Learning and Risk Premiums in an Aribitrage-free Term Structure Model

Marco Giacoletti, Kristoffer T. Laursen
and Kenneth J. Singleton

Stanford University

May 17, 2018

Motivating Learning

- Risk premiums in the U.S. Treasury bond markets vary over time, with the shape of the yield curve and macro conditions.
- These calculations presume that investors know the structure of the economy, and they are based on full-sample estimates.

Motivating Learning

- Risk premiums in the U.S. Treasury bond markets vary over time, with the shape of the yield curve and macro conditions.
- These calculations presume that investors know the structure of the economy, and they are based on full-sample estimates.
- How might market participants *prospectively* form *real-time* views about risks in the Treasury market?
- Views should be adaptive to "regime changes:"
  - e.g., unforeseen changes in monetary and fiscal policies, and transparency in the policy formation process.
  - reflect investor "confusion" in the market?

First Look with "Naive" Learning

- Expected excess returns vary substantially over time in bond markets. Suppose we capture this by linearly projecting onto the first three principal components (PCs) of yields:

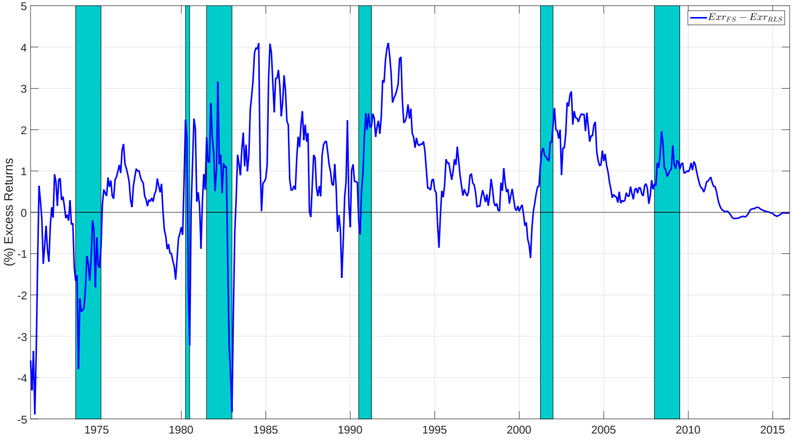
$$Exr^n_{t+h} = \alpha_{n,t} + \mathcal{B}^n_{h\mathcal{P},t}\mathcal{P}_t,$$

  where $\mathcal{P}_t$ includes the low-order PCs of the yield curve.

- Case 1: an econometrician estimates $\mathcal{B}^n_{h\mathcal{P}}$ (fixed over time) using the full sample (no learning).

- Case 2: $\mathcal{RA}$ updates $\mathcal{B}^n_{h\mathcal{P},t}$ in real time using recursive least-squares (Bayesian under special circumstances).

Expected Excess Returns Over 1-Quarter on a 10-Year Zero
Full Sample (FS) Minus Rolling Least Squares (RLS)



What is Known? What Do Investors Learn About?

1. Investors in Treasury bonds have long known:
   - that the cross-sectional distribution of bond yields is well described by a factor model with the low-order PCs;
   - so (plausibly) they observe the risk factors– the relevant "state of the economy"– for pricing Treasury bonds.

What is Known? What Do Investors Learn About?

1. Investors in Treasury bonds have long known:
   - that the cross-sectional distribution of bond yields is well described by a factor model with the low-order PCs;
   - so (plausibly) they observe the risk factors– the relevant "state of the economy"– for pricing Treasury bonds.

2. Bond-market participants cannot foresee the future: Learning about the data-generating process for yields (PCs).

What is Known? What Do Investors Learn About?

1. Investors in Treasury bonds have long known:
   - that the cross-sectional distribution of bond yields is well described by a factor model with the low-order PCs;
   - so (plausibly) they observe the risk factors– the relevant "state of the economy"– for pricing Treasury bonds.
2. Bond-market participants cannot foresee the future: Learning about the data-generating process for yields (PCs).
3. Our Bayesian learner $\mathcal{RA}$:
   - takes as known the pricing distribution for bonds;
   - follows a (constrained) Bayesian learning rule over the parameters of the DGP for the risk factors, under the presumption that these parameters change over time.

The Nature of $\mathcal{RA}$'s Learning Problem

- Agents agree that the one-period riskless rate r_t is given by

$$r_t = \rho_0 + \rho_{\mathcal{P}} \mathcal{P}_t,$$

- The price D_t^m of a zero-coupon bond issued at date t and maturing at date $t+m$ is

$$\begin{aligned} D_t^m &= E_t \left[\underbrace{\mathcal{M}^{\mathcal{B}}(\mathcal{P}_{t+1}, Z_1^t)}_{\text{stochastic discount factor}} D_{t+1}^{m-1} \right] \\ &= \int e^{-r_t} D_{t-1}^{m-1} \underbrace{e^{r_t} \mathcal{M}^{\mathcal{B}}(\mathcal{P}_{t+1}, Z_1^t) \times f(\mathcal{P}_{t+1} | Z_1^t)}_{\text{risk-neutral density} f^{\mathbb{Q}}(\mathcal{P}_{t+1}|Z_1^t)} d\mathcal{P}^{\mathbb{Q}} \\ &= E_t^{\mathbb{Q}} \left[e^{-r_t} D_{t+1}^{m-1} \right]. \end{aligned}$$

The Nature of $\mathcal{RA}$'s Learning Problem

- Agents agree that the one-period riskless rate r_t is given by

$$r_t = \rho_0 + \rho_{\mathcal{P}} \mathcal{P}_t,$$

- The price D_t^m of a zero-coupon bond issued at date t and maturing at date $t+m$ is

$$\begin{aligned} D_t^m &= E_t\left[\mathcal{M}^{\mathcal{B}}(\mathcal{P}_{t+1}, Z_1^t) D_{t+1}^{m-1}\right] \\ &= \int e^{-r_t} D_{t+1}^{m-1} \underbrace{e^{r_t}\, \mathcal{M}^{\mathcal{B}}(\mathcal{P}_{t+1}, Z_1^t) \times f(\mathcal{P}_{t+1} | Z_1^t)}_{\text{risk-neutral density } f^{\mathbb{Q}}(\mathcal{P}_{t+1}|Z_1^t)} \, d\mathcal{P}^{\mathbb{Q}} \end{aligned}$$

Preferences Historical beliefs

- Learning about what? Pricing: $D_t^m = E_t^{\mathbb{Q}}\left[e^{-r_t} D_{t+1}^{m-1}\right].$

Everyone Knows the Pricing Distribution

- "Affine" models imply (Duffie, Pan, and Singleton (2000)):

$$y_t^m = A_m(\Theta^{\mathbb{Q}}) + B_m(\Theta^{\mathbb{Q}})\mathcal{P}_t.$$

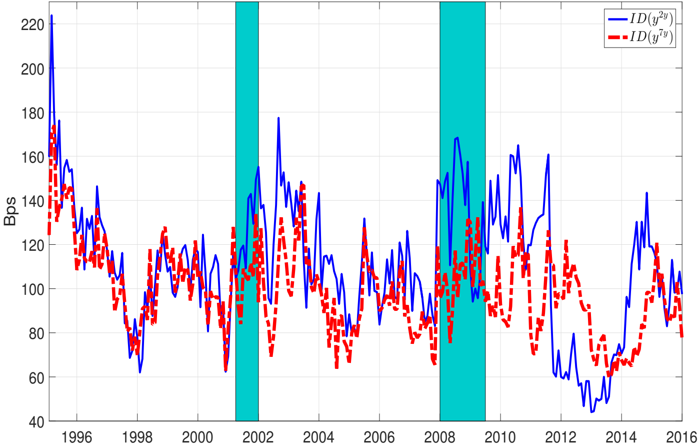
- Market participants reverse engineer the $\mathbb{Q}$ distribution from the prices of traded bonds.
- E.g., suppose, under the pricing distribution, $\mathcal{P}$ is described by

$$\mathcal{P}_{t+1} = K_{0\mathcal{P}}^{\mathbb{Q}} + K_{\mathcal{PP}}^{\mathbb{Q}}\mathcal{P}_t + \Sigma_{\mathcal{PP}}^{1/2} e_{\mathcal{P},t+1}^{\mathbb{Q}}, \; e_{\mathcal{P},t+1}^{\mathbb{Q}} \sim N(0, \Sigma_{\mathcal{P}}).$$

- Then the loadings $B_m(\Theta^{\mathbb{Q}})$ depend only on the 3 eigenvalues of $K_{\mathcal{PP}}^{\mathbb{Q}}$! (Joslin, Singleton, and Zhu (2011)). These loadings can be recovered essentially from cross-sectional regressions.

STANFORD UNIVERSITY

But wait! Investors Disagree: Inter-Decile Ranges of BCFF Forecasts



Forecasts from the Blue Chip Financial Forecasts

Do Blue Chip Financial (BCFF) Forecasters Agree on $\Theta^{\mathbb{Q}}$?

- If all of the BCFF professionals believe that yields are affine in $\mathcal{P}$, then the yield forecasts for horizon h ordered by deciles, $y^h_{t,o_1} < ... < y^h_{t,o_{10}}$, must satisfy

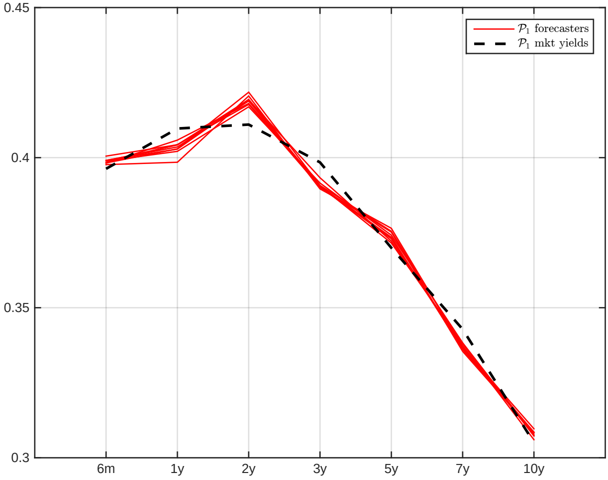
$$\hat{y}^{mh}_{t,o} = \bar{A}^{mh} + \bar{B}^{mh} \hat{\mathcal{P}}^h_{t,o} + e^{mh}_{t,o},$$

  where y^h_{t,o_1} is the tenth percentile forecast, $y^h_{t,o_{10}}$ is the ninetieth percentile, etc.

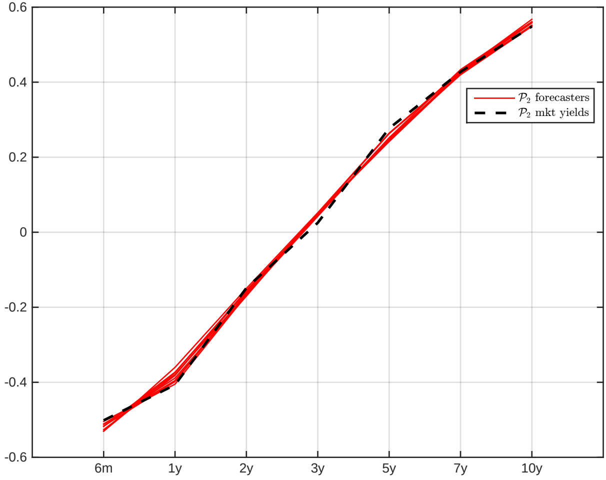
- The loadings should be the same across ordered professionals.

- (Deciles because the forecasters change over time.)

Loadings on $PC1$



Loadings on $PC2$



Learning About the Historical Distribution

- Learning about the jointly Gaussian $\mathbb{P}$ process for Z:

$$Z_{t+1} = K_{0t}^{\mathbb{P}} + K_{Zt}^{\mathbb{P}} Z_t + \Sigma_Z^{1/2} e_{Z,t+1}^{\mathbb{P}}.$$

1. $\Theta_t^{\mathbb{P}}$, is unknown and possibly changing over time:

$$\theta_t^{\mathbb{P}} = \theta_{t-1}^{\mathbb{P}} + \eta_t, \qquad \eta_t \overset{iid}{\sim} N(0, Q_t),$$

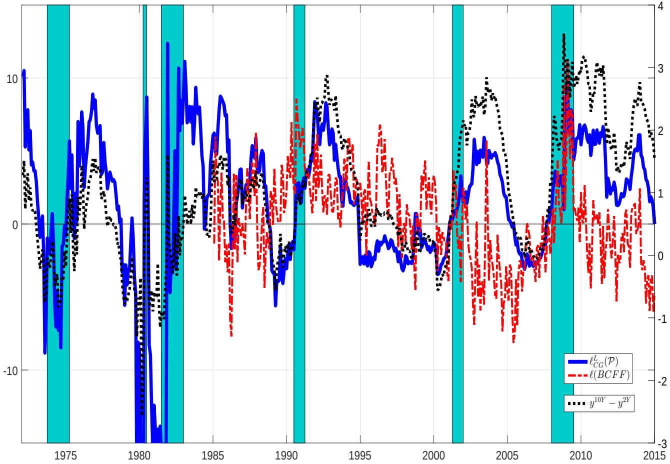
2. $\mathcal{RA}$'s views about $\Theta^{\mathbb{P}}$ revised using a Gaussian posterior distribution $f(\Theta_{t+1}^{\mathbb{P}} | Z_1^t, (\text{yield history})_t)$; $\mathcal{RA}$ does not demand compensation for bearing this parameter risk.

3. Implies a (constrained) version of Bayesian learning is *Constant Gain Learning* with gain coefficient $\gamma \in (0, 1]$. ($\gamma = 1$ is *recursive least-squares (RLS)* learning.)

RMSE's (Basis Points) 3-Month Forecasts Learning From 1985 Through 2014

Rule	6m	1Y	2Y	3Y	5Y	7Y	10Y
$\ell(RW)$	35.6	38.5	40.6	41.1	40.5	39.7	36.7
	(−4.08)	(−3.27)	(−4.52)	(−5.56)	(−5.11)	(−3.93)	(−3.88)
	[]	[]	[]	[]	[]	[]	[]
$\ell(BCFF)$	48.1	48.3	49.2	52.6	47.5	47.1	43.9
	()	()	()	()	()	()	()
	[4.08]	[3.27]	[4.52]	[5.56]	[5.11]	[3.93]	[3.88]
$\ell^L(\mathcal{P})$	34.92	38.30	42.12	41.81	40.55	39.47	37.71
	(−4.29)	(−3.21)	(−4.13)	(−5.88)	(−5.18)	(−4.76)	(−3.32)
	[−0.86]	[−0.23]	[3.39]	[1.34]	[0.10]	[−0.30]	[0.97]
$\ell^L_{CG}(\mathcal{P})$	34.12	38.03	41.75	41.61	40.64	39.79	37.50
	(−4.29)	(−3.14)	(−3.96)	(−5.59)	(−5.11)	(−4.61)	(−3.46)
	[−1.59]	[−0.45]	[3.33]	[1.27]	[0.26]	[0.11]	[1.05]

Expected Excess Returns on 10Y Bond Over 1Y Horizon



Dispersion of Beliefs and Learning

1. $\mathcal{RA}$ represents one view based on rule $\ell^{L}_{CG}(\mathcal{P})$. Investors– including many professional forecasters– disagree.

Dispersion of Beliefs and Learning

1. $\mathcal{RA}$ represents one view based on rule $\ell^{L}_{CG}(\mathcal{P})$. Investors– including many professional forecasters– disagree.
2. Investors (plausibly) agree on the sources of risks for pricing Treasury bonds, summarized by the low-order PCs.

Dispersion of Beliefs and Learning

1. $\mathcal{RA}$ represents one view based on rule $\ell^{L}_{CG}(\mathcal{P})$. Investors– including many professional forecasters– disagree.
2. Investors (plausibly) agree on the sources of risks for pricing Treasury bonds, summarized by the low-order PCs.
3. Learning about the data-generating process for yields using different models/priors.

Dispersion of Beliefs and Learning

1. $\mathcal{RA}$ represents one view based on rule $\ell^L_{CG}(\mathcal{P})$. Investors– including many professional forecasters– disagree.
2. Investors (plausibly) agree on the sources of risks for pricing Treasury bonds, summarized by the low-order PCs.
3. Learning about the data-generating process for yields using different models/priors.
4. Seems natural for $\mathcal{RA}$ to recognize this *dispersion in beliefs*, and to ask whether it is informative about the future.

Aligning Theory with Data/Practice

$$r_t = \underbrace{\delta - \frac{1}{2}\gamma(\gamma+1)\sigma_c^2}_{\text{Lucas Tree}} + \underbrace{\gamma(\omega^a(\eta_t)\hat{g}_t^a + w^b(\eta_t)\hat{g}_t^b)}_{\text{Consensus Belief}} + \underbrace{\frac{\gamma-1}{2\gamma}\omega^a(\eta_t)\omega^b(\eta_t)\Psi_t^2}_{\text{Speculative Demand}},$$

where g is a latent state variable (e.g., "output growth"), $\Psi_t = \sigma_c^{-1}(\hat{g}_t^a - \hat{g}_t^b)$, η_t is the ratio of agents' SDFs.

- Differences of opinion models of Xiong and Yan (2009), Buraschi and Whelan (2016) $\Rightarrow$ high-dimensional "factor space" $\mathcal{P}$ or strong spanning restrictions.
- Data suggests low-dimensional factor structure: the low-order PCs in affine *DTSMs* (Joslin, Singleton, and Zhu (2011)).
  - Danger of over-fitting (Duffee (2010));
  - Spanning by PC's is implausible;
  - Plausibly, these models imply that MPR's depend on disagreement! Disagreement predicts excess returns.

Disagreement is Priced

Disagreement is Correlated with Risk Factors $\mathcal{P}$

	Forecast Horizon			
	1Q	2Q	3Q	4Q
$ID(y^{2y})$	51.18%	58.84%	57.48%	55.11%
$ID(y^{7y})$	41.33%	52.85%	57.22%	56.84%

- Notation: $\mathcal{H}_t = (ID^{2y}_{1y}, ID^{7y}_{1y})$.

Disagreement is Predictive for Annual Excess Returns January 1985 through December 2015

Dependent Variable: One-Year-Ahead Excess Returns

	2y	3y	5y	7y	10y
$\mathcal{P}_1$	0.3221 [5.2630]	0.5335 [4.8375]	0.8024 [4.0640]	1.1068 [3.8154]	1.2961 [3.0806]
$\mathcal{P}_2$	0.3795 [2.5393]	0.7579 [2.6590]	1.7367 [3.5797]	2.7961 [4.2460]	4.3236 [5.0092]
$\mathcal{P}_3$	1.4589 [1.5312]	2.5100 [1.3852]	4.1477 [1.4051]	6.4046 [1.6654]	12.2560 [2.3770]
H_{2y}	0.8851 [1.7784]	2.2796 [2.4121]	4.9092 [2.9237]	6.4796 [2.7830]	8.4242 [2.6414]
H_{7y}	−1.9436 [−3.6679]	−4.1945 [−4.2610]	−7.9459 [−4.5231]	−10.9215 [−4.5408]	−14.2787 [−4.3765]
$adj R^2$	23.76%	23.16%	27.24%	30.14%	31.22%

- Consensus beliefs are redundant: largely spanned by $\mathcal{P}_t$.

STANFORD UNIVERSITY

Models Used for Learning

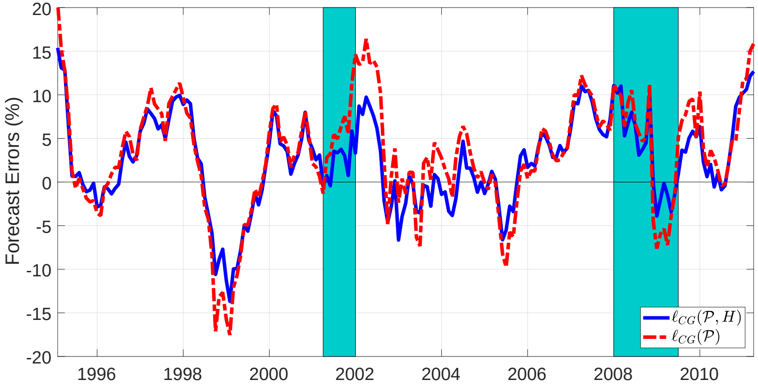
Rule	*DTSM*	State Vector	Restrictions	γ
$\ell(RW)$	No	Own Yield	N/A	N/A
$\ell^L(\mathcal{P})$	Yes	$\mathcal{P}$	No-Arbitrage MPR Constraints	1
$\ell^L_{CG}(\mathcal{P})$	Yes	$\mathcal{P}$	No-Arbitrage + MPR Constraints	0.99
$\ell(\mathcal{P}, \mathcal{H})$	Yes	$(\mathcal{P}, \mathcal{H})$	No-Arbitrage + MPR Constraints	1
$\ell_{CG}(\mathcal{P}, \mathcal{H})$	Yes	$(\mathcal{P}, \mathcal{H})$	No-Arbitrage + MPR Constraints	0.99

RMSE's (Basis Points) for 1 Year-Ahead Forecasts January 1995 through December 2014

	RMSE's (in basis points) for Annual Horizon						
Rule	6m	1Y	2Y	3Y	5Y	7Y	10Y
$\ell(RW)$	118.8 (−1.00) []	115.3 (−0.83) []	103.3 (−1.90) []	94.1 (−2.65) []	84.9 (−2.82) []	78.8 (−2.75) []	70.8 (−2.69) []
$\ell(BCFF)$	128.8 () [1.00]	123.9 () [0.83]	122.1 () [1.90]	122.5 () [2.65]	105.8 () [2.82]	100.6 () [2.75]	88.1 () [2.69]
$\ell_{CG}(\mathcal{P})$	108.4 (−1.60) [−1.32]	105.3 (−1.70) [−1.27]	98.5 (−2.31) [−0.84]	90.7 (−2.96) [−0.58]	82.9 (−3.05) [−0.35]	77.1 (−3.28) [−0.36]	71.2 (−3.02) [0.12]
$\ell_{CG}(\mathcal{P}, \mathcal{H})$	108.5 (−1.43) [−1.29]	104.4 (−1.53) [−1.42]	95.5 (−2.23) [−1.43]	86.3 (−2.98) [−1.33]	76.5 (−3.41) [−1.37]	69.7 (−4.02) [−1.64]	63.3 (−4.01) [−1.73]

STANFORD UNIVERSITY

Errors (% per annum) from Forecasting Realized Excess returns on the 10-year bond over a 1-year horizon: $\ell_{CG}(\mathcal{P}, H)$ (solid) and $\ell_{CG}(\mathcal{P})$ (dashed).



Is Disagreement a Proxy for Macroeconomic Uncertainty? RMSEs of Expected Excess Returns on 10-year Bond

	2Y	3Y	5Y	7Y	10Y
Part A: January, 1995 – December, 2014					
$\ell_{CG}(\mathcal{P})$	1.10%	1.97%	3.39%	4.76%	6.65%
$\ell_{CG}(\mathcal{P}, H)$	1.09%	1.92%	3.17%	4.36%	5.96%
$\ell_{CG}(\mathcal{P}, REA)$	1.07%	1.96%	3.50%	5.02%	7.22%
$\ell_{CG}(\mathcal{P}, REA, INF)$	1.07%	1.97%	3.51%	5.04%	7.24%
$\ell_{CG}(\mathcal{P}, H, REA)$	1.07%	1.92%	3.32%	4.68%	6.60%
$\ell_{CG}(\mathcal{P}, ID(RGDP), ID(INF))$	1.20%	2.16%	3.71%	5.23%	7.14%

RMSEs of Expected Excess Returns on 10-year Bond

	2Y	3Y	5Y	7Y	10Y
Part B: January, 2001 – December, 2007					
$\ell_{CG}(\mathcal{P})$	1.37%	2.44%	3.84%	4.95%	5.72%
$\ell_{CG}(\mathcal{P}, H)$	1.36%	2.39%	3.60%	4.47%	4.79%
$\ell_{CG}(\mathcal{P}, REA)$	1.22%	2.29%	4.02%	5.72%	7.71%
$\ell_{CG}(\mathcal{P}, REA, INF)$	1.23%	2.32%	4.09%	5.84%	7.92%
$\ell_{CG}(\mathcal{P}, H, REA)$	1.21%	2.21%	3.69%	5.08%	6.53%
$\ell_{CG}(\mathcal{P}, ID(RGDP), ID(INF))$	1.48%	2.65%	4.21%	5.48%	6.70%

Buraschi, A., and P. Whelan, 2016, Speculation, Hedging, and Interest Rates, Working paper, Imperial College.

Duffee, G., 2010, Sharpe Ratios in Term Structure Models, Working paper, Johns Hopkins University.

Duffie, D., J. Pan, and K. Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343–1376.

Joslin, S., M. Priebsch, and K. Singleton, 2014, Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks, *Journal of Finance* LXIX, 1197–1233.

Joslin, S., K. Singleton, and H. Zhu, 2011, A New Perspective on Gaussian Dynamic Term Structure Models, *Review of Financial Studies* 24, 926–970.

Ludvigson, S., and S. Ng, 2010, Macro Factors in Bond Risk Premia, *Review of Financial Studies* 22, 5027–5067.

Xiong, W., and H. Yan, 2009, Heterogeneous Expectations and Bond Markets, *Review of Financial Studies*.