Preliminary and incomplete

How Persistent are Monetary Policy Effects at the Zero Lower Bound?

Christopher J. Neely*

This version: January 15, 2014

Abstract

Event studies show that Fed unconventional announcements of forward guidance and large scale asset purchases had large and desired effects on asset prices but such studies cannot directly answer the important question of how long such effects last. Wright (2012) used a structural vector autoregression (SVAR) to argue that unconventional policies have very transient effects on asset prices, wearing off within 2 to 3 months, which would suggest that unconventional policies can have only marginal effects on macroeconomic variables. The present paper shows, however, that the SVAR is unstable, forecasts very poorly and therefore delivers spurious inference about the duration of the unconventional monetary shocks. The predictability of the SVAR greatly exceeds that implied by rational asset pricing models for reasonable levels of risk aversion, which makes its findings implausible for that reason alone. Restricted models are more stable and imply that the unconventional monetary policy shocks were fairly persistent but that our uncertainty about their effects does not decline with forecast horizon. Estimates of the dynamic effects of shocks should respect the limited predictability in asset prices.

JEL classification:E430, E470, E520, C300

Keywords: Federal Reserve, monetary policy, quantitative easing, large-scale asset purchase, VAR, forecasting, structural breaks, good deal

---

*Corresponding author. Send correspondence to Chris Neely, Box 442, Federal Reserve Bank of St. Louis, St. Louis, MO 63166-0442; e-mail: neely@stls.frb.org; phone: +1-314-444-8568; fax: +1-314-444-8731. Christopher J. Neely is an assistant vice president and economist at the Federal Reserve Bank of St. Louis. The views expressed in this paper are those of the author and do not necessarily reflect those of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks.

The financial market turmoil—soaring risk premia, illiquidity—that followed Lehman Brothers' September 2008 bankruptcy prompted extraordinary measures from monetary authorities. The Federal Reserve created special facilities to support lending, expanded swap lines with foreign central banks and reduced the Federal funds rate to essentially zero by December 16, 2008. These measures, however, were insufficient to stem the economic slide and the Federal Reserve soon pursued outright asset purchases to support the economy, especially housing markets. These quantitative easing (QE) purchases occurred in several phases: The Federal Open Market Committee (FOMC) announced QE1 in November 2008 and March 2009, QE2 on November 3, 2010, and a Maturity Extension Program ("Operation Twist") on September 21, 2011. Finally, on September 13, 2012, the FOMC announced a third round of quantitative easing, QE3.[1] Collectively, these programs committed the Fed to trillions of dollars in long-term asset purchases, testing the ability of central banks to stimulate the economy when short-term interest rates are at the zero lower bound.

There has already been substantial research on U.S. and international QE programs, which has mostly focused on event studies of the effect of asset prices. Stroebel and Taylor (2009), Kohn (2009), Meyer and Bomfim (2010), and Gagnon et al. (2011a,b), for example, study the Fed's 2008-09 QE programs. Gagnon et al.'s (2011a,b) announcement study finds that large-scale asset purchase (LSAP) announcements reduced U.S. long-term yields. Joyce et al. (2011) find that the BOE's QE program had bond yield effects quantitatively similar to those reported by Gagnon et al. (2011a,b) for the U.S. program. Hamilton and Wu (2012) indirectly calculate the effects of the Fed's 2008-09 QE programs with a term structure model. Neely (2013)

[1] At the July 31–August 1 FOMC meeting, "many members judged that additional monetary accommodation would likely be warranted fairly soon unless incoming information pointed to a substantial and sustainable strengthening in the pace of the economic recovery"; www.federalreserve.gov/monetarypolicy/files/fomcminutes20120801.pdf.

evaluates the effect of the Fed's 2008-09 QE on international long bond yields and exchange rates, showing that the effects are consistent with a simple portfolio balance model and long-run purchasing power parity. Bauer and Neely (2013) study the relative importance of the signaling and portfolio balance channels on international interest rates with term structure models.

Despite this profusion of research on the immediate effects of asset purchase programs on asset prices, there has been much less work and less certainty about the effect of QE on macroeconomic variables (Baumeister and Benati (2010), Gambacorta, Hofmann and Peersman (2012), and Gertler and Karadi (2013)). Modeling the effect of QE depends on the persistence of the effects of QE and it is inherently difficult to determine the persistence of the effect of shocks to asset prices. The persistence of the estimated effects of QE announcements is very important because transient QE shocks imply that QE is a much less effective policy and central banks must look for other policies or ways to make the effects more permanent.

Indeed, many market observers concluded that QE 1 failed because long yields did not remain low after the March 18, 2009 QE1 buy announcement. Figure 1 shows that from late April through June, long-term nominal Treasury yields rose fairly steadily, gaining more than 100 basis points overall by mid-June. Long-term, sovereign, local currency, nominal yields from Australia, Canada, Germany, and the U.K. similarly rose during this March-to-June period.

Wright (2012) offers a clever solution to determine the persistence of unconventional shocks to interest rates in the form of a structural vector-autoregression (SVAR), in which the contemporaneous effects of unconventional monetary shocks are identified by the heteroskedasticity in interest rates on days of unconventional monetary policy shocks (Rigobon and Sack (2004) and Craine and Martin (2008)). The VAR contains1-lag of 6 daily U.S. interest rates: the 2- and 10-year nominal Treasury yields, the 5-year and 5-10 year forward TIPS breakeven rates, and the Moody's Baa and Aaa corporate bond yield indices, using data from November 3, 2008 to September 30, 2011. Wright concludes that unconventional monetary policy shocks have large but very transient effects on U.S. interest rates. Taken at face value, the impulse responses imply that nearly all of the LSAP effects on 10-year Treasury yields dissipate within 2 or 3 months.

There are reasons, however, to be skeptical of this finding. The methods assume that a simple VAR accurately describes the stable dynamics (i.e., constant coefficients) of the system but asset prices are notoriously difficult to predict, out-of-sample (OOS), and in-sample fit is frequently misleading. Meese and Rogoff (1983) made this point forcefully in the context of structural models of the exchange rate. They showed that good in-sample performance of such models was accompanied by very poor out-of-sample performance. Neely and Weller (2000) show that inferring long-run predictability from VARs with asset prices is not very reliable (Bekaert and Hodrick (1992)). Faust, Rogers and Wright (2003) convincingly argue that Mark's (1995) foreign exchange forecasting model is fragile to data vintage. Goyal and Welch (2008) question the usefulness of many traditional equity premium predictors using both in-sample and OOS forecasting exercises.

The contribution of this paper is to carefully analyze the stability of the data and show that Wright's conclusion on the persistence of unconventional monetary policy shocks is unsupported by the data.[2] Specifically, the present paper replicates Wright's VAR and shows that it forecasts well, in sample, but forecasts very poorly out-of-sample and fails structural stability tests.

[2] This paper does not examine all of Wrights conclusions. In a second part of his paper, Wright also constructs a set of unconventional monetary shocks from high frequency data. This paper does not critique that methodology.

Because the coefficients of the VAR determine the persistence of shocks, any conclusions about persistence are only as reliable as inference about those reduced form coefficients. Because naïve no-change forecasts out-predict the real-time, unrestricted VAR coefficients and one can reject the null of stability for those VAR coefficients, the unrestricted VAR cannot inform us about the persistence of the unconventional shocks.

In addition, this paper argues that the estimated degree of transience of the policy effects is inconsistent with standard thinking about risk-aversion and efficient markets. That is, the transient effects estimated by Wright would create an opportunity for risk-adjusted expected returns that greatly exceed values that are consistent with plausible risk aversion. Restricted VAR models that are consistent with rational asset pricing, however, forecast better than unrestricted VAR models and imply a more plausible structure: monetary policy shocks are probably fairly persistent but we cannot tell exactly how persistent because our uncertainty about them does not shrink with the forecast horizon. Thus, our best guess is that unconventional monetary policy shocks have fairly persistent effects on bond yields.

The next section of the paper describes Wright's methodology. Section 3 describes and replicates Wrights main findings with the SVAR. Section 4 illustrates that the VAR fails to forecast well, OOS, and Section 5 draws conclusions.

2. The Structural VAR Methodology

The reduced form VAR can be written as follows:

$$A(L)y_t = \varepsilon_t \tag{1}$$

where $A(L)$ is a polynomial in the lag operator and ε_t denotes the reduced form errors. OLS regressions with lagged endogenous regressors will generally produce biased coefficients in finite samples. That is, the estimated processes will generally underestimate the persistence of persistent variables. Wright follows Kilian (1998) in correcting this bias with a bootstrapping procedure.[3] The bias-correction appendix to this paper details this method.[4]

The covariance of the reduced form errors, ε_t, is assumed to vary from monetary policy announcement days to non-announcement days. The reduced form errors are related to the structural errors as follows: $\varepsilon_t = \sum_{i=1}^{6} R_i u_{t,i}$, where R_i is a 6 by 1 vector of the initial impacts of the i^{th}structural shock on each of the endogenous variables.

Without loss of generality, the monetary policy shock is assumed to be the first structural shock. The impact of monetary policy shocks are identified under the assumption that the variance of the monetary policy shock, $u_{t,1}$, is different on 28 specific monetary announcement days (o_1^2) than on non-announcement days (o_0^2), in the sample. Under this assumption, the difference in the reduced form residual covariance matrix on announcement and non-announcement days is as follows:

$$\Sigma_1 - \Sigma_0 = R_1 R_1{}'(o_1^2 - o_0^2). \tag{2}$$

Because the terms in the product $R_1 R_1{}'(o_1^2 - o_0^2)$ are not separately identified, Wright normalizes $(o_1^2 - o_0^2)$ to 1 and solves for the estimated elements of R_i by choosing terms to minimize the difference, $\left[\hat{R}_1\hat{R}_1' - \left(\hat{\Sigma}_1 - \hat{\Sigma}_0\right)\right]$, using the covariance matrix of $\left(\hat{\Sigma}_1 - \hat{\Sigma}_0\right)$ to appropriately weight the moments.

---

[3] Mankiw and Shapiro (1986) and Stambaugh (1986) discuss the small sample bias imparted by lagged endogenous regressors. Bekaert, Hodrick and Marshall (1997) use Monte Carlo procedures to correct for such a bias in term structure tests.

[4] Although the present paper follows the bias-correction procedure in the construction of impulse responses and forecasts, it does not seem to make a real difference to any inference.

As Wright is only interested in the impact of unconventional monetary policy shocks, there is no need for additional identifying assumptions. The estimated coefficients of $A(L) = I - A_1 L$ and the reduced form errors come from the VAR estimation, and the elements of $\hat{R}_i$ come from minimizing $\hat{R}_i \hat{R}_i' - (\hat{\Sigma}_1 - \hat{\Sigma}_0)$. These are all one needs to construct impulse response functions and forecast error decompositions for the unconventional monetary policy shocks.

Wrights' (2012) identification scheme does not need to make an assumption about the Federal Reserve's monetary policy instrument or the nature of its reaction function, except to the extent that it assumes that monetary shocks have higher variance on announcement days. It does not identify the system through assumptions about contemporaneous interactions between the variables, which was pioneered by Bernanke (1986), Blanchard and Watson (1986) and Sims (1986). Neither does it use a long-run identification scheme, as introduced by Shapiro and Watson (1988) and Blanchard and Quah (1989). Instead, it follows the spirit of the identification-through-heteroskedasticity procedures laid out by Rigobon and Sack (2004) for their study of the effect of monetary policy shocks on asset prices.

Wright block bootstraps the VAR system to test two hypotheses: 1) that the covariance matrices are the same on announcement and non-announcement days (i.e., $\Sigma_1 = \Sigma_0$), and 2) that there is a single monetary policy shock (i.e., $R_i R_i' = (\Sigma_1 - \Sigma_0)$), as well as to put confidence intervals on the impulse response functions. The bootstrapping tests, which are implicitly conducted under the assumption of a stable VAR system, reject the null that $\Sigma_1 = \Sigma_0$ but fail to reject that there is a single monetary policy shock.

3. Data and replication of SVAR results

Wright (2012) estimates a 1-lag VAR, using the bias-adjusted bootstrap of Kilian (1998), with 6 daily U.S. interest rates: the 2- and 10-year nominal Treasury yield, the 5-year and 5-10 year forward inflation compensation yields, and the Moody's Baa and Aaa corporate bond yield indices, using daily data from November 3, 2008 to September 30, 2011. This paper replicates Wright' VAR results with similar data, estimation procedures and identification scheme.

Wright (2012) identifies the structural response of the system to monetary policy shocks with the assumption that monetary policy shocks have a larger variance on monetary announcement days than on non-announcement days. The 28 announcement days included dates of any FOMC meeting day while short rates were at the zero bound and all other announcements or speeches by the Chairman that were relevant to unconventional monetary policy.[5]

To calculate the impulse responses, Wright normalizes the monetary shock to have an immediate -25 b.p. effect on 10-year yields. Figure 2 illustrates the resulting impulse response patterns, including 90 percent bootstrapped confidence intervals, which are very similar to those found in Wright's paper.[6] As in Wright's analysis, the monetary policy shocks have significant immediate effects: The shock to the 10-year Treasury rate is very significant; the Baa and Aaa yield changes are also significant, being approximately 60 to 80 percent of the Treasury changes.[7] This immediate effect is consistent with the results in a number of event studies of

[5] The monetary announcement dates were as follows: 11/25/2008, 12/1/2008, 12/16/2008, 1/28/2009, 3/18/2009, 4/29/2009, 6/24/2009, 8/12/2009, 9/23/2009, 11/4/2009, 12/16/2009, 1/27/2010, 3/16/2010, 4/28/2010, 6/23/2010, 8/10/2010, 8/27/2010, 9/21/2010, 10/15/2010, 11/3/2010, 12/14/2010, 1/26/2011, 3/15/2011, 4/27/2011, 6/2/2011, 8/9/2011, 8/26/2011, and 9/21/2011.

[6] The bootstrapping procedures used in the results in this paper respect the heteroskedasticity assumption of Wright by bootstrapping separately for announcement and non-announcement days. Experimentation with homoskedastic or block bootstrapping did not seem to make any difference to the results.

[7] The Aaa yield changes are significant with a slight delay.

unconventional Federal Reserve monetary policy, e.g., Gagnon et al. (2011a,b) and Neely (2013).

The significant initial effects, however, are not nearly as remarkable as the fact that the initial effects of the monetary policy shock wear off very quickly. The half-lives of the estimated responses of the 10-year Treasury and corporate yields are only about two or three months. This is a potentially very important result as it suggests that unconventional policy actions should have only very transient effects on yields and therefore very modest effects on macroeconomic variables. Wright (2012) sums this up as follows: "*To the extent that longer term interest rates are important for aggregate demand, unconventional monetary policy at the zero bound has had a stimulative effect on the economy but it might have been quite modest.*" — page F465

4. Analysis of the VAR's Structural Stability

The estimated impulse responses in Figure 2 are functions of the reduced form VAR coefficients and the estimated initial impact of the structural VAR shocks. Specifically, the impacts of the shocks can be written with the moving average representation as

$$y_t = A(L)^{-1}\varepsilon_t = (I - A_1 L)^{-1} R u_t \tag{3}$$

where A_1 is the matrix of reduced form VAR coefficients and R is the 6 by 6 matrix relating the structural error vector, u_t, to the reduced form error vector, ε_t, whose 1st column is R_1.

As discussed previously, a potentially serious difficulty in the inference from the impulse responses concerns the fact that VARs—and other time series relations —are notoriously unstable. Empirical models have failed to forecast asset prices in out of sample exercises with a variety of assets, including exchanges rates (Meese and Rogoff (1983), Faust, Rogers and Wright

(2003)), equities (Goyal and Welch (2008)), interest rates (Thornton and Valente (2012)) and cross-asset studies (Neely and Weller (2000)).

In the present context, the estimated VARs must be stable or the impulse responses in Figure 2 are spurious and the potentially very important inference drawn from them—that unconventional policy has very transient effects— is unsupported by the data. Although VARs are not designed for forecasting, they must describe a stable intertemporal relation between the endogenous variables in order to accurately describe the responses of variables to shocks.

Although one can formally test for structural stability with econometric tests, such as those suggested by Andrews (1993) or Bai and Perron (1998, 2003), OOS forecasting exercises also provide an informal and intuitively attractive test of structural stability. Therefore, to shed light on the structural stability of the VAR system, this paper first considers whether the VAR forecasts outperform a suitably chosen simple benchmark model in-sample and OOS forecasts. As interest rates are highly persistent, a no-change model is a natural benchmark against which to test the VAR forecasts.

Table 1 shows the in-sample root mean squared forecast errors (RMSFE) in basis points for a naïve, no-change model for the interest rates and the bias-adjusted VAR, respectively. The bottom panel of Table 1 shows the ratio of those RMSFEs, Theil U statistics. By construction, the in-sample VAR forecasts should have lower 1-day ahead RMSFEs and they do, although only barely so, with the 1-day ahead Theil U statistics nearly equal to 1. The relative performance of the VAR improves at the 20-day horizon with the Theil statistics declining to the 0.82 – 0.96 range before rising again at longer horizons for most equations.

The bottommost panel of Table 1 shows the proportion of the bootstrapped Theil statistics that exceed the real Theil statistics. These statistics are unremarkable. The statistics are generally well within the distribution, indicating that the in-sample performance of the bias-adjusted VAR is about what one would expect.

Table 2 shows the mean errors and Newey-West t-statistics for the null that the errors have a mean of zero (Newey and West (1994)). Note that the bias-adjusted VAR coefficients do not imply zero forecast errors at the 1-day horizon, as the unadjusted coefficients would. The t-statistics are rarely able to reject the hypothesis that the in-sample forecasts are unbiased for either the naïve or bias-adjusted VARs. Although this unbiasedness should not be surprising for in-sample VAR forecasts, it is a bit more surprising that one cannot reject unbiasedness for the naïve forecasts in this sample.

Even excellent in-sample performance does not necessarily mean that the model will predict OOS returns better than some simple benchmark model. A well-specified VAR with a stable covariance structure should enable us to use the parameter estimates from the period 2008-2011 to forecast asset prices during the out-of-sample period 2011-2013. This is implicitly a test of the structural stability of the VAR structure.

To investigate the out-of-sample forecast performance of the VARs, we carry out the following exercise: We use the coefficient estimates from the sample period 2008:11:03-2011:09:30 to construct forecasts for each of the variables in the system over the out-of-sample period (2011:10:01-2013:11:27), at horizons of 1-, 20-, 60- and 120-days. That is, at each date in the OOS period, we use the actual data at date t and the parameters as estimated over the fixed sample period and project the path of the system at dates $t + 1$ through $t + 120$. We then update the data for the next period's set of forecasts. This gives us a set of 538 one-period-ahead forecasts, 519 overlapping 20-period-ahead forecasts, 479 overlapping 60-period forecasts and 419 overlapping 120-period forecasts.[8]

The OOS forecast performance is poor. The OOS Theil statistics in Table 3 show that a naïve, no-change prediction is superior to the VAR forecasts for 20 of 24 horizon-yield combinations considered. The only cases for which the VAR is competitive with the no-change forecast are for changes in inflation compensation. Even for those inflation-compensation variables, one could not claim that the VAR outperforms the naïve forecast in an economically or statistically significant sense. More importantly, for each of the yield variables, the naïve forecast outperforms the VAR at every horizon. The bootstrapping p-values show that the superiority of the naïve forecast is statistically significant at the 10 percent level at every horizon for each of the yield variables.

Table 4 shows the OOS bias (mean-error) statistics for the naïve and bias-adjusted VAR models. The naïve predictions are not systematically biased in a statistically significant way but the VAR predictions are, especially at longer horizons. All of the VAR forecasts of the yield variables are statistically significantly biased at all horizons.

One might think that modifications to the VAR procedure would improve the forecasting performance and rescue the possibility of constructing informative impulse responses for monetary shocks. There is considerable evidence that combining Bayesian techniques with VARs is helpful in forecasting (Litterman (1986)). Wright (2012), however, already considered such techniques in his robustness checks and found impulse responses that are similar to those in

[8] The overlapping n -period forecast errors will have at least n-1 order serial correlation in their errors which must be taken into account in the tests.

Figure 2, which suggests that the VARs that produced them are also unstable. Of course, one could tighten up the priors on the Bayesian VAR to essentially reproduce the naïve forecasts, but then one would not obtain the mean reverting impulse responses that are the point of Wright's SVAR study. Instead, one would presumably obtain very persistent impulse responses that would be implied by the naïve model.

OOS forecasting exercises are intuitively attractive indicators of lack of structural stability but are not formal econometric tests. This problem, a form of model misspecification, is common in time series regressions. To quantify the extent of the problem we test for a structural break at an unknown date by calculating the standard Wald test statistics for a structural break at each observation in the middle third of each sample. The supremum of these test statistics identifies a possible structural break in the series but will have a nonstandard distribution (Andrews (1993)).[9] The critical values for the supremum are calculated from a Monte Carlo experiment.[10]

Consistent with this poor OOS forecasting performance, Figure 3 plots the Andrews (1993) unknown-point structural break statistics for the null that the VAR parameters are stable over time, along with 1, 5, and 10 percent critical values. The structural break statistics are clearly above the critical values found by Monte Carlo simulation, indicating that one can clearly reject the null of stable VAR parameters. Because the VAR is clearly unstable, the parameters used to construct the impulse response functions are spurious and the conclusions from the impulse responses cannot be trusted.

[9] Bai and Perron (1998, 2003) and Ghysels, Guay and Hall (1998) propose other approaches to examining structural instability with an unknown break point that are less computationally demanding than the Andrews procedures. The computational advantage of this procedure is less helpful in the VAR environment.

[10] The critical values of the Andrews test are potentially dependent on the null model, the error distribution and the estimated coefficients. Neither bootstrapping the errors nor perturbing the estimated coefficients substantially changed the estimated critical values or the inference drawn.

To determine the prevalence of breaks in the six individual VAR equations during the sample period, one can conduct Bai-Perron (1998, 2003) tests for an unknown number of breaks at unknown points. Bai and Perron (2003) recommend that one first test for the presence of any breaks with the *UD max* or *WD max* tests and then evaluate the number of breaks by sequentially testing down from the maximum number of breaks with the SupF tests.[11]

This paper follows those Bai-Perron guidelines under the assumptions of a maximum of 3 breaks with at least 20 percent of the original sample between each break. The first and second rows of Table 5 shows that the *UD max* and *WD max* statistics reject the null of no breaks for nearly all equations at conventional significance levels. The third and fourth rows of that table likewise illustrate that although one cannot reject the null of 2 breaks in favor 3 breaks, one can reject the null of 2 breaks in favor of 1 break for 4 of the 6 equations at the 10 percent level. The Bai-Perron tests indicate that there are certainly breaks and probably multiple breaks in the data.

One should note that the Andrews (1993) and Bai and Perron (1998, 2003) structural stability tests do not require one to specify the date or precise nature of the structural break and so will generally have much less power under the alternative than tests that do so. The fact that they conclusively reject stability is strong evidence against that null hypothesis.

In summary, the VAR that produced Figure 2—evidence for the transient effects of unconventional monetary policy shocks —badly fails fundamental tests of stability: It forecasts very poorly out of sample and one can easily reject structural stability both for the whole VAR and for the individual equations, despite the fact that the stability tests failed to specify the dates or numbers of the breaks and so probably lack power to reject modest deviations from the null.

[11] Denoting the maximum number of breaks permitted by M, the *UD max* statistic tests for a break by considering whether the maximum of all M F-statistics exceeds its critical value, while the *WD max* statistic also tests for breaks with a weighted average of the F-statistics in which the marginal p-values are equalized across statistics. See Bai and Perron (1998, 2003).

These conclusions hold for both the bias-adjusted and the unadjusted VAR coefficients. The data do not support the inference from Figure 2 that monetary shocks are transient. Instead, very persistent shocks appear to be a better approximation to the dynamic structure.

5. Is the Estimated Predictability Consistent with Rational Pricing?

Intuitively, it seems reasonable that arbitrage would bound the amount of predictability in asset returns and that this bound would depend on the degree of risk aversion. That is, if the marginal investor were entirely risk neutral, any predictability would presumably be inconsistent with equilibrium. In the absence of mean reversion, uncertainty about asset prices rises with the forecast horizon, and therefore no one can know the LSAP's long-term effects. The efficient markets hypothesis implies that the market's best guess must have been that the LSAP effects would persist. Otherwise, expectations of a temporary impact would have created a risk-arbitrage opportunity for investors to bet on the reversal of the LSAP effects.

Kirby (1998) formalized this relation to show that the R^2 from a predictive regression must be less than the square of the coefficient of risk aversion multiplied by the variance of the market return:

$$R^2 \leq \left(1 + R_f\right)^2 RRA_V^2 \sigma^2\left(r_{m,t+1}\right) \cong RRA_V^2 \sigma^2\left(r_{m,t+1}\right) \tag{4}$$

where R_f is the riskless rate, RRA_V is the upper bound on relative risk aversion and $\sigma^2\left(r_{m,t+1}\right)$ is the variance of the market excess-return, $r_{m,t+1}$. Levich and Poti (2008) thoroughly and accessibly describe this relation in their application to currency markets.

Are the relations that Wright estimates consistent with the bounds that Kirby derives? Wright's VAR uses yields and inflation compensation, of course, so the Kirby bounds don't directly apply. One can, however, convert the yield and compensation data to returns and changes in inflation compensation, respectively, which enables one to derive bounds on a very similar VAR, one using log gross yields rather than net yields. That is, net yields are approximately the same as log gross yields, which are log prices— $m \times y \cong m \times ln(1+y) = ln(p)$— and returns are differences in log prices.

Of course, the inflation compensation spreads are not exactly asset prices and can take negative values in the sample. Therefore, this paper applies the gross log transform to the Treasury and corporate bond returns but not the inflation compensation variables. Denote the transformed vector as $\tilde{y}_t$, where $\tilde{y}_{i,t} = \ln(1+y_{i,t})$ for i = 1, 2, 5 and 6 and $\tilde{y}_{i,t} = y_{i,t}$ for i = 3 and 4.

One can then regress the constructed returns on all 6 lagged log gross yields, approximating the original VAR, except for the substitution of log gross yields for yields. That is, if one writes the original reduced form VAR as

$$y_t = Ay_{t-1} + c + \varepsilon_t, \tag{1'}$$

then one would write the VAR in gross logs as

$$\tilde{y}_t = \tilde{A}\,\tilde{y}_{t-1} + \tilde{c} + \varepsilon_t, \tag{5}$$

where $\tilde{A}$ and $\tilde{c}$ will be similar to A and c to the extent that the transformation is linear. If one then subtracts $\tilde{y}_{t-1}$ from both sides of (5), one then has a system relating (negative) log returns to lagged log gross yields, which resembles an error correction framework.

$$\tilde{y}_t - \tilde{y}_{t-1} = -r_t = (\tilde{A} - I)\tilde{y}_{t-1} + \tilde{c} + \varepsilon_t. \tag{6}$$

One can then estimate this equation to determine if the in-sample R^2s are too large to be consistent with rational pricing models, for a given level of risk aversion. One can then convert the estimated system in (6) back to a form, (5), that is comparable to the original VAR. That is, (6) contains essentially the same information as the original VAR in yields, (1'), but it enables us to judge the plausibility of the in-sample fit versus a rational asset pricing model.

One can also compare the in-sample R^2 of the regressions in (6) to an OOS R^2 statistic

$$R_{OS}^2 = 1 - \frac{\sum_{t=1}^{T}(r_t - \widehat{r_t})^2}{\sum_{t=1}^{T}(r_t - \overline{r_t})^2}, \quad (7)$$

where $\widehat{r_t}$ is the fitted value from a predictive regression using in-sample coefficients and data through *t-1* and $\overline{r_t}$ is the historical average return estimated with in-sample data. This OOS R^2 is measured in the same units as the in-sample R^2.[12]

Table 6 reports the in-sample R^2, the OOS R^2 and the bounds on these R^2s implied by Kirby's (1998) calculations. The bounds assume a generous estimate of annualized standard deviation of the market return: 20 percent. The bounds are calculated for two possible values for relative risk aversion, 2.5 and 5; Levich and Poti (2008) cite Ross (2005) to argue that 5 is an upper bound on reasonable values of risk aversion. The results are fairly clear: Every in-sample R^2 estimate for bond returns is well above—6 or 7 times—the bounds on R^2s in a rational pricing model, even assuming the upper value, 5, for relative risk aversion. The lowest in-sample R^2 in the bond return regressions is for the 10-year Treasury return, at 0.0233, which is almost 6 times the value of the upper bound for daily R^2, 0.004. In addition, the OOS R^2s are negative for all the bond return regressions. These negative values are consistent with the previous results that show that the original VAR had very poor OOS forecasting performance. In summary, there is too much in-sample predictability in Table 6 to be consistent with a rational pricing model but that the negative OOS R^2s show that this in-sample predictability is spurious.

6. A VAR that Is Consistent with Rational Pricing

One can estimate VAR structures in returns that are more consistent with rational asset pricing than the unrestricted VAR. That is, these structures would restrict the coefficients to

[12] This out-of-sample R^2 is related in concept to the Campbell and Thompson (2008) OOS R^2 statistic, but the latter is usually implemented on an expanding sample, rather than with fixed coefficients.

produce a more plausible R^2 for the bond return regressions.[13] One might hope that such a restricted VAR would also be more stable over time than the unrestricted VAR.

We estimated such a VAR over the in-sample period in an unrestricted form—equation (5)—and under the restrictions that the R^2s implied for the bond yield equations in (6) cannot exceed the values in Table 6 — 0.001 and 0.004. We then examine the forecasting performance and implied impulse responses of these models. None of the restricted models were estimated with a bias correction.

Table 7 shows the OOS RMSE Theil statistics under these three assumptions. Although none of the VARs consistently outperform the naïve model in the OOS period, the point estimates clearly indicate that the restrictions on the R2 improve the OOS forecasting performance. The model with the tightest restrictions on the coefficients does the best (bottom panel) and the model with no restrictions on the coefficients (top panel) does the worst in OOS Theil statistics, particularly for the bond yield regressions.

Figures 4 and 5 show that the impulse responses for the restricted VAR models imply greater persistence and are more intuitively plausible than the original VAR, despite the fact that they are not explicitly corrected for endogenous-regressor bias.[14] In particular, the upper left panel of Figure 4 shows that when the R^2s in the bond yield equations are restricted to 0.004, the point estimate of the monetary shock to the 10-year Treasury is no longer zero after 2 or 3 months but is still non-zero at a horizon of 250 business days and the confidence interval does not narrow

[13] One can use Kuhn-Tucker conditions to determine how to restrict the VAR coefficients. In the case where the constraint binds, it turns out that the restricted regression coefficients are proportional to the unrestricted coefficients: $B = \left(\frac{1}{k}Y'X(X'X)^{-1}X'Y\left((Y-\bar{Y})'(Y-\bar{Y})\right)^{-1}\right)^{-1/2}(X'X)^{-1}X'Y$, where Y denotes the vector of the dependent variable, X denotes the matrix of independent variables and k is the upper bound on the R^2.

[14] Note that the vertical scales of Figures 4 and 5 are about twice those of Figure 2, consistent with greater uncertainty from VARs whose coefficients are restricted to imply less predictability.

but widens as time goes on. Likewise, monetary shocks have similarly more persistent effects on all the variables. Although the effects are often not statistically significant, the effects are more plausible in that the distributions do not become smaller over time, unlike the impulses from the original unrestricted VAR in Figure 2. Figure 5 shows that imposing a tighter bound on the bond yield R^2s increase these effects. Monetary shocks have more persistent effects on bond yields and we are unable to rule out the possibility that the effects are substantial even a year after the shock.

6. Discussion and Conclusion

Event studies show that the Federal Reserve's unconventional monetary policy announcements elicited the desired effects on asset prices, substantially reducing U.S. and foreign long-term yields, as well as the value of the dollar. These immediate, large reductions in long-yields were often followed by weeks or months of increases in yields. For example, U.S. long-term yields rose over 100 b.p. in the three months after the large March 18, 2009 QE1 purchase announcement. Many observers interpreted this pattern to mean that the unconventional shocks had very transient effects on asset prices. If that interpretation were true, it would suggest that unconventional policy has very limited ability to stimulate the economy.

It is very difficult, however, to determine whether post-announcement increases in yields were truly the result of quick dissipation of monetary shocks or simply the result of the impact of other, more positive shocks to long-yields. Wright (2012), however, offers a clever and potentially very helpful method to determine the duration of the effects of monetary shocks: He estimates a structural VAR on interest rate and inflation compensation data that is identified under the assumption that interest rate variance is higher on particular days of monetary announcements. The impulse response functions from this VAR indicate that unconventional monetary shocks have very transient effects on long yields, perhaps 2 or 3 months.

The contribution of the present paper is to show that, unfortunately, the Wright VAR forecasts very poorly, OOS and is structurally unstable. Thus, the impulse response functions from the VAR are spurious and the data do not support the idea that unconventional policy has only very transient effects. In addition, the estimated VAR violates bounds on predictability in rational asset pricing models developed by Kirby (1998).

VARs that are constrained to be consistent with rational asset pricing models tend to forecast better with tighter restrictions and generate much more persistent impulse responses to monetary policy shocks. Preliminary work suggests that such models appear to be more intertemporally stable and provide more realistic inference—more persistence of point estimates and greater uncertainty—regarding the dynamic responses of yields to monetary policy shocks.

What is the significance of the finding that the Wright VAR forecasts worse than a naïve model and is unstable? I am not arguing that one should throw out structural VARs or that one should select models using only forecasting performance or that economists should model all asset price movements as martingales. Structural VARs have advantages in answering interesting economic questions that outweigh the fact that there are likely to be models that forecast better.

Yet, neither can we ignore the structural stability of a VAR, particularly for questions whose answers hinge on dynamic structure. We estimate VAR parameters, rather than simply choosing them, for a reason: We expect past dynamic relations between the variables to represent a true and stable relation, an underlying reality. The failure of the VAR to meet even minimal OOS forecasting performance or to exhibit minimal structural stability indicates that the dynamic relations that the VAR coefficients purport to describe —those that lie behind Figure 2 —do not really exist. They are spurious and therefore the data do not support the claimed lesson of Figure

2—that monetary shocks are transient. Instead, the fact that the restricted and naïve models clearly fit the data better than the unrestricted VAR implies that we should take their implications seriously: monetary shocks appear to be very persistent, although we cannot really know how persistent.

So, how is one to interpret the rise in yields after expansionary unconventional monetary policy shocks? Wright implicitly suggests that markets simply initially overreacted to the quantitative easing actions.

> "*A possible – although optimistic – interpretation is that the economic stimulus provided by these Federal Reserve actions caused the economy to pick up. Another interpretation is that markets initially overreacted to the news of these quantitative easing actions.*" — Wright (2012), page F464

It is certainly possible that the latter interpretation is correct. The rational-bound model for predictability assumes that markets know the structure of the economy, including the distribution of asset returns. The above interpretation explicitly suggests that markets were wrong about the distribution of asset returns.

Such an "overreaction" interpretation would, however, suggest that not only did markets overreact once but that they overreacted systematically to many FOMC announcements that occurred over a period of years. Markets were still strongly reacting to FOMC announcements in the summer of 2012. Do market participants never learn from repeated mistakes? In addition, the fragility of the VAR means that this hypothesis lacks specific empirical support, though it could still be true.

It is also possible that the unconventional monetary policy actions caused the economy to pick up, which raised long yields. To the extent that unconventional actions increased confidence and risk appetites, they sowed the seeds of their own partial reversal; but higher confidence signals success rather than failure. This “stimulus effect” hypothesis, however, would also indicate a lot of predictability in long yields that is not consistent with rational pricing and reasonable degrees of risk aversion.

An alternative explanation is that all sorts of shocks continually influence the economy and asset prices and that the rises in long yields after the unconventional policy actions were the results of partly coincidental non-monetary shocks. For example, Meyer and Bomfim (2009) argue that higher expected growth, new Treasury issuance, and the return of investors’ risk appetite drove the increase in Treasury yields from late March through mid-June 2009. The parallel rise in equity and oil prices over the same March-to-June period corroborates the explanation that higher expected growth and a rise in risk appetites raised long rates. This last interpretation—“other shocks”—would be consistent with the usual assumption that shocks to asset prices are very persistent. Unfortunately, given the statistical problems associated with estimating persistence, we cannot know how persistent with any precision.

Bias Adjustment Appendix

This appendix briefly describes the bias adjustment used in this paper, as well as many previous papers. It follows the discussions in Kilian (1998) and Efron and Tibshirani (1993).

1. Estimate the VAR parameter matrix, A, with the original T by k sample to obtain the OLS estimates of the parameters, A_{OLS}, residuals, ε_{OLS} and the associated covariance matrix.
2. Using the estimated VAR as the data generating process, bootstrap 1000 data samples of size T by k, resampling from the residuals, ε_{OLS}, using coefficients A_{OLS} and drawing initial conditions from the unconditional distribution of the data.
3. Estimate the VAR parameter matrix, A^*, for each simulated data set using OLS, and calculate the average of those matrices, A_{MC}, over the bootstrapped samples.
4. The difference between the true parameters for the simulated data generating process, A_{OLS}, and that of the average estimated VAR coefficient matrix, A_{MC}, is the estimate of the bias in the original VAR on the real data. Therefore, the bias-adjusted coefficient matrix is computed as: $A_{BA} = A_{OLS} + (A_{OLS} - A_{MC})$.
5. The modulus of A_{BA} is checked to ensure that it implies a stationary system. If it does not, the bias correction term, $(A_{OLS} - A_{MC})$, is gradually reduced until the modulus of the implied A_{BA} is less than one. Kilian (1998) describes this procedure in greater detail.

References

Andrews, D. (1993). 'Tests for Parameter Instability and Structural Change with Unknown Change Point', *Econometrica,* vol. 61, pp. 821–856.

Bai, J. and Perron, P. (1998). 'Estimating and Testing Linear Models with Multiple Structural Changes', *Econometrica*, vol. 66, pp. 47–78.

Bai, J. and Perron, P. (2003). 'Computation and Analysis of Multiple Structural Change Models', *Journal of Applied Econometrics*, vol. 18, pp. 1–22.

Bauer, M. and Neely, C. (2013). 'International Channels of the Fed's Unconventional Monetary Policy', Working Paper, Federal Reserve Bank of St. Louis.

Baumeister, C. and Luca B. (2010). 'Unconventional monetary policy and the Great Recession', Working Paper, European Central Bank.

Bekaert, G. and Hodrick, R. (1992). 'Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets', *Journal of Finance,* vol. 47, pp. 467–509.

Bekaert, G., Hodrick, R. and Marshall, D. (1997). 'On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates', *Journal of Financial Economics*, vol. 44, pp. 309–348.

Bernanke, B. (1986). 'Alternative Explanations of the Money-Income Correlation,' *Carnegie-Rochester Conference Series* on Public Policy, Autumn, pp. 49-100.

Blanchard, O. and Quah, D. (1989). 'The Dynamic Effects of Aggregate Demand and Supply Disturbances,' American Economic Review, September, pp. 655-73.

Blanchard, O. and Watson, M. (1986). 'Are Business Cycles All Alike?,' in Robert J. Gordon, ed., The American Business Cycle, University of Chicago Press.

Campbell, J. and Thompson, S. (2008). "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?', *Review of Financial Studies*, vol. 21, pp. 1509–1531.

Craine, R. and Martin, V. (2008). 'International Monetary Policy Surprise Spillovers', *Journal of International Economics*, vol. 75, pp. 180–96.

Efron, B. and Tibshirani, R. (1993). 'An Introduction to the Bootstrap', New York: Chapman and Hall.

Faust, J., Rogers, J. and Wright, J. (2003). 'Exchange rate forecasting: the errors we've really made', *Journal of International Economics* vol. 60, pp. 35–59.

Gagnon, J., Raskin, M., Remache, J. and Sack, B. (2011a). 'The Financial Market Effects of the Federal Reserve's Large-Scale Asset Purchases', *International Journal of Central Banking* vol. 7, pp. 3–43.

Gagnon, J., Raskin, M., Remache, J. and Sack, B, (2011b). 'Large-scale asset purchases by the Federal Reserve: Did they work?', Federal Reserve Bank of New York *Economic Policy Review,* vol. 17, pp. 41–59.

Gambacorta, L. Hofmann, B. and Peersman, G. (2012). 'The Effectiveness of Unconventional Monetary Policy at the Zero Lower Bound: A Cross-Country Analysis', Working Paper, Bank for International Settlements.

Gertler, M. and Karadi, P. (2013). 'QE 1 vs. 2 vs. 3...: A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool', *International Journal of Central Banking*, vol.9, pp. 5–53.

Ghysels, E., Guay, A. and Hall, A. (1998). 'Predictive Tests for Structural Change with Unknown Breakpoint', *Journal of Econometrics*, vol. 82, pp. 209–233.

Goyal, A. and Welch, I. (2008). 'A comprehensive look at the empirical performance of equity premium prediction', *Review of Financial Studies* vol. 21, pp. 1455–1508.

Hamilton, J. and Wu, J. (2012). 'The effectiveness of alternative monetary policy tools in a zero lower bound environment', *Journal of Money, Credit and Banking*, vol. 44, pp. 3–46.

Joyce, Michael, A., Lasaosa, A., Stevens, I. and Tong, M. (2011). 'The Financial Market Impact of Quantitative Easing in the United Kingdom', *International Journal of Central Banking*, vol. 7, pp. 113–61.

Kilian, L. (1998). "Small-sample confidence intervals for impulse response functions," *Review of Economics and Statistics*, vol. 80, pp. 218–30.

Kirby, C., (1998). 'The Restrictions on Predictability Implied by Rational Asset Pricing Models', *Review of Financial Studies* vol. 11, pp. 343–382.

Kohn, D. (2009). 'Monetary policy research and the financial crisis: strengths and shortcomings. Speech at the Federal Reserve Conference on Key Developments in Monetary Policy, Washington D.C.

Levich, R. and Poti, V. (2008). 'Predictability and 'Good Deals' in Currency Markets'. Working Paper, National Bureau of Economic Research.

Litterman, R. (1986). 'Forecasting with Bayesian Vector Autoregressions - Five Years of Experience', *Journal of Business and Economic Statistics*, vol. 4, pp. 25–38.

Mankiw, N., and Shapiro, M. (1986). 'Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models.' *Economics Letters*, vol. 20, pp. 139-145.

Mark, N. (1995). 'Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability', *American Economic Review*, vol. 85, pp. 201–18.

Meese, R. and Rogoff, K. (1983). 'Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?', *Journal of International* Economics, vol. 14, pp. 3–24.

Meyer, L. Bomfim, A. (2009). ‘Were treasury purchases effective? Don’t just focus on treasury yields…’, Monetary Policy Insights: Fixed Income Focus, Macroeconomic Advisers Report.

Meyer, L. and Bomfim, A. (2010). ‘Quantifying the Effects of Fed Asset Purchases on Treasury Yields’, Monetary Policy Insights: Fixed Income Focus, Macroeconomic Advisers Report.

Neely, C. (2013), ‘The Large-Scale Asset Purchases Had Large International Effects’, Working Paper, Federal Reserve Bank of St. Louis.

Newey, W., and West. K. (1994). ‘Automatic lag selection in covariance matrix estimation,’ *The Review of Economic Studies*, vol. 61, pp. 631-53.

Rigobon, R. and Sack, B. (2004). ‘The Impact of Monetary Policy on Asset Prices,’ *Journal of Monetary Economics*, vol. 51, pp. 1553–75.

Ross, S. (2005). ‘Neoclassical Finance’, Princeton University Press: Princeton, New Jersey.

Shapiro, M. D., and Watson, M., (1988). ‘Sources of Business Cycle Fluctuations,’ in Stanley Fischer, ed,, NBER Macroeconomics Annual 1988 (MIT Press), pp. 111-48.

Sims, C. (1986). ‘Are Forecasting Models Usable for Policy Analysis?,’ Federal Reserve Bank of Minneapolis Quarterly Review, Winter, pp. 2-16.

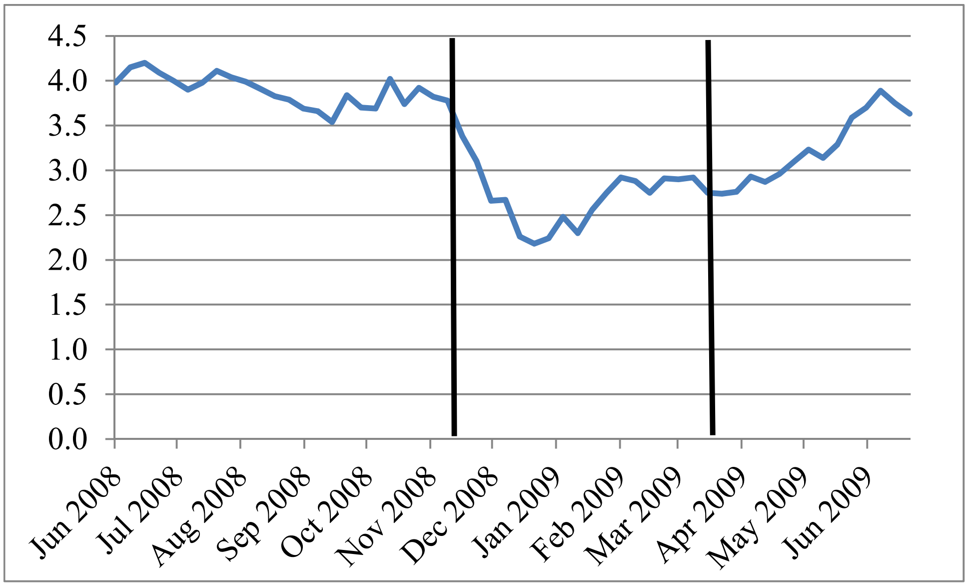
Stambaugh, R. (1986). ‘Bias in Regressions With Lagged Stochastic Regressors’, Working Paper, Center for Research in Security Prices, University of Chicago.

Stroebel, J. and Taylor, J. (2009). ‘Estimated Impact of the Fed’s Mortgage-Backed Securities Purchase Program’, Working Paper, National Bureau of Economic Research.

Thornton, D., and Valente, G. (2012). ‘Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective.’ *Review of Financial Studies*, vol 25, pp. 3141-68.

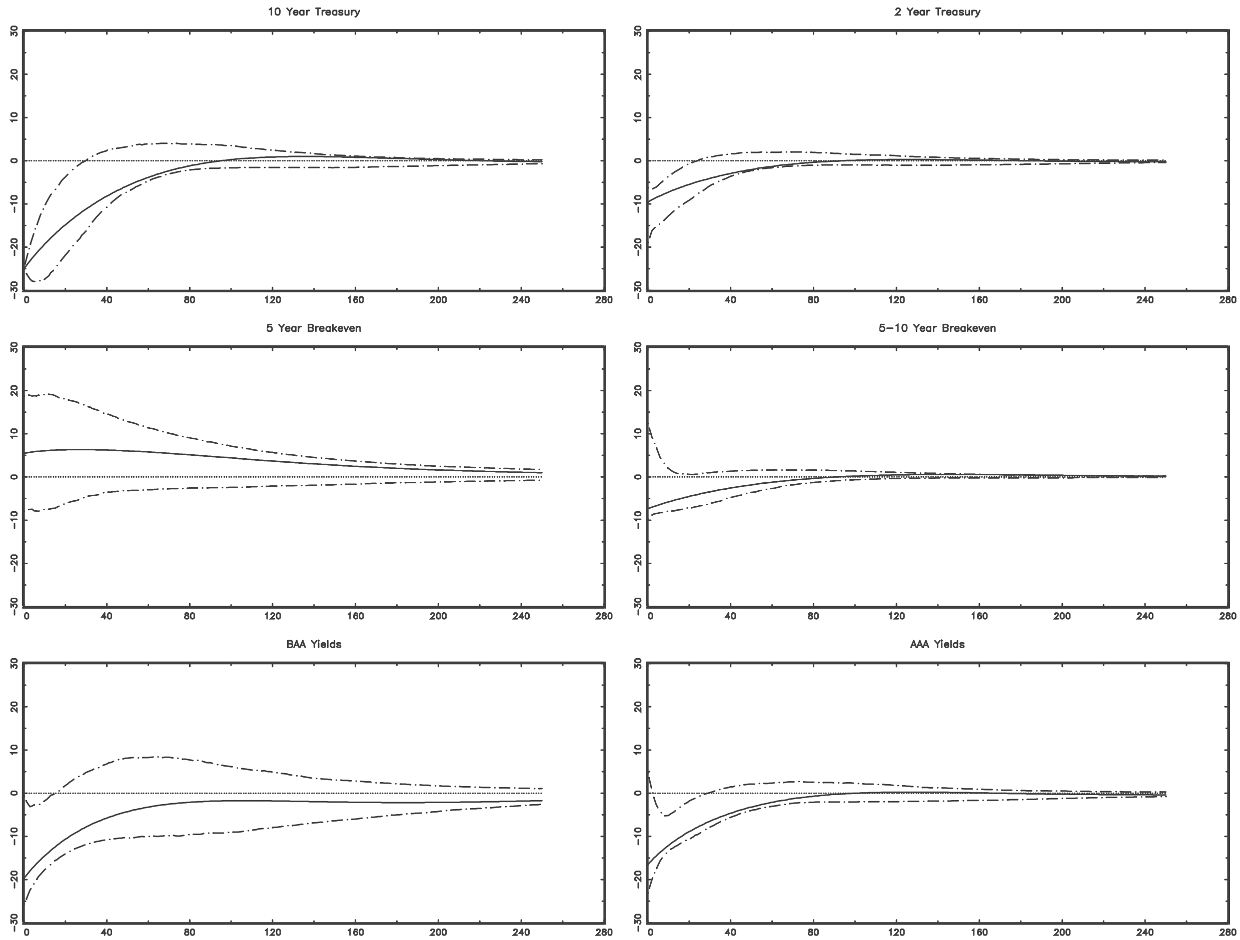
Wright, J. (2012). ‘What does Monetary Policy do to Long-term interest rates at the Zero Lower Bound?’, The Economic Journal, vol. 122, pp. 447–66.

Figure 1: Nominal yields on 10-year Treasuries



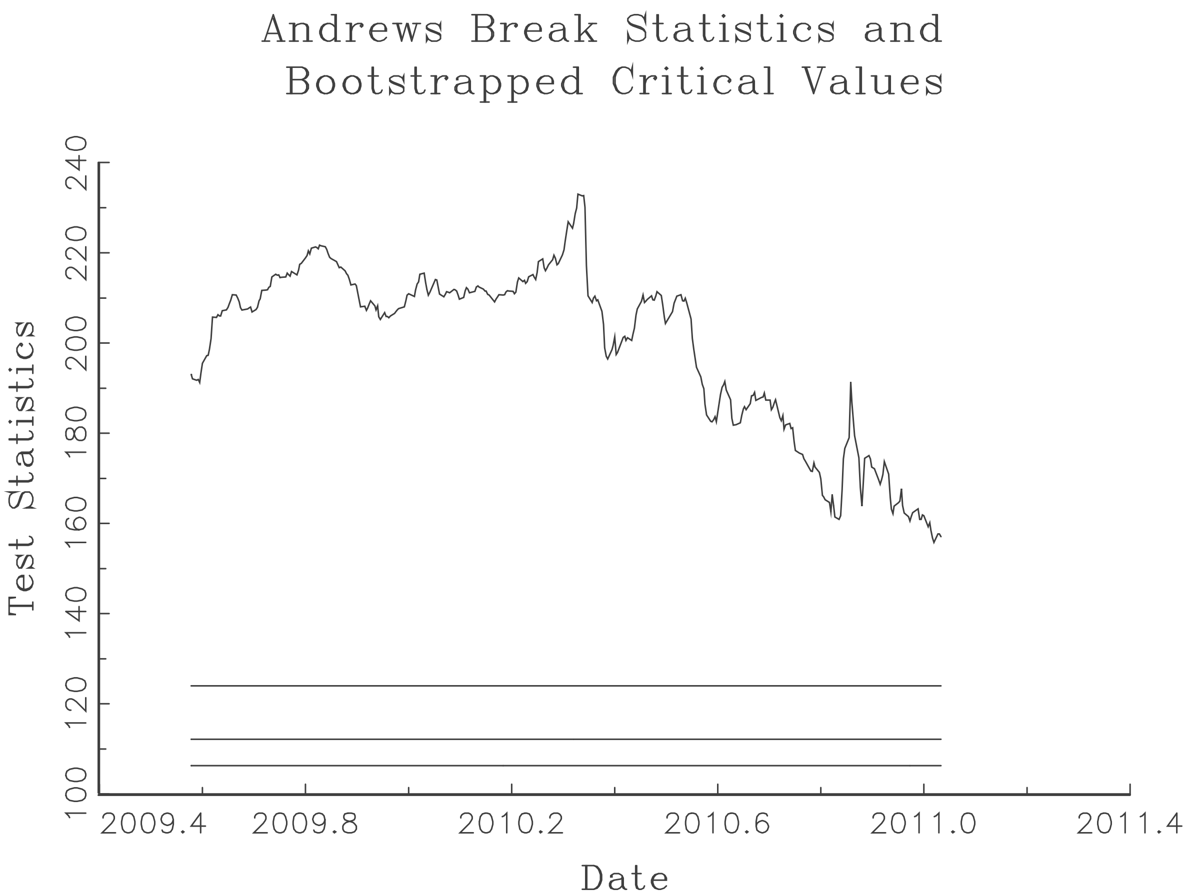
Notes: The figure depicts yields on 10-year Treasuries from June 2008 through June 2009. Vertical lines denote LSAP purchase announcements of November 25, 2008 and March 18, 2009. The source is Fred.

Figure 2: Wright's impulse responses



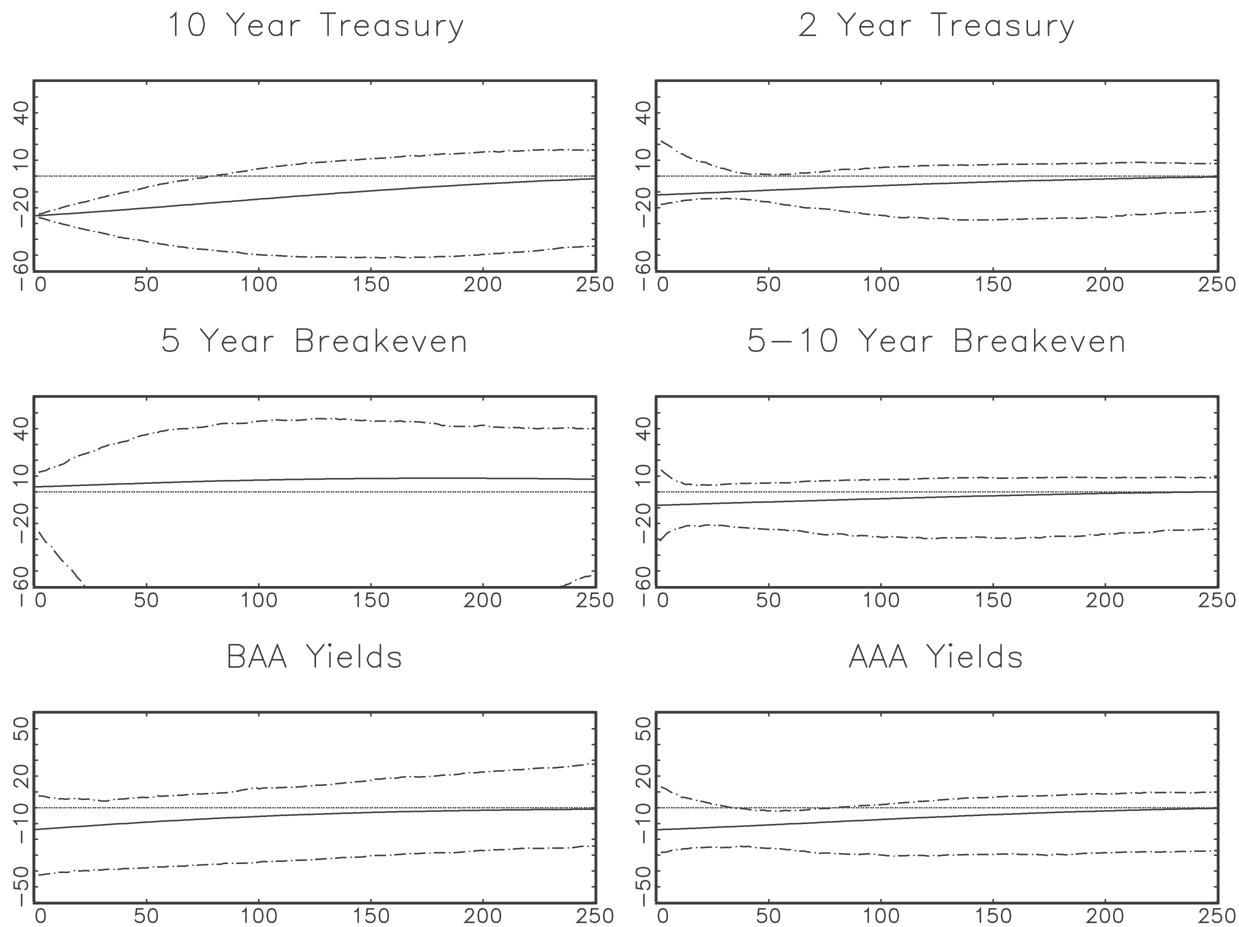
Notes: The figure illustrates impulse responses from monetary policy shocks on daily interest rates in a 6-variable VAR. The impulse responses are structurally identified by the greater variance of interest rates on days of monetary policy announcements. The figure replicates Wright's (2012) study of the impact of unconventional monetary policy shocks. The source of the data is Haver Analytics.

Figure 3: Andrews (1993) structural break statistics for the VAR



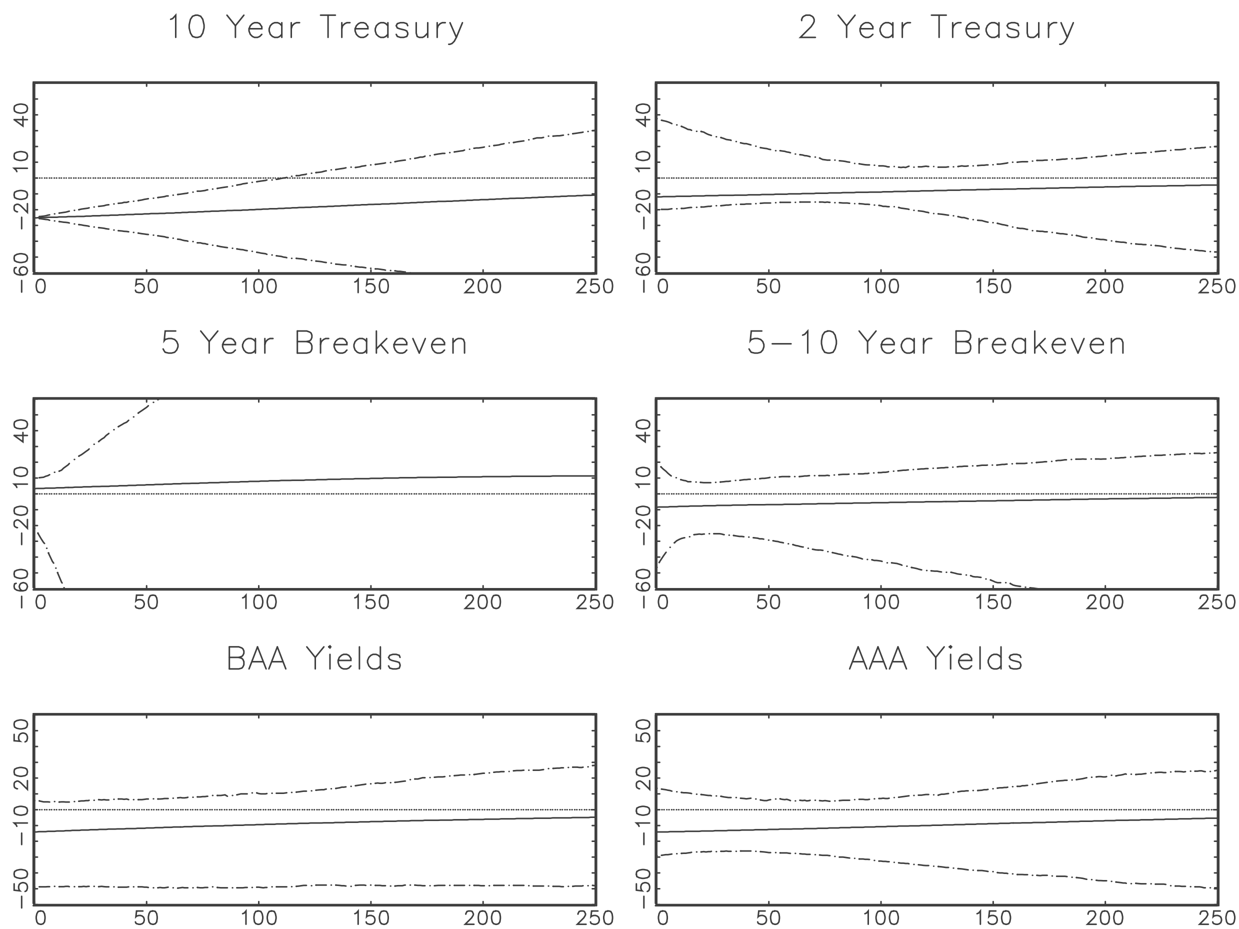
Notes: The figure plots the test statistics for a structural break in the reduced form VAR coefficients from the 25th to the 75th percentile of the sample, including the critical values for significance at the 10, 5, and 1 percent levels. The data clearly rejects any stability in the parameters.

Figure 4: Impulse responses from the VAR with the bond yield R^2s restricted not to exceed 0.004



Notes: The figure illustrates impulse responses from monetary policy shocks on daily interest rates in a restricted 6-variable VAR given by (5). The equations for the bond yields—Treasuries and corporates—are restricted to imply R2s for the returns that do not exceed 0.004. The impulse responses are structurally identified, as in Wright (2012), by the greater variance of interest rates on days of monetary policy announcements.

Figure 5: Impulse responses from the VAR with the bond yield R^2s restricted not to exceed 0.001



The figure illustrates impulse responses from monetary policy shocks on daily interest rates in a restricted 6-variable VAR given by (5). The equations for the bond yields—Treasuries and corporates—are restricted to imply R2s for the returns that do not exceed 0.001. The impulse responses are structurally identified, as in Wright (2012), by the greater variance of interest rates on days of monetary policy announcements.

Table 1: In-Sample root-mean square forecast statistics

	2-Yr Treasury	10-Yr Treasury	5-Yr infl comp	5-10 yr fwd infl comp	Baa Yield	Aaa Yield
1-day Naive RMSFE	4.75	7.97	5.76	6.31	7.21	7.69
20-day Naive RMSFE	17.82	37.10	33.13	25.52	29.32	28.51
60-day Naive RMSFE	23.77	55.77	61.20	31.97	56.30	39.88
120-day Naive RMSFE	33.11	77.19	91.99	41.34	86.94	50.26
1-day VAR RMSFE	4.69	7.88	5.72	6.25	7.11	7.58
20-day VAR RMSFE	16.69	35.78	31.21	23.03	24.17	24.07
60-day VAR RMSFE	25.73	64.35	50.93	32.88	39.55	40.76
120-day VAR RMSFE	34.35	92.36	69.44	43.27	42.54	50.85
1-day Theil	0.99	0.99	0.99	0.99	0.99	0.99
20-day Theil	0.94	0.96	0.94	0.90	0.82	0.84
60-day Theil	1.08	1.15	0.83	1.03	0.70	1.02
120-day Theil	1.04	1.20	0.75	1.05	0.49	1.01
1-day Theil p-value	0.38	0.32	0.10	0.37	0.30	0.35
20-day Theil p-value	0.16	0.25	0.09	0.29	0.27	0.33
60-day Theil p-value	0.05	0.08	0.18	0.17	0.27	0.26
120-day Theil p-value	0.08	0.07	0.29	0.14	0.43	0.26

Notes: This table shows in-sample, root mean square, forecast statistics from the 6-variable VAR. The top panel shows the RMSFE for the naïve (martingale) forecast; the second panel shows the RMSFE for the bias-adjusted VAR predictions; the third panel shows Theil statistics, the ratio of the VAR RMSFE to the naïve RMSFE; the fourth panel shows the bootstrapped proportion of samples in which bootstrapped Theil statistics are greater than the actual Theil statistics in the third panel.

Table 2: In-Sample mean-error (bias) forecast statistics

	2-Yr Treasury	10-Yr Treasury	5-Yr infl comp	5-10 yr fwd infl comp	Baa Yield	Aaa Yield
1-day Naive ME	-0.14	-0.38	0.29	-0.18	-0.59	-0.34
20-day Naive ME	-2.18	-5.79	7.61	-1.13	-11.01	-5.55
60-day Naive ME	-4.21	-8.24	24.91	2.27	-28.13	-8.63
120-day Naive ME	-8.39	-9.40	45.04	6.26	-55.36	-13.07
1-day Naive ME t statistics	-0.80	-1.30	1.26	-0.70	-2.29	-1.26
20-day Naive ME t statistics	-0.98	-1.14	1.71	-0.34	-2.90	-1.48
60-day Naive ME t statistics	-0.79	-0.62	1.75	0.34	-2.42	-1.02
120-day Naive ME t statistics	-1.00	-0.44	1.78	0.50	-2.33	-1.06
1-day VAR ME	-0.01	-0.06	-0.03	-0.06	-0.03	-0.03
20-day VAR ME	-0.12	-0.34	1.19	0.59	-0.36	-0.15
60-day VAR ME	1.10	3.63	7.01	3.28	-0.21	3.18
120-day VAR ME	2.74	11.52	12.80	7.02	-5.91	5.71
1-day VAR ME t statistics	-0.09	-0.22	-0.13	-0.23	-0.10	-0.11
20-day VAR ME t statistics	-0.05	-0.07	0.29	0.18	-0.11	-0.05
60-day VAR ME t statistics	0.17	0.22	0.57	0.41	-0.02	0.32
120-day VAR ME t statistics	0.29	0.46	0.71	0.54	-0.46	0.45

Notes: This table shows in-sample mean-error (ME) forecast statistics from the 6-variable VAR. The top panel shows the ME for the naïve (martingale) forecast; the second panel shows the t statistics for those naïve mean errors, constructed with Newey-West statistics using appropriate lag length for the forecast horizon; the third and fourth panels show the same statistics for the bias-adjusted VAR coefficients.

Table 3: <u>Out-of-Sample</u> root-mean square forecast statistics

	2-Yr Treasury	10-Yr Treasury	5-Yr infl comp	5-10 yr fwd infl comp	Baa Yield	Aaa Yield
1-day Naive RMSFE	1.59	5.48	3.24	4.10	5.10	5.19
20-day Naive RMSFE	6.03	21.76	15.49	13.33	19.21	17.60
60-day Naive RMSFE	7.24	40.35	30.08	17.99	31.02	31.52
120-day Naive RMSFE	8.40	54.25	34.21	21.77	42.96	47.93
1-day VAR RMSFE	1.82	6.02	3.23	4.07	5.21	5.85
20-day VAR RMSFE	11.26	42.12	15.31	12.40	29.25	41.69
60-day VAR RMSFE	14.63	74.82	32.48	18.27	57.78	71.08
120-day VAR RMSFE	16.25	92.09	34.25	21.14	83.19	88.10
1-day Theil	1.15	1.10	1.00	0.99	1.02	1.13
20-day Theil	1.87	1.94	0.99	0.93	1.52	2.37
60-day Theil	2.02	1.85	1.08	1.02	1.86	2.25
120-day Theil	1.93	1.70	1.00	0.97	1.94	1.84
1-day Theil p-value	0.00	0.04	0.05	0.19	0.17	0.10
20-day Theil p-value	0.00	0.00	0.04	0.17	0.17	0.02
60-day Theil p-value	0.00	0.02	0.05	0.11	0.16	0.04
120-day Theil p-value	0.01	0.03	0.07	0.10	0.15	0.09

Table 4: Out-of-Sample mean-error (bias) forecast statistics, constructed in an ex post sample

	2-Yr Treasury	10-Yr Treasury	5-Yr infl comp	5-10 yr fwd infl comp	Baa Yield	Aaa Yield
1-day Naive ME	0.00	0.18	0.09	0.09	0.05	0.15
20-day Naive ME	0.04	2.17	1.13	0.53	-0.01	2.45
60-day Naive ME	1.23	8.64	2.66	2.01	1.61	8.36
120-day Naive ME	2.67	18.33	1.29	2.18	3.85	16.98
1-day Naive ME t statistics	0.05	0.77	0.58	0.47	0.22	0.71
20-day Naive ME t statistics	0.05	0.62	0.44	0.29	0.00	0.91
60-day Naive ME t statistics	0.75	0.81	0.32	0.45	0.19	0.95
120-day Naive ME t statistics	0.99	0.82	0.10	0.29	0.20	0.86
1-day VAR ME	-0.69	-2.08	-0.19	-0.09	-1.21	-2.36
20-day VAR ME	-7.29	-30.85	0.17	-3.35	-21.70	-33.47
60-day VAR ME	-8.12	-54.28	4.87	-7.27	-48.45	-58.33
120-day VAR ME	-10.96	-72.03	-0.76	-10.88	-72.41	-73.96
1-day VAR ME t statistics	-9.06	-8.19	-1.31	-0.49	-5.53	-10.17
20-day VAR ME t statistics	-4.94	-6.05	0.07	-1.86	-6.70	-7.38
60-day VAR ME t statistics	-2.65	-3.60	0.59	-1.74	-5.13	-4.55
120-day VAR ME t statistics	-2.49	-2.92	-0.06	-1.95	-4.12	-3.47

Table 5: Bai-Perron structural stability tests

	2-yr US Treasury yld	10-yr US Treasury yld	5-yr US Infl comp	5-10 yr US Infl comp	Moody's Baa yield	Moody's Aaa yield
UD max	***29.1***	***32.3***	***31.4***	21.4	**24.7**	**22.8**
WD max 5%	**34.2**	**32.4**	**32.3**	**23.1**	**33.9**	**31.3**
SupF(2|1)	*22.8*	16.0	*21.9*	9.3	**27.4**	***28.5***
SupF(3|2)	10.7	11.9	10.5	13.7	16.3	15.7

Table 6: In-sample and out-of-sample R^2 as well as rational bounds

	In-sample R^2	OOS R^2	Bound on R^2 with RRA = 2.5	Bound on R^2 with RRA = 5
2-Yr Treasury return	0.0252	-0.6033	0.0010	0.0040
10-Yr Treasury return	0.0233	-0.4105	0.0010	0.0040
5-Yr Δ inflation comp	0.0150	0.0151	0.0010	0.0040
5-Yr fwd Δ inflation comp	0.0244	0.0147	0.0010	0.0040
Baa return	0.0254	-0.1724	0.0010	0.0040
Aaa return	0.0294	-0.5126	0.0010	0.0040

Notes: This table shows the in-sample and OOS R^2 for the regressions of transformations of the Wright VAR variables into return form and the bounds implied by rational asset pricing and given coefficients of relative risk aversion (RRA). The bounds assume an annualized standard deviation to the market return of 20 percent.

Table 7: In-Sample root-mean square forecast statistics for the restricted VAR

	2-Yr Treasury	10-Yr Treasury	5-Yr infl comp	5-10 yr fwd infl comp	Baa Yield	Aaa Yield
Unrestricted R^2						
1-day Theil	1.27	1.19	0.99	0.99	1.09	1.23
20-day Theil	2.46	2.52	0.93	0.98	2.14	2.99
60-day Theil	3.18	2.44	0.90	1.09	2.47	2.75
120-day Theil	3.70	2.29	0.78	1.14	2.23	2.15
R^2 restricted to 0.004						
1-day Theil	1.05	1.03	0.99	0.99	1.01	1.03
20-day Theil	1.64	1.52	0.96	0.97	1.27	1.62
60-day Theil	2.48	1.74	0.94	0.97	1.54	1.88
120-day Theil	2.67	1.76	0.79	0.91	1.62	1.71
R^2 restricted to 0.001						
1-day Theil	1.01	1.01	0.99	0.99	1.00	1.00
20-day Theil	1.23	1.16	0.99	0.97	1.08	1.18
60-day Theil	1.84	1.30	1.02	0.93	1.20	1.35
120-day Theil	2.24	1.39	0.86	0.82	1.28	1.35

Notes: This table shows in-sample, root mean square, forecast statistics from the 6-variable VAR with the constraint imposed that the R2 not exceed 0.004. The top panel shows the RMSFE for the naïve (martingale) forecast; the second panel shows the RMSFE for the bias-adjusted VAR predictions; the third panel shows Theil statistics, the ratio of the VAR RMSFE to the naïve RMSFE; the fourth panel shows the bootstrapped proportion of samples in which bootstrapped Theil statistics are greater than the actual Theil statistics in the third panel.