NBER WORKING PAPER SERIES

THE NATURAL RATE PUZZLE:
GLOBAL MACRO TRENDS AND THE MARKET-IMPLIED R*

Josh Davis
Cristian Fuenzalida
Alan M. Taylor

Working Paper 26560
http://www.nber.org/papers/w26560

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2019, Revised January 2021

We thank Ben Bernanke, Joachim Fels, and participants in a seminar at the Federal Reserve Bank of San Francisco for helpful discussions. All errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w26560.ack

The Natural Rate Puzzle: Global Macro Trends and the Market-Implied r*
Josh Davis, Cristian Fuenzalida, and Alan M. Taylor
NBER Working Paper No. 26560
December 2019, Revised January 2021
JEL No. C13,C32,E43,E44,E47,G12

ABSTRACT

Benchmark finance models deliver estimates of bond risk premia based on components of Treasury bond yields. Benchmark macroeconomic models deliver estimates of the natural rate of interest based on growth, inflation, and other macro factors. But estimates of the natural rate implied by the former are wildly inconsistent with those of the latter; and estimates of risk premia implied by the latter are wildly inconsistent with those of the former. This is the natural rate puzzle, and we show that it applies not only in the United States but also across several advanced economies. A unified model should not fail such consistency tests. We estimate a unified macro-finance model with long-run trend factors which delivers paths for a market-implied natural rate r* consistent with inflation expectations π* and bond risk premia. These paths are plausible and our factors improve the explanatory power of yield and return regressions. Trading strategies based on signals incorporating both r* and π* trends outperform both yield-only strategies like level and slope and strategies which only add trend inflation. The estimates from our unified model satisfy consistency and deliver a resolution to the puzzle. They show that most of the variation in yields has come from shifts in r* and π*, not from bond risk premia. Our market-implied natural rate differs from consensus estimates, and is typically lower, intensifying concerns about secular stagnation and proximity to the effective lower-bound on monetary policy in advanced economies.

Josh Davis
PIMCO
650 Newport Center Drive
Newport Beach CA 92660
josh.davis@pimco.com

Cristian Fuenzalida
PIMCO
650 Newport Center Drive
Newport Beach, CA 92660
cristian.fuenzalida@pimco.com

Alan M. Taylor
Department of Economics and
Graduate School of Management
University of California
One Shields Ave
Davis, CA 95616-8578
and CEPR
and also NBER
amtaylor@ucdavis.edu

1. Introduction

Like two trains running on different railroads, studies of the natural rate and bond risk premia in the macroeconomic and finance literatures have tended to follow their own line even if ostensibly headed for the same place. The destination is clear and important: rates of interest, across all maturities, matter for saving, investment, capital allocation, economic growth, and monetary policy. But passengers on each route see different landscapes: on the macro track, a panoramic debate about time-series patterns with slow-moving trends in natural rate Wicksellian models and their fundamental drivers (e.g, Holston, Laubach, and Williams, 2017; Rachel and Summers, 2019; Jordà and Taylor, 2019); on the asset pricing track, a more tightly-framed look at cross-section patterns with a no-arbitrage pricing approach using factors typically built from yields (e.g, Litterman and Scheinkman, 1991; Piazzesi, 2010; Adrian, Crump, and Moench, 2013).

But the two tracks converge and a collision has been unavoidable. Workhorse finance models of bond risk premia and inflation expectations generate a path for the natural rate dramatically at odds with the macro literature. Equivalently, workhorse macro models of the natural rate and inflation expectations generate a path for bond risk premia equally at odds with the finance literature. We call this the *natural rate puzzle*. To get on the same track, the two approaches must be somehow shunted together. A consensus unified model should not fail these consistency tests and this is a first-order challenge for macro-finance research. We build on a long literature and make new headway. We explore the international aspect of this problem with newly-constructed data from the U.S. and other advanced economies and we advance a new empirical approach which disciplines estimates of the natural rate and risk premia with both financial market *and* macro information.

We first document the puzzle, for both the U.S. and other countries. For clarity, we do nothing analytically new here: we rely only on off-the-shelf models and data. The analysis revolves around three trend estimates. For the U.S., we construct an estimate of the bond risk premium following the canonical model (Adrian, Crump, and Moench, 2013) used by academic and financial professionals, and also by the Federal Reserve. We estimate inflation expectations following recent research incorporating trend inflation into models of bond yields and risk premia (Cieslak and Povala, 2015). And we construct an estimate of the natural rate following the seminal model in the macroeconomic literature (Laubach and Williams, 2003). We then use directly-observed long-dated forward rate data to show the contradiction. Using bond premia, inflation, and forwards, the implied natural rate is nearly flat over six decades, inconsistent with the rise and fall seen in macro estimates (with the implication that changes in the bond risk premium mostly explain long-yield changes). Conversely, using natural rates, inflation, and forwards, the implied bond risk premium is nearly flat, inconsistent with the rise and fall seen in finance estimates (with the implication that changes in natural rates mostly explain long-yield changes). Both models cannot be true. The puzzle is not an artifact of these particular estimates, and obtains using other well-respected estimates of U.S. bond risk premia, trend inflation, and the natural rate from multiple credible sources. The same puzzle also exists internationally in data we have newly compiled from five other advanced economies.

Why does the puzzle matter for macro and for finance? Because our understanding of recent history hinges on whether one or the other story is more accurate. One narrative (macro) is of steadily declining natural real rates from the 1980s to the present, arguably culminating in a global secular stagnation trap, with active debate over a wide range of causal factors including demography and aging, productivity growth, inequality, and safe-asset demand from emerging economies, and all the attendant problems (Caballero, Farhi, and Gourinchas, 2008; Summers, 2015; Carvalho, Ferrero, and Nechio, 2016; Holston, Laubach, and Williams, 2017; Rachel and Smith, 2017; Rachel and Summers, 2019). The pure finance bond-pricing model undercuts this macro story. The other narrative (finance) is of a sweeping rise and fall in risk premia, from sometime in the 1970s to the 1980s, the backwash of the Great Inflation episode, accounting for most of the trajectory of nominal rate with little or no movement in the natural rate (Kim and Wright, 2005; Wright, 2011; Adrian, Crump, and Moench, 2013; Bernanke, 2015) The pure macro bond-pricing model undercuts this finance story. Something has to give.

To tackle this conflict we set out a unified macro-finance model to ground the empirical work that follows. We build on the idea that term structure models should allow all nominal rates to include a stochastic trend, as seen in early work by Campbell and Shiller (1987) and developed further in the seminal paper of Kozicki and Tinsley (2001). We follow the key contribution of Cieslak and Povala (2015) and allow two trends in a nominal and a real factor, with yields and expected returns to bonds of different maturities derived under no-arbitrage constraints from a short-rate process linked to the two macroeconomic factors, r^* and π^*. The model crystallizes the uncontroversial view—among macroeconomists, at least—that nominal bond returns are not explained simply as a compensation for the compounded benchmark rate, as the failure of the expectations hypothesis shows. Rather, there must be extra compensation for macroeconomic risks linked to real factors and inflation (Ang and Piazzesi, 2003; Ludvigson and Ng, 2009; Cieslak and Povala, 2015).

Next, in the empirical core of the paper, we take the model to the data. Trend inflation is treated as an observable, as in prior work, but the unobservable natural rate is estimated from a unified state-space model with the Kalman filter. However, we make a unifying link to natural rate research, by also including a macroeconomic growth factor in a state equation, in addition to yield and return measurement equations, so our model utilizes information from both macro and financial market data.[1] We therefore refer to our r^* estimate as the *market-implied natural rate*.

We apply the model to the postwar data for the U.S. and five other advanced economies, an historical laboratory as large as any previously explored in the study of these questions as far as we know. We find strong support for the model. Dropping either trend variable significantly worsens

[1] A related paper is the contemporaneous, independent work by Bauer and Rudebusch (2019), which favors a modeling approach based on a single stochastic trend, a nominal natural rate factor i^*. Their estimation uses yield-based factors and restrictions in the model structure, but omits information on macroeconomic variables like growth in the state-space model (see their Appendix C). In contrast, as explained below, we enforce an $r^* = g + z$ equation; that is, we bring in GDP growth data, and so restrict the model by tying it into the canonical latent-variable natural rate formulation in the macro side of the literature. Each approach takes a stand, and these are two different ways to discipline model estimates.

the model fit: the baseline R^2 statistics are relatively high, but fit worsens one or both trends are removed, especially in return forecast regressions. Indeed, the macroeconomic factors subsume much of the relevant information needed to price bonds as compared with benchmark yield-only term-structure models, leaving only detrended yields to play a role, amplifying the insight of Cieslak and Povala (2015), but now for two trends and more countries. Finally, we explore the model's out-of-sample performance. Here, the model is applied recursively, now with one-sided filtering, for a cross-section portfolio of advanced economy bonds in recent years. The model with two macro trends r^* and π^* outperforms all its main rivals in out-of-sample fit.

The main contribution of this paper is a step toward a unified model which bridges the methodological divide and exploits fully all the information used in previous finance and macro approaches. Finance models of unobserved bond risk premia have utilized yield-based factors, macro models of the unobserved natural rate have utilized macroeconomic variables like growth. The two produce inconsistent results and we argue that a unified approach using both sets of information is necessary. To get there, our paper makes a number of specific points along the way, touching on questions that have emerged from distinct literatures. First, we document for many countries, over many decades, an important macro-finance puzzle which the separate paths of risk premia research and natural rate research have often skirted around. Second, to operationalize the model, we apply a one-step joint estimation strategy; though novel, and computationally more difficult, this should be preferred to approaches which draw natural rate and risk premia estimates from disparate models, which can lead to inconsistency. Third, we present estimates from a broader sample of six advanced economies, as this is not just a U.S. story and it allows us to address diverse global trends. Fourth, this matters, as inflation and natural rate factors follow quite distinct paths in other economies, and attract very different yield loadings. Fifth, behind that, our estimation rests on a new, comprehensive database of zero-coupon bond yield time series for the five non-U.S. economies, a valuable data contribution for future researchers in its own right. Sixth, our method produces improved predictions for bond yields and returns in the U.S. and international samples, including out of sample based on the R^2_{OS} statistic (Campbell and Thompson, 2008).

By the end, we are in a position to assay the natural rate puzzle, and we get a clear answer: across advanced economies, most of the long-term variation in yields in recent decades has come from shifts in the natural rate and inflation trend components, not from shifts in bond risk premia. The key takeaway is that macro and finance models can go their separate ways no longer.

2. The natural rate puzzle

The natural rate puzzle is the observation that standard finance models of bond risk premia and inflation expectations generate a path for the natural rate at odds with the macro literature.

To document this we use a general framework. In a wide class of standard affine asset pricing models, the term structures (of bond yields, prices, excess returns, and forwards) are affine functions of the model's vector of risk factors (F_t), which will be made precise in the next section. Then, if

$f_t^{(n,m)}$ is the horizon n, maturity m, forward interest rate at time t in the future, we can write

$$f_t^{(n,m)} = r_t^* + \pi_t^* + \Gamma^{(n,m)}(F_t)\,, \tag{1}$$

where r_t^* is the trend of the real natural rate, π_t^* is the trend of inflation, and $\Gamma^{(n,m)}(F_t)$ is a bond risk premium term, defined implicitly here, and explored in more detail below in a formal model.

This expression is quite intuitive, especially in the case when r_t^* and π_t^* follow processes which are unit root. Investors buying forward rates must be compensated by the sum of the trend real natural rate and trend inflation, plus a term that is by definition the bond risk premium. But the modeling challenge comes in the selection of the factor set F_t and other choices needed to operationalize the idea.

Suppose we naïvely take $\Gamma^{(n,m)}(F_t)$ from benchmark models in the finance literature where F_t is a set of yield factors, take r_t^* and π_t^* from benchmark macro models, and take $f_t^{(n,m)}$ from market data. Having constructed these four terms for multiple countries, we show that the above equation fails to hold. This section documents this fact across the advanced economies and the rest of the paper explores a hybrid macro-finance model which may offer a way out. As might be anticipated, Equation 1) offers only two escape routes. Given that the forward rate is an observed trending variable, and that the inflation trend π_t^* is not subject to large estimation error, or can be treated as quasi-observable, then either the trend in the unobserved natural rate r_t^* is mismeasured, or the trend in the unobserved bond risk premium $\Gamma^{(n,m)}(F_t)$ is mismeasured, or both. It turns out that, while both matter, we argue that mismeasurement of the risk premium has dominated in reality.

2.1. U.S. evidence

To see the puzzle, we take Equation 1 directly to the data. In Figure 1, Panel (a), the U.S. time-series estimates for each of the four terms are shown. We simply take these estimates from canonical models in the finance and macro literatures. The bond risk premium term Γ is constructed as in the baseline five-factor model of Adrian, Crump, and Moench (2013) [henceforth abbreviated ACM]; the inflation expectations term π^* as in Cieslak and Povala (2015) [CiP]; and the (one-sided) real natural rate term r^* as in Laubach and Williams (2003) [LW]. Finally, we have the 10-year, 10-year forward rate (f) which is directly-observed raw data taken from Bloomberg, with $n = m = 120$ months here.

The first version of the consistency test rearranges Equation 1 to obtain a formula for the real natural rate $r_t^* = f_t^{(n,m)} - \pi_t^* - \Gamma^{(n,m)}(F_t)$, and Panel (b) plots both sides of this expression using the above data sources: the left-hand side is taken directly from an LW model and the right-hand side is the implied value using an ACM model. The equality is violated, and the disparity is often quite large. The ACM-implied r^* does not match the LW r^*. The ACM series starts around +2% in the 1960, displays a sharp decline to a level below −2% during the Great Inflation period of the 1970s, returns to +2% in the 1990s, drops to near zero after the financial crisis, and then shows a consistent increase after 2013 to a level close to 2% in 2019. In contrast, the familiar LW estimate of r^* has fallen gradually from a +4% level in the 1960s and 1970s, with the sharpest decline occurring after the mid-2000s, and since 2010 it has sat in the 0.5%–1.0% range, and never turned negative. The difference between the two series, before the last decade, is often large, between 100 and 600 basis points (bps), with the LW r^* much higher than the ACM r^*, on average. Around 2012 the two series intersected and then the difference inverted to about -100 bps in the other direction.

A second, equivalent, version of the test is shown in Panel (c). We rearrange again to obtain a formula for the bond risk premium $\Gamma^{(n,m)}(F_t) = f_t^{(n,m)} - r_t^* - \pi_t^*$, and Panel (b) plots both sides of this expression using the aforementioned data sources. Now the left-hand side is direct from an ACM model and the right-hand side is the implied value using an LW model. This equality is, of course, also violated, and the same large disparity is seen. The ACM bond risk premium starts near zero in the 1960s, rises sharply in the Great Inflation period of the 1970s to about 6%, then gradually falls back, reaching zero again in the mid-2010s. The LW bond risk premium behaves very differently, and is almost flat by comparison. It actually starts at a negative level in the 1960s, rises much later, but only to a modest 2% by the early 1980s, then declines by a small amount up to the mid 2000s. After that the two series cross, with LW signaling a small positive bond risk premium, but ACM turning negative.

The puzzle is vividly apparent in these charts. Persistent inconsistencies of several hundred basis points are quantitatively just too large to ignore. Both approaches cannot be simultaneously right. A substantial contradiction thus emerges from the heart of benchmark macro and finance models once they are studied in unison. The rest of this paper is devoted to building theory and empirics to help resolve the puzzle.

2.2. Alternative trend measures

As a robustness check, Figure 2 examines whether the existence of the puzzle for the U.S. is sensitive to the source data used. For a variety of widely used and respected sources we compute the discrepancy in Equation 1 as $discrepancy = r^* - f + \pi^* + \Gamma$, and plot the series over time.

The same forward rate data f from Bloomberg are used in all cases. The sources of the other three series rotate through all possible combinations, with the sources are abbreviated as follows:

- Natural rate estimates r^*: Laubach and Williams (2003) [LW]; Holston, Laubach, and Williams (2017) [HLW]; Del Negro, Giannone, Giannoni, and Tambalotti (2017) [DGGT]; and Lubik and Matthes (2015) [LM].
- Inflation estimates π^*: Cieslak and Povala (2015) [CiP]; the University of Michigan Inflation Expectations from FRED [MI]; the Survey of Professional Forecasters from the Federal Reserve Bank of Philadelphia [SPF]; and the TIPS 10-Year Breakeven Inflation Rate from FRED [TIPS].
- Bond risk premium estimates Γ: Adrian, Crump, and Moench (2013), 5-factor model [ACM5]; the same authors' 3-factor model [ACM3]; and Kim and Wright (2005), 3-factor model [KW].

Figure 1: *The natural rate puzzle in U.S. data*

This figure displays market data (f) and existing trend data (other variables) based on other studies in Panel (a), and then displays the puzzle in the form of the difference between existing trend data and implied data in Panels (b) and (c). The presentation is based on Equation 1, which we can rewrite in simplified form, omitting subscripts and expectations, and taking them as understood, with the notation $f = r^* + \pi + \Gamma$. The puzzle is that existing benchmark estimates violate this equation.

In Panel (a), the four terms are shown: the bond risk premium Γ from Adrian, Crump, and Moench (2013); inflation expectations π from Cieslak and Povala (2015); and the real natural rate r^* from Laubach and Williams (2003). We also show the 10-year, 10-year forward rate (f) from from Bloomberg. The sample period is June 1961–May 2019. In Panel (b), we compare the real natural rate r^* from Laubach and Williams (2003) to that implied by $r^* = f - \pi^* - \Gamma$. There is a large difference between these two series. In Panel (c), we compare the bond risk premium Γ from Adrian, Crump, and Moench (2013) to that implied by $\Gamma = f - r^* - \pi^*$. There is the same large difference between these two series.

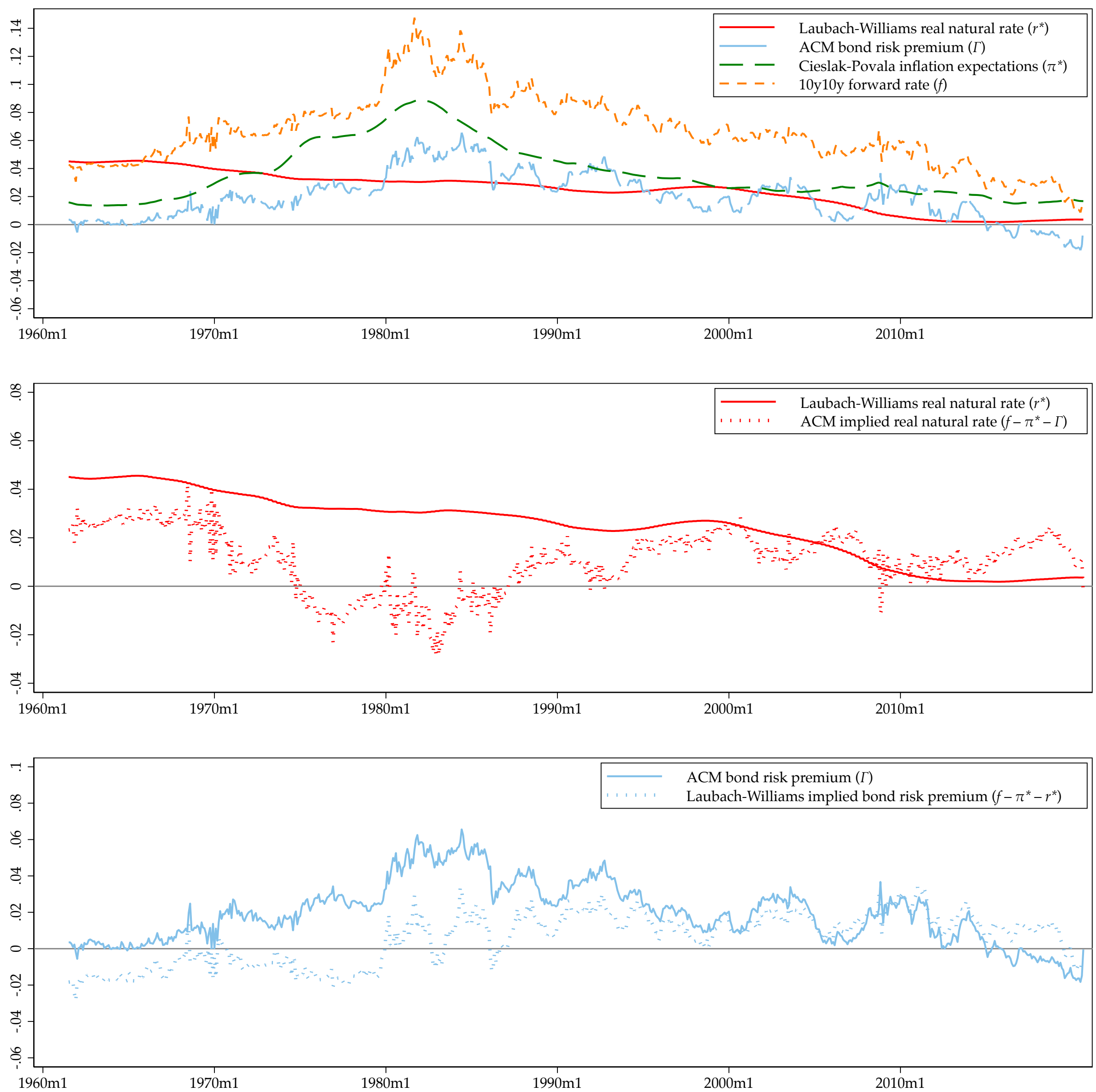
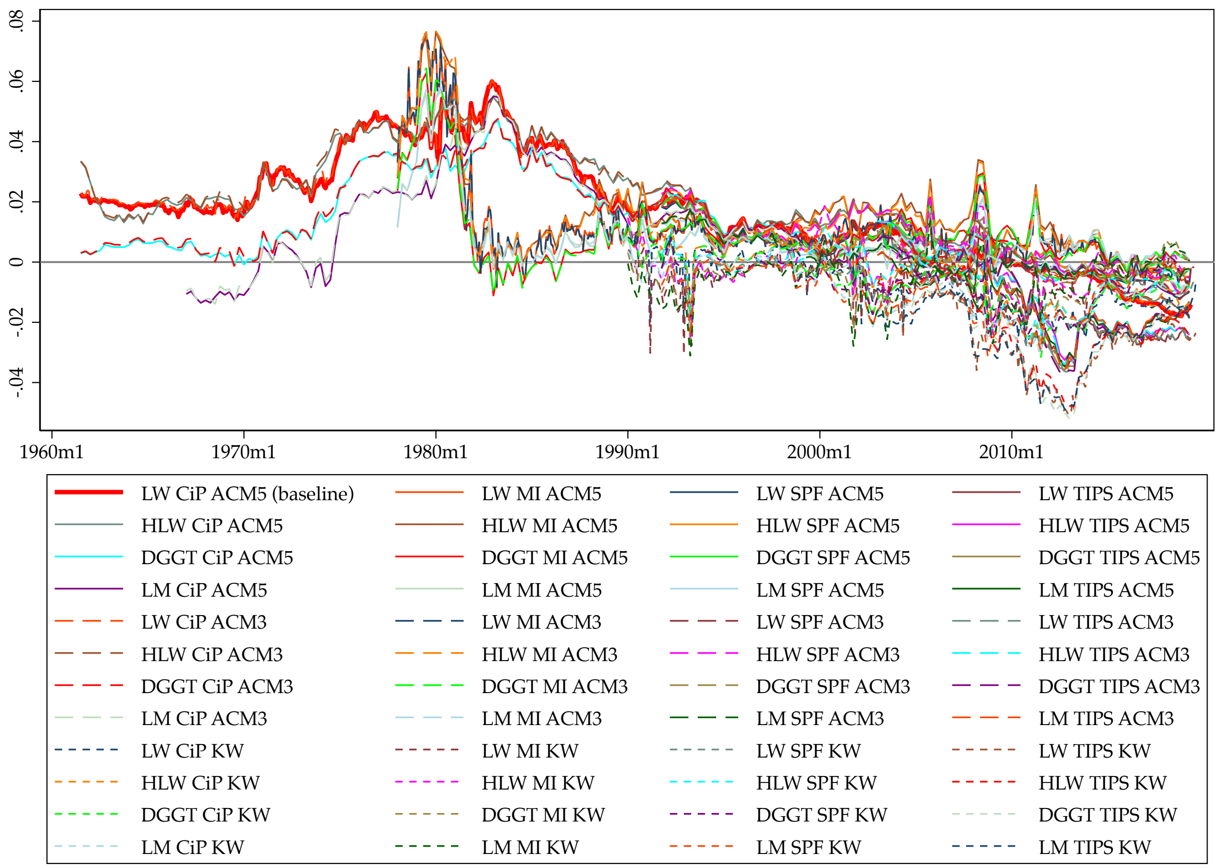


Figure 2: *The natural rate puzzle in U.S. data using alternative trend measures*

This chart displays the discrepancy between implied and existing trend data for the natural rate. The presentation is based on Equation 1 and the series computed is $discrepancy = r^* - f + \pi^* + \Gamma$. The puzzle is that this term is not zero. See text.



Note that because quite a few of these series (e.g., TIPS, KW) are only available for a shorter span of recent years, full-sample comparisons across all trend estimates are not always possible.

The figure reveals that the natural rate puzzle is a quite robust phenomenon in recent U.S. data. A discrepancy arises in all cases. It is often more than 100 bps, and at certain times it exceeds 500 bps. It is present in a wide variety of trend estimates currently used in the macro-finance literatures. The figure shows that, as in the baseline variant above, the extent of the puzzle varies from year to year, and over decades. Most series combinations make errors in one direction, but a few go the other way. The discrepancies are large in the 1970s, and often surge to their highest levels around 1980. The discrepancies are smaller by the late 1990s and early 2000s, but they open up again for some series, in the opposite direction to almost -400 bps, after the global financial crisis.

Table 1: *The natural rate puzzle in international data*

This table applies the approach of Figure 1, Panel (b), extended to a sample of 6 advanced economies, showing sample means. We compare the real natural rate r^* from an LW-type estimation to that implied by the risk premium from an ACM type estimation and inflation expectations from a CiP type estimation plus the forward rate. There is a large difference between these two series, given by $discrepancy = r^* - f + \pi^* + \Gamma$.

Mean (bps)	(1) U.S.	(2) Japan	(3) Germany	(4) U.K.	(5) Canada	(6) Australia
LW r^*	303	104	175	215	236	219
ACM implied r^*	78	−129	−26	−195	−48	−30
Difference	225	232	201	410	284	249
Absolute difference	241	233	211	410	284	249
Observations	695	411	560	472	400	321

2.3. International evidence

We also sought evidence for or against the natural rate puzzle in 5 other advanced economies: Japan, Germany, the U.K., Canada, and Australia. Table 1 and Figure 3 present these findings.

Again, we compare the real natural rate r^* from an LW-type estimation to that implied by the risk premium from an ACM-type estimation, inflation expectations from a CiP-type estimation, and the forward rate. For the LW-type natural rate estimates we use LW itself for the U.S. as above, Holston, Laubach, and Williams (2017) (one-sided estimates) for the Germany, U.K., Canada, Okazaki and Sudo (2018) (two-sided estimates) for Japan, and McCririck and Rees (2017) (two-sided estimates) for Australia. We then replicate the ACM and CiP methodologies and construct forward rates from zero-coupon bonds, as described later in this paper, and compute the discrepancy for all the countries to complete the analysis.

Table 1 shows the mean level of each natural rate estimate, from LW and implied by ACM, along with the mean discrepancy, and the mean absolute discrepancy. The mean absolute discrepancy is 241 bps for the U.S., reaches a maximum of 410 bps for the U.K., and a minimum of 211 bps for Japan. The mean absolute discrepancy is in the range 200–300 bps in all cases.

So the discrepancy can be visualized over time, Figure 3 presents the time-series data for each natural rate estimate. The U.S. pattern is fairly typical: the LW estimates lie well above the ACM implied estimates, and the latter often dips implausibly far into negative territory. In general, the paths are quite far apart and they only get closer, and in rare cases cross, near the end of the sample.

In short, the natural rate puzzle is not simply a U.S. puzzle. It applies to many advanced economies, suggesting a deeper and more general pattern posing problems for standard models.

Figure 3: *The natural rate puzzle in international data*

These charts apply the approach of Figure 1, Panel (b), extended to a sample of 6 advanced economies. We compare the real natural rate r^* from an LW-type estimation to that implied by the risk premium from an ACM type estimation, inflation expectations from a CiP-type estimation, and the forward rate, $f - \pi^* - \Gamma$. There is a large difference between these two series.

3. A term-structure model with two trends

For comparability, we use the model setup of Cieslak and Povala (2015) which features two trends for inflation and the real rate, building on the earlier insights of Kozicki and Tinsley (2001). At time t, we denote the nominal yield on an n-period Treasury bond by $y_t^{(n)}$, trend inflation by π_t, and the trend real natural rate by r_t. (Stars are dropped in this section for clarity and consistency with the prior work.) Nominal yields across all maturities are driven by the two trends and other factors contained in a price-of-risk factor x_t vector, so the full set of factors is $F_t = (\pi_t, r_t, x_t)^\top$.

The core of the model is the specification of the short-rate process and the stochastic discount factor, from which all other pricing relationships follow. The short-rate process is assumed to depend on the factors, which in turn follow independent AR(1) processes, with

$$y_t^{(1)} = \delta_0 + \delta_\pi \pi_t + \delta_r r_t \,, \tag{2}$$

$$r_t = \mu_r + \phi_r r_{t-1} + \sigma_r \epsilon_t^r \,, \tag{3}$$

$$\pi_t = \mu_\pi + \phi_\pi \pi_{t-1} + \sigma_\pi \epsilon_t^\pi \,, \tag{4}$$

where $\delta_\pi > 0, \delta_r > 0$, with $\delta_x = 0$, as shown, and $\epsilon_t^\pi, \epsilon_t^r$ are standard normal, i.i.d.

Concerning Equation 2, a natural benchmark is $\delta_0 = 1, \delta_\pi = \delta_r = 1$, i.e., the Fisher equation, and r_t is the ex-ante real rate. Alternatively, $\delta_\pi > 1, \delta_r < 1$ might reflect a Taylor rule, where the natural rate is dominated by growth shocks at high frequency. Concerning Equation 3 and Equation 4, it is well known that inflation follows a process that is unit root or very close, so we expect $\phi_\pi < 1$ but close to unity. Estimates of the natural rate also tend to be highly persistent, with $\phi_r < 1$ and somewhat close to unity. We find this to be the case in our estimates for all countries.

The price-of-risk factor is assumed to follow its own AR(1) process with i.i.d. normal shocks,

$$x_t = \mu_x + \phi_x x_{t-1} + \sigma_x \epsilon_t^x \,. \tag{5}$$

The model economy is then compactly described by the equations

$$F_t = \mu + \Phi F_{t-1} + \Sigma \epsilon_t \,, \tag{6}$$

$$y_t^{(1)} = \delta_0 + \delta_1^\top F_t \,, \tag{7}$$

with Φ and Σ diagonal, $\delta_1 = (\delta_\pi, \delta_r, 0)^\top$, and $\epsilon_t = (\epsilon_t^\pi, \epsilon_t^r, \epsilon_t^x,)^\top$.

We assume the log nominal stochastic discount factor is exponentially affine in the risk factors,

$$m_{t+1} = -y_t^{(1)} - \frac{1}{2} \Lambda_t^\top \Lambda_t - \Lambda_t^\top \epsilon_{t+1} \,, \tag{8}$$

where Λ_t is the compensation for risk of shock ϵ_{t+1}, with $\Lambda_t = \Sigma^{-1}(\lambda_0 + \Lambda_1 F_t)$.

We need more structure to make progress. In Cieslak and Povala (2015), x_t is taken to be a single yield-based factor, and the loadings in Λ_t are assumed to take the following form

$$\lambda_0 = \begin{pmatrix} \lambda_{0r} \\ \lambda_{0\pi} \\ 0 \end{pmatrix}, \qquad \Lambda_1 = \begin{pmatrix} 0 & 0 & \lambda_{\pi x} \\ 0 & 0 & \lambda_{rx} \\ 0 & 0 & 0 \end{pmatrix}. \tag{9}$$

This baseline setup is motivated by the Cochrane and Piazzesi (2005) finding that a single-factor based on a combination of yields can explain bond pricing quite well, but the x could be expanded to a vector to include widely used three-factor yield models (Nelson and Siegel, 1987; Litterman and Scheinkman, 1991) or even five-factor yield models (Adrian, Crump, and Moench, 2015).

The model can then be solved as a set of affine equations for bond prices, yields, excess returns, and forwards in terms of the factors:

$$y_t^{(n)} = A_n + B_n^\top F_t\,, \tag{10}$$

$$p_t^{(n)} = \mathcal{A}_n + \mathcal{B}_n^\top F_t\,, \tag{11}$$

$$f_t^{(n,m)} = (A_n - A_{n+m}) + (B_n - B_{n+m})^\top F_t\,, \tag{12}$$

$$rx_{t+1}^{(n)} = \mathfrak{B}_n^\top F_t + v_t^n\,, \tag{13}$$

where $A_n = -\frac{1}{n}\mathcal{A}_n$, $B_n = -\frac{1}{n}\mathcal{B}_n$, $v_t^n = \mathcal{B}_{n-1}^\top \Sigma \epsilon_{t+1}$.

Solutions are derived from Riccati equations, where the factor loadings of log bond prices are

$$\mathcal{B}_n^\pi = -\delta_\pi \frac{1-\phi_\pi^n}{1-\phi_\pi}\,, \tag{14}$$

$$\mathcal{B}_n^r = -\delta_r \frac{1-\phi_r^n}{1-\phi_r}\,, \tag{15}$$

$$\mathcal{B}_n^x = -\mathcal{B}_{n-1}^\pi \lambda_{\pi x} - \mathcal{B}_{n-1}^r \lambda_{rx} + \mathcal{B}_{n-1}^x \phi_x, \tag{16}$$

and the factor loadings of excess returns are

$$\mathfrak{B}_n = \mathcal{B}_{n-1}^\top (\lambda_0 + \Lambda_1 \mathbf{1}_3) x_t - \frac{1}{2}\mathcal{B}_{n-1}^\top \Sigma\Sigma^\top \mathcal{B}_{n-1}^x\,. \tag{17}$$

Note that our earlier forward Equation 1 can be recovered here by rewriting Equation 12 in the form $f_t^{(n,m)} = r_t + \pi_t + [(A_n - A_{n+m}) + (\tilde{B}_n - \tilde{B}_{n+m})^\top F_t]$, where $\tilde{B}_n^\pi = B_n^\pi - 1$, $\tilde{B}_n^r = B_n^r - 1$, $\tilde{B}_n^x = B_n^x$, and the term in brackets represents the bond risk premium term $\Gamma^{(n,m)}(F_t)$.

As is common in the literature, one could choose to define $x_t = \bar{y}_t$, so the price-of-risk factor is the average level of yields, $\bar{y}_t = \frac{1}{N}\sum_1^N y_t^{(n)}$. But to better describe the role of the trends in driving bond pricing we build on the key innovation in Cieslak and Povala (2015), who switch to yield factors x which have been *detrended* to orthogonalize them relative to the trends, with their focus being on the inflation trend. We extend this idea here to apply *both* trends, and we will define the detrended yield by $c_t^{(n)} = y_t^{(n)} - \hat{A}_n - \hat{B}_n^r r_t - \hat{B}_n^\pi \pi_t$, which is the residual from the regression defined by Equation 10 with the yield factor x suppressed. Now let the average of this detrended yield be $\bar{c}_t = \frac{1}{N}\sum_1^N c_t^{(n)}$. The model can then be expressed in our preferred form in terms of $x_t = \bar{c}_t$ and the full set of factors consists of the two trends and the detrended average yield, and so, henceforth we will be assuming $F_t = (\pi_t, r_t, \bar{c}_t)^\top$, except in a few cases where otherwise indicated.

4. Estimation and model evaluation

An innovation in this paper is to use two sets of information rather than one to extract long run trends. In state-space models of the natural rate typically a single measurement equation on bond yields is employed to estimate the single latent trend. We argue that information on yield curve disturbances will also be embedded, as risk premia, in forward-looking excess bond returns, which provides a second measurement equation to discipline the estimation of latent trends.

Formally, we proceed as follows, where from now on, we restore the stars, and denote by r_t^* the trend natural rate, and by π_t^* trend inflation. We will extract r^* from average bond yields and bond excess returns by using state-space estimation with two affine measurement equations of the form

$$\overline{y}_t = a_y + b_\pi \pi_t^* + b_r r_t^* + \epsilon_t^{cyc}\,, \tag{18}$$

$$\overline{rx}_{t+1} = d_0 + d_\pi \pi_t^* + d_r r_t^* + d_{cyc}\epsilon_t^{cyc} + \epsilon_{t+1}^{rx}\,, \tag{19}$$

where π_t^* is trend inflation, a variable which is treated as an observable, and is set equal to the Cieslak and Povala (2015) measure $\pi_t^* = (1-\nu)\sum_{i=0}^{t-1}\nu^i \pi_{t-i}$, where π_t denotes year-on-year CPI inflation reported in month t. We include the detrended yields ϵ_t^{cyc} in the excess return equation to account for the effect of a cyclical factor as driver of bond returns. Going beyond the U.S., we compute exactly the same constant-gain learning estimate π^* for each one of our six economies.

We further assume that the error terms ϵ_{t+1}^{rx} and ϵ_{t+1}^{cyc} follow AR(1) processes of the form

$$\epsilon_{t+1}^{rx} = \rho_{rx}\epsilon_t^{rx} + e_{t+1}^{rx}\,, \qquad e_{t+1}^{rx} \sim N(0, \sigma_{rx}^2)\,, \tag{20}$$

$$\epsilon_{t+1}^{cyc} = \rho_y \epsilon_t^{cyc} + e_{t+1}^{y}\,, \qquad e_{t+1}^{cyc} \sim N(0, \sigma_{cyc}^2)\,. \tag{21}$$

Now let g_t denote trend GDP growth. We also treat this variable as observable, set equal to the exogenously detrended rate of GDP growth using a Hodrick-Prescott filter. We then define the state variable z_t as a "headwinds" factor related to the natural rate through the state transition equation

$$r_t^* = z_t + g_t\,, \tag{22}$$

as is standard in state-space models of the natural rate, such as Laubach and Williams (2003).

Finally, we also assume that the headwinds factor follows an AR(1) process, so that

$$z_{t+1} = \rho_z z_t + e_{t+1}^z\,, \qquad e_{t+1}^z \sim N(0, \sigma_z^2)\,. \tag{23}$$

Including two equations linking r_t^* to bond market data, Equation 18 and Equation 19, and one equation linking r_t^* to growth, Equation 22, is the distinctive feature of our unified empirical macro-finance model, as we bring information from both financial and macroeconomic data to bear on estimating the natural rate.

The Kalman system is thus defined by the following state equation,

$$\begin{pmatrix} z_t \\ \epsilon_t^{rx} \\ \epsilon_t^{cyc} \end{pmatrix} = \begin{pmatrix} \rho_z & 0 & 0 \\ 0 & \rho_{rx} & 0 \\ 0 & 0 & \rho_{cyc} \end{pmatrix} \begin{pmatrix} z_{t-1} \\ \epsilon_{t-1}^{rx} \\ \epsilon_{t-1}^{cyc} \end{pmatrix} + \begin{pmatrix} e_t^z \\ e_t^{rx} \\ e_t^{ye} \end{pmatrix} . \tag{24}$$

The associated measurement equation is

$$\begin{pmatrix} \overline{y}_t \\ \overline{rx}_{t+1} \end{pmatrix} = \begin{pmatrix} a_y \\ d_0 \end{pmatrix} + \begin{pmatrix} b_\pi & b_{r^*} \\ d_\pi & d_{r^*} \end{pmatrix} \begin{pmatrix} \pi_t^* \\ g_t \end{pmatrix} + \begin{pmatrix} b_{r^*} & 0 & 1 \\ d_{r^*} & 1 & d_{cyc} \end{pmatrix} \begin{pmatrix} z_t \\ \epsilon_t^{rx} \\ \epsilon_t^{cyc} \end{pmatrix} . \tag{25}$$

This fully describes the state-space model, which has then to be estimated. The estimation algorithm is described in the Appendix.

4.1. Construction of new zero-coupon yields

We estimate our bond pricing model using monthly data for zero-coupon yields for six advanced economies: the U.S., Japan, Germany, the U.K., Canada, and Australia. For this purpose, we need estimates of zero-coupon yields at all monthly maturities, from 1 to 180 months, in all countries, and these have to be recovered from market data on government bond yields.

For the U.S. we use a standard source, the estimates of Gürkaynak, Sack, and Wright (2007), comprising data from 1961 published by the Federal Reserve Board, and extended to the present. We then extend their approach to other countries as follows. We use data for a subset of maturities as an input to estimate time-varying parameters $\beta_0, \beta_1, \beta_2, \beta_3, \tau_1$, and τ_2 of a Svensson (1994) model that expresses the yield $y_t^{(n)}$, at any given time t, of a maturity n zero-coupon bond as

$$y_t^{(n)} = \beta_0 + \beta_1 \frac{1 - e^{-n/\tau_1}}{n/\tau_1} + \beta_2 \left(\frac{1 - e^{-n/\tau_1}}{n/\tau_1} - e^{-n/\tau_1} \right) + \beta_3 \left(\frac{1 - e^{-n/\tau_2}}{n/\tau_2} - e^{-n/\tau_2} \right) . \tag{26}$$

Obtaining the parameters of a Svensson model allows us to generate zero-coupon yields for all maturities at each point in the time series, circumventing the problem of data sparsity in some parts of the curve. (Note that this model is estimated separately on the cross-section of yields at every date t, but for notational clarity the time indices of the parameters have been suppressed here.)[2]

[2] There are alternatives to the influential Svensson model. In recent work, Andreasen, Christensen, and Rudebusch (2019) propose an Arbitrage-Free Nelson-Siegel (AFNS) model which may approximate the true curve better at the very short end, below two years, a region of interest to central bank policymakers. However,

For the five-country international yield data, we employ sources as follows:

- For the U.K., we use Bank of England data on the yield curve; this allows us to recover complete yield curves from January 1980 to the present.
- For Japan, yields come from the Ministry of Finance starting in September 1974, and comprising maturities from 1 to 40 years, in yearly maturity increments.
- For Germany, we use the parameters of the Svensson model estimated by the Deutsche Bundesbank from 1972 to 2019.
- For Canada, we use of Bank of Canada data comprising estimates of yield curves for maturities ranging from 0.25 years to 30 years and covering the period January 1986 to the present.
- For Australia, we employ data from the Reserve Bank of Australia dating from August 1992, where yields are available from 0 to 10 years in quarterly maturity increments.

To the best of our knowledge, these estimations provide a new and unique set of zero-coupon data unmatched in the literature by extending the Gürkaynak, Sack, and Wright (2007) methodology consistently to other developed markets and over many more years.[3]

4.2. Short rate process

In this subsection, we take the first step in the empirical assessment of the model by asking whether it can provide a useful description of the short-rate process, as specified by by Equation 2, as an affine function of two factors, the observable expected inflation rate π_t^* and the latent model-implied real natural rate r_t^*, which we write as $y_t^{(1)} = \delta_0 + \delta_r r_t^* + \delta_\pi \pi_t^*$. We first report results for the U.S., and then for the international sample.

4.2.1 U.S. short rate

Table 2 reports the ordinary least squares (OLS) estimates of the short-rate equation for the U.S. Here, Column (1) contains results for the baseline specification with r_t^* and π_t^*, and Column (2) with only π_t^* included as a factor ($\delta_r = 0$) for comparison with the earlier literature.

The results show a good fit for the baseline model, with an R^2 of 0.644, and both factors significant at the 1% level. The coefficient on the natural rate is 1.180, and the coefficient on inflation is 0.986. A null Fisher hypothesis of the both coefficients equal to one could not be rejected here.

The model using only the inflation trend does not fit the data quite as well. The R^2 is only 0.484. The coefficient on inflation is somewhat larger, at 1.192. In prior work, Kozicki and Tinsley (2001)

our interest is primarily in bond pricing and excess return forecasts in maturity buckets in the belly of the curve, from 5 to 15 years. In that range the Svensson model outperforms to AFNS (see their Table 3).

[3]The closest prior work was a decade ago. Wright (2011) compiled a 10-country panel of zero-coupon yields with the data series ending in 2009, using Svensson, Nelson-Siegel, and spline models.

Table 2: *U.S. short rate,* $y^{(1)}$

The table reports OLS estimates on U.S. monthly data of the short rate process $y_t^{(1)} = \delta_0 + \delta_r r_t^* + \delta_\pi \pi_t^*$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample is 1961/6 to 2020/9.

	(1) U.S. $y^{(1)}$	(2) U.S. $y^{(1)}$
r^*	1.180***	
	(0.146)	
π^*	0.986***	1.192***
	(0.131)	(0.116)
Constant	−0.012***	0.005
	(0.004)	(0.005)
N	669	711
R^2	0.644	0.484

reported an inflation coefficient of 1.44 in this specification ($N = 41$, quarterly, 1980–1990, based on Hoey survey inflation measures), and Cieslak and Povala (2015) reported an inflation coefficient of 1.43, with an R^2 of 0.71 ($N = 470$, monthly, 1971–2011), using shorter samples of data.[4]

We conjecture that a lower inflation coefficient in our longer sample may in part reflect the inclusion in our estimation window of more observations from eras of low and stable inflation (the 1960s plus the recent decade or so), in contrast to, say, the 1980s Volcker-Greenspan era when short policy rates were made to respond more aggressively in a period of dogged inflation fighting.

Our findings for the restricted specification in Column (2) still echo these earlier works, but the restriction is clearly rejected in Column (1). The real natural rate trend adds important predictive information for the short rate, and the fact that the coefficient on inflation changes little between these two specifications, shows that this information is distinct and largely orthogonal to the information contained in the inflation trend.

4.2.2 International short rate

Table 3 presents estimates of the short-rate process for the six-country sample. As explained above, sample periods vary by country given the available zero-coupon yield data. Panel (a) shows estimates with the natural rate and inflation trends, and Panel (b) with inflation trends only. The U.S. results are reproduced for comparison.

These results confirm the importance of allowing for both trends, rather than the inflation trend only, when modeling the short-rate process. In Panel (a) both trends are statistically significant in 11 out of 12 cases, the exception being the natural rate for Japan. The results also show that loadings

[4] Note that Bauer and Rudebusch (2019) do not report a short-rate regression, and their assumed short-rate process is expressed in a different form.

on the two trends vary quite a bit by country. The Fisher hypothesis would be rejected in general.

Comparing Panel (a) with Panel (b) we again see a significant improvement in model fit for the model that allows for two trends rather than one, again with the exception of Japan. In Germany and the U.K., the model R^2 increases by a factor of almost 1.5 times when the natural rate trend is added, and by somewhat less for Canada and Australia.

4.3. Bond pricing

In this subsection, we now apply the affine bond pricing model as given by Equation 10, $y_t^{(n)} = \mathcal{A}_n + \mathcal{B}_n^r r_t^* + \mathcal{B}_n^\pi \pi_t^* + \mathcal{B}^{\bar{c}} \bar{c}_t$. This is the specification in Cieslak and Povala (2015), but with a second trend for the estimated natural rate. We also report results that add a short-rate regressor, which was used by Cieslak and Povala (2015) to proxy real trend shifts. These results are not greatly different and this short-rate term contributes little, confirming that the inflation and real rate trends do most of the work. We first report results for the U.S., and then move to the international sample.

4.3.1 U.S. yields

Table 4 presents estimates of U.S. yields at the 2-, 5-, and 10-year points. In Panel (a), we take the traditional approach and use a raw yield factor $\bar{y}$, and in Panel (b) we employ our preferred approach and use a detrended yield factor $\bar{c}$, as in Cieslak and Povala (2015). Obviously, these two regressions are equivalent: the detrended yield factor is just the projection of the raw yield factor on the other two factors. In other words, this is an attribution exercise, where movements in yields not associated with the two rate trends are captured in the detrended ("cyclical") yield factor.

In Panels (a) and (b) the first three columns show these results, and the last three columns augment the regression with the short-rate term as an extra factor in raw form as $y^{(1)}$ or in detrended form as $c^{(1)}$. This provides a point of comparison with the specification in Cieslak and Povala (2015), who include this extra factor, but have no natural rate term.

We find that once the natural rate term is included, the marginal explanatory power of the short-rate term is small for yields, but not zero, and later we will see that is virtually zero for excess returns. For the basic three-factor model, the model fit gives an R^2 of 0.973, 0.998, and 0.991 at the 2-, 5-, and 10-year points, respectively. This improves only marginally to 0.981, 0.998, and 0.994, with the inclusion of the short-rate term as a fourth factor, with no new information at the 5-year point.

At first glance, the loadings in Panel (a) might seem to suggest that the trend factors r^* and π^* play a weak role, but that is because their indirect impact—via the shifts that they induce in the entire yield curve—are not properly accounted for in the trend coefficients. The key insight in Cieslak and Povala (2015) was to orthogonalize yields to get the correct attribution. Thus our preferred specification in Panel (b), in the first three columns, uses detrended yields and shows that large and statistically significant loadings now attach to both the natural rate and inflation trend at all maturities. This result is displayed more clearly in Figure 4, which shows the coefficient loadings

Table 3: *International short rate, $y^{(1)}$, baseline*

The table reports OLS estimates on six-country monthly data of the short rate process $y_t^{(1)} = \delta_0 + \delta_r r_t^* + \delta_\pi \pi_t^*$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample varies by country.

(a) With natural rate and inflation trends

	(1) U.S. $y^{(1)}$	(2) Japan $y^{(1)}$	(3) Germany $y^{(1)}$	(4) U.K. $y^{(1)}$	(5) Canada $y^{(1)}$	(6) Australia $y^{(1)}$
r^*	1.180***	−1.514	1.329***	1.393***	0.751***	1.162***
	(0.146)	(1.288)	(0.110)	(0.112)	(0.129)	(0.080)
π^*	0.986***	6.832*	1.018***	0.451***	1.252*	0.458***
	(0.131)	(2.853)	(0.207)	(0.064)	(0.517)	(0.137)
Constant	−0.012***	−0.012	−0.014***	−0.014***	−0.013	0.011**
	(0.004)	(0.006)	(0.003)	(0.003)	(0.009)	(0.004)
N	669	453	549	453	382	351
R^2	0.644	0.198	0.674	0.838	0.288	0.758

(b) With inflation trend only

	(1) U.S. $y^{(1)}$	(2) Japan $y^{(1)}$	(3) Germany $y^{(1)}$	(4) U.K. $y^{(1)}$	(5) Canada $y^{(1)}$	(6) Australia $y^{(1)}$
π^*	1.192***	3.630**	2.108***	0.938***	1.603***	1.541***
	(0.116)	(1.278)	(0.211)	(0.075)	(0.224)	(0.230)
Constant	0.005	−0.009	−0.012*	0.011*	−0.008	−0.003
	(0.005)	(0.008)	(0.005)	(0.005)	(0.005)	(0.007)
N	711	477	576	489	417	382
R^2	0.484	0.143	0.395	0.624	0.303	0.648

Table 4: *U.S. yields*

The table reports OLS estimates on U.S. monthly data of the yield equation $y_t^{(n)} = \tilde{\mathcal{A}}_n + \tilde{\mathcal{B}}_n^r r_t^* + \tilde{\mathcal{B}}_n^\pi \pi_t^* + \tilde{\mathcal{B}}^{\bar{c}} \bar{c}_t$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample is 1961/6 to 2020/9.

(a) With yield factors

	(1) U.S. $y^{(2)}$	(2) U.S. $y^{(5)}$	(3) U.S. $y^{(10)}$	(4) U.S. $y^{(2)}$	(5) U.S. $y^{(5)}$	(6) U.S. $y^{(10)}$
r^*	-0.147^{*}	0.018	0.067	-0.105^{*}	0.020	0.043
	(0.069)	(0.020)	(0.035)	(0.052)	(0.019)	(0.024)
π^*	-0.300^{***}	-0.045^{*}	0.142^{***}	-0.183^{***}	-0.040^{*}	0.075^{***}
	(0.053)	(0.019)	(0.026)	(0.048)	(0.019)	(0.022)
$\bar{y}$	1.321^{***}	1.063^{***}	0.850^{***}	1.051^{***}	1.050^{***}	1.004^{***}
	(0.049)	(0.015)	(0.024)	(0.067)	(0.016)	(0.032)
$y^{(1)}$				0.183^{***}	0.008	-0.104^{***}
				(0.039)	(0.006)	(0.019)
Constant	-0.011^{***}	-0.004^{***}	0.005^{***}	-0.009^{***}	-0.004^{***}	0.004^{***}
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
N	669	669	669	669	669	669
R^2	0.973	0.998	0.991	0.981	0.998	0.994

(b) With detrended yield factors

	(1) U.S. $y^{(2)}$	(2) U.S. $y^{(5)}$	(3) U.S. $y^{(10)}$	(4) U.S. $y^{(2)}$	(5) U.S. $y^{(5)}$	(6) U.S. $y^{(10)}$
r^*	1.115^{***}	1.033^{***}	0.879^{***}	1.115^{***}	1.033^{***}	0.879^{***}
	(0.046)	(0.014)	(0.024)	(0.035)	(0.013)	(0.017)
π^*	1.156^{***}	1.126^{***}	1.079^{***}	1.156^{***}	1.126^{***}	1.079^{***}
	(0.025)	(0.007)	(0.013)	(0.017)	(0.007)	(0.008)
$\bar{c}$	1.321^{***}	1.063^{***}	0.850^{***}	1.051^{***}	1.050^{***}	1.004^{***}
	(0.049)	(0.015)	(0.024)	(0.067)	(0.016)	(0.032)
$c^{(1)}$				0.183^{***}	0.008	-0.104^{***}
				(0.039)	(0.006)	(0.019)
Constant	-0.014^{***}	-0.006^{***}	0.003^{***}	-0.014^{***}	-0.006^{***}	0.003^{***}
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
N	669	669	669	669	669	669
R^2	0.973	0.998	0.991	0.981	0.998	0.994

for yields at points along the zero-coupon curve from 1 to 15 years (i.e., 12, 24, 36,. . . , 180 months).

Moving to robustness checks, Table 5 revisits our preferred specification but alternates between dropping the trend and cycle factors to see where the bulk of the explanatory power lies. The lesson is clear. The model fit gives an R^2 of around 0.86–0.93 when only the trend factors are used, as in the last three columns, but only 0.07–0.12 when just the cyclical factor is used, as in the first three columns. That is, about 90% of the predictive power of the U.S. bond pricing model stems from correctly accounting for just two factors, the slow moving trends in inflation and the natural rate. In contrast, the remaining cyclical movements in yields, cleansed of these trend factors, contribute about 10% of predictive power.

Finally, Table 6 revisits the specification with raw yields but drops the trend factors to see whether the different forms of detrending matter. This choice will be important for excess return prediction (see below). In the case of fitting the yield curve, here the costs are not as great in terms of worse fit when one or both factor trends are omitted, but the R^2 is certainly reduced when the yields are detrended only by inflation as in the first three columns, or not detrended at all as in the last three columns. But the attribution is clear: it is the trend factors subsumed in yields at work.

Figure 4: *U.S. yield loadings on macro factors, using detrended yields*

The figure shows loading estimates $\hat{\mathcal{B}}_n^r$ and $\hat{\mathcal{B}}_n^\pi$ at maturity n from Equation 10 using U.S. data.

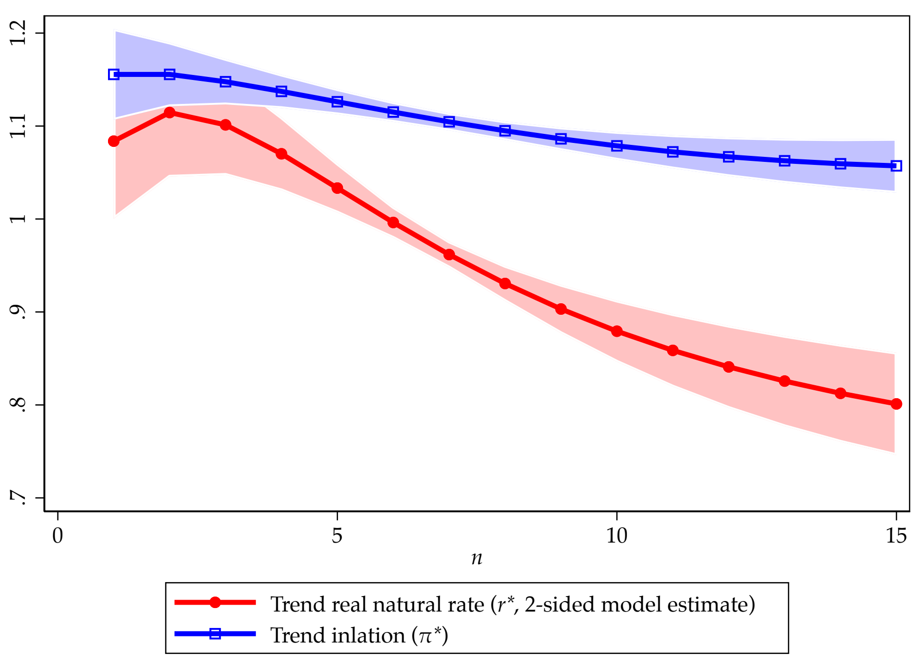


Table 5: *U.S. yields, additional results*

The table reports OLS estimates on U.S. monthly data of the yield equation $y_t^{(n)} = \tilde{\mathcal{A}}_n + \tilde{\mathcal{B}}_n^r r_t^* + \tilde{\mathcal{B}}_n^\pi \pi_t^* + \tilde{\mathcal{B}}^{\bar{c}} \bar{c}_t$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample is 1961/6 to 2020/9.

	(1) U.S. $y^{(2)}$	(2) U.S. $y^{(5)}$	(3) U.S. $y^{(10)}$	(4) U.S. $y^{(2)}$	(5) U.S. $y^{(5)}$	(6) U.S. $y^{(10)}$
$\bar{c}$	1.051^*	1.050^*	1.004^*			
	(0.429)	(0.417)	(0.393)			
$c^{(1)}$	0.183	0.008	−0.104			
	(0.144)	(0.132)	(0.125)			
r^*				1.115^{***}	1.033^{***}	0.879^{***}
				(0.088)	(0.066)	(0.056)
π^*				1.156^{***}	1.126^{***}	1.079^{***}
				(0.072)	(0.055)	(0.050)
Constant	0.053^{***}	0.057^{***}	0.062^{***}	-0.014^{***}	-0.006^{**}	0.003^*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
N	669	669	669	669	669	669
R^2	0.122	0.084	0.068	0.860	0.913	0.927

Table 6: *U.S. yields, additional results*

The table reports OLS estimates on U.S. monthly data of the yield equation $y_t^{(n)} = \tilde{\mathcal{A}}_n + \tilde{\mathcal{B}}_n^r r_t^* + \tilde{\mathcal{B}}_n^\pi \pi_t^* + \tilde{\mathcal{B}}^{\bar{y}} \bar{y}_t$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample is 1961/6 to 2020/9.

	(1) U.S. $y^{(2)}$	(2) U.S. $y^{(5)}$	(3) U.S. $y^{(10)}$	(4) U.S. $y^{(2)}$	(5) U.S. $y^{(5)}$	(6) U.S. $y^{(10)}$
$\bar{y}$	1.232^{***}	1.079^{***}	0.890^{***}	1.095^{***}	1.037^{***}	0.956^{***}
	(0.032)	(0.011)	(0.016)	(0.018)	(0.006)	(0.009)
π^*	-0.223^{***}	-0.068^{***}	0.108^{***}			
	(0.039)	(0.015)	(0.020)			
Constant	-0.011^{***}	-0.003^{***}	0.005^{***}	-0.011^{***}	-0.004^{***}	0.005^{***}
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
N	711	711	711	712	712	712
R^2	0.972	0.997	0.991	0.968	0.997	0.989

4.3.2 International yields

We next show results for the international six-country sample in Table 7, again keeping the U.S. results for reference. For brevity we report results at the 10-year maturity point, but similar findings apply at the 2- and 5-year points. And for reasons of space we here report results using only $\bar{c}$, and omit $c^{(1)}$, but the results are not sensitive to this choice.

Supportive results obtain in all six economies with an R^2 ranging from 0.989 to 0.996. Yields load strongly on the two trend factors, inflation and the natural rate. Coefficients are positive, usually greater than one, and highly statistically significant. We see that the cyclical factor also attracts statistically significant loadings, but here again its explanatory power is not as strong. To see this, Table 8 repeats the exercise of dropping the trend factors and keeping only the cyclical factor. Once more, as in the U.S. case, the explanatory power is poor, with an R^2 less than 0.1 in all cases.

For a fuller picture, Figure 5 again shows coefficient loadings for yields along the curve from 1 to 15 years (12, 24, 26,..., 180 months). The main takeaway from this section is that, all across the curve, and all around the world, bond yields are largely driven by investors' best estimates of the two key slow-moving trend factors, inflation and the natural rate. In contrast, the cyclical factor in yields, summarized by the detrended average yield, is of relatively little importance.

Table 7: *International yields*

The table reports OLS estimates on international monthly data of yields $y_t^{(n)} = \tilde{\mathcal{A}}_n + \tilde{\mathcal{B}}_n^r r_t^* + \tilde{\mathcal{B}}_n^\pi \pi_t^* + \tilde{\mathcal{B}}^{\bar{c}} \bar{c}_t$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample varies by country.

	(1) U.S. $y^{(10)}$	(2) Japan $y^{(10)}$	(3) Germany $y^{(10)}$	(4) U.K. $y^{(10)}$	(5) Canada $y^{(10)}$	(6) Australia $y^{(10)}$
r^*	0.879***	1.159***	1.311***	1.032***	0.938***	1.065***
	(0.024)	(0.024)	(0.022)	(0.021)	(0.024)	(0.013)
π^*	1.079***	0.458***	0.784***	0.524***	1.181***	1.002***
	(0.013)	(0.025)	(0.024)	(0.009)	(0.039)	(0.013)
$\bar{c}$	0.850***	0.576***	0.739***	0.851***	0.844***	0.926***
	(0.024)	(0.128)	(0.041)	(0.042)	(0.045)	(0.025)
Constant	0.003***	0.002***	0.004***	0.002***	0.001	0.003***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
N	669	453	549	453	382	351
R^2	0.991	0.989	0.992	0.992	0.992	0.996

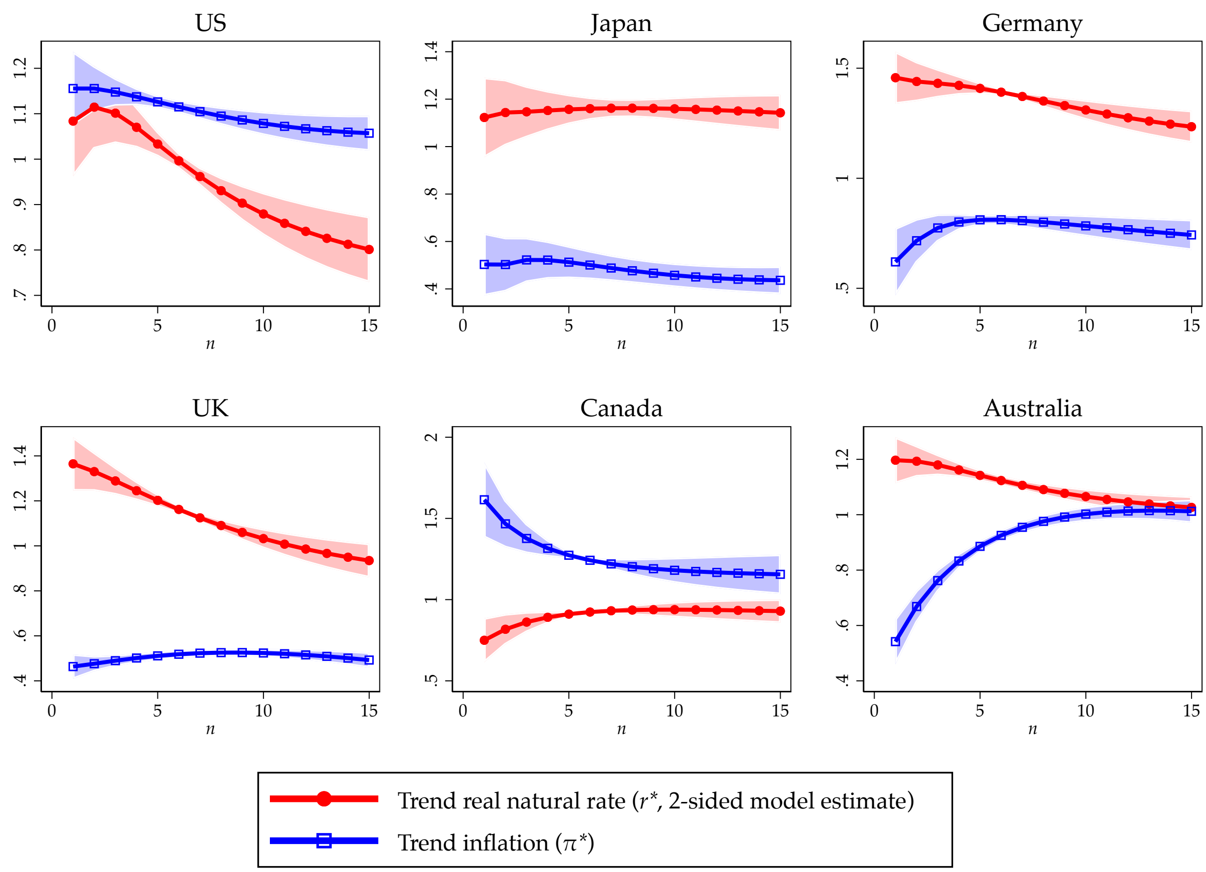
Table 8: *International yields, additional results*

The table reports OLS estimates on international monthly data of yields $y_t^{(n)} = \tilde{\mathcal{A}}_n + \tilde{\mathcal{B}}^{\bar{c}} \bar{c}_t$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample varies by country.

	(1) U.S. $y^{(10)}$	(2) Japan $y^{(10)}$	(3) Germany $y^{(10)}$	(4) U.K. $y^{(10)}$	(5) Canada $y^{(10)}$	(6) Australia $y^{(10)}$
$\bar{c}$	0.850*	0.576	0.739	0.851	0.844	0.926*
	(0.365)	(0.621)	(0.457)	(0.472)	(0.604)	(0.409)
Constant	0.062***	0.027***	0.052***	0.057***	0.049***	0.053***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
N	669	453	549	453	382	351
R^2	0.064	0.012	0.023	0.032	0.035	0.087

Figure 5: *International loadings on macro factors, using detrended yields*

The figure shows loading estimates $\hat{\mathcal{B}}_n^r$ and $\hat{\mathcal{B}}_n^\pi$ at maturity n from Equation 10 using international data.



4.4. Return predictability

We just saw that accounting for trends can make some improvements in modeling yields, but we now see how they matter a great deal for predicting bond returns. This was shown for the U.S. case in Cieslak and Povala (2015) with just an inflation trend extracted, and also in the contemporaneous work of Bauer and Rudebusch (2019) with trends for inflation and the real natural rate, or a nominal natural rate trend. We show that the same applies more generally at the international level.

The intuition is quite straightforward. The trend factors, being slow moving and near unit-root, are mainly priced in one-period ahead and contain little useful information about short-run returns. In contrast, the cyclical factor, being the driver of the high-frequency error-correction part of the bond price process, is very informative about how bond prices revert to trend in the short run.

Formally, in this section we will be presenting estimates for the excess return Equation 13, $rx_{t+1}^{(n)} = \mathfrak{B}_n^\top F_t + v_t^n$. These one-step ahead predictions are noisy but we shall see that their small explanatory power is almost entirely due to the role of the detrended, or cyclical, yield factor $\bar{c}$.

4.4.1 U.S. excess returns

Table 9 and Table 10 show excess return regressions for the U.S. case. We show the 1-, 2-, and 5-year maturities, and an average return across all maturities. Starting with Table 9, Panel (a) uses yield factors, and Panel (b) uses detrended yield factors. Again, these two sets of regressions are identical models, with the same fit, predictions, residuals, etc. They differ only in that the detrended yields are orthogonalized relative to the two trends to give full attribution of trend movements to inflation and natural rate movements, again following Cieslak and Povala (2015).

Columns (1) to (4) include the additional short rate term, but Columns (5) to (8) contain only the average yield factor, our baseline model. The latter is our preferred specification because the short-rate term is not statistically significant in any specifications. In our preferred specification in Panel (a) excess returns can load on all three terms, but the orthogonalization of yields in Panel (b) makes clear that this is rather an illusion. The trend terms in the specifications using raw yield in Panel (a) merely serve to soak up the trends in yields. But once the trends are projected out, with the yields detrended in Panel (b), it is only the cyclical component of yields $\bar{c}$ that has highly significant explanatory power for excess bond returns in all cases. The loading is positive: a cyclically high yield curve, relative to the trends, will be expected to revert down to trend at all maturities; thus, yields are predicted fall, and bond returns are expected to be higher. In terms of attribution, all the explanatory power comes from the average detrended yield, as we see in the last four columns.

Finally, Table 10 shows that the form of detrending matters. The baseline preferred model uses both inflation and natural rates to detrend yields. Here we explore excess return forecasts using inflation-only detrending of yields in Panel (a), and using yields with no detrending in Panel (b). To achieve this parsimoniously we simply perform regressions with raw yield factors as in Table 9, Panel (a), and then omit the trend terms, first just the inflation trend and then both trends. In the

Table 9: *U.S. excess returns*

The table reports OLS estimates on U.S. monthly data of the excess return equation $rx_{t+1}^{(n)} = \mathfrak{B}_n^\top F_t + v_t^n$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample is 1961/6 to 2020/9.

(a) With yield factors

	(1) U.S. $rx^{(1)}$	(2) U.S. $rx^{(2)}$	(3) U.S. $rx^{(5)}$	(4) U.S. $\overline{rx}$	(5) U.S. $rx^{(1)}$	(6) U.S. $rx^{(2)}$	(7) U.S. $rx^{(5)}$	(8) U.S. $\overline{rx}$
r^*	-0.063^*	-0.069^{**}	-0.071^{***}	-0.068^{**}	-0.061^*	-0.065^*	-0.065^{**}	-0.063^{**}
	(0.027)	(0.026)	(0.021)	(0.022)	(0.028)	(0.027)	(0.022)	(0.023)
π^*	-0.078^{**}	-0.087^{**}	-0.090^{***}	-0.085^{***}	-0.073^*	-0.077^{**}	-0.074^{**}	-0.073^{**}
	(0.030)	(0.029)	(0.023)	(0.024)	(0.029)	(0.029)	(0.023)	(0.024)
$\bar{y}$	0.082^*	0.096^{**}	0.102^{***}	0.095^{***}	0.070^{**}	0.072^{**}	0.067^{***}	0.067^{**}
	(0.032)	(0.030)	(0.025)	(0.026)	(0.026)	(0.025)	(0.020)	(0.021)
$y^{(1)}$	−0.007	−0.017	−0.024	−0.019				
	(0.016)	(0.015)	(0.013)	(0.013)				
N	668	668	668	668	668	668	668	668
R^2	0.021	0.025	0.035	0.031	0.020	0.021	0.023	0.024

(b) With detrended yield factors

	(1) U.S. $rx^{(1)}$	(2) U.S. $rx^{(2)}$	(3) U.S. $rx^{(5)}$	(4) U.S. $\overline{rx}$	(5) U.S. $rx^{(1)}$	(6) U.S. $rx^{(2)}$	(7) U.S. $rx^{(5)}$	(8) U.S. $\overline{rx}$
r^*	0.007	0.004	−0.001	0.001	0.007	0.004	−0.001	0.001
	(0.012)	(0.013)	(0.012)	(0.011)	(0.012)	(0.013)	(0.012)	(0.012)
π^*	0.005	0.003	−0.001	0.001	0.005	0.003	−0.001	0.001
	(0.015)	(0.013)	(0.011)	(0.012)	(0.015)	(0.014)	(0.012)	(0.012)
$\bar{c}$	0.082^*	0.096^{**}	0.102^{***}	0.095^{***}	0.070^{**}	0.072^{**}	0.067^{***}	0.067^{**}
	(0.032)	(0.030)	(0.025)	(0.026)	(0.026)	(0.025)	(0.020)	(0.021)
$c^{(1)}$	−0.007	−0.017	−0.024	−0.019				
	(0.016)	(0.015)	(0.013)	(0.013)				
N	668	668	668	668	668	668	668	668
R^2	0.021	0.025	0.035	0.031	0.020	0.021	0.023	0.024

Table 10: *U.S. excess returns, additional results*

The table reports OLS estimates on U.S. monthly data of the excess return equation $rx_{t+1}^{(n)} = \mathfrak{B}_n^\top F_t + v_t^n$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample is 1961/6 to 2020/9.

(a) With yield factors, omit natural rate trend

	(1) U.S. $rx^{(1)}$	(2) U.S. $rx^{(2)}$	(3) U.S. $rx^{(5)}$	(4) U.S. $\overline{rx}$	(5) U.S. $rx^{(1)}$	(6) U.S. $rx^{(2)}$	(7) U.S. $rx^{(5)}$	(8) U.S. $\overline{rx}$
π^*	−0.034	−0.039	−0.041*	−0.037*	−0.031	−0.033	−0.030	−0.030
	(0.022)	(0.021)	(0.017)	(0.018)	(0.020)	(0.020)	(0.017)	(0.017)
$\bar{y}$	0.036	0.046*	0.050**	0.045*	0.030*	0.029*	0.023*	0.025*
	(0.023)	(0.022)	(0.019)	(0.020)	(0.013)	(0.013)	(0.012)	(0.011)
$y^{(1)}$	−0.004	−0.013	−0.020	−0.015				
	(0.015)	(0.014)	(0.011)	(0.012)				
N	700	700	700	700	700	700	700	700
R^2	0.011	0.012	0.017	0.014	0.010	0.009	0.008	0.009

(b) With yield factors, omit natural rate and inflation trends

	(1) U.S. $rx^{(1)}$	(2) U.S. $rx^{(2)}$	(3) U.S. $rx^{(5)}$	(4) U.S. $\overline{rx}$	(5) U.S. $rx^{(1)}$	(6) U.S. $rx^{(2)}$	(7) U.S. $rx^{(5)}$	(8) U.S. $\overline{rx}$
$\bar{y}$	0.010	0.016	0.020	0.016	0.011	0.009	0.005	0.007
	(0.016)	(0.015)	(0.012)	(0.013)	(0.009)	(0.008)	(0.007)	(0.007)
$y^{(1)}$	0.000	−0.007	−0.014	−0.010				
	(0.013)	(0.012)	(0.010)	(0.011)				
N	700	700	700	700	700	700	700	700
R^2	0.005	0.005	0.007	0.005	0.005	0.004	0.002	0.003

former case, this regression is identical to using yields projected onto inflation to construct the cyclical components, and in the latter case it amounts to no detrending at all.

The results are clear. Note that the short-rate terms continue to be statistically insignificant. Looking back for comparison, when we used both the inflation and natural rate to detrend yields, the excess return predictions had moderately good fit. In Table 9, Panel (b), Columns (1) to (3), the R^2 values are 0.021, 0.025, and 0.035 at the 1, 2 and 5 year maturities, respectively. This is a respectably good fit for a return forecast model. Table 10 shows that when only an inflation trend is used as in Panel (a), the measures of fit decline to 0.011, 0.012, and 0.017, respectively. That is, the R^2 is cut in half when we only use the inflation trend of Cieslak and Povala (2015). Finally, when no detrending is allowed as in Panel (b), the measures of fit collapse even more to 0.005, 0.005, and 0.007, respectively. Using only the single level factor $\bar{y}$ the fit falls even further, to 0.005 or less.

Not accounting for the important macro trends thus destroys about three-quarters of the model's explanatory power. Conversely, accounting for the trends improves return predictability more than fourfold for each of these bond maturities, with inflation and the natural rate each making similarly substantive contributions to the fit.

4.4.2 International excess returns

We now take the bond return forecast model to international data. Our preferred results in Table 11 are the baseline results for six countries. These estimates omit the short-rate factor. Table 12 and Table 13 report additional results. The former shows that the short-rate factor, if included is never statistically significant, confirming our baseline choice, as in the U.S. case. The latter shows that omitting the natural rate trend, or omitting both trends, comes at the cost of much worse model performance, with R^2 statistics collapsing to near zero when trends are removed, again as we saw for the U.S. In all tables the U.S. results are shown for comparability.

The detrending approach is also supported. In Table 12, Panel (a) shows again how loadings attach to the trend terms (6 out of 12 coefficients are statistically significant), but Panel (b) again confirms that this is an artifact of failing to detrend the yield factor. Once that is done, only the detrended average yield term has consistent predictive power, and this is best seen in our preferred results in Table 11 with the insignificant short-rate term omitted.

Returning then to the baseline results in Table 11, we find that the most reliable predictor of excess bond returns is again the cyclical component of yields $\bar{c}$ (5 out of 6 coefficients are statistically significant, and z-score for Germany is 1.2). Residual loadings on the trend terms are generally not important (only 1 out of 12 coefficients is statistically significant). Measures of fit range from a low R^2 of 0.010 for Germany, up to 0.043 for Australia. Overall, the average fit here is again around 3%, a respectable in-sample fit for a return forecasting model.

Table 11: *International excess returns*

The table reports OLS estimates on international data of the excess return equation $rx_{t+1}^{(n)} = \mathfrak{B}_n^\top F_t + v_t^n$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample varies by country.

	(1) U.S. $\overline{rx}$	(2) Japan $\overline{rx}$	(3) Germany $\overline{rx}$	(4) U.K. $\overline{rx}$	(5) Canada $\overline{rx}$	(6) Australia $\overline{rx}$
r^*	0.001	0.007	−0.008	−0.004	0.005	−0.016
	(0.012)	(0.016)	(0.008)	(0.008)	(0.009)	(0.010)
π^*	0.001	−0.005	0.012	0.002	0.001	0.032*
	(0.012)	(0.018)	(0.015)	(0.006)	(0.021)	(0.015)
$\bar{c}$	0.067**	0.088*	0.032	0.066*	0.073**	0.054*
	(0.021)	(0.036)	(0.027)	(0.026)	(0.023)	(0.023)
N	668	453	549	453	382	351
R^2	0.024	0.032	0.010	0.026	0.029	0.043

Table 12: *International excess returns, additional results*

The table reports OLS estimates on international data of the excess return equation $rx_{t+1}^{(n)} = \mathfrak{B}_n^\top F_t + v_t^n$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample varies by country.

(a) With yield factors

	(1) U.S.	(2) Japan	(3) Germany	(4) U.K.	(5) Canada	(6) Australia
	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$
r^*	−0.068**	−0.103*	−0.061	−0.082**	−0.061*	−0.083**
	(0.022)	(0.042)	(0.039)	(0.029)	(0.024)	(0.026)
π^*	−0.085***	−0.058*	−0.015	−0.036*	−0.092*	0.000
	(0.024)	(0.026)	(0.018)	(0.015)	(0.036)	(0.024)
$\bar{y}$	0.095***	0.098*	0.051	0.085**	0.075**	0.012
	(0.026)	(0.038)	(0.029)	(0.031)	(0.023)	(0.036)
$y^{(1)}$	−0.019	0.000	−0.012	−0.013	−0.002	0.047
	(0.013)	(0.001)	(0.007)	(0.017)	(0.003)	(0.032)
N	668	453	549	453	382	351
R^2	0.031	0.033	0.019	0.029	0.030	0.060

(b) With detrended yield factors

	(1) U.S.	(2) Japan	(3) Germany	(4) U.K.	(5) Canada	(6) Australia
	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$	$\overline{rx}$
r^*	0.001	0.007	−0.008	−0.004	0.005	−0.016
	(0.011)	(0.016)	(0.008)	(0.008)	(0.009)	(0.009)
π^*	0.001	−0.005	0.012	0.002	0.001	0.032*
	(0.012)	(0.017)	(0.015)	(0.006)	(0.021)	(0.014)
$\bar{c}$	0.095***	0.098*	0.051	0.085**	0.075**	0.012
	(0.026)	(0.038)	(0.029)	(0.031)	(0.023)	(0.036)
$c^{(1)}$	−0.019	0.000	−0.012	−0.013	−0.002	0.047
	(0.013)	(0.001)	(0.007)	(0.017)	(0.003)	(0.032)
N	668	453	549	453	382	351
R^2	0.031	0.033	0.019	0.029	0.030	0.060

Table 13: *International excess returns, additional results*

The table reports OLS estimates on international data of the excess return equation $rx_{t+1}^{(n)} = \mathfrak{B}_n^\top F_t + v_t^n$. Newey-West standard errors, 6 lags, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample varies by country.

(a) With yield factors, omit natural rate trend

	(1) U.S. $\overline{rx}$	(2) Japan $\overline{rx}$	(3) Germany $\overline{rx}$	(4) U.K. $\overline{rx}$	(5) Canada $\overline{rx}$	(6) Australia $\overline{rx}$
π^*	-0.037^*	-0.025	0.011	-0.004	-0.045^*	0.028
	(0.018)	(0.015)	(0.016)	(0.010)	(0.021)	(0.016)
$\bar{y}$	0.045^*	0.015	-0.001	0.008	0.025^{**}	-0.013
	(0.020)	(0.012)	(0.011)	(0.018)	(0.009)	(0.028)
$y^{(1)}$	-0.015	0.001	0.000	-0.001	-0.003	0.014
	(0.012)	(0.001)	(0.008)	(0.016)	(0.003)	(0.025)
N	700	467	566	478	406	371
R^2	0.014	0.021	0.002	0.002	0.016	0.033

(a) With yield factors, omit natural rate and inflation trends

	(1) U.S. $\overline{rx}$	(2) Japan $\overline{rx}$	(3) Germany $\overline{rx}$	(4) U.K. $\overline{rx}$	(5) Canada $\overline{rx}$	(6) Australia $\overline{rx}$
$\bar{y}$	0.016	-0.001	0.003	0.003	0.006	0.006
	(0.013)	(0.006)	(0.010)	(0.017)	(0.005)	(0.026)
$y^{(1)}$	-0.010	0.001	0.000	0.000	-0.003	0.009
	(0.011)	(0.001)	(0.008)	(0.015)	(0.004)	(0.025)
N	700	468	567	479	407	372
R^2	0.005	0.011	0.001	0.002	0.003	0.025

4.5. Out-of-sample performance

The results so far rely on in-sample measures of model performance, principally the R^2 measure of fit. For example, in return regressions, we typically find an R^2 in the two-trend models utilizing estimates of r^* and π^* that is about twice as high as in a one-trend π^* model (Cieslak and Povala, 2015); the latter is in turn higher than in models using only yield-based factors (Cochrane and Piazzesi, 2005). But in-sample measures are not the gold standard for predictive performance, so we now ask if the two-trend model also excels in out-of-sample performance tests. We find that it does.

To construct out-of-sample forecasted excess returns we re-estimate our baseline model each period using a one-sided Kalman filter to estimate r^* recursively. To allow for a burn-in period of reasonable sample size, and given the short span of the Australian data, the out-of-sample window starts in 2004/1. We include in the baseline model, as before, inflation expectations π^* (one-sided by construction), and also $\bar{y}$ and $y^{(1)}$ (real-time yield information) as regressors. For the analog of the CiP type model, a level-slope model, and a level-only model, we do likewise and in each case omit the natural rate trend (r^*), or both trends (r^*, π^*), or both trends and the short-rate ($r^*, \pi^*, y^{(1)}$). For a broader comparison, we also construct out-of-sample forecasted excess returns using a recursively estimated ACM3 three-factor model (Adrian, Crump, and Moench, 2013) using lagged principal components of yields $F_t = (PC1_t, PC2_t, PC3_t)$ and estimated VAR(1) residuals $\hat{v}$ from a first stage $F_t = \mu + \Phi F_{t-1} + v_t$. The comparison group of five out-of-sample recursive forecasts is then

$$\widehat{rx}_{t+1}^{\text{Baseline}} = \hat{d}_0 + \hat{d}_1 \bar{y}_t + \hat{d}_2 y_t^{(1)} + \hat{d}_\pi \pi_t^* + \hat{d}_r r_t^* \,, \tag{27}$$

$$\widehat{rx}_{t+1}^{\text{CiP}} = \hat{d}_0 + \hat{d}_1 \bar{y}_t + \hat{d}_2 y_t^{(1)} + \hat{d}_\pi \pi_t^* \,, \tag{28}$$

$$\widehat{rx}_{t+1}^{\text{Level+Slope}} = \hat{d}_0 + \hat{d}_1 \bar{y}_t + \hat{d}_2 y_t^{(1)}, \tag{29}$$

$$\widehat{rx}_{t+1}^{\text{Level}} = \hat{d}_0 + \hat{d}_1 \bar{y}_t \,, \tag{30}$$

$$\widehat{rx}_{t+1}^{\text{ACM3}} = \hat{a}_0 + \hat{b} \hat{v}_t + \hat{c} F_{t-1} \,. \tag{31}$$

An initial sign of good performance by the baseline two-trend model is given in Table 14. The table shows simple correlation coefficients for all model forecasts and the actual ex-post return rx_{t+1}, for the pooled sample of 6 countries and out-of-sample window 2004/1 to 2020/8. Two things stand out. First, in panel (a), the correlation between forecast and ex-post return is higher for the baseline model (0.21) than in any competing model (0.18 for ACM, 0.05 or less for others). Second, in panel (b), the baseline forecast has typically lower correlation with competing model forecasts (0.40 or less), evidence that the second trend contains additively-valuable predictive information orthogonal to that already present in the ensemble of models in the existing literature. Looking at competing model forecasts, Level+Slope is highly correlated with Level (0.95), consistent with the Cochrane and Piazzesi (2005) finding that adding more yield factors contributes little new information. We also see that both CiP and ACM forecasts correlate less with each other, but more so with Level and Level+Slope. The baseline forecast has some correlation with CiP (0.40), which is unsurprising as it nests the one-trend model, but it is virtually uncorrelated with the other models.

Table 14: *Out-of-sample performance: Correlations of model forecasts and actual returns*

This table shows the correlation matrix of actual returns rx_{t+1} and five out-of-sample model forecasts $\widehat{rx}_{t+1}$, including our baseline model and four alternatives. See text. Pooled sample of 6 countries, 2004/1 to 2020/8.

	(a) with actual returns	(b) Correlations with other model forecasts				
	rx_{t+1}	$\widehat{rx}^{\text{Baseline}}_{t+1}$	$\widehat{rx}^{\text{CiP}}_{t+1}$	$\widehat{rx}^{\text{Level+Slope}}_{t+1}$	$\widehat{rx}^{\text{Level}}_{t+1}$	$\widehat{rx}^{\text{ACM3}}_{t+1}$
rx_{t+1}	1.00					
$\widehat{rx}^{\text{Baseline}}_{t+1}$	0.22	1.00				
$\widehat{rx}^{\text{CiP}}_{t+1}$	0.03	0.40	1.00			
$\widehat{rx}^{\text{Level+Slope}}_{t+1}$	0.05	0.06	0.49	1.00		
$\widehat{rx}^{\text{Level}}_{t+1}$	0.04	0.15	0.52	0.95	1.00	
$\widehat{rx}^{\text{ACM3}}_{t+1}$	0.18	0.10	0.18	0.43	0.41	1.00
N	1200					

However, whilst suggestive of performance improvement in the baseline two-trend model, these correlation results are not conclusive. Thus we turn to a strict out-of-sample testing procedure, using the clean concept of an *out-of-sample R-squared* or R^2_{OS} proposed by Campbell and Thompson (2008). Using their definition, let the model's recursive return forecast at time t be denoted by $\widehat{rx}_{t+1}$, and let the recursive return forecast at time t using purely historical mean returns be denoted by $\overline{rx}_{t+1}$. Then at any time $t \in \{1, \ldots, T\}$ in the out-of-sample period we define

$$R^2_{OS}(t) = 1 - \frac{\sum_{\tau=1}^{t}(rx_{\tau+1} - \widehat{rx}_{\tau+1})^2}{\sum_{\tau=1}^{t}(rx_{\tau+1} - \overline{rx}_{\tau+1})^2} .$$

This is then a rolling, cumulative measure of fit, starting at the beginning of the out-of-sample evaluation window at time 1 and running to its end at time T. Of course, one can also compute an in-sample analog of this statistic $R^2_{IS}(t)$ where full-sample estimates are used, but the same cumulation is employed in the window $t \in \{1, \ldots, T\}$. We present both.

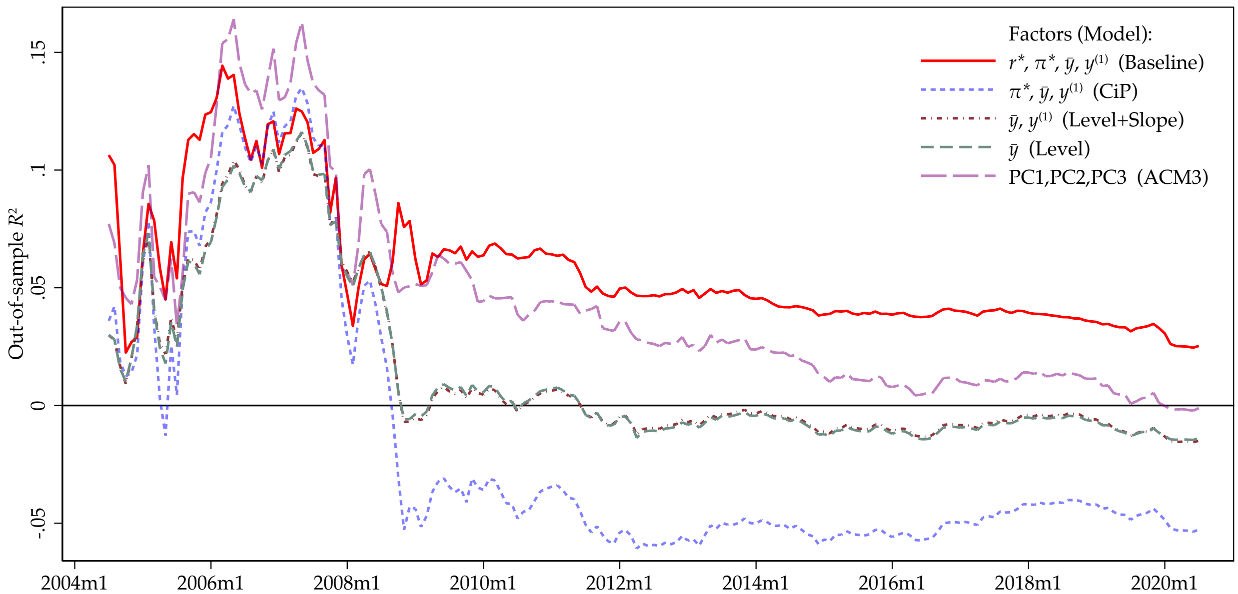
The measures of fit R^2_{OS} and R^2_{IS} for the five models are shown in Figure 6. Note that both measures can be *negative*, unlike conventional R-squared statistics, as shown by Campbell and Thompson (2008), and as happens here. One reason is because the cumulation window is a subset of the full-sample model window, so fit can be inferior to the historical mean in some periods, even for the latter IS measure. This is even more likely at the start of the expanding window, of course, and for this reason we discard the first 6 months of the window to exclude erratic and often wildly negative values that are uninformative in such small samples. In addition, note that the OS measure excludes future information, and so at any time, unlike the correlations discussed above, it is possible for the model to perform even worse than a historical mean extrapolation.

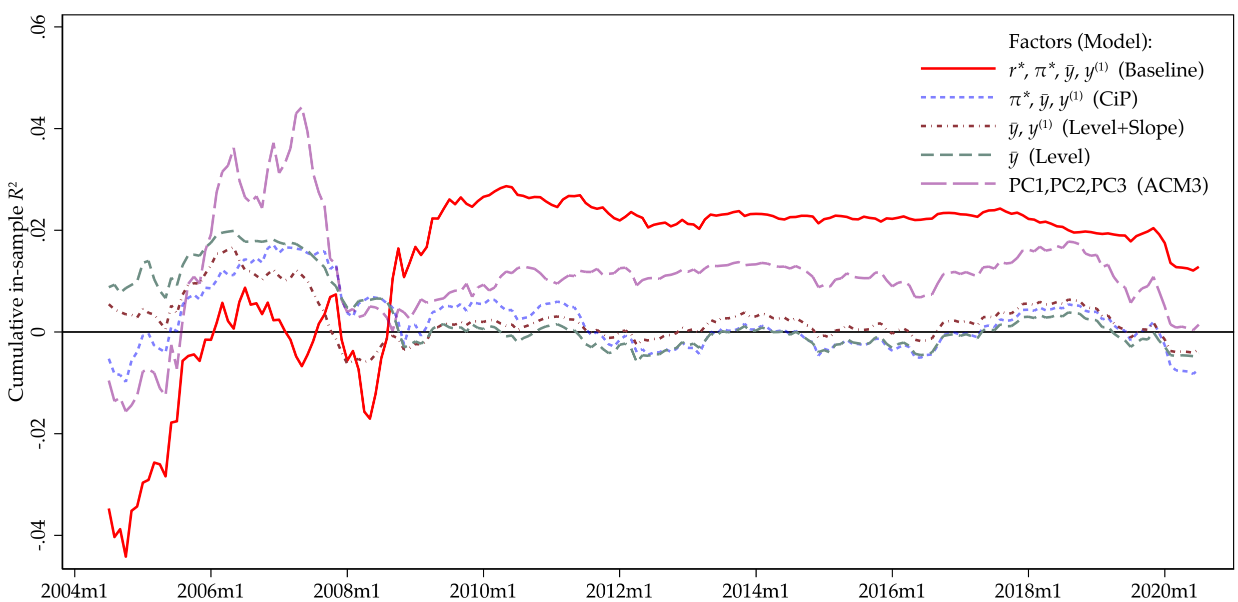
Figure 6: *Out-of-sample performance: R^2_{OS} and R^2_{IS} statistics of Campbell and Thompson (2008)*

These charts show, for our baseline model and four alternatives, the *out-of-sample R-squared* or R^2_{OS} proposed by Campbell and Thompson (2008),

$$R^2_{OS}(t) = 1 - \frac{\sum_{\tau=1}^{t}(rx_{\tau+1} - \widehat{rx}^k_{\tau+1})^2}{\sum_{\tau=1}^{t}(rx_{\tau+1} - \overline{rx}_{\tau+1})^2}\,.$$

We present these recursive out-of-sample estimates in the upper panel, and in the lower panel the in-sample analog $R^2_{IS}(t)$ using full-sample estimates. Each one is a rolling, cumulative measure of fit, starting at beginning of the out-of-sample evaluation window at time 1 and running to its end at time T. See text. Pooled sample of 6 countries, 2004/1 to 2020/8.





What do these results show? The more important R^2_{OS} statistics in the upper chart show that all the models perform similarly at the start of the evaluation window from 2004 to 2007. The trajectories all overlap, with no clear winner. Forecast performance deteriorates for all models when the Global Financial Crisis period starts in mid-2007, but more or less equally. The stark differences emerge after mid-2008. After that, our baseline two-trend model is the only model to deliver a consistently positive and stable R^2_{OS} statistic in the entire evaluation window. The closest rival is the ACM3 model, which loses explanatory power consistently as more of the sample encompasses the post-2008 period. All other models fall below zero in 2009, indicating that out-of-sample performance has not held up well in the post–Global Financial Crisis period, as they fail to outperform simple historical mean forecasts. The in-sample R^2_{IS} statistics in the lower chart confirm that more flattering results are achieved using full-sample estimation but even here there are still many negative values, and again it is still the baseline two-trend model which performs best out of all five models.

Our findings align with Sarno, Schneider, and Wagner (2016). As we do at Equation 18 and Equation 19, they propose an estimation strategy for affine term structure models that jointly fits yields and bond excess returns to capture predictive information otherwise hidden to standard estimations. However, they do not use any macro trends, and fit pure yield-based models. They find no performance improvement in forecasting bond returns with such models, as we do here. Our contribution is to show that performance *is* improved with a two-trend macro-finance model.

How can we interpret and make sense of these findings? Why does a noticeable outperformance by the baseline model emerge after 2008 and why should a model which adds a natural rate trend display such behavior? These results reveal how the fundamental drivers of bond pricing shifted after 2008. In the previous two decades, π^* was falling and volatile as advanced economies converged on their nominal anchors, typicaly an inflation target near 2% per annum. Yields also fell. But the natural rate r^* was more stable and less volatile. After the global financial crisis, π^* was more sticky, at or below target, with central banks struggling to prop it up, and yields languished near the ZLB, especially at the short end. Meanwhile, the natural rate r^* collapsed rapidly as poor growth prospects (and other factors, like demography) brought on a putative secular stagnation scenario. Perhaps not surprisingly, models of the bond market that ignored the natural rate coped better with the former epoch than the latter. But once the times changed, our two-factor model came into its own as natural rate shocks emerged as the main drivers of bond market trends, a shift that comes to the fore in the next, concluding section.

5. Conclusion: Resolving the puzzle

Benchmark finance models imply a natural rate inconsistent with estimates from benchmark macro models. We call this the natural rate puzzle.

We presented a general equilibrium macro-finance model with real and nominal factors in which bond yields and excess returns are explained by two slow-moving latent trend factors, the real natural rate trend r^* and the inflation trend π^*, in an arbitrage-free affine term structure model.

Empirically, we take the model to the data using state-space estimation and the Kalman filter. The model succeeds on multiple dimensions. The pricing regressions for yields improve somewhat and estimates of excess returns are far more accurate than when one or both macro factors are excluded, in and out of sample. Moreover, in our approach the model is forced to estimate "correct" paths of bond risk premia, the natural rate, and inflation that are consistent with forward rates. Thus our approach delivers a resolution of the natural rate puzzle.

Looking again at the U.S. case where we began this paper, Figure 7 displays our model-consistent, market-implied estimates of the real natural rate r^* and the bond risk premium Γ. In panel (a), we compare our real natural rate with those from LW and implied by ACM, as in Figure 1. In panel (b), we compare our bond risk premium with those from ACM and implied by LW, also as in Figure 1.

The various approaches provide radically different historical narratives of the postwar history of the U.S. bond market. Note the important differences here: the ACM model (like any yield-only model) attributes the big rise and fall of interest rates in the 1970–2000 period to large up and down shifts in the bond risk premium, which peaks in the 1980s; but our market-implied estimates produce no such dramatic shifts, and instead the movements in interest rates are attributed to changes in macro trends, r^* and π^*. When we turn to the international data in Figure 8, we see that broadly the same result obtains in other advanced economies. In all cases, we are much closer to the LW estimates, in levels and trends, over the whole sample window, although here there are some noticeable differences relative to the LW estimates as well. Our our macro-finance approach clearly differs strongly from the pure finance approach, and somewhat from the pure macro approach.

Finally, in Figure 9 we better see how our macro-finance approach is meaningfully different from the pure macro approach, by looking at the path of the headwinds factor $z = r^* - g$ for six countries. This residual difference between the natural rate and the real growth rate accounts for shifts in any non-growth factors that move the natural rate. In our model estimation results this factor moves very differently than in the LW model. It falls considerably less steeply in the U.S. (around 100 bps in 60 years, not 200 bps). It falls considerably more steeply in the other five countries (about 200–300 bps in 60 years, not 0–150 bps). We conjecture that our z is picking up starker non-growth differences between the U.S. and other advanced economies; this may reflect in part trends of demography and aging which are most acute in this sample in Europe and Japan.

In sum, our results differ from both the prior macro and finance approaches, but they resolve the puzzle by coming down closer to the macro view. Our market-implied estimates of the natural rate are closer to those of benchmark macro models and further from those of benchmark finance models. Our market-implied natural rate has trends and turning points much like consensus macro estimates, but differs in being typically somewhat lower in the most recent years—and negative in all six countries in 2020—intensifying concerns about secular stagnation and proximity to the effective lower-bound on monetary policy in advanced economies.

The canonical finance approach to bond pricing and return forecasting using term structure models traditionally excludes macro factors. Our findings suggest a new track is needed, and the powerful effects around the world of macro factors should play an important part in future studies.

Figure 7: *Market-implied natural rate r^* and bond risk premium Γ in U.S. data*

The top chart displays our market-implied estimates of the U.S. real natural rate r^*. Our real natural rate estimate is close to the LW level throughout the sample, and far from the ACM-implied level. The bottom chart displays our market-implied estimates of the U.S. bond risk premium Γ. By construction, this is close to the LW-implied level throughout the sample, and far from the ACM level.

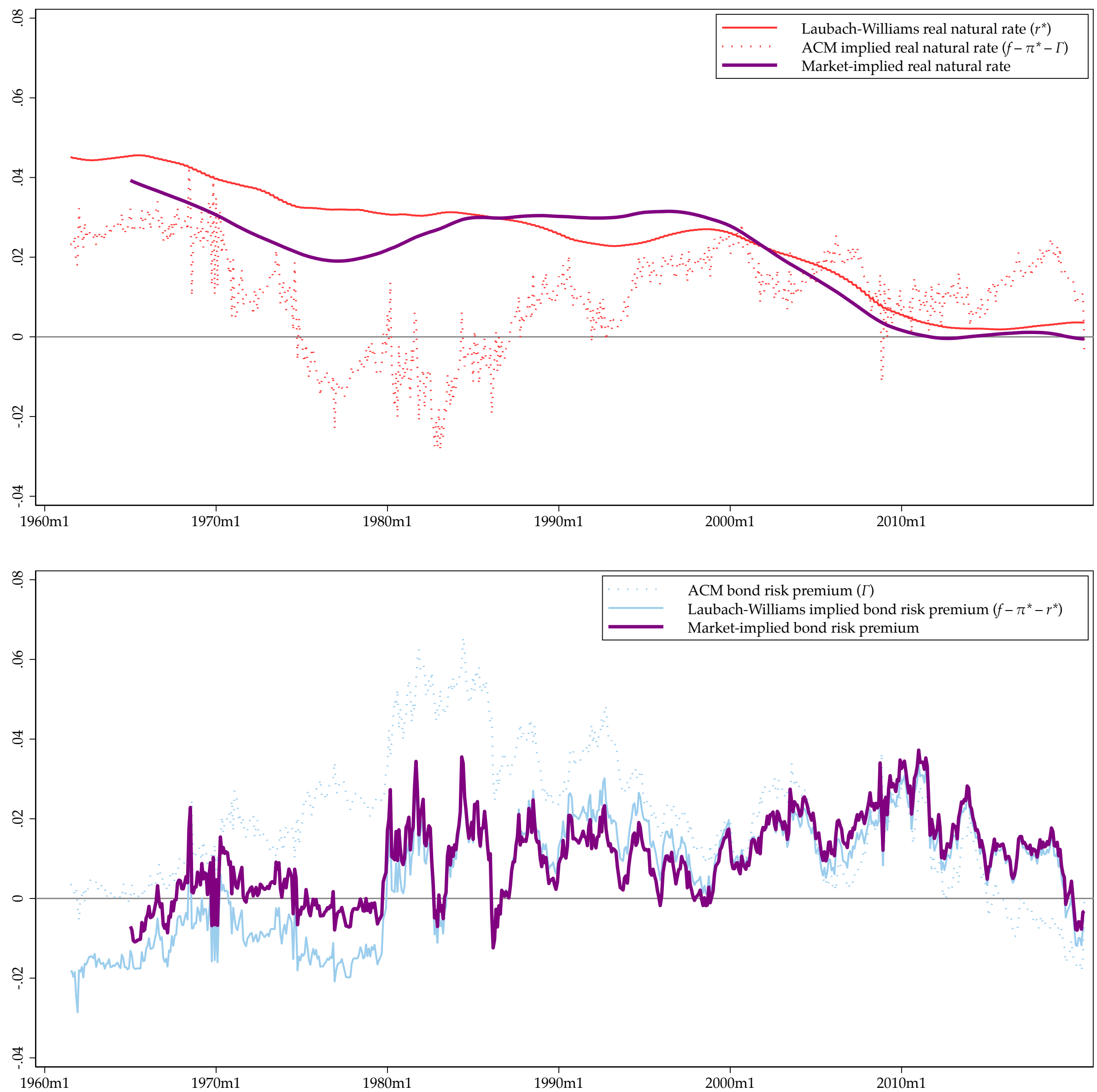


Figure 8: *Market-implied natural rates r^* in international data*

The charts display our market-implied estimates of the real natural rate r^* for six countries. Our real natural rate estimate is close to LW level throughout the sample, and far from the ACM implied level.

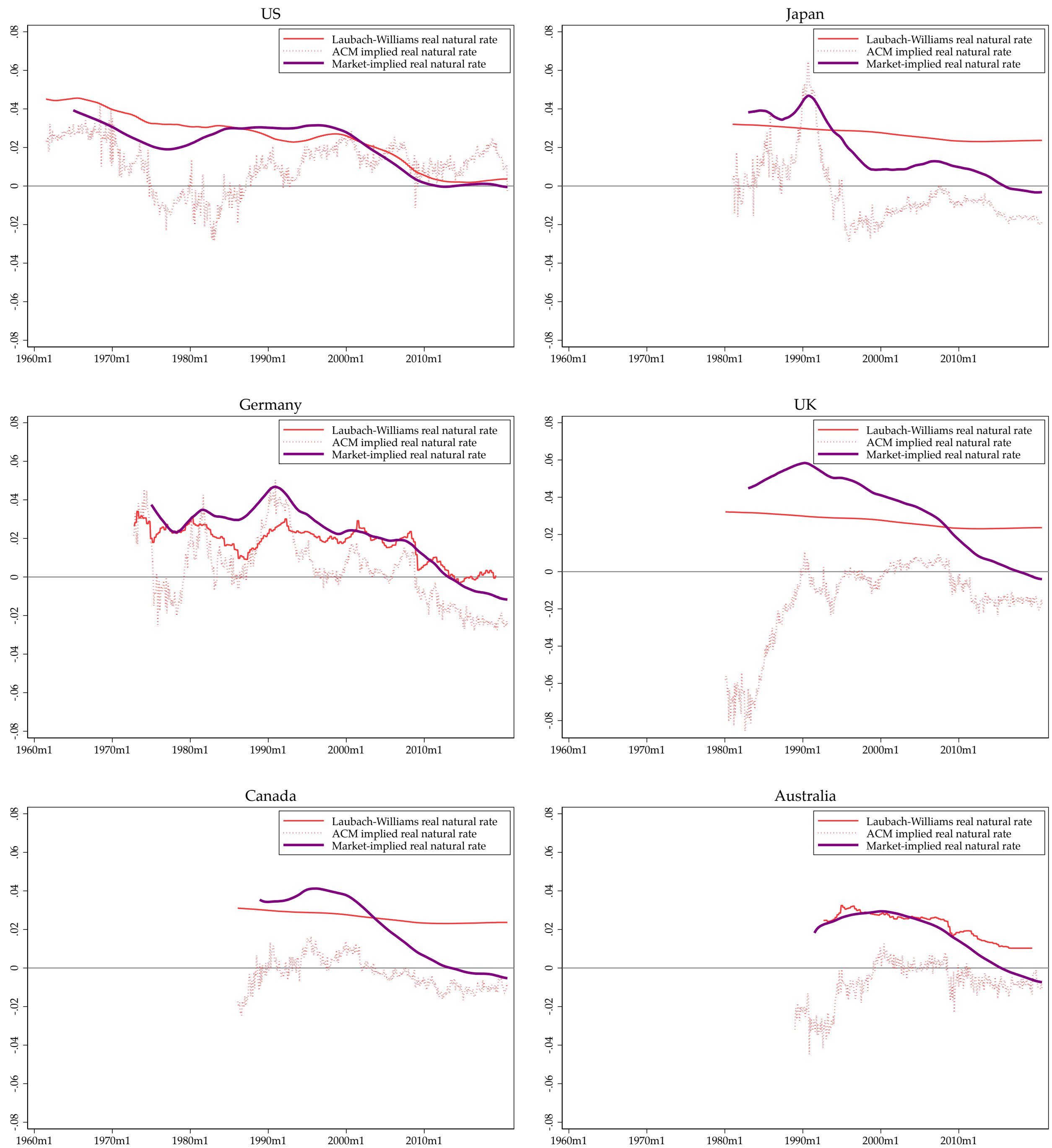
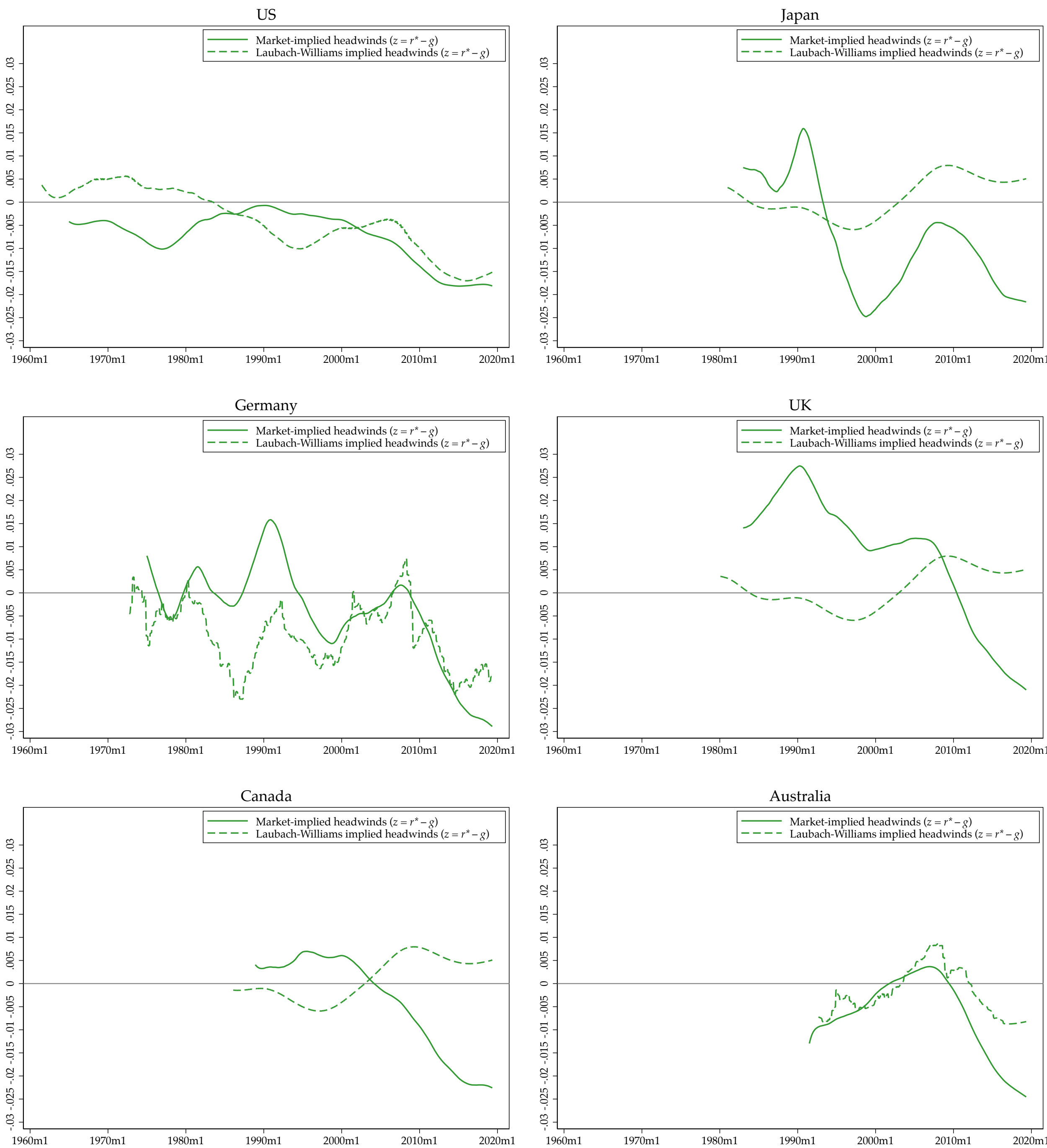


Figure 9: *Market-implied headwinds factor z in international data*

This chart displays our market-implied estimates of the headwinds factor $z = r* - g$ for six countries. Our headwinds estimate drops less steeply than the LW-implied headwinds for the U.S. but drops more steeply in the other five cases.



References

Adrian, Tobias, Richard K. Crump, and Emanuel Moench. 2013. Pricing the term structure with linear regressions. *Journal of Financial Economics* 110(1): 110–138.

Adrian, Tobias, Richard K. Crump, and Emanuel Moench. 2015. Regression-based estimation of dynamic asset pricing models. *Journal of Financial Economics* 118(2): 211–244.

Andreasen, Martin M., Jens H.E. Christensen, and Glenn D. Rudebusch. 2019. Term Structure Analysis with Big Data: One-Step Estimation Using Bond Prices. *Journal of Econometrics* 212(1): 26–46.

Ang, Andrew, and Monika Piazzesi. 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50(4): 745–787.

Bauer, Michael D., and Glenn D. Rudebusch. 2019. Interest Rates Under Falling Stars. Working Paper Series 2017-16, Federal Reserve Bank of San Francisco.

Bernanke, Ben. 2015. Why are interest rates so low, part 4: Term premiums. Https://www.brookings.edu/blog/ben-bernanke/2015/04/13/why-are-interest-rates-so-low-part-4-term-premiums/.

Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas. 2008. An Equilibrium Model of "Global Imbalances" and Low Interest Rates. *American Economic Review* 98(1): 358–93.

Campbell, John Y., and Robert J. Shiller. 1987. Cointegration and Tests of Present Value Models. *Journal of Political Economy* 95(5): 1062–1088.

Campbell, John Y., and Samuel B. Thompson. 2008. Predicting Excess Stock Returns out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies* 21(4): 1509–1531.

Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio. 2016. Demographics and real interest rates: Inspecting the mechanism. *European Economic Review* 88: 208 – 226. SI: The Post-Crisis Slump.

Cieslak, Anna, and Pavol Povala. 2015. Expected returns in Treasury bonds. *Review of Financial Studies* 28(10): 2859–2901.

Cochrane, John H., and Monika Piazzesi. 2005. Bond risk premia. *American Economic Review* 95(1): 138–160.

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti. 2017. Safety, Liquidity, and the Natural Rate of Interest. *Brookings Papers on Economic Activity* 48(1 (Spring): 235–316.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright. 2007. The US Treasury yield curve: 1961 to the present. *Journal of Monetary Economics* 54(8): 2291–2304.

Holston, Kathryn, Thomas Laubach, and John C. Williams. 2017. Measuring the natural rate of interest: International trends and determinants. *Journal of International Economics* 108: S59–S75. 39th Annual NBER International Seminar on Macroeconomics.

Jordà, Òscar, and Alan M. Taylor. 2019. Riders on the Storm. In *Challenges for Monetary Policy*, Proceedings of a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyo., August 23–24, 2019, 17–59. Kansas City, Mo.: Federal Reserve Bank of Kansas City.

Kim, Don H., and Jonathan H. Wright. 2005. 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, Board of Governors of the Federal Reserve System.

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

Laubach, Thomas, and John C. Williams. 2003. Measuring the natural rate of interest. *Review of Economics and Statistics* 85(4): 1063–1070.

Litterman, Robert B., and José Scheinkman. 1991. Common Factors Affecting Bond Returns. *Journal of Fixed Income* 1(1): 54–61.

Lubik, Thomas A., and Christian Matthes. 2015. Time-Varying Parameter Vector Autoregressions: Specification, Estimation, and an Application. *Federal Reserve Bank of Richmond Economic Quarterly* (4Q): 323–352.

Ludvigson, Sydney C., and Serena Ng. 2009. Macro factors in bond risk premia. *Review of Financial Studies* 22(12): 5027–5067.

McCririck, Rachael, and Daniel Rees. 2017. The Neutral Interest Rate. *RBA Bulletin* (September): 9–18.

Nelson, Charles, and Andrew F. Siegel. 1987. Parsimonious Modeling of Yield Curves. *Journal of Business* 60(4): 473–89.

Okazaki, Yosuke, and Nao Sudo. 2018. Natural Rate of Interest in Japan: Measuring its size and identifying drivers based on a DSGE model. Working Paper Series 18-E-6, Bank of Japan.

Piazzesi, Monika. 2010. Affine Term Structure Models. In *Handbook of Financial Econometrics: Tools and Techniques*, edited by Aït-Sahalia, Yacine, and Lars Peter Hansen, volume 1 of *Handbooks in Finance*, 691–766. San Diego: North-Holland.

Rachel, Lukasz, and Thomas D. Smith. 2017. Are Low Real Interest Rates Here to Stay? *International Journal of Central Banking* 13(3): 1–42.

Rachel, Łukasz, and Lawrence H. Summers. 2019. On Secular Stagnation in the Industrialized World. *Brookings Papers on Economic Activity* (Spring): 1–54.

Sarno, Lucio, Paul Schneider, and Christian Wagner. 2016. The economic value of predicting bond risk premia. *Journal of Empirical Finance* 37: 247–267.

Summers, Lawrence H. 2015. Have we Entered an Age of Secular Stagnation? IMF Fourteenth Annual Research Conference in Honor of Stanley Fischer, Washington, DC. *IMF Economic Review* 63(1): 277–280.

Svensson, Lars E.O. 1994. Estimating and Interpreting Forward Interest Rates: Sweden 1992–1994. NBER Working Paper 4871.

Wright, Jonathan H. 2011. Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset. *American Economic Review* 101(4): 1514–34.

A. Appendix

Bayesian estimation of the state-space model

We estimate the model using Bayesian methods. We run a random-walk Metropolis-Hastings (RWMH) algorithm on the model parameter vector for each country. This algorithm starts from a drawing from a prior distribution as specified in the next section, and builds a Markov chain to approximate the distribution of the posterior. At each step i of this algorithm, conditional on the parameter vector θ^{i-1} in the previous step, we accept a drawing ϑ with probability

$$\alpha\left(\vartheta|\theta\right) = \min\left\{\frac{p(Y|\vartheta)p(\vartheta)/q(\vartheta|\theta^{i-1})}{p(Y|\theta^{i-1})p(\theta^{i-1})/q(\theta^{i-1}|\vartheta)}\right\}, \tag{32}$$

where $p(Y|\vartheta)$ can be computed using the Kalman filter on the state-space representation of the model determined by equations Equation 24 and Equation 24. The RWMH version of the algorithm sets the proposal distribution to

$$q\left(\cdot|\theta^{i-1}\right) = N\left(\theta^{i-1}|\mathfrak{c}\cdot\widehat{\Sigma}\right). \tag{33}$$

Our algorithm sets the constant $\mathfrak{c} = 0.075$ and $\widehat{\Sigma}$ to a constant diagonal matrix equal to a diagonal matrix with entries calculated from the empirical variance of the drawn parameters for the first 1000 draws. The chosen value of $\mathfrak{c}$ yields an acceptance rate of 23.4 percent.

Our estimation is calculated with a number of simulations equal to 50,000, and assumes the same prior distribution for all countries. We run our RWMH in four parallel chains, which gives us a total of 200,000 iterations for the posterior distribution characterization.

For each country, we use as input information the inflation expectations measure π_t^* and the real GDP series. We obtain the trend GDP series by applying the HP filter over quarterly GDP data with a smoothing parameter equal to $25,600 = 1600\cdot 16$. The series thus obtained is interpolated to monthly data, and the trend growth rate g_t is then calculated. We use the time series of the estimated parameters of the yield curve (discussed below) to recover zero-coupon curves for all maturities ranging from 1 to 180 months, in a monthly grid. The time span covered is country-specific and determined by data availability.

Our identification assumption is r^* cointegrates with the trend GDP series, and that the difference has mean reversion properties compatible with a half-life within business cycle frequency, and not higher. This avoids the problem of r^* acting as a residual term that would capture high frequency oscillations in bond markets.

Prior specification

We specify tight priors for the parameters of the headwinds factor z_t, and relatively loose priors for other parameters in the model, as detailed in Table A.1. The prior distribution is common for all countries.

- The persistence parameter ρ_z is chosen to be close to unit root, but within the unitary circle to prevent an explosive solution. Its magnitude is taken to match a half life of around 60 months. This keeps the headwinds factor within business cycle frequency and not higher. In a freer setting, this term would act as a residual in the Kalman system and would tend to capture high frequency variations that are hard to identify given the data.

Table A.1: *Prior specification for model parameters*

Parameter	Distribution	Mean	Variance
ρ_z	Beta	0.997	2.00×10^{-6}
ρ_{rx}	Normal	1.00×10^{-12}	1.00×10^{-14}
ρ_y	Normal	0.9	0.025
σ_z	Log-Normal	1.00×10^{-4}	1.00×10^{-5}
σ_{rx}	Log-Normal	1.00×10^{-4}	1.00×10^{-4}
σ_y	Log-Normal	1.00×10^{-3}	5.00×10^{-4}
d_{r^*}	Normal	0.00	0.10
b_{r^*}	Normal	0.30	0.05
d_0	Normal	0.00	5.00×10^{-3}
a_y	Normal	0.00	7.50×10^{-4}
d_{cyc}	Normal	0.28	0.01
b_π	Normal	1.25	0.10
d_π	Normal	0.00	0.01
σ_{b_r}	Log-Normal	0.10	0.05
σ_{a_y}	Log-Normal	0.03	5.00×10^{-3}
σ_{d_r}	Log-Normal	0.01	7.50×10^{-3}
σ_{d_0}	Log-Normal	0.00	1.00×10^{-4}
$\sigma_{d_{cyc}}$	Log-Normal	0.10	0.01

- The volatility of the headwinds factor σ_z is set to match a relatively small range of variation for the frequency of the headwinds factors. While the drawing process allows for certain range around the 60 months business frequency, it imposes the view that this variable should not deviate largely from that horizon. Similarly, we impose relatively tight priors on σ_{rx} and σ_y
- The volatility terms σ_{b_r}, σ_{a_y}, σ_{d_r}, σ_{d_0}, and $\sigma_{d_{cyc}}$ are drawn from a log-normal distribution as it is standard in the literature, with relatively variances that make them largely uninformative.
- The model coefficients d_{r^*}, b_{r^*}, d_0, a_y, d_{cyc}, b_π, and d_π are set to have large variances with coefficients centered around values that calibrate the U.S. experience relatively well in a simple regression context.