Comment on "Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset"*

Michael D. Bauer[†], Glenn D. Rudebusch[‡], Jing Cynthia Wu[§]

May 15, 2013
Forthcoming in *American Economic Review*

Abstract

Term premia implied by maximum likelihood estimates of affine term structure models are misleading because of small-sample bias. We show that accounting for this bias alters the conclusions about the trend, cycle, and macroeconomic determinants of the term premia estimated in Wright (2011). His term premium estimates are essentially acyclical, and often just parallel the secular trend in long-term interest rates. In contrast, bias-corrected term premia show pronounced countercyclical behavior, consistent with theoretical and empirical arguments about movements in risk premia.

Keywords: term premia, dynamic term structure model, small-sample bias

JEL Classifications: E43, E44

*The views expressed herein are those of the authors and not necessarily shared by others at the Federal Reserve Bank of San Francisco or in the Federal Reserve System. We thank the editor Marty Eichenbaum, three anonymous referees, Alan Bester, and Drew Creal for helpful comments, and Jonathan Wright for providing us with his estimation code.

[†]Federal Reserve Bank of San Francisco, michael.bauer@sf.frb.org
[‡]Federal Reserve Bank of San Francisco, glenn.rudebusch@sf.frb.org
[§]The University of Chicago Booth School of Business, cynthia.wu@chicagobooth.edu

1 Introduction

In an important recent paper, Jonathan Wright (2011) presents a new international interest rate data set and estimates term premia in far-ahead forward rates using an affine Gaussian dynamic term structure model (DTSM) as well as a survey-based method. He then relates these empirical term premia to macroeconomic conditions, notably inflation uncertainty. In particular, he shows that the marked secular decline in estimated forward term premia across countries is correlated with decreases in measures of uncertainty around forecasts of future inflation. In this comment, we demonstrate that his model-based term premium estimates are distorted by small-sample bias in the estimated model parameters. Accounting for this bias changes the behavior of the resulting term premium estimates in terms of their trend and cyclical behavior.

In an affine Gaussian DTSM, bond yields are determined by a small number of risk factors, and the dynamic evolution of these risk factors follows a Gaussian vector autoregression (VAR). If the risk factors in such a model are observable, and if the model is "maximally-flexible," i.e., if it has no parameter restrictions beyond those required for identification, we know from Joslin et al. (2011) that the maximum likelihood estimates of the VAR parameters can be found by means of ordinary least squares (OLS). However, such estimates of autoregressive parameters generally suffer from small-sample bias, which can be serious if the data sample is relatively short or the persistence of the series is high, both arguably the case in Wright's paper. Term premia are calculated as the difference between long-term interest rates and the corresponding risk-neutral rates, i.e., forecasts of (average) future short-term interest rates. These are based on VAR forecasts of the risk factors. Inaccurate parameter estimates for the underlying VAR translate into inaccurate and potentially misleading risk-neutral rates and term premium estimates. The intuition is straightforward. The bias typically goes in the direction of making the estimated system less persistent than the data-generating process. Due to this upward bias in the speed of mean reversion, forecasts of future short rates revert to their unconditional mean too quickly and consequently risk-neutral rates are too stable.

Here we quantify the small-sample bias in the DTSM estimates of Wright, and demonstrate the economic consequences of correcting for it, using the methodology laid out in Bauer et al. (2012). The bias is large, and bias correction substantially alters the estimated risk-neutral rates and term premia. Wright's risk-neutral forward rates are close to constant for several countries, so movements in long-term yields are entirely accounted for by changes in term premia. Hence, the long downward trend in the long-term yields of most developed countries is attributed almost exclusively to a secular decline in term premia. In contrast, bias-corrected risk-neutral rates have a distinct downward trend in most countries, so the associated term premia show a much less pronounced secular decline. Consequently, after bias correction, a larger share of the decline in long-term yields is attributed to declining risk-neutral rates, which is consistent with the decline in inflation expectations and expected policy rates over the past two decades described by Kozicki and Tinsley (2001), Kim and Orphanides (2012), and others. Furthermore, we document that Wright's term premia are essentially acyclical in all countries, as is common for conventional DTSM-based term premia. In contrast, bias-corrected term premia show distinctly countercyclical behavior. Accordingly, the bias-corrected premia are more consistent with much empirical and theoretical research that has found support for countercyclical risk premia, including, among many others, Campbell and Cochrane (1999), Wachter (2006), and Cochrane and Piazzesi (2005).

Finally, a key result of Wright's analysis is that inflation uncertainty is an important driver of nominal term premia. Using panel regressions, he shows that the estimated term premia across countries have a positive relationship with measures of inflation uncertainty. This is an appealing result, since uncertainty about future inflation is often cited by financial market participants as an important source of risk in nominal bond returns. We find that bias-corrected estimates of term premia show a similar relationship with measures of inflation uncertainty, so Wright's key conclusion remains intact.

The conclusion of our analysis is that small-sample bias is a serious concern when using a DTSM to estimate policy expectations and term premia. The bias in the estimated DTSM parameters is often large, and can have important economic consequences. Notably, in the context of the model and data set used in Wright's paper, correcting for small-sample bias leads to term premia that are more plausible from a macroeconomic perspective. However, it should be noted that, like Wright, our focus is on the point estimates of term premia and how they relate to measures of inflation uncertainty. An important caveat to our analysis (and Wright's) is that the estimation uncertainty is large. Indeed, we calculate confidence intervals and show that the differences between the OLS and bias-corrected point estimates over time are invariably smaller than the width of the estimated confidence intervals. In large part, this uncertainty is an unavoidable result of the relatively short data sample used by Wright. Of course, estimating DTSMs is not the only way to infer policy expectations and term premia. One alternative is to include the information from surveys of market participants' expectations of future short-term interest rates, as in Wright. However, some researchers find the reliability of survey data problematic.[1]

[1]For example, Christensen and Rudebusch (2012) examine the interest rate forecasts and term premia of the Kim and Wright (2005) model, which includes survey data in an empirical DTSM. They find that the Kim and Wright model forecasts worse out of sample than a yields-only DTSM. In particular, the model with

2 Model specification, data, and estimation

The affine DTSM that is the focus of Wright's paper and of our own analysis includes three observable risk factors, the first three principal components of yields, corresponding to level, slope, and curvature, and two observable macro factors, inflation and GDP growth. The macro factors are treated as unspanned, meaning that they help to predict future yields, but do not have additional explanatory power for contemporaneous yields beyond the yield factors. The model imposes absence of arbitrage by positing the existence of a pricing kernel, with risk prices taken to be essentially affine. There are no over-identifying restrictions, i.e., the model is maximally-flexible. The data used for estimation consist of quarterly zero-coupon yields, inflation and growth, from the first quarter of 1990 to the first quarter of 2009. We use the exact same data and model specification as in Wright's paper.[2]

Maximum likelihood estimation of the model follows the approach suggested by Joslin et al. (2011), applied to a macro-finance DTSM with unspanned risk factors in Joslin et al. (2012). First, the VAR parameters are estimated using OLS. Then, in a second step the remaining model parameters are estimated by maximizing the likelihood function, taking the VAR parameters as given. We replicate Wright's model estimates and denote this set of DTSM estimates and term premia in the following as "OLS estimates."

The OLS estimator is biased for the coefficients in a VAR, because the regressions include lagged dependent variables and strict exogeneity is violated. This small-sample bias will typically tend to make the estimated system less persistent than the data-generating process, i.e., the speed of mean reversion will be overestimated. This issue is closely related to small-sample bias in predictive regressions with persistent regressors, discussed in Stambaugh (1999) and elsewhere. We will see that in the DTSM context bias has the same economic implication as in the return regression context (for a discussion of the latter refer to Bekaert et al., 1997): It spuriously attributes too much variation to the risk premium and too little to the expectations component in long-term interest rates.

Because Wright's data sample here spans a relatively short time period (77 quarterly observations) and because interest rates and macroeconomic variables are highly persistent, small-sample bias is a serious concern in this application. One solution to the bias problem is to directly correct the small-sample bias in the DTSM estimates using the methodology developed in Bauer et al. (2012).[3] Our estimation approach consists in replacing the OLS

survey data generates expectations for future short rates that revert to the mean too quickly, so estimates of the term premium are too variable.

[2]The data and an online appendix with detailed parameter estimates for Wright's paper are available at http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.4.1514.

[3]Another possibility is to restrict the risk pricing as in Joslin et al. (2012) and Bauer (2011), and in this

estimates of the VAR parameters in the first step by bias-corrected (BC) estimates. For this purpose, we use an indirect inference estimator. The idea is to use repeated simulations to find those data-generating VAR parameters that yield a mean of the OLS estimator equal to the actual OLS estimates obtained from the data.[4] We ensure that the resulting BC estimates imply a stationary VAR system by using the stationarity adjustment suggested by Kilian (1998): Bias-corrected VAR estimates with explosive eigenvalues are shrunk toward the OLS estimates until the VAR is stable. In the second estimation step, the remaining parameters are again estimated by maximizing the likelihood function for given VAR parameters. Further details for our estimation approach as well as Monte Carlo evidence that this method has desirable small-sample properties are provided in Bauer et al. (2012).

Table 1 displays summary statistics for the OLS and BC model estimates for each of the ten countries. The first column shows the root-mean-squared fitting errors across yields, in annualized percentage points. Both sets of estimates lead to a very tight cross-sectional fit of yields. The BC estimates do not worsen the cross-sectional fit relative to OLS, since in both cases the second estimation step minimizes a weighted sum of squared pricing errors, and is barely affected by the choice of VAR parameters of the first step.

The next four columns of Table 1 show measures of the persistence of the estimated VAR system under both the real-world probability measure ($\mathbb{P}$) and under the risk-neutral probability measure ($\mathbb{Q}$). For intuition, note that higher persistence under the $\mathbb{P}$-measure will generally translate into higher variability of expectations of future short-term interest rates, i.e., risk-neutral interest rates, whereas the persistence under the $\mathbb{Q}$-measure is related to the variability of fitted interest rates.

The results show that the persistence of the VAR under the $\mathbb{P}$-measure is generally increased substantially through bias correction. The second and third column show the largest absolute eigenvalue of the mean-reversion matrix under each measure. The $\mathbb{P}$-eigenvalue is always larger, and quite significantly so, for the BC estimates than for the OLS estimates.[5] The $\mathbb{Q}$-eigenvalue is barely affected by bias correction. Comparing $\mathbb{P}$- and $\mathbb{Q}$-eigenvalues shows that for the OLS estimates, the persistence under the risk-neutral measure is always larger than the persistence under the real-world measure. Bias correction reduces this difference in persistence, and in some cases pushes the persistence under $\mathbb{P}$ higher than the persistence under $\mathbb{Q}$.

way to use the information in the cross section of interest rates to pin down the VAR parameters.

[4]A simpler alternative would be to use simple bootstrap bias correction—this is theoretically less appealing, because it only removes first-order bias, but in most practical applications will lead to similar results.

[5]For the U.S., Germany, Canada, and the U.K. the restriction that the largest $\mathbb{P}$-eigenvalue remain below unity is binding.

Because the largest eigenvalue of the mean reversion matrix does not fully capture the persistence of a VAR, we also report the value of the impulse-response function of the level factor in response to a level shock at the five-year horizon. It measures by how much five-year-ahead forecasts of the level of interest rates are revised in response to a unit shock to the level factor today. Comparing the value of the impulse response between the OLS and BC estimates reveals that bias correction increases persistence by an economically sizeable magnitude. In several countries including the U.S., a unit level shock essentially dies out over the course of the five-year-horizon for the case of OLS, whereas it leads to a substantial revision of forecasts at that horizon for the case of the BC estimates. Again, bias correction barely affects the persistence under the $\mathbb{Q}$-measure, and generally brings the persistence of the real-world dynamics closer to the persistence of the VAR under the risk-neutral pricing measure.

3 Estimated risk-neutral rates

We follow Wright in decomposing five-to-ten year forward rates into an expectations component and a forward term premium. The risk-neutral forward rate is calculated as the average expected three-month interest rate from five to ten years horizon. The forward term premium is calculated as the difference between the fitted forward rate and the risk-neutral rate. The last three columns of Table 1 report the volatilities of actual forward rates, risk-neutral rates, and forward term premia. Figure 1 shows fitted forward rates, as well as OLS and BC risk-neutral forward rates. Also shown are 90%-confidence intervals for the BC risk-neutral forward rates, obtained using a bootstrap method similar to the one in Kilian (1998).[6] In this section, we focus on how bias correction affects the variability and trend behavior of the expectations component of forward rates.

Risk-neutral interest rates implied by the OLS estimates are extraordinarily stable over time. Figure 1 shows that OLS measures of average future short rates for many countries, including the U.S., are essentially constant. Unreasonably stable short-rate expectations—in light of likely variation in expectations about the inflation target on the equilibrium real short rate—are a common problem with conventional DTSM estimates (Kozicki and Tinsley, 2001; Kim and Orphanides, 2012). As evident from Table 1, the standard deviation of the

[6] In each bootstrap replication, we impose stationarity using the adjustment proposed in Kilian (1998), shrinking any non-stationary BC estimates toward the OLS estimates until they are stationary. We also choose the VAR intercept to exactly match the sample mean of the risk factors—that is, we effectively work with demeaned series. Taking into account the uncertainty about the mean would make the confidence intervals even wider.

expectations component of forward rates in most countries is substantially smaller than the standard deviation of actual forward rates. Hence, expectations of future short-term interest rates barely contribute to explaining the movements in forward rates, and consequently term premia simply mirror the evolution in forward rates. In particular, the secular decline in long-term interest rates that most developed countries exhibited over the recent past, and the cyclical variation in long-term interest rates is attributed almost entirely to movements in term premia.

Bias correction dramatically alters the variability and trend behavior of risk-neutral rates. Most importantly, risk-neutral rates become significantly more volatile when we account for small-sample bias in the model estimation. The intuition for this is that the higher persistence of the VAR implies that forecasts of future short-term interest rates revert to their mean more slowly. Therefore, these forecasts respond more strongly to shocks to the current yield curve. Volatilities of risk-neutral rates rise by a multiple with bias correction, doubling at the least. For the countries where expectations under the OLS estimates are implausibly stable, bias correction leads to more meaningful variation. Consequently, BC risk-neutral rates contribute significantly to the secular decline and cyclical patterns in far-ahead forward rates, and the resulting term premia differ from forward rates and have a life of their own.

One concern about bias-corrected estimation is that it can increase the variance of the estimator. Importantly, the mean-squared error could either increase or decrease through bias correction, depending on whether the decrease in the bias outweighs a potential increase in variance. Here we focus on estimation uncertainty for the objects of interest, the expectations of future short-term interest rates. As Figure 1 shows, the confidence intervals for the BC risk-neutral rates are quite wide. However, we have found that confidence intervals are not substantially larger than those for OLS (not shown). Because the OLS and BC confidence intervals do not completely overlap, bias correction can alter statistical conclusions. On a more general level, the fact that the confidence intervals are wide in both cases reflects the fact that parameters of a VAR with very persistent series are hard to estimate. This argument can motivate the use of additional information—such as survey expectations (Kim and Orphanides, 2012) or parameter restrictions (Bauer, 2011).

Of particular interest is the trend behavior over the last two decades. The decline in long-term interest rates was accompanied by substantial decreases in long-term inflation expectations (see, for example, Wright's figure 2). Table 2 takes a closer look at this trend behavior, and answers the question how much of the secular decline in far-ahead forward rates rates is attributed to short rate expectations and term premia, respectively. For those six countries for which survey expectations are available, it shows the changes in the Consensus forecast of five-to-ten-year CPI inflation and real GDP growth, in the five-to-ten-year forward rate, and in the model-based fitted forward rates, risk-neutral rates, and term premia. Changes are calculated for averages across the available observations during the two calendar years at the beginning and end of the sample, and are reported in basis points.

Based on the OLS estimates, short-rate expectations (risk-neutral rates) play only a small role in explaining the secular decline in far-ahead forward rates over the sample. With the exception of Germany, the majority of the decline in forward rates is attributed to decreasing term premia. While this in itself is not necessarily implausible, the stability of OLS-based risk-neutral rates becomes a puzzle when comparing them to inflation expectations: For the U.S., the U.K., and Australia, risk-neutral rates declined by much less than survey-based inflation expectations over the same period. This has the counter-factual implication that expectations of future real short-term rates increased over a period of time when interest rates generally decreased. In addition, Kim and Orphanides (2012) document for the U.S. that survey-based expectations of future policy rates declined over the past two decades, in contrast to the nearly constant OLS-based policy expectations. In light of this evidence, the trend behavior of OLS-based risk-neutral rates appears implausible.

The decomposition of far-ahead forward rates based on the BC estimates gives a significantly larger role to the expectations component. The declines in risk-neutral rates over the sample are larger than the declines in inflation expectations, which implies that both inflation expectations and expectations of future real interest rates contributed to the secular decline in interest rates. From a macroeconomic perspective, this type of trend behavior generally appears more plausible—both the decreasing trend in interest rates and decreasing long-term growth expectations indicate that real rates likely decreased over the last two decades. Hence, for most countries expectations of future nominal short-term interest rates have almost certainly decreased by more than the OLS estimates suggest.

For the U.S. and for Germany, the BC decompositions imply very large decreases in risk-neutral rates, relative to the decreases in inflation expectations. For the U.S., implied real rates fall by more than 2.5 percentage points, over a period when real growth expectations have remained relatively stable and Blue Chip long-run expectations of inflation and nominal interest rates (not shown) both declined by a similar magnitude. For Germany, the implied decline in real rates is over five percentage points, and while real growth expectations also declined, this magnitude appears implausibly large. This indicates that for these particular data samples, the BC estimation probably over-correct for bias, and that the true persistence of the VAR, and the true decrease in far-ahead policy expectations lies somewhere in between the OLS and BC results. Additional BC estimates for the U.S. based on longer, more commonly used sample periods starting in the early 1980s, lead to more plausible magnitudes of declines in policy expectations and real rates.[7] More generally, bias correction always adjusts by the estimated mean bias, and for a particular sample, the realization of the estimator might be closer to or further from the true parameters than this estimate implies. Therefore it is particularly revealing to see how bias correction fares across samples, as in this application which considers a number of countries. Overall, BC estimation substantially alters the trend behavior of risk-neutral rates, and for the majority of countries the implications of the BC estimates are more closely in line with economic intuition than those of conventional DTSM estimates.

4 Estimated term premia and the macroeconomy

From a macroeconomic perspective, past research has led to some received wisdom about the behavior of government bond premia, notably their relationship to economic uncertainty and to the business cycle. In this section, we assess whether the OLS and BC term premia conform to these views following the methods of Wright.

4.1 Macroeconomic uncertainty and term premia

High uncertainty about future inflation implies additional risk for investors in nominal bonds. Hence inflation uncertainty should help drive term premia, and a key contribution of Wright's paper is to document a significant and robust relationship between different proxies for inflation uncertainty and his estimated term premia. We run the same panel regressions to assess the impact of bias correction on the relationship between term premia and uncertainty proxies. The results are displayed in Table 3. The top panel contains slope coefficients and bootstrapped p-values for regressions with the OLS term premium as the dependent variable—comparable to the top panel of Wright's Table 5—and the bottom panel shows the same for the BC term premium. Bias correction does not alter the fact that estimated term premia have a statistically and economically significant relationship with proxies of uncertainty about future inflation and growth. In regressions for the BC term premia, the measures of inflation uncertainty retain their positive coefficients and remain significant, with slightly smaller magnitudes of the coefficients. For the dispersion of growth forecasts, essentially the same observation holds. Overall, the relationship between these uncertainty measures and term premia appears to be robust to the correction for small-sample bias in the DTSM estimates.

[7]We estimated DTSMs using a sample starting in 1985, and calculated decompositions of the decline in far-ahead forward rates over the same period as in Table 2. Results are omitted for sake of brevity.

4.2 Cyclical behavior of term premia

One notable difference between the top and bottom panels of Table 3 concerns the recession dummy included by Wright in the last two specifications. The coefficient on the dummy is negative for the OLS term premia, whereas for the BC term premia, the dummy coefficient is positive and statistically significant. This result is evident in Figure 2, which shows the estimated term premia in each country with shaded recession bars according to Wright's recession dummy.[8] On average across all recessions, the OLS premia are lower at the trough of the recession than they are at the previous business cycle peak. Indeed, for no recession in any country, do the OLS term premia really increase. This raises a stark questions about the OLS term premia, since much theoretical (Campbell and Cochrane, 1999; Bansal and Yaron, 2005; Wachter, 2006) and empirical research (Harvey, 1989; Cochrane and Piazzesi, 2005; Lustig and Verdelhan, 2007; Lustig et al., 2010) supports the view that term premia evolve in a countercyclical fashion. Both the quantity and the price of risk are generally viewed as being higher during recessions than expansions. From the figure, the BC term premia are more likely to rise during a recession—for example, in Germany and Switzerland.

One countervailing force to the rise in risk and risk aversion during a recession is that government bonds may provide a safe haven for investors during financial crises, and the resulting "flight-to-quality" demand could put downward pressure on term premia. For the most prominent safe haven—nominal U.S. Treasury securities—this effect is evident in Figure 2 at the very end of the U.S. sample. After the failure of Lehman Brothers in September 2008, fears of credit and liquidity risks jumped, and many financial markets, including equity and currency markets, encountered intense selling pressure. The jump in risk aversion also helped create a heightened global demand for safe assets, which favored highly liquid nominal Treasury securities and led to a sharp decline in their yields (Christensen et al., 2010). However, recessions don't always coincide with financial crises and, even when they do, the bad economic conditions typically outlive the financial market dislocations, so it seems very unlikely that any procyclical safe-haven flows could completely offset the true countercyclical pattern of risk.

To provide a detailed assessment of the cyclical behavior of our term premia, we estimate univariate panel regressions, using several different cyclical indicators, and report the results in Table 4. We include the recession dummy used in Wright's study, real GDP growth (quarter-over-quarter), and the OECD's Composite Leading Indicator (CLI).[9] The CLI, which

[8]The recession dummy is not available for Norway. Figure 2 does not include confidence intervals since these would have the exact same width, at each point in time, as the ones in Figure 1. The reader should keep in mind that the estimation uncertainty around these term premia is sizeable.

[9]GDP growth and the CLI are obtained from http://stats.oecd.org.

is prominently used in policy discussions to time international business cycles, is particularly appealing for our application for two reasons: First, it indicates the business cycle in a consistent manner, while different trend growth rates make business cycle measurement across countries using growth rates problematic. Second, it is designed to provide early signals of turning points, and hence leads the business cycle. This is desirable for estimating the cyclical behavior of term premia, since these rise ahead of and early in recessions, and vice versa fall around the business cycle trough and early in expansions (Cochrane and Piazzesi, 2005).

The coefficients in Table 4 all have the expected signs—positive on the recession dummy, negative on the business cycle indicators—but only BC term premia display a statistically significant relationship, with significant coefficients on the recession dummy and the CLI. For the relationship with GDP growth, bias adjustment increases the magnitude of the coefficient, but it remains insignificant. Overall, these results are evidence of a pronounced countercyclical pattern for BC term premia, and the lack thereof for OLS term premia. The estimates suggest that on average, term premia in far-ahead forward rates are 75 basis points higher in recessions than in expansions.

To understand the panel regression results, it is instructive to consider each country separately by running individual time series regressions. We estimate univariate regressions of the term premium estimates on the cyclical indicators, and report the slope coefficients, p-values based on Newey-West standard errors (with four lags) and regression R^2 in Table 5. For most countries, bias correction makes the estimated term premia distinctly countercyclical, leading to the right sign or increasing the magnitudes of the coefficients, and raising significance levels and explanatory power of the regressions. This effect is particularly strong for the U.S., Germany, the U.K., Canada, and Switzerland. Focusing on the CLI as the cyclical indicator, OLS term premia are acyclical for all countries, whereas the BC term premia are significantly countercyclical (at the 10% level) for four of the ten countries.

The panel and country-level regressions are used to estimate the relationship between highly persistent variables. This is problematic from a statistical perspective since it can lead to a spurious regression problem. The panel regressions account for this by using bootstrapped p-values, and the country-level regressions employ Newey-West standard errors. While this does not completely eliminate the persistent regressor problem, it goes a long way to make the inference results more reliable.

The conclusion of this analysis is that there is a pronounced countercyclical pattern of term premia in long-term interest rates across developed countries. To clearly reveal this pattern in DTSM-based term premium estimates, it is important to address the issue of small-sample bias.

5 Conclusion

The small-sample bias in estimates of commonly used affine DTSMs, as in Wright's paper, has important economic consequences. Conventional estimates lead to term premia that are potentially unreliable and misleading. For many of the countries considered by Wright, term premia simply mirror the trend and cyclical behavior of long-term interest rates, since the bias makes model-based short rate forecasts and hence the expectations component artificially stable at long horizons. The estimated term premia do not show any countercyclical behavior. Neither a recession dummy, as considered by Wright, nor the cyclical indicators of economic activity that we consider in addition, has significant regression coefficients of the right sign.

If small-sample bias is accounted for in term premium estimation, some economic conclusions change. The estimated expectation component in far-ahead forward rates displays economically meaningful variation. Hence, the resulting term premia do not simply mirror the movements of long-term interest rates alone, but instead become more stable and mainly vary over the business cycle, tending to be high during recessions and low in expansions. In panel regressions similar to the ones of Wright but using our estimated term premia, business cycle indicators have significant coefficients of the right sign, which they do not for conventional term premia. We also drill down to the country level and show for each country the relationship between term premia and the macroeconomic variables of interest. While OLS term premia generally have no significant relationship with the business cycle, the opposite is true for BC term premia. Their pronounced countercyclical behavior is a pattern that holds across countries. Overall, bias-corrected term premium estimates appear to be more consistent with conventional macroeconomic wisdom and our economic intuition about risk premia. However, given the short data samples employed by Wright, estimation uncertainty, reflected by wide confidence intervals around our estimates, constitutes an important caveat to his results and ours.

References

Bansal, R. and A. Yaron, "Risks for the long run: A potential resolution of asset pricing puzzles," *Journal of Finance*, 2005, *59* (4), 1481–1509.

Bauer, Michael D., "Bayesian Estimation of Dynamic Term Structure Models under Restrictions on Risk Pricing," Working Paper 2011-03, Federal Reserve Bank of San Francisco November 2011.

— , **Glenn D. Rudebusch, and Jing Cynthia Wu**, "Correcting Estimation Bias in Dynamic Term Structure Models," *Journal of Business and Economic Statistics*, July 2012, *30* (3), 454–467.

Bekaert, G., R.J. Hodrick, and D.A. Marshall, "On biases in tests of the expectations hypothesis of the term structure of interest rates," *Journal of Financial Economics*, 1997, *44* (3), 309–348.

Campbell, John Y. and John H. Cochrane, "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy*, 1999, *107* (2), 205–251.

Christensen, Jens H. E. and Glenn D. Rudebusch, "The Response of Interest Rates to US and UK Quantitative Easing," *Economic Journal*, November 2012, *122* (564), F385–F414.

— , **Jose A. Lopez, and Glenn D. Rudebusch**, "Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields," *Journal of Money, Credit and Banking*, 2010, *42*, 143–178.

Cochrane, John H. and Monika Piazzesi, "Bond Risk Premia," *American Economic Review*, March 2005, *95* (1), 138–160.

Harvey, Campbell R., "Time-varying conditional covariances in tests of asset pricing models," *Journal of Financial Economics*, 1989, *24* (2), 289–317.

Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu, "A New Perspective on Gaussian Dynamic Term Structure Models," *Review of Financial Studies*, 2011, *24* (3), 926–970.

— , **Marcel Priebsch, and Kenneth J. Singleton**, "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," working paper October 2012.

Kilian, Lutz, "Small-sample confidence intervals for impulse response functions," *Review of Economics and Statistics*, 1998, *80* (2), 218–230.

Kim, Don H. and Athanasios Orphanides, "Term Structure Estimation with Survey Data on Interest Rate Forecasts," *Journal of Financial and Quantitative Analysis*, February 2012, *47* (1), 241–272.

— **and Jonathan H. Wright**, "An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates," Finance and Economics Discussion Series 2005-33, Federal Reserve Board of Governors 2005.

Kozicki, S. and P.A. Tinsley, "Shifting endpoints in the term structure of interest rates," *Journal of Monetary Economics*, 2001, *47* (3), 613–652.

Lustig, H. and A. Verdelhan, "The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk," *American Economic Review*, 2007, *97* (1), 89–117.

Lustig, Hanno N., Nikolai L. Roussanov, and Adrien Verdelhan, "Countercyclical Currency Risk Premia," NBER Working Papers 16427, National Bureau of Economic Research September 2010.

Stambaugh, Robert F., "Predictive regressions," *Journal of Financial Economics*, December 1999, *54* (3), 375–421.

Wachter, J.A., "A consumption-based model of the term structure of interest rates," *Journal of Financial Economics*, 2006, *79* (2), 365–399.

Wright, Jonathan H., "Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset," *American Economic Review*, 2011, *101* (4), 1514–1534.

Table 1: Summary statistics

			max. eigenv.		IRF(5y)		volatility		
		RMSFE	$\mathbb{P}$	$\mathbb{Q}$	$\mathbb{P}$	$\mathbb{Q}$	Forw.	Exp.	FTP
U.S.	OLS	0.050	0.9310	0.9751	0.18	0.61	1.26	0.13	1.14
	BC	0.052	0.9995	0.9738	0.99	0.61	1.26	1.45	0.60
Japan	OLS	0.025	0.9386	0.9874	0.32	0.85	1.77	0.34	1.47
	BC	0.025	0.9816	0.9872	0.78	0.83	1.77	1.27	1.09
Germany	OLS	0.043	0.9704	0.9755	0.57	0.66	1.48	1.20	0.34
	BC	0.043	0.9979	0.9737	0.87	0.66	1.48	2.24	0.82
U.K.	OLS	0.074	0.9585	0.9923	0.29	0.86	2.12	0.54	1.70
	BC	0.074	0.9986	0.9930	0.61	0.86	2.12	1.73	1.34
Canada	OLS	0.043	0.9744	0.9974	0.51	0.82	1.98	0.92	1.12
	BC	0.043	0.9993	0.9973	0.84	0.82	1.98	1.93	0.62
Norway	OLS	0.031	0.7785	0.9990	-0.01	0.30	0.59	0.01	0.59
	BC	0.031	0.8598	0.9995	0.13	0.30	0.59	0.13	0.52
Sweden	OLS	0.041	0.9328	0.9993	0.23	0.98	2.18	0.34	1.85
	BC	0.041	0.9964	0.9991	0.91	0.98	2.18	2.21	0.62
Switzerland	OLS	0.055	0.8784	0.9891	0.13	0.55	1.14	0.07	1.12
	BC	0.055	0.9773	0.9892	0.66	0.55	1.14	1.02	0.83
Australia	OLS	0.040	0.9304	0.9995	0.20	0.99	2.25	0.19	2.07
	BC	0.040	0.9879	0.9997	0.67	0.99	2.26	1.00	1.27
New Zealand	OLS	0.022	0.8972	0.9994	0.04	0.80	1.55	0.07	1.50
	BC	0.022	0.9753	0.9994	0.54	0.80	1.55	0.75	0.84

Summary statistics for OLS and BC model estimates: root-mean-square fitting errors (RMSFE) in annualized percentage points, persistence of estimated VAR system under real-world ($\mathbb{P}$) and risk-neutral ($\mathbb{Q}$) measure – largest eigenvalue of mean-reversion matrix and impulse-response function at horizon of five years of level factor to level shock – and volatility of fitted forward rates, expectations component, and forward term premium (FTP).

Table 2: Changes from 1990-91 to 2008-09

	Survey exp.		Forward				
	Inflation	Growth	rate		Fitted	Exp.	FTP
U.S.	-170	23	-391	OLS	-390	-41	-349
				BC	-390	-438	48
Japan	-130	-237	-470	OLS	-474	-121	-353
				BC	-473	-446	-27
Germany	-120	-137	-391	OLS	-408	-367	-41
				BC	-407	-688	281
U.K.	-117	-13	-557	OLS	-566	-73	-493
				BC	-568	-212	-356
Canada	-130	-53	-605	OLS	-629	-268	-360
				BC	-628	-521	-107
Australia	-237	2	-630	OLS	-621	-53	-567
				BC	-622	-278	-344

Changes from the early part of the sample (1990-91) to the late part of the sample (2008-09), in basis points, of inflation expectations, real GDP growth expectations (both 5-10 years hence against survey date), forward rates, and model-based expectations and term premium components of forward rates. For the survey expectations the change is the difference between the mean across the first three available surveys, in April 2008, October 2008, and April 2009, and the mean across the last three surveys, October 1990, April 1991, and October 1991, of the expectations of five-to-ten-year ahead CPI inflation and real GDP growth from the Consensus forecasts. For forward rates, short rate expectations, and forward term premia, the change is the difference between the mean of observations from Q1-2008 to Q1-2009, and the mean of the observations from Q3-1990 to Q3-1991.

Table 3: Panel regressions – uncertainty

Regressor						
Dependent variable: OLS term premium estimate						
PERM-UCSV	6.92					6.28
	(0.00)					(0.00)
Survey dispersion		5.36		4.90	5.21	
(Inflation)		(0.00)		(0.00)	(0.00)	
Survey dispersion			3.09	1.34	1.56	2.52
(Growth)			(0.01)	(0.23)	(0.11)	(0.00)
Recession dummy					-0.25	-0.39
					(0.31)	(0.01)
Dependent variable: BC term premium estimate						
PERM-UCSV	3.36					2.63
	(0.00)					(0.00)
Survey dispersion		2.33		2.02	1.97	
(Inflation)		(0.02)		(0.05)	(0.04)	
Survey dispersion			1.62	0.90	0.61	0.95
(Growth)			(0.00)	(0.09)	(0.20)	(0.03)
Recession dummy					0.37	0.30
					(0.06)	(0.00)

Slope coefficients for fixed effects panel regressions of OLS and bias-corrected (BC) term premium estimates on macroeconomic variables, with bootstrap p-values in parentheses. For details on explanatory variables and estimation methodology, see Wright (2011).

Table 4: Panel regressions – cyclicality

Regressor	OLS	BC
Recession dummy	0.80	0.75
	(0.25)	(0.00)
GDP (QoQ)	-0.06	-0.20
	(0.94)	(0.22)
CLI	-0.06	-0.08
	(0.59)	(0.02)

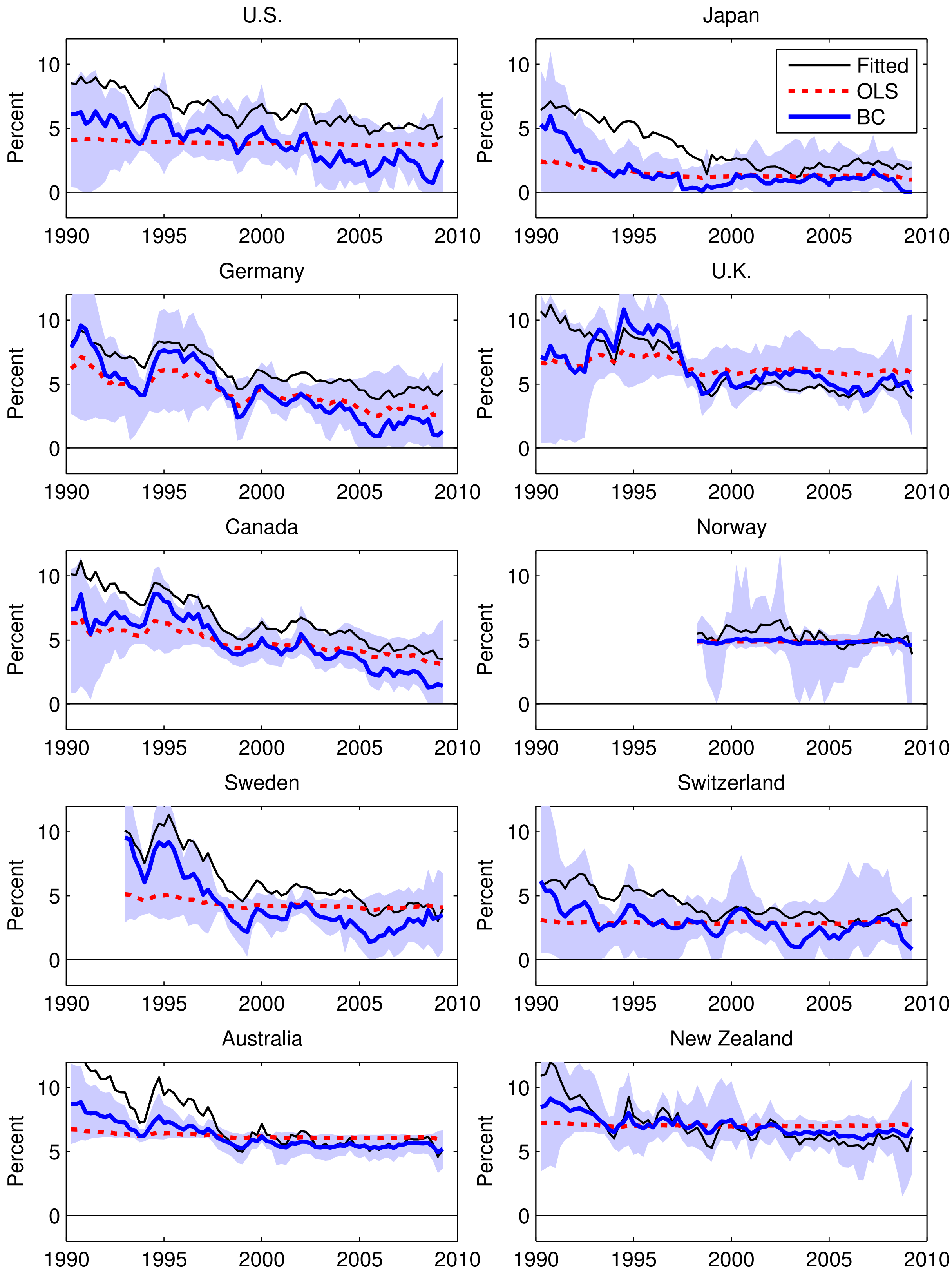
Slope coefficients for fixed effects panel regressions of OLS and BC term premium estimates on business cycle indicators, with bootstrap p-values in parentheses. The estimation methodology is the same as in Table 3 and in Wright (2011). Additional explanatory variables are quarter-over-quarter (QoQ) real GDP growth and the OECDs CLI measure.

Table 5: Time series regressions

		Rec. dummy		GDP (QoQ)		CLI	
		OLS	BC	OLS	BC	OLS	BC
U.S.	Coef.	-0.20	0.72	0.35	-0.27	-0.02	-0.03
	p	(0.80)	(0.01)	(0.23)	(0.02)	(0.84)	(0.32)
	R^2	0.4%	16.6%	4.6%	10.1%	0.3%	2.3%
Japan	Coef.	-0.60	0.11	0.06	-0.13	-0.01	-0.05
	p	(0.30)	(0.80)	(0.51)	(0.20)	(0.92)	(0.33)
	R^2	4.1%	0.3%	0.4%	1.8%	0.0%	3.8%
Germany	Coef.	0.19	0.31	-0.07	-0.20	-0.01	-0.08
	p	(0.16)	(0.29)	(0.16)	(0.07)	(0.42)	(0.02)
	R^2	7.3%	3.3%	3.4%	6.9%	2.5%	13.3%
U.K.	Coef.	2.14	2.38	-0.26	-0.57	-0.17	-0.25
	p	(0.08)	(0.01)	(0.66)	(0.22)	(0.50)	(0.19)
	R^2	21.0%	42.1%	1.2%	9.5%	3.5%	12.2%
Canada	Coef.	1.01	1.17	-0.35	-0.52	-0.08	-0.10
	p	(0.25)	(0.00)	(0.40)	(0.00)	(0.42)	(0.01)
	R^2	11.5%	51.4%	4.6%	34.4%	4.1%	20.8%
Norway	Coef.	n.a.	n.a.	0.09	0.06	-0.04	-0.07
	p	n.a.	n.a.	(0.29)	(0.38)	(0.38)	(0.08)
	R^2	n.a.	n.a.	2.3%	1.4%	2.5%	9.0%
Sweden	Coef.	0.01	-1.45	0.40	0.34	-0.02	0.04
	p	(0.99)	(0.00)	(0.11)	(0.00)	(0.82)	(0.28)
	R^2	0.0%	52.5%	4.5%	30.3%	0.2%	6.3%
Switzerland	Coef.	1.37	0.89	-0.46	-0.72	-0.07	-0.17
	p	(0.00)	(0.00)	(0.19)	(0.00)	(0.35)	(0.00)
	R^2	35.7%	27.3%	7.0%	32.1%	4.1%	48.8%
Australia	Coef.	4.19	2.57	-0.63	-0.40	-0.38	-0.24
	p	(0.00)	(0.00)	(0.26)	(0.23)	(0.15)	(0.13)
	R^2	29.9%	30.0%	4.2%	4.6%	13.5%	15.0%
New Zealand	Coef.	1.25	0.75	-0.14	-0.06	-0.14	-0.07
	p	(0.31)	(0.30)	(0.60)	(0.70)	(0.37)	(0.44)
	R^2	10.6%	12.1%	0.7%	0.4%	4.5%	3.5%

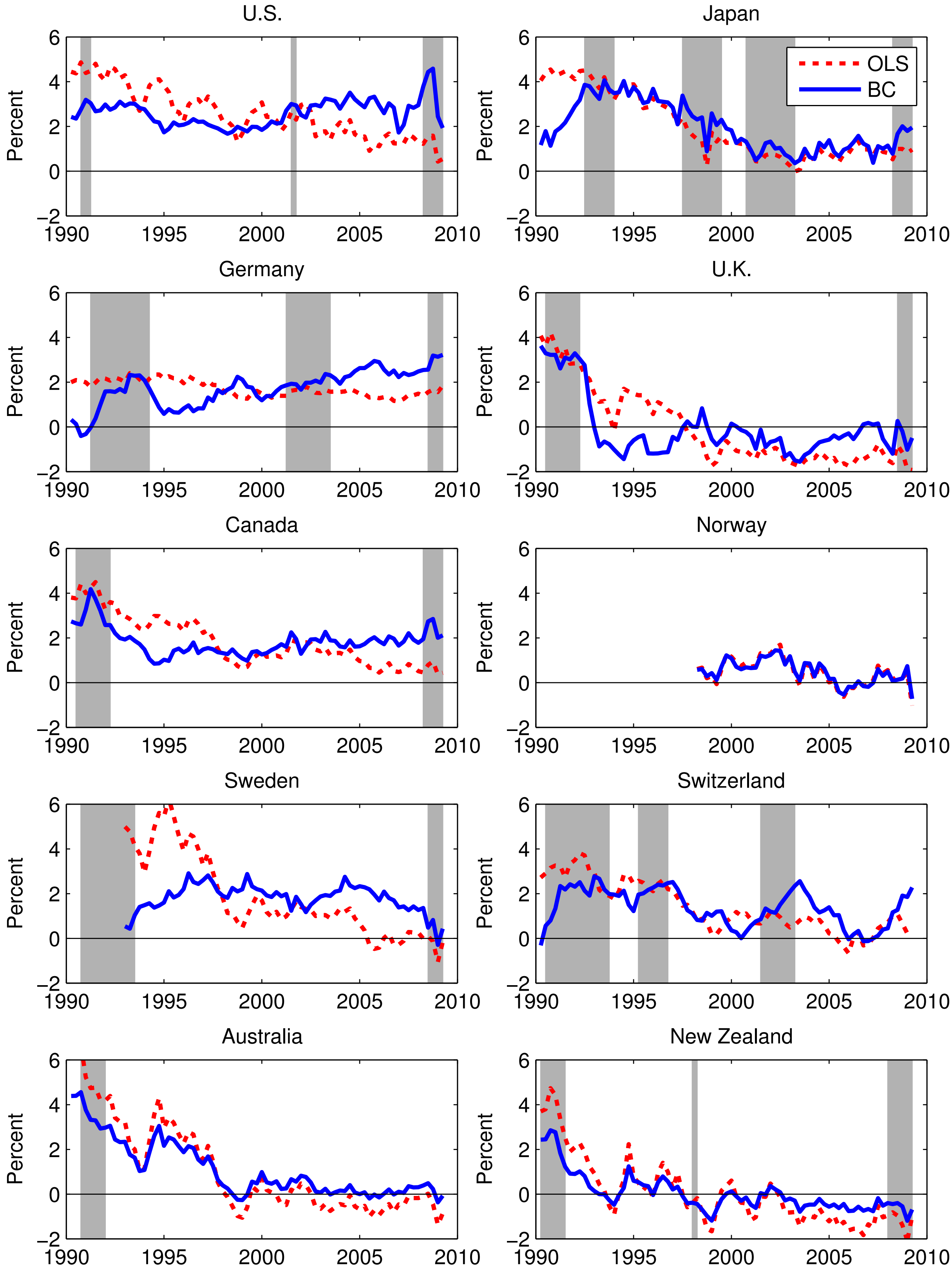
Results for univariate time series regressions of the five-by-five-year forward term premium, based on OLS and BC model estimates, on cyclical indicators. The table shows slope coefficients, p-values based on Newey-West standard errors (using four lags), and regression R^2.

Figure 1: Forward rates and risk-neutral rates



Fitted five-by-five-year forward rates and risk-neutral forward rates, based on OLS and BC model estimates, with bootstrapped 90%-confidence intervals for BC risk-neutral rates.

Figure 2: Term premia



OLS and BC term premia in five-by-five-year forward rates. Shaded areas indicate quarters where the recession dummy is equal to one.