A.1 Different short-term rates

Table A.1 shows that similar results hold using using different proxies for the short rate in equation (1.1) of the main text—i.e., using changes in 3-month, 6-month, or 2-year Treasury yields as the independent variable in equation (1.1). (The results in Table 1 of the text correspond to those reported in Table A.1 for the 1-year Treasury rate.) For all of these short rate proxies, the sensitivity of 10-year yields to changes in short rates was similar irrespective of frequency prior to 2000. After 2000, the sensitivity at high frequencies increases while the sensitivity at low frequencies declines.

A.2 Long-term private yields

Table A.2 shows that we obtain very similar results using a host of long-term private yields as the dependent variable in equation (1.1). We report results for both Aaa and Baa seasoned corporate bond yields from Moodys, the 10-year swap yield, and the yield on current coupon Fannie Mae MBS (FNCL). For of all these long-term private yields, the sensitivity to changes in 1-year Treasury rate is similar at high- and low- frequencies in the pre-2000 sample. Post-2000, the sensitivity at high frequencies increases while the sensitivity at low frequencies declines significantly.

A.3 Dating the break

Here we use an alternate procedure to date the break in the sensitivity of long-term rates to movements in short-term rates. As opposed to focusing simply on the break the low-frequency
Table A.1: Regressions of changes in long-term rates on short-term rates. This table reports the estimated regression coefficients from equation (1.1) in the main text for each reported sample. The dependent variable is the change in the 10-year nominal U.S. Treasury yield or forward rate, either nominal. The independent variable is alternately the change in the 3-month, 6-month, 1-year, and 2-year nominal U.S. Treasury yield. Changes are considered with daily data, and with monthly data using monthly \((h = 1)\) and annual \((h = 12)\) horizons. In the 1971-1999 monthly sample, time \(t\) runs from 1971m8 to 1999m12. In the 2000-2017 monthly sample, \(t\) runs from 2000m1 to 2017m12. For \(h > 1\), we report Newey-West (1987) standard errors in brackets, using a lag truncation parameter of \([1.5 \times h]\); for \(h = 1\), we report heteroskedasticity robust standard errors. Significance: *\(p < 0.1\), **\(p < 0.05\), ***\(p < 0.01\). Significance is computed using the asymptotic theory of Kiefer and Vogelsang (2005) which has better finite sample properties than traditional asymptotic theory.

<table>
<thead>
<tr>
<th></th>
<th>Pre-2000</th>
<th>Post-2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily</td>
<td>Monthly</td>
</tr>
<tr>
<td>3-month</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16***</td>
<td>0.26***</td>
<td>0.39***</td>
</tr>
<tr>
<td>[0.02]</td>
<td>[0.03]</td>
<td>[0.05]</td>
</tr>
<tr>
<td>6-month</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.41***</td>
<td>0.37***</td>
<td>0.46***</td>
</tr>
<tr>
<td>[0.02]</td>
<td>[0.03]</td>
<td>[0.06]</td>
</tr>
<tr>
<td>1-year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.56***</td>
<td>0.46***</td>
<td>0.56***</td>
</tr>
<tr>
<td>[0.02]</td>
<td>[0.04]</td>
<td>[0.05]</td>
</tr>
<tr>
<td>2-year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.65***</td>
<td>0.57***</td>
<td>0.68***</td>
</tr>
<tr>
<td>[0.01]</td>
<td>[0.04]</td>
<td>[0.05]</td>
</tr>
</tbody>
</table>

sensitivity, \(\beta_{12}\), here we seek to date the emergence of the frequency-dependent sensitivity of long-term rates. Specifically, we consider the following system of regression equations:

\[
y_{t+1}^{(10)} - y_{t}^{(10)} = a_{1}^{Pre} + \varphi_{1}^{Post} \cdot Post + \beta_{1}^{Pre} \cdot \left( y_{t+1}^{(1)} - y_{t}^{(1)} \right)
+ \Delta_{1}^{Post} \cdot \left( y_{t+1}^{(1)} - y_{t}^{(1)} \right) \times Post + \varepsilon_{t+1}
\] (A.1a)

\[
y_{t+12}^{(10)} - y_{t}^{(10)} = a_{12}^{Pre} + \varphi_{12}^{Post} \cdot Post + \beta_{12}^{Pre} \cdot \left( y_{t+12}^{(1)} - y_{t}^{(1)} \right)
+ \Delta_{12}^{Post} \cdot \left( y_{t+12}^{(1)} - y_{t}^{(1)} \right) \times Post + \varepsilon_{t+12}
\] (A.1b)

where \(Post\) is an indicator variable that switches of after some pre-specified break date. Thus, for \(h = 1\) and 12-month changes, the pre-break estimate of \(\beta_{h}\) is \(\beta_{h}^{Pre}\) and the post-break estimate of \(\beta_{h}\) is \(\beta_{h}^{Pre} + \Delta_{h}^{Post}\).

To date the emergence of the frequency-dependent sensitivity of long-term rates, we estimate equations (A.1a) and (A.1b) imposing the restriction that \(\beta_{1}^{Pre} = \beta_{12}^{Pre}\). For a given break date, we estimate this system of equations using the generalized method of moments and compute standard errors using a Newey-West variance-covariance matrix with a lag truncation parameter of 18 months. We use a standard two-step GMM estimator that uses 8 moment conditions—namely, the four least-squares normal equations for both equations (A.1a) and (A.1b)—to identify the 7
Table A.2: Regression of changes in corporate bond, swap and secondary mortgage market rates on short-term rates. This table reports the estimated slope coefficients from equation (1.1) in the main text for each reported sample. Specifically, the dependent variables are long-term corporate bond yields with Moody’s ratings of Baa and Aaa (labeled BAA and AAA), the 10-year swap yield (SWAP10), and the yield on current-coupon Fannie Mae mortgage-backed-securities (FNCL). The independent variable in all regressions is the change in the 1-year nominal Treasury yield. Changes are considered with daily data, and with monthly data using monthly \((h = 1)\), quarterly \((h = 3)\), semi-annual \((h = 6)\) and annual \((h = 12)\) horizons. We report Newey-West (1987) standard errors in brackets, using a lag truncation parameter of \(1.5 \times (h - 1)\) (rounded to the nearest integer). Significance: \(^* p < 0.1\), \(^{**} p < 0.05\), \(^{***} p < 0.01\). Significance is computed using the asymptotic theory of Kiefer and Vogelsang (2005) which has better finite sample properties than traditional asymptotic theory.

<table>
<thead>
<tr>
<th></th>
<th>(1) BAA</th>
<th>(2) AAA</th>
<th>(3) SWAP10</th>
<th>(4) FNCL</th>
<th>(5) BAA</th>
<th>(6) AAA</th>
<th>(7) SWAP10</th>
<th>(8) FNCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>0.48***</td>
<td>0.51***</td>
<td>0.74***</td>
<td>0.95***</td>
<td>0.55***</td>
<td>0.57***</td>
<td>0.94***</td>
<td>0.88***</td>
</tr>
<tr>
<td></td>
<td>[0.02]</td>
<td>[0.02]</td>
<td>[0.04]</td>
<td>[0.04]</td>
<td>[0.04]</td>
<td>[0.04]</td>
<td>[0.03]</td>
<td>[0.04]</td>
</tr>
<tr>
<td>Monthly</td>
<td>0.52***</td>
<td>0.59***</td>
<td>0.81***</td>
<td>0.92***</td>
<td>0.21</td>
<td>0.34***</td>
<td>0.81***</td>
<td>0.73***</td>
</tr>
<tr>
<td></td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.07]</td>
<td>[0.06]</td>
<td>[0.13]</td>
<td>[0.11]</td>
<td>[0.10]</td>
<td>[0.11]</td>
</tr>
<tr>
<td>Quarterly</td>
<td>0.49***</td>
<td>0.57***</td>
<td>0.80***</td>
<td>0.84***</td>
<td>0.04</td>
<td>0.21***</td>
<td>0.61***</td>
<td>0.52***</td>
</tr>
<tr>
<td></td>
<td>[0.06]</td>
<td>[0.06]</td>
<td>[0.07]</td>
<td>[0.07]</td>
<td>[0.12]</td>
<td>[0.07]</td>
<td>[0.09]</td>
<td>[0.09]</td>
</tr>
<tr>
<td>Yearly</td>
<td>0.41***</td>
<td>0.55***</td>
<td>0.67***</td>
<td>0.73***</td>
<td>-0.06</td>
<td>0.08</td>
<td>0.36***</td>
<td>0.31***</td>
</tr>
<tr>
<td></td>
<td>[0.08]</td>
<td>[0.09]</td>
<td>[0.08]</td>
<td>[0.10]</td>
<td>[0.12]</td>
<td>[0.06]</td>
<td>[0.05]</td>
<td>[0.06]</td>
</tr>
</tbody>
</table>

Figure A.1: Break test for the emergence of horizon-dependent sensitivity. This figure plots the Wald test statistic for each possible break date in equations (A.1a) and (A.1b) from a fraction 15% of the way through the sample to 85% of the way through the sample. The figure shows Wald test statistics for a break in equations (A.1a) and (A.1b) jointly, imposing the restriction that $\beta_{1}^{pre} = \beta_{12}^{pre}$. We estimate equations (A.1a) and (A.1b) jointly using a standard two-step efficient GMM estimator. The horizontal red dashed lines denote 10%, 5% and 1% critical values for the maximum of these Wald statistics as in Andrews (1993). Our Wald tests use a Newey and West (1987) variance matrix with a lag truncation parameter of 18. To address the tendency for tests based on the Newey-West variance estimator to over-reject in finite samples, we use the Cho and Vogelsang (2017) critical values for a null of no structural break. The Cho and Vogelsang (2017) critical values are based on the asymptotic theory of Kiefer and Vogelsang (2005) and are slightly larger than the traditional critical values from Andrews (1993).

Parameters of the system. For a given break date, we test the null that $\varphi_{1}^{post} = \Delta_{1}^{post} = \varphi_{12}^{post} = \Delta_{12}^{post} = 0$: this test is equivalent to asking whether there is a change in the low-frequency sensitivity ($\Delta_{12}^{post} \neq 0$) that differs from the change in the high-frequency sensitivity ($\Delta_{12}^{post} \neq \Delta_{1}^{post}$). To date the emergence of horizon-dependent sensitivity, we then use the test of Andrews (1993) who conducts a Chow (1960) test at all possible break dates and then takes the maximum of the Wald test statistics.

Figure ?? plots the Wald test statistic for each possible annual break date in equations (A.1a) and (A.1b) along with the Cho and Vogelsang (2017) critical values for a null of no structural break. Using this procedure, we find that the strongest evidence for emergence of horizon-dependent sensitivity is in 1999 or 2000.
Here we briefly discuss our GMM estimates of equations (A.1a) and (A.1b) using a break date of 2000 as in the main text. We obtain $\beta_{Pre}^1 = \beta_{Pre}^{12} = 0.54^{***} [0.05]$, $\Delta_{Post}^1 = 0.11 [0.14]$, and $\Delta_{Post}^{12} = -0.33^{***} [0.06]$. Hansen’s $J$-test does not reject the single over-identifying restriction, yielding $J = 1.07$ ($p$-val. = 0.32). However, the difference between $\beta_{Post}^1 = 0.64$ and $\beta_{Post}^{12} = 0.20$ is highly statistically significant ($p$-val. < 0.001).

Finally, we estimate equations (A.1a) and (A.1b) allowing $\beta_{Pre}^1$ and $\beta_{Pre}^{12}$ to differ and assuming a break date in 2000. In this case, our GMM estimates correspond to equation-by-equation OLS estimation of (A.1a) and (A.1b) as in Table 1 of the main text. (This has no effect on $\beta_{Post}^1$ and $\beta_{Post}^{12}$ relative to the constrained estimator that imposes $\beta_{Pre}^1 = \beta_{Pre}^{12}$, but it does affect the estimates of $\Delta_{Post}^1$ and $\Delta_{Post}^{12}$.) Using this unconstrained estimator, we obtain $\beta_{Pre}^1 = 0.46^{***} [0.05]$, $\beta_{Pre}^{12} = 0.56^{***} [0.05]$, $\Delta_{Post}^1 = 0.18 [0.14]$, and $\Delta_{Post}^{12} = -0.36^{***} [0.07]$ just as in Table 1 of the main text. Here the difference between $\beta_{Pre}^1$ and $\beta_{Pre}^{12}$ is economically small and is not statistically significant ($p$-val. = 0.13). However, the difference between $\beta_{Post}^1$ and $\beta_{Post}^{12}$ is economically large and is highly statistically significant ($p$-val. < 0.001).

### A.4 Trading strategies

As another way of assessing the resulting return predictability documented in Section 2.2 of the paper, we consider simple market-timing strategies in which an investor decides to take either a long or short position in the slope-mimicking portfolio—i.e., in a “curve steepener” trade—every month.

Specifically, we consider strategies that take a long (short) position in the slope-mimicking portfolio from month $t$ to month $t + 1$ if $L_t < L_{t-h}$ ($L_t > L_{t-h}$). Alternatively, we consider strategies that take a position in the slope-mimicking portfolio from month $t$ to month $t + 1$ that is proportional to $-(L_t - L_{t-h})$. Table ?? computes the annualized Sharpe ratios of these two trading strategies for different choices of $h$, in the pre- and post-2000 samples.

As shown in Table ?? the implied annualized Sharpe ratios for these strategies range between about 0.5 to 0.7 in the post-2000 sample but were negligible in the pre-2000 sample.

### A.5 Implications for affine term-structure models

Affine term-structure models (ATSMs) are a widely-used, reduced-form tools for understanding the term structure of bond yields. A standard discrete-time affine term-structure model (Duffee, 2002; Duffie and Kan, 1996) starts from the assumption that there is a $m \times 1$ state vector $x_t$ that follows a VAR(1) under the physical or P-measure:

$$x_t = \mu + \Phi x_{t-1} + \Sigma \varepsilon_t,$$

(A.2)
Table A.3: Sharpe ratios for slope-mimicking portfolios
This table reports the annualized Sharpe ratios since 2000 of the strategy of going long (short) the slope-mimicking portfolio if the level fell (rose) over the previous $h$ months and also the strategy of taking a position in the slope-mimicking portfolio that is proportional to $- (L_t - L_{t-h})$, and holding the position from $t$ to $t+1$. The position is rebalanced each month. Annualized Sharpe ratios are computed as the sample average monthly excess returns multiplied by $\sqrt{12}$ and divided by the standard deviation of those monthly excess returns.

<table>
<thead>
<tr>
<th>Strategy with $h$:</th>
<th>1</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$2 \times I(L_t - L_{t-h} &lt; 0) - 1$</td>
<td>0.22</td>
<td>0.03</td>
<td>-0.09</td>
<td>-0.00</td>
</tr>
<tr>
<td>$-(L_t - L_{t-h})$</td>
<td>0.08</td>
<td>-0.01</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>Post-2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$2 \times I(L_t - L_{t-h} &lt; 0) - 1$</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>$-(L_t - L_{t-h})$</td>
<td>0.52</td>
<td>0.62</td>
<td>0.68</td>
<td>0.64</td>
</tr>
</tbody>
</table>

where the error term is Gaussian with mean zero and identity variance-covariance matrix. The short-term riskless interest rate between time $t$ and $t+1$ is an affine function of the state vector:

$i_t = \delta_0 + \delta_1' x_t$. Meanwhile, the pricing kernel or stochastic discount factor is

$M_{t+1} = \exp(-i_t - \lambda_t^0 - \frac{1}{2} \lambda_t^0 \lambda_t), \quad (A.3)$

where the prices of factor risk, $\lambda_t = \lambda_0 + A_1 x_t$, are also an affine function of the state vector. The price of an $n$-period zero-coupon bond, $P_t^{(n)}$, satisfies the recursion

$P_t^{(n)} = E_t^P[M_{t+1}P_t^{(n-1)}] = \exp(-i_t)E_t^Q[P_t^{(n-1)}]. \quad (A.4)$

Here $E_t^P[\cdot]$ denotes expectations under the physical measure or $P$-measure and $E_t^Q[\cdot]$ denotes expectations under the risk-neutral pricing measure or $Q$-measure. (For any random variable $X_{t+1}$, $E_t^Q[X_{t+1}] = E_t^P[M_{t+1}X_{t+1}]/E_t^P[M_{t+1}]$.) Under the $Q$-measure, the state variables evolve according to

$x_t = \mu^* + \Phi^* x_{t-1} + \Sigma \varepsilon_t, \quad (A.5)$

where $\mu^* = \mu - \Sigma \lambda_0$ and $\Phi^* = \Phi - \Sigma A_1$.

After extensive, but well-known algebra, it follows that

$P_t^{(n)} = \exp(a_{(n)} + b_{(n)}'x_t), \quad (A.6)$

where $a_{(n)}$ is a scalar and $b_{(n)}$ is an $m \times 1$ vector that satisfy the recursions:

$a_{(n+1)} = -\delta_0 + a_{(n)} + b_{(n)}' \mu^* + \frac{1}{2} b_{(n)}' \Sigma \Sigma' b_{(n)} \quad (A.7)$

$b_{(n+1)} = \Phi^* b_{(n)} - \delta_1, \quad (A.8)$
starting from \( a_{(1)} = -\delta_0 \) and \( b_{1} = -\delta_{(1)} \). The continuously compounded yield on an \( n \)-period zero-coupon bond, \( y_t^{(n)} \), is in turn given by
\[
y_t^{(n)} = -n^{-1} \log(P_t^{(n)}) = -n^{-1} a_{(n)} - n^{-1} b_{(n)}' x_t.
\] (A.9)

Applying the estimation methodology of Adrian et al. (2013), we fit affine term-structure models with monthly data using the first \( K \) principal components of 1- to 10-year yields as the state variables \( x_t \). We do this in the pre-2000 and post-2000 samples separately for \( K = 2 \) to 5.

We then take the estimated model parameters and work out the model-implied \( \beta_h \) regression coefficients. Specifically, let \( \Gamma(j) = E[(x_{t+j} - E[x_{t+j}]) (x_t - E[x_t])'] \) denote the autocovariance function of the state vector, which can be obtained from the equations \( \text{vec}(\Gamma(0)) = (I - \Phi \otimes \Phi)^{-1} \text{vec}(\Sigma \Sigma') \) and \( \Gamma(j) = \Phi^j \Gamma(0) \) for \( j \geq 1 \). The population coefficient in a regression of \( h \)-month changes in 120-month yields on \( h \)-month changes in 12-month yields is then
\[
\beta_h = \frac{E[(y_{t+h}^{(120)} - y_t^{(120)})(y_{t+h}^{(12)} - y_t^{(12)})]}{E[(y_{t+h}^{(12)} - y_t^{(12)})^2]} = \frac{1}{10} \frac{b_{(120)}' [2\Gamma(0) - \Gamma(h) - \Gamma(h)' \Gamma(h)] b_{(12)}'}{b_{(12)}' [2\Gamma(0) - \Gamma(h) - \Gamma(h)' \Gamma(h)] b_{(12)}}.
\] (A.10)

These model-implied regression coefficients are shown in Panel A of Table A.4. Even if we include a large number of factors as state variables, these standard affine models fail to match the low-frequency decoupling between short- and long-term yields that we observe in the post-2000 data.

We next consider an alternate affine model that augments the state vector \( x_t \) to include not just \( K \) principal components of yields, but also \( L - 1 \) additional lags of these principal components. Thus, by construction, the model is non-Markovian with respect to the filtration given by the current principal components. We furthermore treat these lagged principal components as “unspanned state variables.” In standard affine models, if the true model is known, one can obtain the full set of state variables by inverting an appropriate set of yields—i.e., the state variables are spanned by current yields. An unspanned state variable is a variable that is useful for forecasting future bond yields and returns but that has no impact on the current yield curve—i.e., it is not “spanned” by current yields—and cannot be recovered in this way. Formally, this means that if the first \( K \) elements of the state vector \( x_t \) are the current principal components, all but the first \( K \) elements of \( \delta_t \) are zero, and the upper right \( K \times K(L - 1) \) block of \( \Phi^* \) is a matrix of zeros. As Duffee (2011) explains, a state variable will be unspanned if it has perfectly offsetting effects on the evolution of future short rates and future term premia.

Economically speaking, this is a rather unusual model. However, it allows us to parsimoniously capture our key finding that past changes in the level of rates are useful for forecasting future yields. In addition, similar models have been considered in Joslin et al. (2013). To be clear, we do not
believe that the past increase in the level of rates is literally unspanned—i.e., that it has no effect on the current yield curve. Instead, we would argue that this variable is close to being unspanned. Specifically, like any factor that has a short-lived impact on bond risk premia, past increases in the level of rates should have only a small effect on the yield curve. Thus, in practice it will likely be quite difficult to recover information about this variable from current yields alone—e.g., because yields are measured with a tiny amount of error or because the true data-generating model evolves over time—so conditioning on this variable will add information beyond that revealed by current yields.

Again, the parameters of this augmented model can be estimated, imposing the restriction that the lagged principal components are unspanned factors as in Adrian et al. (2013). The model-implied $\beta_h$ regression coefficients can again be derived from equation (A.10). These model-implied coefficients are shown in Panel B of Table A.4. The augmented model is able to get reasonably close to matching the empirical regression coefficients at both high- and low-frequencies and in both samples.

### Table A.4: Affine Term Structure Model-Implied coefficients in regression of monthly/yearly changes in 10-year yields on changes in 1-year yields

This table reports the slope coefficients in equation (A.10) corresponding to the parameters in an affine term structure model estimated as proposed by Adrian et al. (2013) over August 1971-December 2000 and January 2001-December 2017 subsamples at monthly ($h = 1$) and yearly ($h = 12$) frequencies. The term structure model uses $K$ principal components of yields as state variables in panel A, and adds $L−1$ additional lags of these principal components (for a total of $LK$ state variables) in panel B. $p$-values are also reported; these are two-sided bootstrap $p$-values comparing the sample value of $\beta_{12}/\beta_1$ with the bootstrap distribution using that affine model. As memo items the results of the regressions using actual yields are included—these are simply transcribed from Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Pre-2000</th>
<th></th>
<th></th>
<th>Post-2000</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>$\beta_{12}$</td>
<td>$p$-value</td>
<td>$\beta_1$</td>
<td>$\beta_{12}$</td>
<td>$p$-value</td>
</tr>
<tr>
<td>Panel A: ATSM with $K$ principal components of yields as factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 2$</td>
<td>0.42</td>
<td>0.49</td>
<td>0.89</td>
<td>0.83</td>
<td>0.74</td>
<td>0.000</td>
</tr>
<tr>
<td>$K = 3$</td>
<td>0.46</td>
<td>0.52</td>
<td>0.65</td>
<td>0.80</td>
<td>0.59</td>
<td>0.006</td>
</tr>
<tr>
<td>$K = 4$</td>
<td>0.47</td>
<td>0.52</td>
<td>0.62</td>
<td>0.74</td>
<td>0.55</td>
<td>0.005</td>
</tr>
<tr>
<td>$K = 5$</td>
<td>0.47</td>
<td>0.52</td>
<td>0.52</td>
<td>0.71</td>
<td>0.49</td>
<td>0.009</td>
</tr>
<tr>
<td>Panel B: ATSM with $L−1$ lags as additional unspanned factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 2$, $L = 6$</td>
<td>0.42</td>
<td>0.50</td>
<td>0.84</td>
<td>0.85</td>
<td>0.32</td>
<td>0.59</td>
</tr>
<tr>
<td>$K = 3$, $L = 6$</td>
<td>0.46</td>
<td>0.54</td>
<td>1.00</td>
<td>0.81</td>
<td>0.34</td>
<td>0.42</td>
</tr>
<tr>
<td>$K = 2$, $L = 12$</td>
<td>0.42</td>
<td>0.51</td>
<td>0.76</td>
<td>0.90</td>
<td>0.25</td>
<td>0.87</td>
</tr>
<tr>
<td>$K = 3$, $L = 12$</td>
<td>0.46</td>
<td>0.55</td>
<td>0.90</td>
<td>0.85</td>
<td>0.29</td>
<td>0.78</td>
</tr>
<tr>
<td>Memo: Estimates in data (from Table 1)</td>
<td>0.46</td>
<td>0.56</td>
<td>0.64</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As another way of looking at this, we use a bootstrap to test the hypothesis that each estimated affine term structure model is correctly specified. Our test uses the ratio of yearly to monthly coefficients ($\beta_{12}/\beta_1$) as the test statistic; the test rejects if this ratio is too low or too high to have
been generated by the estimated Q-measure model. This is a test in the spirit of Giglio and Kelly (2018), who test the hypothesis that a given affine model is correctly specified by checking whether the comovement of rates at different points on the term structure is consistent with the estimated Q-measure dynamics.

To implement this test, we simulate the bootstrap distribution of the ratio $\beta_{12}/\beta_1$ for each estimated affine term-structure model. To do this for a given model, we first generate bootstrapped time series of the state vector $\mathbf{x}$, using a residual-based bootstrap based on our estimates of the P-measure dynamics in equation (A.2). Using our estimates of $a(n)$ and $b(n)$ for that same model—which reflect the estimated dynamics under the Q-measure—in combination with a set of bootstrapped draws of the corresponding yield measurement errors, we obtain a bootstrapped time series of yields $y_{t(n)}$ for $n = 12$ and 120 months. We then compute $\beta_{12}$, $\beta_1$, and the ratio $\beta_{12}/\beta_1$ for each bootstrapped time series. Repeating this exercise many times, we obtain a bootstrap distribution for the ratio $\beta_{12}/\beta_1$.

For each model, we then compute the two-sided p-value of the observed ratio $\beta_{12}/\beta_1$ with respect to this bootstrap distribution. These p-values are also reported in Table A.4. We find that in the pre-2000 sample, none of the models are rejected. In the post-2000 sample, the Markovian models are decisively rejected: if these standard models were correctly specified it would be highly unlikely to observe a value of $\beta_{12}/\beta_1$ as small as we do in the data. However, the non-Markovian models are not rejected in the post-2000 sample. In this way, we again conclude that a non-Markovian term structure model is required to match the yield-curve dynamics in the post-2000 data.

To summarize, our conclusion is that affine term-structure models need to include lagged yield-curve factors to match the frequency-dependent sensitivity of long-term rates we observe in recent years. A large number of static yield curve factors will not do the job.

B Model solution

This Appendix provides additional details on our economic model. In particular, we provide additional details on how we solve for the rational expectations equilibrium of our model.

B.1 Long-term nominal bonds

In this subsection, we derive the Campbell-Shiller (1988) approximation to the return on a default-free perpetuity.

Consider a perpetual default-free nominal bond which pay a nominal coupon of $K$ each period. Let $P_t$ denote the nominal price of the long-term bond at time $t$. Thus, the nominal return on the long-term bond from $t$ to $t+1$ is

$$1 + R_{t+1} = \frac{P_{t+1} + K}{P_t}.$$
Defining $\theta \equiv 1/(1 + K) < 1$, the one-period log return on the bond from time $t$ to $t+1$ is approximately

$$r_{t+1} \equiv \ln(1 + R_{t+1}) \approx \frac{D}{1 - \theta} y_t - \frac{D - 1}{1 - \theta} y_{t+1}, \quad (A.11)$$

where $y_t$ is the log yield-to-maturity at time $t$ and

$$D = \frac{1}{1 - \theta} = \frac{K + 1}{K} \quad (A.12)$$

is the Macaulay duration when the bond is trading at par. The log-linear approximation for default-free coupon-bearing bonds in equation (A.11) appears in Chapter 10 of Campbell et al. (1996).

To derive this approximation, note that the Campbell-Shiller (1988) approximation of the 1-period log return on the long-term bond is

$$r_{t+1} = \ln(P_{t+1} + K) - p_t \approx \varphi + \theta p_{t+1} + (1 - \theta) k - p_t, \quad (A.13)$$

where $p_t = \log(P_t)$ is the log price, $k = \log(K)$ is the log coupon, and where $\theta = 1/(1 + \exp(k - \bar{p}))$ and $\varphi = -\log(\theta) - (1 - \theta) \log(\theta^{-1} - 1)$ are parameters of the log-linearization. Iterating equation (A.13) forward, we find that the log bond price is

$$p_t = (1 - \theta)^{-1} \varphi + k - \sum_{i=0}^{\infty} \theta^i E_t[r_{t+i+1}] \quad (A.14).$$

Applying this approximation to the yield-to-maturity, defined as the constant return that equates bond price and the discounted value of promised cashflows, we obtain

$$p_t = (1 - \theta)^{-1} \varphi + k - (1 - \theta)^{-1} y_t. \quad (A.15)$$

Equation (A.11) then follows by substituting the expression for $p_t$ in equation (A.15) into the Campbell-Shiller return approximation in equation (A.13).

Assuming the steady-state price of the bonds is par ($\bar{p} = 0$), we have $\theta = 1/(1 + K)$. Thus, bond duration is $D = -\partial p_t/\partial y_t = (1 - \theta)^{-1} = (1 + K)/K$. Since $-\partial p_t/\partial y_t = - (\partial P_t/\partial Y_t)(1 + Y_t)/P_t = (Y_t' + 1)/Y_t$ this corresponds to Macaulay duration when the bonds are trading at par ($Y_t = K$).

Let $i_t$ denote the interest rate on short-term nominal bonds from $t$ to $t+1$ and let $r x_{t+1} \equiv r_{t+1} - i_t$ denote the excess return on long-term nominal bonds over short-term nominal bonds from $t$ to $t+1$. Then, iterating equation (A.11) forward and taking expectations, the yield on long-term nominal bonds is given by:

$$y_t = (1 - \theta) \sum_{j=0}^{\infty} \theta^j E_t[r_{t+j} + rx_{t+j+1}] \quad (A.16).$$
B.2 Market participants

There are two types of risk-averse arbitrageurs in the model, each with identical risk tolerance $\tau$, who differ solely in the frequency with which they can rebalance their bond portfolios.

The first group of arbitrageurs are fast-moving arbitrageurs who are free to adjust their holdings of long-term and short-term bonds each period. Fast-moving arbitrageurs are present in mass $q$ and we denote their demand for long-term bonds at time $t$ by $b_t$. Fast-moving arbitrageurs have mean-variance preferences over 1-period portfolio log returns. Thus, their demand for long-term bonds at time $t$ is given by

$$b_t = \tau \frac{E_t [r_{x_{t+1}}]}{\text{Var}_t [r_{x_{t+1}}]}, \quad (A.17)$$

where

$$r_{x_{t+1}} \equiv r_{t+1} - i_t = \frac{1}{1-\theta} y_t - \frac{\theta}{1-\theta} y_{t+1} - i_t \quad (A.18)$$

is the excess return on long-term bonds from $t$ to $t+1$.

The second group of arbitrageurs is a set of slow-moving arbitrageurs who can only adjust their holdings of long-term and short-term bonds every $k$ periods. Slow-moving arbitrageurs are present in mass $1-q$. A fraction $1/k$ of these slow-moving arbitrageurs is active each period and can reallocate their portfolios. However, they must then maintain this same portfolio allocation for the next $k$ periods. As in Duffie (2010), this is a reduced-form way to model the frictions that limit the speed of capital flows. Since they only rebalance their portfolios every $k$ periods, slow-moving arbitrageurs have mean-variance preferences over their $k$-period cumulative portfolio excess return. Thus, the demand for long-term bonds from the subset of slow-moving arbitrageurs who are active at time $t$ is given by

$$d_t = \tau \frac{E_t [\sum_{j=1}^{k} r_{x_{t+j}}]}{\text{Var}_t [\sum_{j=1}^{k} r_{x_{t+j}}]}, \quad (A.19)$$

B.3 Risk factors

Short-term nominal interest rates: Short-term nominal bonds are available in perfectly elastic supply. At time $t$, arbitrageurs learn that short-term bonds will earn a riskless log return of $i_t$ in nominal terms between time $t$ and $t+1$. We assume that the short-term nominal interest rate is the sum of a highly persistent component $i_{P,t}$ and a more transient component $i_{T,t}$:

$$i_t = i_{P,t} + i_{T,t}. \quad (A.20)$$

We assume that the persistent component $i_{P,t}$ follows an exogenous AR(1) process:

$$i_{P,t+1} = \bar{i} + \rho_P (i_{P,t} - \bar{i}) + \varepsilon_{P,t+1}, \quad (A.21)$$
where $0 < \rho_P < 1$ and $\text{Var}_t[\varepsilon_{P,t+1}] = \sigma_P^2$. Similarly, we assume that the transient component $i_{T,t}$ follows an exogenous AR(1) process:

$$i_{T,t+1} = \rho_T i_{T,t} + \varepsilon_{T,t+1}, \quad (A.22)$$

where $0 < \rho_T \leq \rho_P < 1$ and $\text{Var}_t[\varepsilon_{T,t+1}] = \sigma_T^2$.

**Supply of long-term bonds:** We assume that the long-term nominal bond is available in an exogenous, time-varying net supply $s_t$ that must be held in equilibrium by fast arbitrageurs and slow-moving arbitrageurs. This net supply equals the gross supply of long-term bonds minus the demand for long-term bonds from any unmodeled agents who have inelastic demand for these bonds. Formally, we assume that $s_t$ follows an AR(1) process:

$$s_{t+1} = \bar{s} + \rho_s (s_t - \bar{s}) + \varepsilon_{s,t+1} + C\varepsilon_{P,t+1} + C\varepsilon_{T,t+1}, \quad (A.23)$$

where $0 < \rho_s < 1$, $C > 0$, and $\text{Var}_t[\varepsilon_{s,t+1}] = \sigma_s^2$. The $\varepsilon_{s,t+1}$ shocks capture other forces that impact the net supply of long-term bonds. We have made no assumptions about the correlation structure among the three underlying shocks $\varepsilon_{P,t+1}$, $\varepsilon_{T,t+1}$, and $\varepsilon_{s,t+1}$. Indeed, the model can be solved for any arbitrary correlation structure among these shocks.

To understand the implied process for net bond supply, let $L$ denote the time-series lag operator and note

$$(1 - \rho_s L) (s_t - \bar{s}) = \varepsilon_{s,t} + C\varepsilon_{P,t} + C\varepsilon_{T,t} \quad (A.24)$$

$$= \varepsilon_{s,t} + C (1 - \rho_P L) (i_{P,t} - \bar{i}) + C (1 - \rho_T L) i_{T,t}.$$

Working out the lag polynomial, we see that

$$s_t = \bar{s} + C[(i_{P,t} - \bar{i}) - (\rho_P - \rho_s) \sum_{j=0}^{\infty} \rho_P^j (i_{P,t-j-1} - \bar{i})]$$

$$+ C[i_{T,t} - (\rho_T - \rho_s) \sum_{j=0}^{\infty} \rho_T^j i_{T,t-j-1}] + \sum_{j=0}^{\infty} \rho_s^j \varepsilon_{s,t-j}. \quad (A.25)$$

which follows from the fact that

$$(1 - \rho_s L)^{-1} (1 - \rho_x L) x_t = \sum_{j=0}^{\infty} \rho_s^j L^j (1 - \rho_x L) x_t$$

$$= x_t - \rho_x x_{t-1} + \rho_x \rho_x x_{t-2} + \rho_x^2 x_{t-2} - \rho_x^3 x_{t-2} + \cdots$$

$$= x_t - (\rho_x - \rho_s) \sum_{j=0}^{\infty} \rho_s^j x_{t-j-1}. \quad (A.26)$$

**B.4 Equilibrium Conjecture**

For the sake of concreteness, suppose that $k = 4$. We conjecture that equilibrium yields take the form

$$y_t = \alpha_0 + \alpha'_1 x_t, \quad (A.27)$$
and that the demands of active slow-moving arbitrageurs are of the form

\[ d_t = \delta_0 + \delta_1' x_t, \quad (A.28) \]

where the \( k + 2 \) dimensional state vector is

\[ x_t = \begin{bmatrix} i_{P,t} - \bar{i} \\ i_{T,t} \\ s_t - \bar{s} \\ d_{t-1} - \delta_0 \\ d_{t-2} - \delta_0 \\ d_{t-3} - \delta_0 \end{bmatrix}. \quad (A.29) \]

These assumptions imply that the state vector follows an AR(1) process. Critically, the transition matrix \( \Gamma \) is a function of the parameters \( \delta_1 \) governing slow-moving arbitrageur demand so we write \( \Gamma = \Gamma(\delta_1) \). Specifically, we have

\[ x_{t+1} = \Gamma(\delta) x_t + \epsilon_{t+1} \quad (A.30) \]

where \( \Sigma \equiv Var[\epsilon_{t+1}] \). Assuming for simplicity that \( \epsilon_{P,t+1}, \epsilon_{T,t+1}, \) and \( \epsilon_{s,t+1} \) are mutually orthogonal, we have

\[ \Sigma = \begin{bmatrix} \sigma_P^2 & 0 & C\sigma_P^2 & 0 & 0 & 0 \\ 0 & \sigma_T^2 & C\sigma_T^2 & 0 & 0 & 0 \\ C\sigma_P^2 & C\sigma_T^2 & \sigma_s^2 + C^2\sigma_P^2 + C^2\sigma_T^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_P^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_T^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (A.31) \]

We adopt the convention that \( e \) is the vector with a 1 corresponding to \( i_{P,t} - \bar{i} \) and \( i_{T,t} \) and 0s elsewhere, i.e., \( e = [1 \ 1 \ 0 \ 0 \ 0 \ 0]' \); \( e_s \) is the basis vector with a 1 corresponding to \( s_t - \bar{s} \) and 0s elsewhere, i.e., \( e_s = [0 \ 0 \ 1 \ 0 \ 0 \ 0]' \); and \( e_d \) is a 1 corresponding to \( d_{t-1} - \delta_0, d_{t-2} - \delta_0, \cdots, d_{t-(k-1)} - \delta_0 \) and 0s elsewhere, i.e., \( e_d = [0 \ 0 \ 0 \ 1 \ 1 \ 1]' \).

Finally, we denote \( C_{[t+i,t+j]} = Cov[ x_{t+i}, x_{t+j} | x_t ] \) and note that

\[ C_{[t+i,t+j]} = \sum_{s=1}^{\min(i,j)} [\Gamma^s]_2 [\Sigma]_2 [\Gamma^{j-s}]_2', \quad (A.32) \]

so \( C_{[t+i,t+j]} = C'_{[t+i,t+j]} \).
B.5 Arbitrageur demands

Fast-moving arbitrageurs’ demand: Given this conjecture, we work out fast-moving arbitrageurs’ demand for long-term bonds. Given the conjectured form of equilibrium yields, the realized 1-period excess returns on long bonds from \( t \) to \( t + 1 \) is

\[
rx_{t+1} = \frac{1}{1 - \theta} y_{t} - \frac{\theta}{1 - \theta} y_{t+1} - i_{t} \\
= (\alpha_{0} - \bar{r}) + \left( \frac{1}{1 - \theta} \alpha_{1} - \bar{e} \right) x_{t} - \left( \frac{\theta}{1 - \theta} \alpha_{1} \right) x_{t+1},
\]

which implies

\[
E_{t}[rx_{t+1}] = (\alpha_{0} - \bar{r} - \bar{\pi}) + \left( \frac{1}{1 - \theta} \alpha_{1} - \bar{e} \right) x_{t} - \left( \frac{\theta}{1 - \theta} \alpha_{1} \right) \Gamma x_{t}
\]

and

\[
Var_{t}[rx_{t+1}] = \left( \frac{\theta}{1 - \theta} \right)^{2} \alpha' \Sigma \alpha
\]

Thus, fast-moving arbitrageurs’ demand for long-term bonds is

\[
b_{t} = \tau \frac{E_{t}[rx_{t+1}]}{Var_{t}[rx_{t+1}]} = \left[ \tau \frac{\alpha_{0} - \bar{r}}{\left( \frac{1}{1 - \theta} \alpha_{1} - \bar{e} \right) x_{t} - \left( \frac{\theta}{1 - \theta} \alpha_{1} \right) \Gamma x_{t}} \right] x_{t}.
\]

Slow-moving arbitrageurs’ demand: We next work out slow-moving arbitrageurs’ demand for long-term bonds. Given our conjecture, the realized \( k \)-period cumulative excess returns on long bonds from \( t \) to \( t + k \) is

\[
\sum_{j=1}^{k} rx_{t+j} = \sum_{j=0}^{k-1} (y_{t+j} - i_{t+j}) - \frac{\theta}{1 - \theta} (y_{t+k} - y_{t})
\]

Thus, expected \( k \)-period cumulative returns are

\[
E_{t}[\sum_{j=1}^{k} rx_{t+j}] = k (\alpha_{0} - \bar{r}) + \left( (\alpha_{1} - \bar{e})'(I - \Gamma)^{-1} + \frac{\theta}{1 - \theta} \alpha_{1}' \Gamma \right) x_{t},
\]

and the variance of \( k \)-period cumulative excess returns is

\[
Var_{t}[\sum_{j=1}^{k} rx_{t+j}] = Var_{t}[\left( (\alpha_{1} - \bar{e})'(\sum_{j=0}^{k-1} x_{t+j}) - \left( \frac{\theta}{1 - \theta} \right) \alpha_{1}' x_{t+k} \right)]
\]

\[
= (\alpha_{1} - \bar{e})' \left( \sum_{l=1}^{k-1} \sum_{j=1}^{k-1} C_{[t+l,t+j]} \right) (\alpha_{1} - \bar{e}) + \left( \frac{\theta}{1 - \theta} \right)^{2} \alpha_{1}' C_{[t+k,t+k]} \alpha_{1}
\]

\[- 2 \left( \frac{\theta}{1 - \theta} \right) (\alpha_{1} - \bar{e})' \sum_{j=1}^{k-1} C_{[t+j,t+k]} \alpha_{1}.
\]

Slow-moving arbitrageurs’ demand long long-term bonds is

\[
d_{t} = \tau \frac{E_{t}[\sum_{j=1}^{k} rx_{t+j}]}{Var_{t}[\sum_{j=1}^{k} rx_{t+j}]}.
\]
Thus, given our conjectures, slow-moving arbitrageurs demands will indeed take a linear form. Specifically, we have
\[ \delta_0 = \tau \frac{k (\alpha_0 - \overline{i})}{V^{(k)}} \] (A.41)
where \( V^{(k)} = Var_x \sum_{j=1}^{k} r_{x+t+j} \) and
\[ \delta_1' = \tau \left( \frac{(\alpha_1 - e)' (I - \Gamma)^{-1} + \frac{\varrho}{1-\varrho} \alpha_1'}{V^{(k)}} \right) (I - \Gamma^k) \] (A.42)

### B.6 Equilibrium solution

To solve for the equilibrium, we need to clear the market for bonds in a way that is consistent with optimization on the part of fast-moving arbitrageurs and slow-moving arbitrageurs. The market-clearing condition is
\[ \text{Active demand} = \text{Active supply} \]
\[ (1 - q)k^{-1} d_t + q b_t = s_t - (1 - q)(k^{-1} \sum_{i=1}^{k-1} d_{t-i}) . \] (A.43)

Letting \( V^{(1)} = Var_x [r_{x+t+1}] = \left( \frac{\varrho}{1-\varrho} \right)^2 \alpha_1' \Sigma \alpha_1 \), denote the variance of 1-period excess returns, active demand is
\[ (1 - q)k^{-1} d_t + q b_t = \left[ (1 - q)k^{-1} \delta_0 + q \tau \left( \frac{\alpha_0 - \overline{i}}{V^{(1)}} \right) \right] + \left[ (1 - q)k^{-1} \delta_1' + q \tau \left( \frac{\frac{1}{1-\varrho} \alpha_1 - e}' - \frac{\varrho}{1-\varrho} \alpha_1' \Gamma}{V^{(1)}} \right) \right] x_t \] (A.44)

Active supply is
\[ s_t - (1 - q)k^{-1} \sum_{i=1}^{k-1} d_{t-i} = \left[ \overline{s} - (1 - q) \left( \frac{k-1}{k} \right) \delta_0 \right] + \left[ (e_s - (1 - q)k^{-1} e_d)' \right] x_t . \] (A.45)

Matching constants terms, we obtain
\[ \alpha_0 = \overline{i} + \frac{V^{(1)}}{\tau q} (\overline{s} - (1 - q)\delta_0) \] (A.46)

Matching slope coefficients, we have
\[ \alpha_1 = (1 - \theta) \left[ I - \theta \Gamma' \right]^{-1} e + (1 - \theta) \frac{V^{(1)}}{\tau q} \left[ I - \theta \Gamma' \right]^{-1} [e_s - k^{-1}(1 - q) (1_d + \delta_1)] \] (A.47)
\[ = \frac{1 - \theta}{1 - \theta \rho_P} e_P + \frac{1 - \theta}{1 - \theta \rho_T} e_T + \frac{V^{(1)}}{\tau q} \left[ \frac{1 - \theta}{1 - \theta \rho_s} e_s - k^{-1}(1 - q) (1 - \theta) [I - \theta \Gamma']^{-1} (1_d + \delta_1) \right] \]

where \( e_P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}' \) and \( e_T = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}' \).
Thus, equilibrium yields take the form

\[ y_t = \alpha_0 + \alpha_1' x_t \]  

(A.48)

\[
= \tilde{z} + \frac{1 - \theta}{1 - \theta \rho_p} (i_{p,t} - \tilde{z}) + \frac{1 - \theta}{1 - \theta \rho_T} i_{p,t}
\]

Expected future short real rates

\[
+ \frac{V^{(1)}(\tau q)}{\tau q} (\bar{s} - (1 - q)\delta_0)
\]

Unconditional term premia

\[
+ \left[ \frac{V^{(1)}(\tau q)}{\tau q} \left( \frac{1 - \theta}{1 - \theta \rho_s} (s_t - \bar{s}) - (1 - \theta) (1 - q) k^{-1} (e_d + \delta_1)' [I - \theta \Gamma]^{-1} x_t \right) \right].
\]

Conditional term premia

Equilibrium excess returns are given by

\[
E_t [r x_{t+1}] = \frac{V^{(1)}(\tau q)}{\tau q} (\bar{s} - (1 - q)\delta_0) + \frac{V^{(1)}(\tau q)}{\tau q} \left[ (s_t - \bar{s}) - (1 - q) k^{-1} (1_d + \delta_1)' x_t \right]
\]

(A.49)

**B.7 Equilibrium existence and uniqueness**

A rational expectations equilibrium of our model is a fixed point of a specific operator involving the “price-impact” coefficients, \((\alpha_1')\), which show how bond supply and inactive slow-moving arbitrageur demand impact bond yields, and the “demand-impact” coefficients, \((\delta_1')\), which show how bond supply and inactive demand impact the demand of active slow-moving arbitrageurs. Specifically, let \(\omega = (\alpha_1', \delta_1')'\) and consider the operator \(f(\omega_0)\) which gives (i) the price-impact coefficients that will clear the market for long-term bonds and (ii) the demand-impact coefficients consistent with optimization on the part of active slow-moving arbitrageurs when arbitrageurs conjecture that \(\omega = \omega_0\) at all future dates. Thus, a rational expectations equilibrium of our model is a fixed point \(\omega^* = f(\omega^*)\).

Specifically, an equilibrium solves the following system of equations

\[
\alpha_1 = (1 - \theta) \left[ I - \theta \Gamma(\delta_1) \right]^{-1} \left[ e + \frac{V^{(1)}(\alpha_1)}{\tau q} (e_d - k^{-1} (1 - q) (e_d + \delta_1)) \right]
\]

(A.50)

and

\[
\delta_1' = \tau - \frac{\left( (\alpha_1 - e_d)' (I - \Gamma(\delta))^{-1} + \frac{\theta}{1 - \rho_\delta} \alpha_1' \right) (I - \Gamma(\delta_1)^k)}{V^{(k)}(\alpha_1, \delta_1)}
\]

(A.51)

where we write \(V^{(1)}(\alpha_1)\) to emphasize that the 1-period return variance depends on \(\alpha_1\); \(\Gamma(\delta_1)\) to emphasize that the transition matrix depends on \(\delta_1\); and \(V^{(k)}(\alpha_1, \delta_1)\) to emphasize that the \(k\)-period return variance depends on \(\alpha_1\) and \(\delta_1\). We can write this system of non-linear equations more compactly as

\[
\alpha_1 = f_{\alpha_1}(\alpha_1, \delta_1) \quad \text{and} \quad \delta_1 = f_{\delta_1}(\alpha_1, \delta_1)
\]

(A.52)
or simply as \( \omega = f(\omega) \) where \( \omega = (\alpha'_1, \delta'_1)' \).

This is a system of \( 2(k+1) \) equations in \( 2(k+1) \) unknowns. However, in any rational expectations equilibrium of our model, bond yields always reflect the expected path of future short rates. As a result, equilibrium bond holdings do not depend on future short rates. Formally, it is easy to see that, in any equilibrium, active slow-moving arbitrageur demand does not depend on \( i_{P,t} \) and \( i_{T,t} \), so the first two elements of \( \delta'_1 \) are zeros and the first two elements are \( \alpha'_1 \) are \((1 - \theta) / (1 - \theta \rho_P)\) and \((1 - \theta) / (1 - \theta \rho_T)\), respectively. This implies that an equilibrium of our model is a solution to a system of \( 2k \) nonlinear equations in \( 2k \) unknowns. Specifically, we need to determine how equilibrium yields and active slow-moving demand respond to shifts in the supply of bonds: this generates \( 2 \) unknowns and \( 2 \) corresponding equations. We also need to determine how equilibrium yields and active slow-moving demand respond to the holdings of inactive slow-moving arbitrageurs: this generates \( 2(k - 1) \) unknowns and \( 2(k - 1) \) corresponding equations.

We solve the relevant system of \( 2k \) nonlinear equations numerically using the Powell hybrid algorithm which performs a quasi-Newton search to find solutions to a system of nonlinear equations starting from an initial guess.\(^1\) To find all of the solutions, we apply this algorithm by sampling over 10,000 different initial guesses. Once a solution for \( \alpha_1 \) and \( \delta_1 \) is in hand, we can compute \( V^{(1)} \) and \( V^{(k)} \) and can then solve for \( \alpha_0 \) and \( \delta_0 \) using

\[
\alpha_0 = i + \frac{V^{(1)}}{q} \left[ \bar{s} - (1 - q) \delta_0 \right] \quad \text{and} \quad \delta_0 = \frac{k \left( \alpha_0 - i \right)}{V^{(k)}},
\]

which yields

\[
\alpha_0 = i + \frac{\bar{s}}{\tau \left[ q \frac{1}{V^{(1)}} + (1 - q) \frac{k}{V^{(k)}} \right]} \quad \text{and} \quad \delta_0 = \frac{k}{q \frac{1}{V^{(1)}} + (1 - q) \frac{k}{V^{(k)}}} \times \bar{s}.
\]

When asset supply is stochastic, an equilibrium solution only exists if arbitrageurs are sufficiently risk tolerant (i.e., for \( \tau \) sufficiently large). When an equilibrium exists, there are multiple equilibrium solutions. Equilibrium non-existence and multiplicity of this sort arise in overlapping-generations, rational-expectations models such as ours where risk-averse arbitrageurs with finite investment horizons trade an infinitely-lived asset that is subject to supply shocks.\(^2\) Different equilibria correspond to different self-fulfilling beliefs that arbitrageurs can hold about the price-impact of supply shocks and, hence, the risks associated with holding long-term bonds. See Greenwood et al. (2018) for an extensive discussion of these issues.

The intuition for equilibrium multiplicity can be understood most clearly in the simple case when there are only fast-moving arbitrageurs. If arbitrageurs are sufficiently risk tolerant there are

\(^1\)Rational expectations models with noisy asset supply can only be solved in closed-form in very simple, special cases. For instance, we have only been able to obtain closed-form solutions to our model if we turn off the key asset pricing friction—namely, slow-moving capital—that is the focus of our paper.

\(^2\)For previous treatments of these issues, see Spiegel (1998), Bacchetta and van Wincoop (2003), Watanabe (2008), Banerjee (2011), Greenwood and Vayanos (2014), Albagli (2015), and Greenwood, Hanson, and Liao (forthcoming).
two equilibria in this special case: a low price impact (or low return volatility) equilibrium and a high price impact (or high return volatility) equilibrium. If arbitrageurs believe that supply shocks will have a large impact on long-term bonds prices, they will perceive bonds as being highly risky. As a result, arbitrageurs will only absorb a positive supply shock if they are compensated by a large decline in bond prices, making the initial belief self-fulfilling. However, if arbitrageurs believe that bond prices will be less sensitive to supply shocks, they will perceive bond as being less risky and will absorb a supply shock even if they are only compensated by a modest decline in bond prices.

Things are slightly more complicated in our general model with slow-moving capital. Specifically, the introduction of slow-moving capital can give rise to additional unstable equilibria. However, we always find a unique equilibrium that is stable in the sense that equilibrium is robust to a small perturbation in arbitrageurs’ beliefs regarding the equilibrium that will prevail in the future.

Formally, letting $\omega(1) = \omega^* + \xi$ for some small $\xi$ and defining $\omega^{(n)} = f(\omega^{(n-1)})$, an equilibrium $\omega^*$ is stable if $\lim_{n \to \infty} \omega^{(n)} = \omega^*$ and is unstable if $\lim_{n \to \infty} \omega^{(n)} \neq \omega^*$. Let $\{\lambda_i\}$ denote the eigenvalues of the Jacobian $D_\omega f(\omega^*)$. If $\max_i |\lambda_i| < 1$, then $\omega^*$ is stable; if $\max_i |\lambda_i| > 1$, then $\omega^*$ is unstable.

Consistent with Samuelson’s (1947) “correspondence principle,” which says that the comparative statics of stable equilibria have certain properties, this unique stable equilibrium has comparative statics that accord with standard economic intuition. By contrast, the unstable equilibria have comparative statics that run contrary to standard intuition.\footnote{For instance, in the special case where there is no slow-moving capital, the low price-impact equilibrium is stable and the high price-impact equilibrium is unstable. At the stable equilibrium, an increase in the volatility of short rates or the volatility of supply shocks is associated with an increase in the price-impact coefficient and an increase in the volatility of returns. By contrast, these comparative statics take the opposite sign at the unstable equilibrium.}

We focus on this unique stable equilibrium in our numerical illustrations.

### B.8 Solution in the special case without slow-moving capital

In this subsection, we show how to solve the model the special case in which there is no slow-moving capital (i.e., if either $q = 1$ or $k = 1$). For simplicity, we also assume that the three underlying shocks—i.e., $\varepsilon_{P,t+1}$, $\varepsilon_{T,t+1}$, and $\varepsilon_{s,t+1}$—are mutually orthogonal. In this special case, the equilibrium yield on long-term bonds is

$$y_t = \left(\bar{i} + \tau^{-1} V^{(1)} \bar{s}\right) + \frac{\alpha_0}{\tau} \bar{y}_{t+1} + \frac{\alpha_{1P}}{1 - \rho_P \theta} \left(i_{P,t} - \bar{i}\right) + \frac{\alpha_{1T}}{1 - \rho_T \theta} \bar{i}_{T,t} + \tau^{-1} V^{(1)} \frac{1 - \theta}{1 - \rho_s \theta} \left(s_t - \bar{s}\right),$$  \hspace{1cm} (A.55)

where

$$V^{(1)} = \text{Var}_t \left[\frac{\theta}{1 - \theta} y_{t+1}\right]$$  \hspace{1cm} (A.56)

$$= \text{Var}_t \left[\frac{\theta}{1 - \theta} \varepsilon_{P,t+1} + \frac{\theta}{1 - \rho_T \theta} \varepsilon_{T,t+1} + \tau^{-1} V^{(1)} \frac{\theta}{1 - \rho_s \theta} \left(\varepsilon_{s,t+1} + C\varepsilon_{P,t+1} + C\varepsilon_{T,t+1}\right)\right]$$
is the smaller root of the following quadratic equation:

\[ 0 = \left( \tau^{-1} \frac{\theta}{1 - \rho_s \theta} \sigma_s \right)^2 + \left( \tau^{-1} \frac{\theta}{1 - \rho_p \theta} C \sigma_P \right)^2 + \left( \tau^{-1} \frac{\theta}{1 - \rho_s \theta} \sigma_T \right)^2 \times (V^{(1)})^2 \]  
(A.57)

\[ + 2 \left( \frac{\theta}{1 - \rho_p \theta} \sigma_P \right) \left( \tau^{-1} \frac{\theta}{1 - \rho_s \theta} C \sigma_P \right) + 2 \left( \frac{\theta}{1 - \rho_T \theta} \sigma_T \right) \left( \tau^{-1} \frac{\theta}{1 - \rho_s \theta} C \sigma_T \right) - 1 \times V^{(1)} \]

\[ + \left[ \left( \frac{\theta}{1 - \rho_p \theta} \sigma_P \right)^2 + \left( \frac{\theta}{1 - \rho_T \theta} \sigma_T \right)^2 \right]. \]

In this case, the model-implied regression coefficient is

\[ \beta_h = \frac{\text{Cov}[y_{t+h} - y_t, i_{t+h} - i_t]}{\text{Var}[i_{t+h} - i_t]} \]  
(A.58)

\[ = \frac{\alpha_{1P} \text{Var}[\Delta_h i_{P,t}] + \alpha_{1T} \text{Var}[\Delta_h i_{T,t}] + \alpha_{1s} (\text{Cov}[\Delta_h i_{P,t}, \Delta_h s_t] + \text{Cov}[\Delta_h i_{T,t}, \Delta_h s_t])}{\text{Var}[\Delta_h i_{P,t+h}] + \text{Var}[\Delta_h i_{T,t+h}]} \]

where for \( X \in \{P, T\} \) we have \( \text{Var}[\Delta_h i_{X,t}] = 2 \left[ (1 - \rho_X^h) / (1 - \rho_X^2) \right] \sigma_X^2 \) and \( \text{Cov}[\Delta_h i_{X,t}, \Delta_h s_t] = C \left[ (2 - \rho_X^h - \rho_X^p) / (1 - \rho_X \rho_P) \right] \sigma_X \sigma_P \).

We assume that throughout that \( C \geq 0 \) and \( \rho_s \leq \rho_T \leq \rho_P \). For simplicity, in the following discussion, we will also assume that \( \sigma_X^2 = 0.4 \).

We first consider the level of \( \beta_h \) irrespective of horizon \( h \). Inspecting equation (A.58), it is easy to see that:

- **When \( C = 0 \), the level of \( \beta_h \) is increasing in \( \sigma_P \) for all \( h \).** An increase in \( \sigma_P \) raises the fraction of total short-rate variation at all horizons that is due to movements in the more persistent component (i.e., raises \( \text{Var}[\Delta_h i_{P,t+h}] / (\text{Var}[\Delta_h i_{P,t+h}] + \text{Var}[\Delta_h i_{T,t+h}]) \) for all \( h \)). Since shocks to the more persistent component of short rates have larger impact on long-term yields via a straightforward expectations hypothesis channel (i.e., since \( \alpha_{1P} > \alpha_{1T} \)), an increase in \( \sigma_P \) raises the level of \( \beta_h \) at all horizons. Thus, if \( \sigma_P \) declined between the pre-2000 and post-2000 periods as we have argued, this would lead \( \beta_h \) to decline at all horizons \( h \).

We next consider the way \( \beta_h \) behaves as a function of horizon \( h \). Again, using equation (A.58), it is easy to show that:

- **When \( C = 0 \) and \( \rho_T = \rho_P \), \( \beta_h \) is a constant that is independent of \( h \).** These assumptions imply that the expectations hypothesis holds—i.e., there is no excess sensitivity—and that all shocks to short rates have the same persistence. In this benchmark case, \( \beta_h = \alpha_{1P} = \alpha_{1T} \) for all \( h \)—i.e., the sensitivity of long rates to short rates is the same at all horizons.

\(^4\text{This is without loss of generality since } \sigma_X^2 \text{ only impacts the level of } \alpha_{1s}, \text{ and does not otherwise affect } \beta_h.\)
• When $C = 0$ and $\rho_T < \rho_P$, $\beta_h$ is an increasing function of $h$. These assumptions imply that the expectations hypothesis holds, but there are now transient and persistent shocks to short rates. In this case, $\beta_h$ rises with $h$ since (i) movements in the more persistent component of short rates are associated with larger movement in long-term yields (i.e., $\alpha_{1P} > \alpha_{1T}$) and (ii) because the persistent component dominates changes in short rates at longer horizons (i.e., $\text{Var} [\Delta_h i_{P,t+h}] / (\text{Var} [\Delta_h i_{P,t+h}] + \text{Var} [\Delta_h i_{T,t+h}])$ rises with $h$ when $\rho_T < \rho_P$).

• When $C > 0$ and $\rho_s = \rho_T = \rho_P$, $\beta_h$ is a constant that is independent of $h$. In this case, there is excess sensitivity—shifts in short rates lead to shifts in the term premium on long-term bonds—but the excess sensitivity is the same irrespective of horizon $h$. This is because $\Delta_h s_{t+h} = C \Delta_h i_{P,t+h} + C \Delta_h i_{T,t+h}$ when $\rho_s = \rho_T = \rho_P$ (see equation (A.25)) and $\text{Var} [\Delta_h i_{P,t+h}] / (\text{Var} [\Delta_h i_{P,t+h}] + \text{Var} [\Delta_h i_{T,t+h}]) = \sigma_P^2 / (\sigma_P^2 + \sigma_T^2)$ when $\rho_T = \rho_P$.

• When $C > 0$ and $\rho_s < \rho_T = \rho_P$, $\beta_h$ is a decreasing function of $h$. In this case, long-term interest rates exhibit excess sensitivity to movements in short rates that declines with horizon $h$. Intuitively, if the supply shocks induced by shocks to short rates are more transient than the underlying shocks to short rates, then term premia will react more in the short run than in the long run. Thus, there will be greater excess sensitivity in the short run.

References


