeneous trading technologies play an important role not only in the patterns of asset pricing but also in the distributional effects of consumption. In the benchmark economy, the households with better trading technologies earn a higher return on their wealth, while the households with less sophisticated trading technologies earn a lower return. This phenomenon generates a distributional impact on consumption and eventually widens the welfare gap across households. However, the welfare inequality caused by heterogeneous trading technologies vanishes in an economy without business cycles. The reason for this is quite simple: Since all assets are risk

free, the portfolio choice between equity and risk-free bonds does not affect the return on the portfolios. There is no investment advantage for a household that has an advanced trading technology. The returns on wealth between active traders and passive traders are identical in an environment without aggregate risk.

As discussed earlier, the large welfare cost is mainly driven by the consumption inequality caused by heterogeneous trading technologies. A reasonable question is "To what extent does the consumption inequality in our model compare with that in the data?" This is especially important since the model is not calibrated to match the consumption dispersion observed in the data. Krueger and Perri (2006) reported a narrow variation of the Gini index in U.S. consumption inequality ranging from 0.23 to 0.26 between 1980 and 2003. The simulated data of the benchmark model generate a Gini index of 0.248, which is in the range of the reported data even though the consumption inequality is not targeted in calibration. This finding enhances the confidence of the welfare calculation.

No Heterogeneous Trading Economy. In my second exercise, where all households are active traders, the welfare cost is only 1.45 percent.[2] This low welfare cost is consistent with the findings in the literature on the cost of business cycles. This result suggests that the welfare cost of business cycles is less significant in an environment where all agents have sophisticated trading technologies and make no investment mistakes. This outcome can easily be understood: Because all households have the same trading technologies, there is no heterogeneity in portfolio choice. The income and consumption inequality are greatly reduced in this case. The aggregate risk no longer amplifies the distributional impact on consumption, so the welfare cost of business decreases considerably.

The amount of reduction in the welfare cost of business cycles can be seen as the average welfare gain from preventing the investment mistakes made by passive traders in my model. Clearly, the results show that the average welfare loss resulting from these investment errors is large: 5.04 percent of per-period consumption (the welfare cost difference between the benchmark economy and the economy with no heterogeneity in trading technologies). This number implies that the welfare cost of inferior trading technologies is sizable. My findings also shed light on the importance of understanding the investment mistakes made by passive traders, since avoiding them can improve the average welfare of the society.

Sensitivity Analysis

Risk Aversion Coefficient. The benchmark calibration sets the risk aversion coefficient to 4. Although my choice of risk aversion is in the range considered in many macroeconomics models, it is different from the choice made by Lucas (1987), who uses a log utility. More importantly, the welfare cost of business cycles might be sensitive to the risk aversion rate. Here, I investigate the sensitivity of the results to changes in the risk aversion coefficient by conducting two sensitivity analyses with respect to changes in the risk aversion rate. In each analysis, I vary the risk aversion coefficient from 3 to 1.

The first analysis considers only changes in the risk aversion rate while keeping all other parameters unchanged. Table 2 reports the results of my first analysis. The decrease in the risk aversion rate lowers the risk premium as well as the welfare cost. The risk premium drops

Table 2

Results of Sensitivity Analysis 1

	Panel A	Panel B	Panel C
Risk aversion rate (α)	3	2	1
Active traders (%)	10	10	10
Passive equity traders (%)	40	40	40
Nonparticipants (%)	50	50	50
Asset pricing			
Market price of risk: $\dfrac{\sigma(m)}{E(m)}$	0.2856	0.1868	0.0872
SD of market price of risk: $Std\left[\dfrac{\sigma_t(m)}{E_t(m)}\right]$ (%)	3.2615	2.1882	1.0307
Equity premium: $E\left(R^D_{t+1,t}-R^f_{t+1,t}\right)$ (%)	5.1849	3.0434	1.2799
SD of equity premium: $\sigma\left(R^D_{t+1,t}-R^f_{t+1,t}\right)$ (%)	18.3498	16.4103	14.6313
SR	0.2826	0.1855	0.0875
Risk-free rate: $E\left(R^f_{t+1,t}\right)$ (%)	2.8186	3.8275	4.8322
SD of risk-free rate: $\sigma\left(R^f_{t+1,t}\right)$ (%)	1.6744	1.0946	0.5360
Approximation			
R^2	0.9997	0.9997	0.9998
Welfare cost			
Welfare cost of business cycle (%)	5.27	4.22	0.6

NOTE: Based on Storesletten, Telmer, and Yaron's (2007) calibration of idiosyncratic shocks without countercyclical variation risk and Alvarez and Jermann's (2001) calibration of aggregate consumption growth shocks. Parameters: $\beta = 0.95$ and the collateralized share of income is 10 percent. The results are generated by simulating an economy with 12,000 agents and 10,000 periods.

substantially, from 5.18 percent with a risk aversion coefficient of 3 to 1.28 percent in the log utility case. In addition, the welfare cost of eliminating business cycles decreases in a nonlinear pattern. With a risk aversion rate of 3 or 2, the welfare costs are still very significant: 5.27 percent and 4.22 percent, respectively. However, the costs are sharply reduced—to 0.6 percent—when I consider the case of log utility. This analysis demonstrates a close relationship between the risk premium and the welfare cost of business cycles. This is not surprising, because the welfare cost of business cycles in my article depends critically on the magnitude of consumption dispersion, which is based on the return difference between equity and risk-free bonds. As the risk premium decreases, the heterogeneity in wealth returns is reduced along with the welfare cost.

The first analysis indicates that when households become less risk-averse, the model misses the calibration target, the equity premium, by a wide margin. Therefore, I conduct a second

Table 3

Results of Sensitivity Analysis 2

	Panel A	Panel B	Panel C
Risk aversion rate (α)	3	2	1
Active traders (%)	3	1	1
Passive equity traders (%)	47	49	49
Nonparticipants (%)	50	50	50
Asset pricing			
Market price of risk: $\dfrac{\sigma(m)}{E(m)}$	0.3957	0.3461	0.2238
SD of market price of risk: $Std\left[\dfrac{\sigma_t(m)}{E_t(m)}\right]$ (%)	7.5114	9.6565	9.5149
Equity premium: $E\left(R_{t+1,t}^{D}-R_{t+1,t}^{f}\right)$ (%)	7.3828	5.7938	3.0422
SD of equity premium: $\sigma\left(R_{t+1,t}^{D}-R_{t+1,t}^{f}\right)$ (%)	19.2098	17.7495	15.6695
SR	0.3843	0.3264	0.1941
Risk-free rate: $E\left(R_{t+1,t}^{f}\right)$ (%)	2.2503	3.0917	4.3834
SD of risk-free rate: $\sigma\left(R_{t+1,t}^{f}\right)$ (%)	1.6619	1.0823	0.5280
Approximation			
R^2	0.9995	0.9995	0.9997
Welfare cost			
Welfare cost of business cycle (%)	9.37	9.56	3.81

NOTE: Based on Storesletten, Telmer, and Yaron's (2007) calibration of idiosyncratic shocks without countercyclical variation risk and Alvarez and Jermann's (2001) calibration of aggregate consumption growth shocks. Parameters: $\beta = 0.95$ and the collateralized share of income is 10 percent. The results are generated by simulating an economy with 12,000 agents and 10,000 periods.

sensitivity analysis. For each risk aversion rate considered earlier, I adjust the composition between active traders and passive equity traders to match the historical risk premium as much as possible, while keeping all other parameters fixed. The results of the second analysis are shown in Table 3.

Panel A of Table 3 reports the results when the risk aversion coefficient is 3. To match the high historical risk premium, the fractions of active traders and passive equity traders are adjusted to be 3 percent and 47 percent, respectively. The asset pricing results are similar to those in my benchmark economy. The risk premium is high (7.38 percent) and volatile (SD of 19.21 percent), while the risk-free rate is low (2.25 percent) and stable (SD of 1.66 percent). Most importantly, the welfare cost of business cycles increases to 9.37 percent. The higher welfare cost result can be understood as follows: First, active traders are those who respond to the change in state-contingent prices and bear extra aggregate risk. Put simply, they are mar-

ginal traders who price the risk premium. Second, if these active traders still bear the same amount of aggregate risk as in the benchmark case, then the risk premium will drop since their risk aversion rate is lower now. To maintain the same high risk premium while having a lower risk aversion rate, a larger amount of aggregate risk has to be concentrated and borne by a smaller fraction of active traders. As the fraction of active traders is adjusted from 10 percent to 3 percent, each active trader bears more aggregate risk but is able to enjoy an even higher level of consumption in terms of compensation. The reduction in the fraction of active traders worsens the consumption inequality and, consequently, increases the welfare cost of business cycles.

Panels B and C of Table 3 report the results of the case with $\alpha = 2$ and 1, respectively. In both cases, I am unable to match the high risk premium shown in the data even when the fraction of active traders is set to be only 1 percent of the total population. The risk premiums of both cases are significantly smaller: 5.79 percent with $\alpha = 2$ and only 3.04 percent with log utility. Nevertheless, the welfare cost of business cycles is even higher, 9.56 percent, when the risk aversion coefficient is 2. The reason for this is simply that there is a higher inequality in consumption. Although the lower risk premium reduces the inequality of consumption by decreasing the heterogeneity of wealth returns across the population, the smaller fraction of active traders amplifies the consumption inequality even more. The second effect on consumption inequality caused by the diminishing size of active traders dominates the first effect resulting from the lower risk premium. Consequently, the welfare cost increases slightly. The last panel reports the results for the log utility case. The welfare cost drops substantially from 9.56 percent to 3.81 percent when the risk aversion coefficient changes from 2 to 1. This result is not surprising given that the composition of traders is the same in both Panels B and C.

The second sensitivity analysis demonstrates that the welfare cost of business cycles is even larger with a lower risk aversion coefficient whenever the historical, high risk premium can be matched in my calibration economy. Additionally, for the log utility case, the welfare cost of business cycles is still significant even if my calibration fails to match the risk premium. The welfare cost is 3.81 percent when active traders comprise 1 percent of the total population.

CONCLUSION

This article demonstrates that heterogeneous trading technologies can play an important role not only in the patterns of asset pricing but also in the welfare cost of business cycles. In my calibrated model, a large amount of aggregate risk is borne by a small fraction of households, while a large fraction of households bear little or even no aggregate risk. The concentration of risk in a limited set of households drives the large risk premium in my model. As a result, sophisticated investors who hold a large fraction of equity in their portfolios are compensated with a much higher return on wealth, while less sophisticated investors earn a lower return on their wealth. A larger wealth return difference worsens the income and consumption inequality. In addition, the new feature of my model—heterogeneous trading technologies—has no distributional effect on consumption in an economy without aggregate shocks because the return difference between stocks and bonds vanishes. Cessation of aggregate shocks can

greatly improve the consumption inequality caused by the heterogeneity in investment behavior. Therefore, the welfare cost of business cycles is more pronounced in an economy with heterogeneous trading technologies.

For economies with homogeneous trading technologies, the results show an insignificant welfare cost of business cycles. This result implies a large welfare difference between economies with and those without heterogeneous trading technologies, which can be thought of as the welfare cost of investment mistakes made by passive traders. These mistakes include relinquishing high risk premiums and missing the market timing. The significant welfare cost of investment errors highlights the importance of the study of household finance. If a way can be found to avoid these investment mistakes, the average welfare of the society can be improved considerably. Additionally, the results indicate that the welfare improvement from avoiding these investment errors is comparable to that of eliminating business cycles. Therefore, if the elimination of aggregate output volatility is infeasible or extremely expensive, then concentrating more resources on preventing household investment mistakes may be a reasonable alternative. ∎

NOTES

[1] The SR estimated from the data is enormous and highly countercyclical. My model still falls short of matching the data quantitatively. However, Chien, Cole, and Lustig (2012) extend a similar version of this model by introducing inertia investment behavior among some of the households. Their work shows that the inertia investment behavior helps significantly to explain the large countercyclical variation in the SR.

[2] This welfare cost number is significantly larger than those in the standard complete market literature for two reasons. First, the endowment growth shock is assumed to be permanent and hence has an infinite variance, which gives the largest uncertainty for future aggregate consumption. Second, the risk aversion parameter is higher than those in the standard literature.

REFERENCES

Abel, Andrew B. "Risk Premia and Term Premia in General Equilibrium." *Journal of Monetary Economics*, February 1999, *43*(1), pp. 3-33.

Alvarez, Fernando and Jermann, Urban J. "Quantitative Asset Pricing Implications of Endogenous Solvency Constraints." *Review of Financial Studies*, Winter 2001, *14*(4), pp. 1117-51.

Alvarez, Fernando and Jermann, Urban J. "Using Asset Prices to Measure the Cost of Business Cycles." *Journal of Political Economy*, December 2004, *112*(6), pp. 1223-56.

Bansal, Ravi and Yaron, Amir. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, August 2004, *59*(4), pp. 1481-509.

Campbell, John Y. "Household Finance." *Journal of Finance*, August 2006, *61*(4), pp. 1553-604.

Cecchetti, Stephen G.; Lam, Pok-sang and Mark, Nelson C. "Mean Reversion in Equilibrium Asset Prices." *American Economic Review*, June 1990, *80*(3), pp. 398-418.

Chien, YiLi; Cole, Harold and Lustig, Hanno. "A Multiplier Approach to Understanding the Macro Implications of Household Finance." *Review of Economic Studies*, 2011, *78*(1), pp. 199-234.

Chien, YiLi; Cole, Harold and Lustig, Hanno. "Is the Volatility of the Market Price of Risk Due to Intermittent Portfolio Rebalancing?" *American Economic Review*, October 2012, *102*(6), pp. 2859-96.

Krebs, Tom. "Job Displacement Risk and the Cost of Business Cycles." *American Economic Review*, June 2007, *97*(3), pp. 664-86.

Krueger, Dirk and Perri, Fabrizio. "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory." *Review of Economic Studies*, 2006, *73*(1), pp. 163-93.

Krusell, Per; Mukoyama, Toshihiko; Şahin, Ayşegül and Smith, Anthony A. Jr. "Revisiting the Welfare Effects of Eliminating Business Cycles." *Review of Economic Dynamics*, July 2009, *12*(3), pp. 393-402.

Krusell, Per and Smith, Anthony A. Jr. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, October 1998, *106*(5), pp. 867-96.

Krusell, Per and Smith, Anthony A. Jr. "On the Welfare Effects of Eliminating Business Cycles." *Review of Economic Dynamics*, January 1999, *2*(1), pp. 245-72.

Lucas, Robert. *Models of Business Cycles*. New York: Blackwell, 1987.

Lucas, Robert. "Macroeconomic Priorities." *American Economic Review*, March 2003, *93*(1), pp. 1-14.

Mehra, Rajnish and Prescott, Edward C. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, March 1985, *15*(2), pp. 145-61.

Mendoza, Enrique G.; Quadrini, Vincenzo and Ríos-Rull, José-Víctor. "Financial Integration, Financial Development, and Global Imbalances." *Journal of Political Economy*, June 2009, *117*(3), pp. 371-416.

Parker, Jonathan A. and Vissing-Jorgensen, Annette. "Who Bears Aggregate Fluctuations and How?" NBER Working Paper No. 14665, National Bureau of Economic Research, January 2009; http://www.nber.org/papers/w14665.pdf.

Storesletten, Kjetil; Telmer, Chris I. and Yaron, Amir. "The Welfare Cost of Business Cycles Revisited: Finite Life and Cyclical Variation in Idiosyncratic Risk." *European Economic Review*, June 2001, *45*(7), pp. 1311-39.

Storesletten, Kjetil; Telmer, Chris I. and Yaron, Amir. "Cyclical Dynamics of Idiosyncratic Labor Market Risk." *Journal of Political Economy*, June 2004, *112*(3), pp. 695-717.

Storesletten, Kjetil; Telmer, Chris I. and Yaron, Amir. "Asset Pricing with Idiosyncratic Risk and Overlapping Generations." *Review of Economic Dynamics*, October 2007, *10*(4), pp. 519-48.