Pontifícia Universidade Católica
do Rio de Janeiro

Fernando Luiz Macedo Cardoso

The Expectations Hypothesis Holds. At Times.

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia.

Advisor : Prof. Carlos Viana de Carvalho
Co-advisor: Prof. Ruy Monteiro Ribeiro

Rio de Janeiro
April 2020

Pontifícia Universidade Católica
do Rio de Janeiro

Fernando Luiz Macedo Cardoso

The Expectations Hypothesis Holds. At Times.

Dissertation presented to the Programa de Pós–graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia. Approved by the Examination Committee.

Prof. Carlos Viana de Carvalho
Advisor
Departamento de Economia – PUC-Rio

Prof. Ruy Monteiro Ribeiro
Co-advisor
Departamento de Economia – PUC-Rio

Prof. Marcelo Cunha Medeiros
Departamento de Economia – PUC-Rio

Prof. Emanuel Moench
Research Centre – Deutsche Bundesbank

Rio de Janeiro, April 17th, 2020

Fernando Luiz Macedo Cardoso

B.A., Economics, Insper Intituto de Ensino e Pesquisa (Insper), 2014.

Bibliographic data

Macedo Cardoso, Fernando Luiz

The Expectations Hypothesis Holds. At Times. / Fernando Luiz Macedo Cardoso; advisor: Carlos Viana de Carvalho; co-advisor: Ruy Monteiro Ribeiro. – Rio de janeiro: PUC-Rio, Departamento de Economia, 2020.

v., 62 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia.

Inclui bibliografia

1. Economia – Teses. 2. Macroeconomia – Teses. 3. Finanças – Teses. 4. Teoria das Expectativas;. 5. Estrutura a Termo da Taxa de Juros;. 6. Curva de Juros;. 7. Taxas de Juros de Títulos;. 8. Prêmio de Risco.. I. Viana de Carvalho, Carlos. II. Monteiro Ribeiro, Ruy. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 620.11

Acknowledgments

I thank all of my family, especially my parents, Monica and Fábio, for their sacrifices and invaluable support and guidance, and my grandmother, Maria Luiza, for her wise words and loving care. I also thank my girlfriend, Marcela, for the beautiful partnership and insightful exchanges.

I thank my advisors, Profs. Carlos Viana de Carvalho and Ruy Monteiro Ribeiro, for their continued support and always making themselves available as well as the friendship built over this period. I thank Prof. Alexandre Schwartsman for advising me to pursue this degree. Finally, I thank Profs. Tiago Berriel, Svetlana Bryzgalova, Robin Greenwood, Christian Julliard, Marcelo Medeiros, and seminar participants at PUC-Rio and the 2019 International Workshop in Financial Econometrics for helpful comments and suggestions.

This study was carried out with the support of CNPq, National Council for Scientific and Technological Development - Brazil and the XV ANBIMA Capital Markets Prize.

Abstract

Macedo Cardoso, Fernando Luiz; Viana de Carvalho, Carlos (Advisor); Monteiro Ribeiro, Ruy (Co-Advisor). **The Expectations Hypothesis Holds. At Times.**. Rio de Janeiro, 2020. 62p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

The yield curve literature typically decomposes long-term interest rates into expected future short-term rates and a risk premium. We show that the relative importance of the expectational component vis-à-vis the risk premium component can be time-varying and state-dependent. Further, the likelihood of an "Expectations Hypothesis (EH) State" has a clear relation to the business cycle. Moreover, our results indicate that incorporating the probability of these EH states boosts the predictive power of the benchmark yield curve measure, the term spread, both for future excess bond returns and economic activity.

Keywords

Expectations Hypothesis; Term Structure of Interest Rates; Yield Curve; Bonds Yields; Term Premium.

Resumo

Macedo Cardoso, Fernando Luiz; Viana de Carvalho, Carlos; Monteiro Ribeiro, Ruy. **A Teoria das Expectativas Vale. Ocasionalmente.**. Rio de Janeiro, 2020. 62p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Tipicamente, a literatura de curva de juros assume que taxas de juros para horizontes longos são compostas por expectativas de taxas de juros curtas que devem vigorar nesse horizonte longo e/ou um prêmio de risco. O objetivo deste trabalho é mostrar evidência de que o peso relativo de um componente expectacional vis-à-vis um componente de prêmio de risco pode depender do tempo e do estado da economia. Ademais, a probabilidade de um "Regime da Teoria das Expectativas" mostra-se relacionado ao ciclo de negócios. Ainda, os resultados indicam que ao se incorporar a probabilidade destes regimes, é possível intensficar o poder preditivo do diferencial entre os juros longos e o curto tanto para excesso de retornos quanto para atividade econômica.

Palavras-chave

Teoria das Expectativas; Estrutura a Termo da Taxa de Juros; Curva de Juros; Taxas de Juros de Títulos; Prêmio de Risco.

Table of contents

List of figures

List of tables

List of Abreviations

ACM – Adrian et al. (2013)
APW – Ang and Piazzesi (2003)
CRSP – Center for Research in Security Prices
CS – Campbell and Shiller (1991)
CW – Clark and West (2007)
DM – Diebold and Mariano (1995)
EH – Expectations Hypothesis
FB – Fama and Bliss (1987)
GDP – Gross Domestic Product
GSW – Gürkaynak et al. (2007)
HH – Hansen and Hodrick (1980)
ML – Maximum Likelihood
NBER – National Bureau of Economic Research
OLS – Ordinary Least Squares
PC – Principal Component
RMPE – Root Mean Prediction Error
TV – Time-Varying
UK – United Kingdom
US – United States

Yet good, or even competent, economists are the rarest of birds. An easy subject, at which very few excel! The paradox finds its explanation, perhaps, in that the master-economist must possess a rare combination of gifts. He must reach a high standard in several different directions and must combine talents not often found together. He must be mathematician, historian, statesman, philosopher – in some degree. He must understand symbols and speak in words. He must contemplate the particular in terms of the general, and touch abstract and concrete in the same flight of thought. He must study the present in the light of the past for the purposes of the future. No part of man's nature or his institutions must lie entirely outside his regard. He must be purposeful and disinterested in a simultaneous mood; as aloof and incorruptible as an artist, yet sometimes as near the earth as a politician.

John Maynard Keynes, *Alfred Marshall, 1842-1924 in The Economic Journal, Vol. 34, No. 135.*

1
Introduction

The yield curve portraits market interest rates payed at any given moment by bonds with different maturities. A number of different measures of the yield curve have proven valuable in economic tracking and forecasting. The term spread, the difference between long- and short-term rates, is frequently used in indicators of economic and financial conditions and has become ubiquitous in forecasting exercises. The literature typically assumes that longer-maturity rates are comprised of expectations for the path of shorter-maturity rates plus a risk premium for the long-term commitment. When the yield curve moves upwards (downwards), this can be driven by higher (lower) expectations of future short-term rates and/or a higher (lower) risk premium. Explanations for its predictive power in the media usually center on the expectational component of long-term rates. This, however, seems strange in light of papers as Fama and Bliss (1987) which show evidence that *all* of the variation in long-term rates comes from the term premium. Understanding this composition is of paramount importance for investors and policy-makers. Former Fed Chairman Bernanke (2006), for instance, dismissed a growing attention to lower long rates relative to shorter rates citing, among other things, that "to the extent that the flattening or inversion of the yield curve is the result of a smaller term premium, the implications for future economic activity are positive rather than negative". In hindsight, this diagnostic was proved wrong by the greatest recession to hit the United States since the Great Depression. In fact, every recession since 1969 was preceded by an inverted yield curve, with only one false positive in 1966 – although it was followed by a period of slow growth.

In this paper we expand the analysis in (Campbell and Shiller, 1991, henceforth CS), in which they test the Expectations Hypothesis (EH) using simple linear regressions of future changes of longer rates on a maturity-specific proportion of the term spread, and incorporate regime switching to allow for an "Expectations Hypothesis state", i.e. one in which the term premium is constant (but not necessarily zero), in an effort to identify periods in which the EH has greater support in the data. These are periods in which all of the variation in yields comes from the expectational component. We then analyze the probabilities of this EH state and discuss their relation to the business cycle, first by plotting the probabilities against US recessions as determined by the NBER's Business Cycle Dating Committee, then by regressing against macroeconomic variables commonly used for business cycle tracking and forecasting. Finally, we argue that the correct characterization of our regimes implies that it should be useful as a conditioning variable in forecasting models of future economic activity and bond excess returns based on the term spread, and provide in- and out-of-sample results to confirm this.

We show how the model-implied probabilities of the EH state are very high, or even one, for several periods. This seems to happen when all is well and the economy is growing, and then come to an end right before a recession starts. The estimates for the unrestricted regime are close to those in CS, with volatility in the EH regime considerably smaller, and high persistence for both regimes. We quantify this relation between the EH probabilities and the business cycle in two ways. First we extend the model to allow time-varying transition probabilities which are taken to be a function of the short rate, and our results indicate that higher levels of interest rates, usually seen when the Fed is actively trying to restrain economic activity, imply a lower probability of remaining at or transitioning to the EH state. The second way is to look at the relation between the probabilities, taken as observable, and a vast number of variables typically used for monitoring economic activity using logit regressions, which shows considerable differences on the number of relevant relationships between groups: almost every variable in the employment group, for example, is related to the probabilities of at least one maturity, while not one of the 36 variables in the prices group seems relevant – an indication that these probabilities capture a real, rather than nominal phenomenon. We then compare benchmark models based on the term spread with extensions which add an interaction between the term spread and the probabilities of the EH state, which not only prove significant in-sample, but also significantly outperform the benchmark in out-of-sample exercises both for predicting future economic activity and bond excess returns. The in-sample results also imply that, when it is known that we are in an EH state, we should more than double the weight given to the information in the term spread. Finally, we compare the out-of-sample gains of forecasting excess returns from using the EH probabilities versus a decomposition implied by an affine model, with mixed results: the EH probabilities have an advantage for shorter maturities and the affine model in the longer ones – however, these differences are statistically insignificant. The appendix offers robustness checks using alternative data and specifications.

This apparent relation between the yield curve and economic activity has received a lot of attention in the literature. Fama (1984) discusses forward rates as predictors of premiums and future spot rates. Harvey (1988) was among the first to point out how a lower (real) term spread could indicate decreasing future consumption by exploring the first-order conditions of the representative agent problem. Campbell et al. (2017) and Viceira (2012) discuss the empirical relations between bond and stock returns. Estrella and Hardouvelis (1991) document the predictive power of the term spread for economic activity, both in predicting future GDP growth and fitting discrete choice models for the probability of oncoming recessions. The term spread is also one of the most important series in Stock and Watson (1989)'s leading economic index, who go on to discuss in Stock and Watson (2003) its predictive power relative to a number of asset prices and state that while "no single asset price is a reliable predictor of output growth across countries over multiple decades[, t]he term spread perhaps comes closest to achieving this goal". They discuss that its predictive power has been unstable, as do Estrella et al. (2003), Rossi and Sekhposyan (2010) and Hännikäinen (2017), depending on the time period, monetary policy regime and the phase of the business cycle. Ang et al. (2006) address the predictive content of the yield curve for GDP growth and make the case that the short rate, and not the term spread, offers the best predictions. Further, they argue that the predictive power of the term spread comes exclusively from the expectational component, and subtracting the risk premium from yields could improve forecasting. In spite of all the evidence, however, Rudebusch and Williams (2009) show that professional forecasters underweight the relevance of the yield curve and underperform simple models based on the term spread.

Our work naturally relates to the part of the literature that tests the Expectations Hypothesis of the term structure of interest rates. In its weaker form, the EH posits that the risk premium in longer-term rates must be constant through time. On one side of this debate we have contributions by Mankiw and Miron (1986), who discuss how the evidence against the EH becomes stronger after the founding of the Federal Reserve System in 1915; Fama and Bliss (1987) regress excess returns on a forward spread and find that all of the variation in yields comes from the risk premium component; Campbell and Shiller (1991) also find counterintuitive coefficients in regressions of future changes in the short rates on a linear transformation of the term spread; while Pflueger and Viceira (2011) again reject the EH looking at inflation-linked bonds in the US and UK. On the other side, Froot (1989) argues that the rejection of EH tests on longer maturity bonds should instead be attributed to a failure of the rational expectations hypothesis; Gerlach and Smets (1997) study Euro-rates for 17 countries and find little evidence against the EH; meanwhile Bekaert and Hodrick (2001) argue that, when small-sample properties are taken into account, the evidence against the EH is "much less strong than under asymptotic inference".

We also touch the part of literature that looks at understanding the information in the yield curve. Litterman and Scheinkman (1991) started the practice of condensing the information of a cross-section of yields on three components, typically understood as the level, slope and curvature. Cochrane and Piazzesi (2005), in the spirit of Fama and Bliss (1987), construct a tent-shaped return-forecasting factor widely used in this literature. They go on in Cochrane and Piazzesi (2009) to decompose the yield curve into expectations and risk premia components using an affine-class model, as do Cieslak and Povala (2010) who discuss a decomposition relating to inflation expectations and interest-rate cycles. Crump et al. (2018) perform a decomposition based on professional forecasts and obtain the risk premium as the residual.

Finally, our specification is based on the regime switching literature that goes back to Hamilton (1989) and has examples in Ang and Bekaert (1998), in which the authors model the short rate and term spread dynamics and discuss the regime classifications and relation to the business cycle; Bansal and Zhou (2002) and Bansal et al. (2004), who in the context of affine-models, show that incorporating regime shifts in the dynamics of market prices of risk allows the model to fit the data better than the alternatives. Singleton (2006) offers a good overview. Lastly, we must cite Cao et al. (2014), who also use regime switching models based on a specification of CS and discuss the regimes' relation to the business cycle. Complementary to, but unlike this paper, they restrict themselves in classifying the regimes according to volatility, and do not go into the matter of the predictive power of the term spread.

2
Model

The strong form of the expectations hypothesis states that interest rates payed on an n-period zero-coupon bond, $y_t^{(n)}$, must be given by the weighted average of expected future rates on an m-period bond with the same characteristics, for $n > m$. By allowing for risk aversion, the weaker form includes a premium to compensate for the long-term commitment as opposed to rolling the short-term contract each period, i.e.

$$y_t^{(n)} = \mathbb{E}_t \left[\frac{1}{k} \sum_{i=0}^{k-1} y_{t+mi}^{(m)} \right] [+\text{term premium}], \quad k = n/m \tag{2-1}$$

where $\mathbb{E}_t$ denotes the expectation conditional on information available at time t. The crucial assumption is that the term premium may vary for different m and n, but must be constant through time. Our focus here is on the weaker form.

Campbell and Shiller (1991) test this by regressing changes in the short rate against a linear transformation of the term spread:

$$y_{t+m}^{(n-m)} - y_t^{(n)} = \gamma_0^{(n)} + \gamma_1^{(n)} \frac{m}{n-m} \left(y_t^{(n)} - y_t^{(m)} \right) + \varepsilon_{t+m}^{(n)} \tag{2-2}$$

Under the EH, the coefficient $\gamma_1^{(n)}$ should equal 1.[1] Table 2.1 reports estimates for regression (2-2) for maturities $n = 2, .., 7$ years and $m = 1$ year (as for the rest of this paper) using (Gürkaynak et al., 2007, GSW) data at a quarterly frequency from 1961Q2 through 2019Q3. (Hansen and Hodrick, 1980, HH) standard errors are reported in parenthesis. As the original, the updated results not only reject the EH for every maturity, but indicate that the term spread in fact gives the wrong direction of future changes in the short rate.

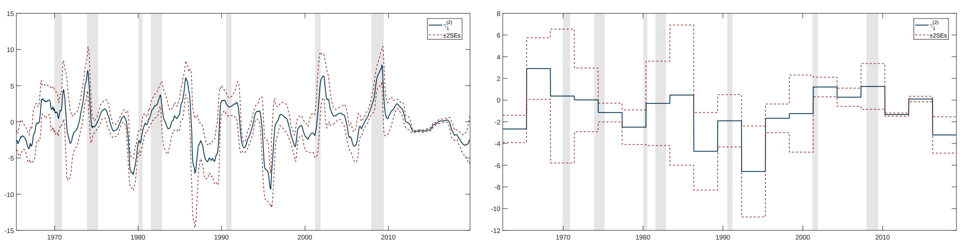
A look at Figure 2.1 instigates a question as to the stability of the $\gamma_1^{(n)}$ coefficients through time. Figure 2.1a plots the rolling estimate for maturity $n = 2$ years using a 3-year overlapping window of weekly data, such that the

[1]To see this, consider $m = 1$ and $n = 2$. By imposing $\gamma_1^{(2)} = 1$, rearranging and taking time-t expectations we have that $y_t^{(2)} = \frac{1}{2}\mathbb{E}_t \left[y_{t+1}^{(1)} + y_t^{(1)} \right] - \frac{1}{2}\gamma_0^{(2)}$, with the intercept proportional to the negative of the term premium. The term premia for $n > 2$ are given recursively by $TP^{(n)} = -\frac{(n-1)}{n}\gamma_0^{(n)} + \frac{(n-1)}{n}TP^{(n-1)}$.

Campbell-Shiller Regression						
	Maturity (yrs)					
	2	3	4	5	6	7
$\gamma_0^{(n)}$	-0.08	-0.00	0.05	0.09	0.13	0.15
	(0.23)	(0.21)	(0.19)	(0.17)	(0.16)	(0.15)
$\gamma_1^{(n)}$	-0.72	-0.99	-1.25	-1.48	-1.70	-1.89
	(0.42)	(0.35)	(0.32)	(0.32)	(0.34)	(0.36)
Obs	231	231	231	231	231	231

Table 2.1: Results for regression (2-2) using quarterly data by Gürkaynak et al. (2007) from 1961Q2 through 2019Q3. Hansen and Hodrick (1980) standard errors in parenthesis.

time t rolling estimate spans data from $t-3$ years to t. In the notation of (2-2), the coefficient represents $\gamma_{t+1}^{(2)}$, which is to say that it looks at past values of the term spread on the current short rate. While the coefficient remains in the negative territory for most of the sample, it does present important spikes towards positive values, and therefore towards the EH implied value (and higher). Further, these spikes occur in relevant stages of the business cycle: every recession in the sample is preceded by a spike and sees a subsequent drop in the coefficient. Of course, one may expect that at a high enough frequency there must be periods in which the variation in long rates should come exclusively from expectations, and therefore that this coefficient *should* present variation towards the EH-implied value. However, the fact that these spikes represent 3 years of data (and are present for windows as long as 5 years) and happen at specific times in the business cycle is intriguing. Figure 2.1b plots the same exercise except with non-overlapping windows and tells a similar story.



(a) (b)

Figure 2.1: Rolling estimates of $\gamma_1^{(2)}$ using GSW weekly data with 3-year windows, overlapping (2.1a) and non-overlapping (2.1b). Red dashed lines indicate $\pm$ 2 HH standard errors and shaded regions correspond to recessions.

This parameter instability motivates the use of a model that allows for variation in the $\gamma_1^{(n)}$ coefficient. Specifically, we investigate if this instability is related to the composition of the term spread and the dynamics of expectations relative to the term premium component. A simple, parsimonious way to look at this is to expand the CS regression (2-2) and assume that there exists two states of nature $S = \{0, 1\}$: one in which the EH holds and one in which it may not. The model may be summarized as:

$$\begin{aligned} y_{t+m}^{(n-m)} - y_t^{(n)} &= \gamma_0^{(n)} + \gamma_{1,s_{t+m}=S}^{(n)} \frac{m}{n-m} \left(y_t^{(n)} - y_t^{(m)} \right) + \varepsilon_{t+m}^{(n)} \\ \varepsilon_{t+m}^{(n)} &\sim \mathbf{N}\left(0, \sigma_{s_{t+m}=S}^{2,(n)}\right) \\ \mathbb{P}(s_t = 0 | s_{t-1} = 0) &= q^{(n)} \\ \mathbb{P}(s_t = 1 | s_{t-1} = 1) &= p^{(n)} \end{aligned} \tag{2-3}$$

Coefficients $\gamma_{1,s_{t+m}=S}^{(n)}$ therefore takes the value of 1 when $S = 0$ and is left unrestricted for $S = 1$, with volatility $\sigma_{s_t=S}$ also regime-specific. We choose to aggregate the intercept across regimes, since it has negligible effect in the estimated parameters and regime probabilities and only addresses the mean term premium for each state.[2] The intuition behind this model is that, if we are looking for a regime that has no support in the data then we expect the estimation to attribute negligible probabilities to that state while the unrestricted regime converges to the OLS estimates.[3] Our goal is to see how much weight does the data attribute to a specific, EH-implied regime – and when. Appendix B discusses a Bayesian approach with collapsed priors for $\gamma_{1,S=0}^{(n)} = 1$ while Appendix C allows for time-varying coefficients in the unrestricted regime. Both approaches corroborate with the main analysis.

[2]The curious reader can find these estimates in Appendix A.

[3]We performed a few exercises with $\gamma_{1,s=0}^{(n)}$ taking values higher than 10 and the results were as expected.

3
Results

Estimates are based on data from Gürkaynak et al. (2007) from 1961Q2 through 2019Q3 at a quarterly frequency and we restrict our attention to the case that $m = 1$ and maturities $n = 2, .., 7$ years. Table 3.1 reports the maximum likelihood (ML) estimates for model (2-3). First, although the standard errors are not comparable[1], the point estimates of $\gamma^{(n)}_{1,s_t=1}$ are typically within two (HH) standard errors of the estimates in (2-2), in line with the earlier intuition. Second, for all maturities the volatility of the EH state is smaller than that of the second regime, also intuitive since there is the added variance from an unstable term premium. Third, from the estimates of p and q, both regimes are highly persistent and have an expected duration of between 13 and 20 quarters for the EH state and 20 to 27 quarters for the second regime. We must point out, however, that a likelihood ratio test dismisses the EH restriction on $\gamma_{1,s_t=0}$. Our understanding is that this is driven by a stronger classification in terms of regime variances, which tend to dominate the estimation, and resonates with previous findings[2]. To circumvent this, we could add more regimes to better fit the data. However, not only is there a critical small-sample problem for estimating too many regimes, but this would steer away from our original point.

Figure 3.1 plots the probabilities of the EH state according to the ML estimates of model 2-3. Panel 3.1a compares the filtered and smoothed probabilities for maturity $n = 2$ years while Panel 3.1b gathers the smoothed probabilities for all maturities $n = 2, .., 7$ years. As the paper's title hints at, there are several periods in which the model attributes probabilities close to (or exactly) 1 to the EH state, or, in other words, periods in which the expectations hypothesis has greater support in the data. It is worth noting that the probabilities are very robust across maturities.[3] Further, the shaded regions representing US recessions indicate how these probabilities appear to be related

[1] We did not find an obvious correction in the spirit of HH for the switching regressions.

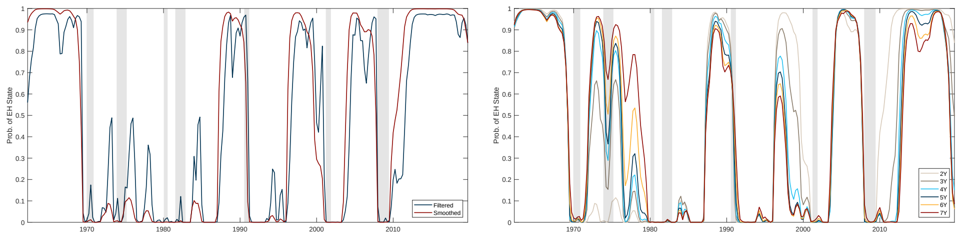
[2] This problem was overcome in the Bayesian approach by assuming prior distributions for $\sigma^{2,(n)}_{S_t=S}$ with more weight on lower values for $S = 0$.

[3] Appendix A offers estimates using the Fama-Bliss database from the Center for Research in Security Prices of the Booth School of Business at the University of Chicago, as well as an alternative specification based on Fama and Bliss (1987). The results are essentially the same.

Campbell-Shiller Switching Regression						
	Maturity (yrs)					
	2	3	4	5	6	7
$\gamma_0^{(n)}$	-0.13	-0.06	-0.00	0.01	0.02	0.03
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
$\gamma_{1,s_t=1}^{(n)}$	-1.21	-1.65	-1.96	-2.15	-2.32	-2.47
	(0.027)	(0.031)	(0.030)	(0.032)	(0.034)	(0.036)
$\sigma_{s_t=0}^{(n)}$	0.63	0.61	0.56	0.54	0.55	0.56
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$\sigma_{s_t=1}^{(n)}$	1.91	1.66	1.50	1.39	1.31	1.27
	(0.009)	(0.008)	(0.007)	(0.006)	(0.006)	(0.006)
$p^{(n)}$	0.95	0.93	0.92	0.92	0.92	0.93
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$q^{(n)}$	0.96	0.95	0.95	0.95	0.95	0.95
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obs	231	231	231	231	231	231

Table 3.1: Parameter estimates for model (2-3). Data is from GSW from 1961Q2-2019Q3 at quarterly frequency. Standard errors reported in parenthesis.

to the business cycle. In general, every recession is preceded by a period of high probability of the EH state. Moreover, the EH probabilities seem to lower around the onset of a recession and rise as the economy stabilizes. Recessions are routinely classified as a different regime in this literature. These shifts might be related to agents missing their forecasts, which may induce changes to how they manage risk. This resonates with Cieslak (2018), who argues that investors underestimate easing cycles by the monetary authority, which induces excess returns not due to time-varying risk premium. It could also be related to Campbell and Cochrane (1999) and their discussion on how agents care more about how assets behave in recessions as opposed to the covariance with consumption growth *per se*. Regarding the period of the 1970s and 1980s, we stress that these were times of great uncertainty regarding both monetary policy and inflation dynamics, which translates into erratic estimates. Piazzesi and Schneider (2006) discuss a possible relation between inflation shocks and the yield curve and illustrate with the early 1980s.



(a) (b)

Figure 3.1: Figure 3.1a plots filtered (blue) and smoothed (red) probabilities of the EH state according to ML estimates of model (2-3) for maturity $n = 2$. Figure 3.1b plots the smoothed probabilities for maturities $n = 2, ..., 7$. Shaded regions indicate US recessions.

3.1 EH Probabilities and the Macroeconomy

One simple way to look at the relation between the EH state probabilities and the macroeconomy is to allow for time-varying (TV) transition probabilities. Keeping the same framework as in (2-3), we can define

$$\begin{aligned}\mathbb{P}(s_t = 0|s_{t-1} = 0) &= \frac{\exp\left\{p_0^{(n)} + p_1^{(n)} y_{t-1}^{(1)}\right\}}{1 + \exp\left\{p_0^{(n)} + p_1^{(n)} y_{t-1}^{(1)}\right\}} \\ \mathbb{P}(s_t = 1|s_{t-1} = 1) &= \frac{\exp\left\{q_0^{(n)} + q_1^{(n)} y_{t-1}^{(1)}\right\}}{1 + \exp\left\{q_0^{(n)} + q_1^{(n)} y_{t-1}^{(1)}\right\}}\end{aligned} \tag{3-1}$$

At a quarterly frequency, the first principal component (PC) explains over 99% of the variation in yields in our sample. Commonly known as the level factor, the first PC also has a 98% correlation with the short rate. Thus, we follow Ang et al. (2006) in using the short rate as proxy for the the first PC and thus condensing most of the variation in the term structure while maintaining parsimony. Table 3.2 reports the ML results. The negative estimates of p_1 as well as the positive estimates of q_1 indicate that higher levels of the short rate imply a lower probability of remaining at or transitioning to the EH state. One caveat is that the estimation for maturity $n = 3$ stands relatively apart from the rest in this specification.

Another approach is to take the probabilities as observed from model (2-3) and regress on macroeconomic variables using a logit model. We use the FRED-QD database, a quarterly frequency companion to FRED-MD by McCracken and Ng (2016) with 248 variables used for business cycle tracking and forecasting. Figure 3.2 plots t-statistics from logit regressions of the EH state probabilities on individual variables for maturities $n = 2, 3, 5$ and 7. Shading indicates variable groups as determined in the appendix to the FRED-

Campbell-Shiller Switching Regression - TV Probabilities						
	Maturity (yrs)					
	2	3	4	5	6	7
$\gamma_0^{(n)}$	-0.13	-0.23	-0.01	0.01	0.02	0.02
	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
$\gamma_{1,s_t=1}^{(n)}$	-1.21	-0.65	-1.98	-2.17	-2.32	-2.48
	(0.028)	(0.022)	(0.031)	(0.032)	(0.034)	(0.037)
$\sigma_{s_t=0}^{(n)}$	0.64	0.33	0.56	0.53	0.52	0.54
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
$\sigma_{s_t=1}^{(n)}$	1.93	1.47	1.49	1.38	1.30	1.25
	(0.009)	(0.005)	(0.007)	(0.006)	(0.006)	(0.006)
$p_0^{(n)}$	5.07	4.15	4.71	4.60	4.56	4.46
	(0.101)	(0.122)	(0.109)	(0.099)	(0.098)	(0.104)
$p_1^{(n)}$	-0.43	-0.58	-0.46	-0.46	-0.45	-0.40
	(0.017)	(0.039)	(0.019)	(0.018)	(0.018)	(0.019)
$q_0^{(n)}$	1.60	-0.66	2.33	2.36	2.36	2.34
	(0.062)	(0.240)	(0.047)	(0.047)	(0.047)	(0.048)
$q_1^{(n)}$	0.29	11.92	0.08	0.07	0.07	0.08
	(0.015)	(1.025)	(0.009)	(0.009)	(0.009)	(0.010)
Obs	231	231	231	231	231	231

Table 3.2: Parameter estimates for model (3-1). Data is from GSW from 1961Q2-2019Q3 at quarterly frequency. Standard errors reported in parenthesis.

QD project and listed in Table 3.3. Variables with t-statistics greater than 2 in absolute value are represented with ticks on the top and bottom axes for reference.

The distribution of t-statistics with absolute value greater than 2 is further evidence of the strong relation between the EH state probabilities and the business cycle. Groups of particular interest are Employment and Unemployment (EMPL), Industrial Production (IP), Non-Household Balance Sheets (N-HH BAL SHTS), and National Income and Product Accounts (NIPA), not to mention the trivial Interest and Exchange Rates groups. Housing, Inventories, Orders, & Sales, Earnings & Productivity, Stock Markets, Money & Credit and Household Balance Sheets also show some level of interrelation. The only group that has no related variables is Prices, with not one of its 36 variables with a t-statistic greater than 2 in absolute value, which indicates that the EH state probabilities are capturing a real, rather than nominal phenomenon. Appendix E discusses results of variable selection using Lasso.

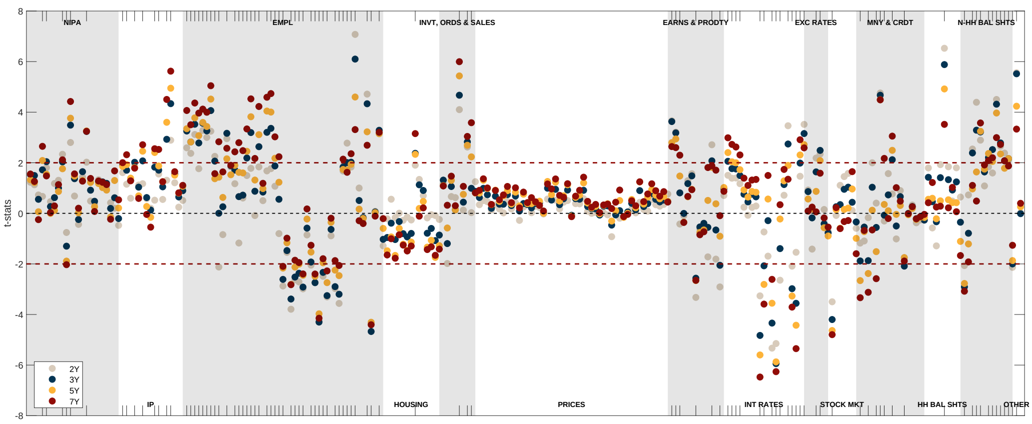


Figure 3.2: Markers indicate t-statistics from logit regressions of the EH state probabilities for maturities $n = 2, 3, 5$ and 7 on 248 individual macroeconomic variables from the FRED-QD database. Red dashed line indicates $\pm$ 2 for reference. Shading indicates variable groups as defined in the appendix to McCracken and Ng (2016) (See table 3.3).

Abbreviation	**Group**
NIPA	National Income and Product Accounts
IP	Industrial Production
EMPL	Employment and Unemployment
HOUSING	Housing
INVT, ORDS & SALES	Inventories, Orders, and Sales
PRICES	Prices
EARNS & PRODTY	Earnings and Productivity
INT RATES	Interest Rates
EXC RATES	Exchange Rates
STOCK MKT	Stock Markets
MNY & CRDT	Money and Credit
HH BAL SHTS	Household Balance Sheets
N-HH BAL SHTS	Non-Household Balance Sheets
OTHER	Other

Table 3.3: FRED-QD group names and abbreviations.

4
Predictability

The evidence that the expectations hypothesis holds at certain time periods dialogues with two strands of the literature. First, it relates to the evidence in Stock and Watson (2003), Estrella et al. (2003), Rossi and Sekhposyan (2010), and more recently Hännikäinen (2017) regarding the unstable predictive power of the term spread as a forecasting variable for growth. Second, to the discussion in Ang et al. (2006) as to the source of its predictive power when considering the decomposition into an expectational component and a risk premium implied by their affine model. Their results show that not only does the decomposition allow for better in-sample R^2s, but also that the coefficient relative to the term premium component is always insignificant. They thus argue that "we should subtract the [risk premium] component from the spread[, o]therwise the expectations contained in the term spread are contaminated by the [risk premium] component, which blurs the GDP forecasts."

An implication therefore is that, if the EH state probabilities correctly identify periods in which the risk premium is constant, then conditioning of these probabilities should boost the predictive power of the term spread. The first panel in Table 4.1 reports in-sample results for regressions of the h-period ahead annualized GDP growth $g_{t,t+h} \equiv \frac{1}{h}\left(\log GDP_{t+h} - \log GDP_t\right)$ on the term spread, as in (4-1), while the second panel includes the interaction between the term spread and EH state probabilities implied by a joint model across maturities. Hodrick (1992) standard errors are reported in parenthesis. The joint specification stacks all yields available and imposes the EH state on all maturities. These probabilities, reported in Figure 4.1, are more unstable, but tell the same story (higher in stable times, lower in and around recessions) as the independent estimates and are discussed in Appendix D. Excess return predictability results are robust for the independent estimations, but those for GDP predictability are mostly insignificant.

$$g_{t,t+h} = \beta_0 + \beta_1\left(y_t^{(n)} - y_t^{(1)}\right)\left[+\beta_2\left(y_t^{(n)} - y_t^{(1)}\right) * \mathbb{P}(\mathrm{EH})\right] + \upsilon_{t+h} \qquad (4\text{-}1)$$

Two comments should accompany Table 4.1: first, not only is the inter-

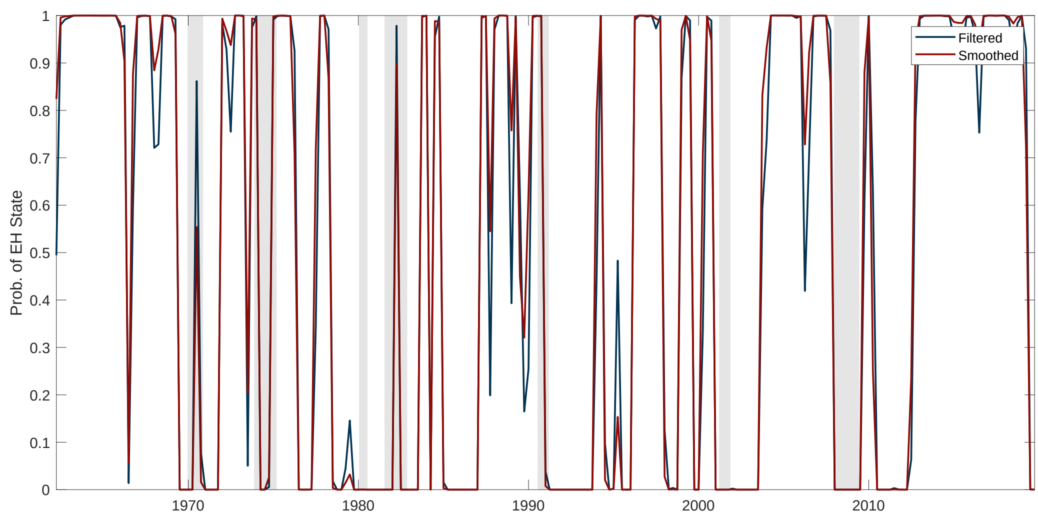


Figure 4.1: Probability of the Expectations Hypothesis State for a joint model of all maturities. Shading indicates US recessions.

action significant at 5% for $n = 2, ..., 6$ and at 10% for $n = 7$, but the point estimates imply that when it is known that we are in an EH state the weight given to the information in the term spread should more than double. Second, there are some gains in R^2, although they only range from 1-2 percentage points.

Predicting One-Quarter-Ahead GDP Growth						
	Maturity (yrs)					
	2	3	4	5	6	7
Term spread	8.42	4.58	3.11	2.30	1.80	1.45
	(1.71)	(1.02)	(0.78)	(0.66)	(0.59)	(0.54)
R^2	10%	8%	7%	5%	4%	3%
Term spread	6.39	3.52	2.39	1.76	1.35	1.07
	(1.86)	(1.10)	(0.85)	(0.72)	(0.64)	(0.59)
Term spread * $\mathbb{P}$(EH)	7.09	3.93	2.75	2.13	1.75	1.51
	(3.08)	(1.73)	(1.27)	(1.03)	(0.88)	(0.78)
R^2	12%	10%	8%	7%	5%	4%
Obs	229	229	229	229	229	229

Table 4.1: In-sample results for term spread predictability conditioned on the EH state probabilities from a joint estimation of model (2-3). The first panel refers to the benchmark model while the second adds the interaction with the EH state probabilities. Hodrick (1992) standard errors in parenthesis. Coefficients and standard errors $\times 10^{-3}$.

Table 4.2 reports the out-of-sample results using half the sample as a training period and then incorporating the new information. Panel 4.2a displays the ratio of root mean prediction errors (RMPE) for the model with the interaction over those for the benchmark model. Results lower than 1 indicate that the forecasts when conditioning on the EH state probabilities outperform the benchmark, and constitute the majority of Panel 4.2a. The out-of-sample gains of up to 2.2% are in line with the in-sample gains in R^2 in Table 4.1. Panel 4.2b reports (Clark and West, 2007, CW) one-sided statistics for comparing nested models' forecasting power, with critical values of 1.282 (10%) and 1.645 (5%). It is important to note that the CW statistic corrects for the fact that the larger model estimates a parameter that is known under the null, and therefore can reject it even if the RMPE ratio is close to or even above 1. The evidence is that the EH state boosts the predictive power of the term spread out-of-sample for up to 4 quarters ahead.

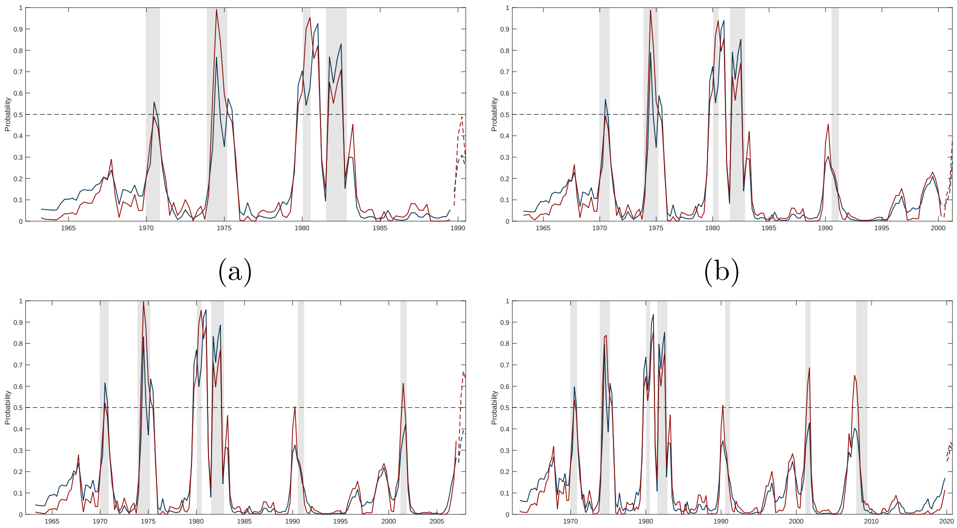
Pseudo Out-of-Sample Forecasting GDP Growth

Horizon (Q)	Maturity (yrs)					
	2	3	4	5	6	7
(a) RMPE Ratio						
1	0.982	0.985	0.989	0.993	0.995	0.997
2	0.978	0.981	0.986	0.991	0.994	0.997
3	0.982	0.985	0.991	0.996	1.000	1.003
4	0.983	0.987	0.994	0.999	1.003	1.006
(b) CW Statistic						
1	1.894	1.862	1.802	1.758	1.733	1.722
2	1.767	1.686	1.612	1.571	1.553	1.546
3	1.444	1.415	1.372	1.354	1.348	1.346
4	1.304	1.276	1.232	1.212	1.203	1.195
Obs	115	115	115	115	115	115

Table 4.2: Out-of-sample results for term spread predictability of GDP conditioned on the EH state probabilities from a joint estimation of model (2-3). Panel (a) reports the ratio of RMPE for the model conditioned on the EH state probabilities relative the benchmark. Panel (b) reports CW one-sided statistics for comparing RMPE of nested models. Critical values are 1.282 (10%) and 1.645 (5%).

4.1 Recession Logits

Figure 4.2 plots the fitted values for logit regressions of the probability of a recession 4 quarters ahead on the same right-hand side variables in (4-1), using the longest term spread in our sample with $n = 7$, as is recommended in APW[1]. The blue and red lines represent the benchmark and conditional models, respectively. Shading represents actual recessions and dashed lines indicate out-of-sample forecasts. Sample periods for Figures 4.2a through 4.2c end in 1989Q3, 2000Q2, 2007Q1, which corresponds to 1 year prior to the 1990, 2001, and 2008 recessions. Figure 4.2d uses the full sample up to 2019Q3.



(a) (b)

(c) (d)

Figure 4.2: Fitted values for logit regressions of the probability of a recession 4 quarters ahead on the the right-hand side variables in (4-1) for $n = 7$. The blue and red lines represent the benchmark and conditional models, respectively. Shading represents actual recessions and dashed lines indicate out-of-sample projections. Sample periods for Figures 4.2a through 4.2c end in 1989Q3, 2000Q2, 2007Q1, which corresponds to 1 year prior to the 1990, 2001, and 2008 recessions. Figure 4.2d uses the full sample up to 2019Q3.

We may judge the logit models on (i) how high the fitted probabilities are in past recessions and (ii) by considering that the model signals a coming recession if the fitted probabilities are higher than a threshold, which we can take to be 50%, and count successes. By both of these metrics, the logit conditional on the EH state probabilities outperforms the benchmark. According to the first metric, the conditional model fits a higher probability in all but the 1970 and 1981-82 recessions looking in-sample, and in all of the

[1]Results are again robust across maturities, with an intriguing outperformance of the 2Y for the beginning of the sample.

out-of-sample recessions, while it does not consistently fit higher probabilities for non-recession periods. By the second metric, while the conditional model correctly signals all of the in-sample recessions, the benchmark only signals those up to 1981-82. In the out-of-sample exercises, it correctly signaled the 2008 recession, while the benchmark misses all three.

4.2 Predicting Excess Bond Returns

Tables 4.3 and 4.4 report in- and out-of-sample results for bond excess return forecasting. In fact, the in-sample results show that, as expected, excess returns are not predictable when conditioning the term spread on the EH state probabilities, whereas by conditioning in the non-EH state probabilities they are[2]. Further, the conditioned model has R^2 gains between 3-4%. The RMPE ratios of the conditioned model relative to the term spread alone in Table 4.4 also indicate gains up to 3%, most of which are significant at 5%.

Predicting One-Quarter-Ahead Excess Returns						
	Maturity (yrs)					
	2	3	4	5	6	7
Term spread	3.91	3.21	3.11	3.14	3.21	3.30
	(1.71)	(1.36)	(1.24)	(1.19)	(1.17)	(1.17)
R^2	6%	6%	6%	6%	6%	6%
Term spread * $(1 - \mathbb{P}(\text{EH}))$	5.39	4.43	4.22	4.18	4.20	4.25
	(2.00)	(1.55)	(1.40)	(1.34)	(1.30)	(1.29)
Term spread * $\mathbb{P}(\text{EH})$	0.19	-0.09	-0.02	0.14	0.33	0.52
	(1.71)	(1.42)	(1.32)	(1.30)	(1.31)	(1.34)
R^2	10%	10%	10%	10%	10%	9%
Obs	230	230	230	230	230	230

Table 4.3: In-sample results for predicting excess returns conditioned on the EH state probabilities from a joint estimation of model (2-3). Hodrick (1992) standard errors in parenthesis.

[2]Results for one-quarter excess returns calculated by using the GSW parameters and generating artificial quarterly maturities.

Pseudo Out-of-Sample Forecasting Excess Returns						
Horizon (Q)	Maturity (yrs)					
	2	3	4	5	6	7
(a) RMPE Ratio						
1	0.971	0.978	0.987	0.995	1.003	1.010
2	0.972	0.972	0.979	0.988	0.997	1.004
3	0.976	0.975	0.979	0.984	0.990	0.995
4	0.979	0.979	0.981	0.984	0.988	0.991
(b) CW Statistic						
1	3.339	3.018	2.516	2.033	1.580	1.159
2	3.053	3.226	2.805	2.250	1.706	1.199
3	2.545	2.836	2.688	2.334	1.911	1.460
4	2.243	2.316	2.282	2.112	1.884	1.620
Obs	115	115	115	115	115	115

Table 4.4: Out-of-sample results for predicting excess returns conditioned on the EH state probabilities from a joint estimation of model (2-3). Panel (a) reports the ratio of RMPE for the model conditioned on the EH state probabilities relative the benchmark. Panel (b) reports CW one-sided statistics for comparing RMPE of nested models. Critical values are 1.282 (10%) and 1.645 (5%).

4.2.1 Predicting Returns: EH Probabilities vs ACM

Finally, we compare the out-of-sample performance for predicting excess returns of the term spread conditioned on the EH state probabilities relative to the decomposition of the term spread into expectational and term premium components implied by the affine model in (Adrian et al., 2013, ACM)[3]. Table 4.5 reports root mean prediction error ratios and (Diebold and Mariano, 1995, DM) statistics (since models aren't nested, CW isn't appropriate here). The results show that, according to the DM statistics, the two approaches are equally accurate. It is worth noting that the EH Probabilities improve the forecasts for the shorter maturities and longer horizon, with the 2-year term spread overperforming the ACM decomposition by 9% for 4-quarters ahead, while the 5-year underperforms by 7%.

[3]For each period, we use the first 3 principal components and excess returns from the fitted yield curve using GSW parameters.

Pseudo Out-of-Sample Forecasting Excess Returns EH Probabilities vs. ACM Decomposition						
Horizon (Q)	Maturity (yrs)					
	2	3	4	5	6	7
(a) RMPE Ratio						
1	0.994	1.011	1.020	1.026	1.031	1.033
2	0.971	1.011	1.037	1.054	1.066	1.073
3	0.940	0.991	1.027	1.053	1.070	1.081
4	0.911	0.962	1.003	1.034	1.056	1.072
(b) DM Statistic						
1	0.122	-0.313	-0.619	-0.805	-0.906	-0.955
2	0.390	-0.177	-0.614	-0.899	-1.073	-1.178
3	0.604	0.103	-0.322	-0.637	-0.854	-0.996
4	0.743	0.348	-0.027	-0.329	-0.553	-0.709
Obs	115	115	115	115	115	115

Table 4.5: Out-of-sample results for predicting excess returns conditioned on the EH state probabilities from a joint estimation of model (2-3) vs. the ACM decomposition. Panel (a) reports the ratio of RMPE for the model conditioned on the EH state probabilities relative to ACM. Panel (b) reports DM statistics.

5
Concluding Thoughts

Expanding on the work in Campbell and Shiller (1991), we have provided evidence that the expectations hypothesis is a good approximation for certain time periods, and its empirical failures may be concentrated in specific, turbulent times. Moreover, these EH states further our understanding of yield curve dynamics and its relation to the business cycle. We relate the probability of an "Expectations Hypothesis state" to the level of the yield curve and to numerous macroeconomic variables, and include in the appendix a discussion of the robustness of these EH state probabilities. Further, we provide in- and out-of-sample results of how the probability of the EH state accounts for some of the instability in the predictive power of the benchmark yield curve measure, the term spread, and how it can boost its power as a conditioning variable, outperforming affine-model decompositions for longer horizon excess returns forecasting. Understanding the mechanisms behind these dynamics remains a challenge.

Bibliography

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

Ang, A. and Bekaert, G. (1998). Regime Switches in Interest Rates. *SSRN Electronic Journal*.

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

Ang, A., Piazzesi, M., and Wei, M. (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics*, 131(1-2):359–403.

Bansal, R., Tauchen, G., and Zhou, H. (2004). Regime Shifts, Risk Premiums in the Term Structure, and the Business Cycle. *Journal of Business & Economic Statistics*, 22(4):396–409.

Bansal, R. and Zhou, H. (2002). Term Structure of Interest Rates with Regime Shifts. *The Journal of Finance*, 57(5):1997–2043.

Bekaert, G. and Hodrick, R. J. (2001). Expectations Hypotheses Tests. *The Journal of Finance*, 56(4):1357–1394.

Bernanke, B. S. (2006). Reflections on the Yield Curve and Monetary Policy. In *Economic Club of New York*, New York.

Campbell, J. Y. and Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2):205–251.

Campbell, J. Y. and Shiller, R. J. (1991). Yield Spreads and Interest Rate Movements: A Bird's Eye View. *The Review of Economic Studies*, 58(3):495.

Campbell, J. Y., Sunderam, A., and Viceira, L. M. (2017). Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds. *Critical Finance Review*, 6(2):263–301.

Cao, R., Zhu, X., and Rahman, S. (2014). Yield factors, The expectations hypothesis and regime shifts. *Annals of Economics and Finance*, 15(2):411–431.

Cieslak, A. (2018). Short-Rate Expectations and Unexpected Returns in Treasury Bonds. *The Review of Financial Studies*, 31(9):3265–3306.

Cieslak, A. and Povala, P. (2010). Understanding Bond Risk Premia. *SSRN Electronic Journal*.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cochrane, J. H. and Piazzesi, M. (2005). Bond Risk Premia. *American Economic Review*, 95(1):138–160.

Cochrane, J. H. and Piazzesi, M. (2009). Decomposing the Yield Curve. *SSRN Electronic Journal*, (18).

Crump, R. K., Eusepi, S., and Moench, E. (2018). The Term Structure of Expectations and Bond Yields. Technical Report 775, Federal Reserve Bank of New York.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263.

Estrella, A. and Hardouvelis, G. A. (1991). The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance*, 46(2):555–576.

Estrella, A., Rodrigues, A. P., and Schich, S. (2003). How Stable is the Predictive Power of the Yield Curve? Evidence from Germany and the United States. *Review of Economics and Statistics*, 85(3):629–644.

Fama, E. F. (1984). The information in the term structure. *Journal of Financial Economics*, 13(4):509–528.

Fama, E. F. and Bliss, R. R. (1987). The Information in Long-Maturity Forward Rates. *American Economic Review*, 77(4):680–692.

Froot, K. A. (1989). New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates. *The Journal of Finance*, 44(2):283–305.

Gerlach, S. and Smets, F. (1997). The term structure of Euro-rates: some evidence in support of the expectations hypothesis. *Journal of International Money and Finance*, 16(2):305–321.

Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304.

Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, 57(2):357.

Hännikäinen, J. (2017). When does the yield curve contain predictive power? Evidence from a data-rich environment. *International Journal of Forecasting*, 33(4).

Hansen, L. P. and Hodrick, R. J. (1980). Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis. *Journal of Political Economy*, 88(5):829–853.

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics*, 22(2):305–333.

Hodrick, R. J. (1992). Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement. *Review of Financial Studies*, 5(3):357–386.

Kadiyala, K. R. and Karlsson, S. (1997). Numerical Methods for Estimation and Inference in Bayesian VAR-Models. *Journal of Applied Econometrics*, 12(2):99–132.

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

Mankiw, N. G. and Miron, J. A. (1986). The Changing Behavior of the Term Structure of Interest Rates. *The Quarterly Journal of Economics*, 101(2):211.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Pflueger, C. and Viceira, L. (2011). Inflation-Indexed Bonds and the Expectations Hypothesis. Technical report, National Bureau of Economic Research, Cambridge, MA.

Piazzesi, M. and Schneider, M. (2006). Equilibrium Yield Curves. Technical report, National Bureau of Economic Research, Cambridge, MA.

Rossi, B. and Sekhposyan, T. (2010). Have economic models' forecasting performance for US output growth and inflation changed over time, and when? *International Journal of Forecasting*, 26(4):808–835.

Rudebusch, G. D. and Williams, J. C. (2009). Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve. *Journal of Business & Economic Statistics*, 27(4):492–503.

Singleton, K. J. (2006). *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. Princeton University Press, Princeton.

Stock, J. H. and Watson, M. W. (1989). New Indexes of Coincident and Leading Economic Indicators. *NBER Macroeconomics Annual*, 4:351.

Stock, J. H. and Watson, M. W. (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature*, 41(3):788–829.

Viceira, L. M. (2012). Bond risk, bond return volatility, and the term structure of interest rates. *International Journal of Forecasting*, 28(1):97–117.

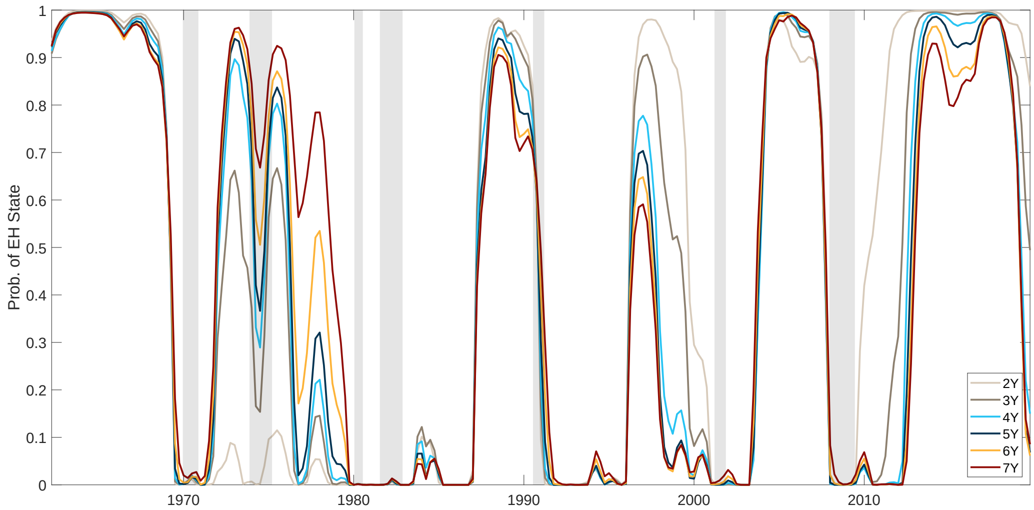
A
Independent Model Robustness

This section reports various estimates of the probability of an Expectations Hypothesis state using GSW (same as paper) and CRSP data (from 1952Q2-2019Q3, with maturities of 1,...,5 years), as well as a specification analogous to (3) but based on Fama and Bliss (1987), where excess returns are regressed on a forward spread such that, if the EH holds, the slope coefficient should be zero (excess returns should not be predictable).

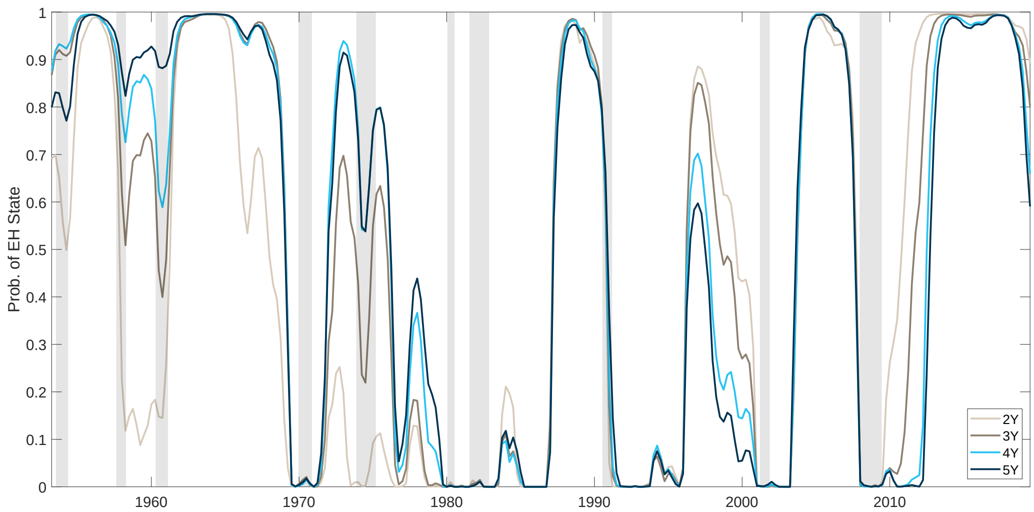
Figure A.1 plots the smoothed probabilities of the Expectations Hypothesis state according to model (3) using GSW (A.1a) and CRSP (A.1b) data. Figure A.2 plots the smoothed probabilities of the Expectations Hypothesis state according to the FB specification using GSW (A.2a) and CRSP (A.2b) data. Estimation results are reported in Tables A.1 and A.2. Tables A.3 and A.4 report the estimation results for a model with regime-specific intercepts.

Campbell-Shiller Switching Regression										
	Maturity (yrs)									
	GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5
$\gamma_0^{(n)}$	-0.13	-0.06	0.00	0.01	0.02	0.03	-0.06	-0.05	0.00	0.03
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)	(0.007)	(0.005)	(0.004)	(0.004)
$\gamma_{1,s_t=1}^{(n)}$	-1.21	-1.65	-1.96	-2.15	-2.32	-2.47	-1.19	-1.71	-2.18	-2.20
	(0.027)	(0.031)	(0.030)	(0.032)	(0.034)	(0.036)	(0.025)	(0.029)	(0.029)	(0.032)
$\sigma_{s_t=0}^{(n)}$	0.63	0.61	0.56	0.54	0.55	0.56	0.69	0.69	0.65	0.63
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.004)	(0.003)
$\sigma_{s_t=1}^{(n)}$	1.91	1.66	1.50	1.39	1.31	1.27	1.84	1.69	1.51	1.40
	(0.009)	(0.008)	(0.007)	(0.006)	(0.006)	(0.006)	(0.010)	(0.008)	(0.007)	(0.006)
$p^{(n)}$	0.95	0.93	0.92	0.92	0.92	0.93	0.94	0.94	0.94	0.94
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$q^{(n)}$	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.94	0.95	0.95
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obs	231	231	231	231	231	231	263	263	263	263

Table A.1: Parameter estimates for model (3). Data is from GSW from 1961Q2-2019Q3 and CRSP from 1952Q2 at quarterly frequency. Standard errors reported in parenthesis.

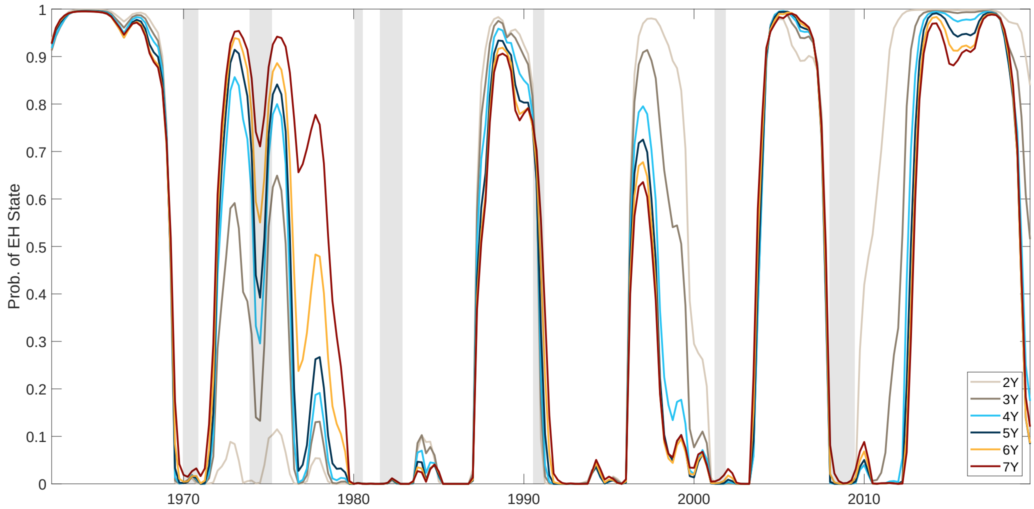


(a)

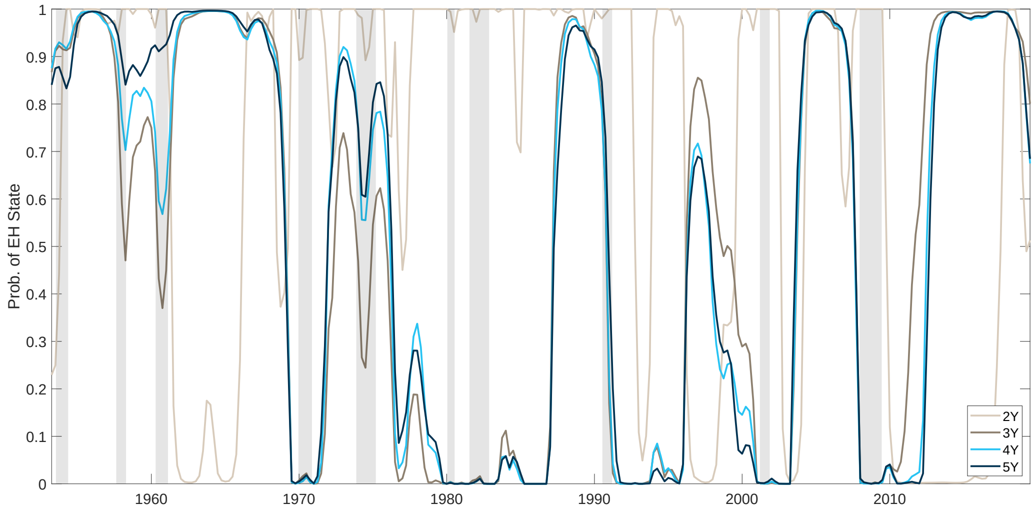


(b)

Figure A.1: Figure A.1a plots smoothed probabilities of the EH state according to ML estimates of model (3) for GSW data from 1961Q2-2019Q3, and A.1b uses CRSP data from 1952Q2-2019Q3. Shaded regions indicate US recessions.



(a)



(b)

Figure A.2: Figure A.2a plots smoothed probabilities of the EH state according to the FB specification for GSW data from 1961Q2-2019Q3, and A.2b uses CRSP data from 1952Q2-2019Q3. Shaded regions indicate US recessions.

Fama-Bliss Switching Regression										
	Maturity (yrs)									
	GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5
$\gamma_0^{(n)}$	0.13	0.13	0.03	-0.01	-0.03	-0.03	-0.19	0.09	-0.01	0.01
	(0.005)	(0.010)	(0.013)	(0.018)	(0.023)	(0.028)	(0.005)	(0.011)	(0.013)	(0.015)
$\gamma_{1,s_t=1}^{(n)}$	1.10	1.39	1.61	1.75	1.87	1.98	1.45	1.50	1.75	1.56
	(0.014)	(0.017)	(0.017)	(0.019)	(0.021)	(0.023)	(0.006)	(0.016)	(0.016)	(0.019)
$\sigma_{s_t=0}^{(n)}$	0.63	1.22	1.69	2.18	2.76	3.40	1.86	1.38	1.95	2.52
	(0.003)	(0.008)	(0.011)	(0.014)	(0.020)	(0.022)	(0.006)	(0.012)	(0.012)	(0.013)
$\sigma_{s_t=1}^{(n)}$	1.91	3.35	4.56	5.68	6.76	7.86	0.35	3.37	4.57	6.01
	(0.009)	(0.016)	(0.020)	(0.025)	(0.031)	(0.038)	(0.003)	(0.017)	(0.020)	(0.027)
$p^{(n)}$	0.95	0.93	0.92	0.92	0.92	0.93	0.95	0.94	0.94	0.95
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
$q^{(n)}$	0.96	0.95	0.95	0.95	0.95	0.95	0.88	0.94	0.95	0.95
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Obs	231	231	231	231	231	231	263	263	263	263

Table A.2: Parameter estimates for Fama-Bliss specification. Data is from GSW from 1961Q2-2019Q3 and CRSP from 1952Q2 at quarterly frequency. Standard errors reported in parenthesis.

Campbell-Shiller Switching Regression, Regime-Specific Intercept										
	Maturity (yrs)									
	GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5
$\gamma_{0,s_t=0}^{(n)}$	-0.08	0.00	0.26	0.33	0.35	-0.74	0.40	0.27	0.25	-1.48
	(0.005)	(0.006)	(0.007)	(0.008)	(0.006)	(0.012)	(0.007)	(0.007)	(0.008)	(0.008)
$\gamma_{0,s_t=1}^{(n)}$	-0.36	-0.25	-0.50	-0.39	-0.31	0.71	-0.80	-0.54	-0.36	0.60
	(0.014)	(0.013)	(0.012)	(0.011)	(0.010)	(0.007)	(0.012)	(0.010)	(0.009)	(0.006)
$\gamma_{1,s_t=1}^{(n)}$	-0.95	-1.42	-2.01	-2.05	-2.15	-3.26	-1.16	-1.86	-2.33	-0.69
	(0.031)	(0.034)	(0.034)	(0.036)	(0.036)	(0.037)	(0.026)	(0.028)	(0.031)	(0.022)
$\sigma_{s_t=0}^{(n)}$	0.64	0.62	0.70	0.64	0.60	1.30	0.78	0.72	0.67	0.83
	(0.003)	(0.005)	(0.004)	(0.004)	(0.003)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)
$\sigma_{s_t=1}^{(n)}$	1.90	1.65	1.29	1.18	1.11	0.54	1.52	1.41	1.28	0.79
	(0.009)	(0.008)	(0.007)	(0.006)	(0.005)	(0.003)	(0.007)	(0.007)	(0.006)	(0.003)
$p^{(n)}$	0.95	0.93	0.91	0.89	0.89	0.90	0.90	0.90	0.90	0.86
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
$q^{(n)}$	0.96	0.95	0.90	0.91	0.91	0.92	0.90	0.89	0.89	0.92
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Obs	231	231	231	231	231	231	263	263	263	263

Table A.3: Parameter estimates for model with regime-specific intercepts. Data is from GSW from 1961Q2-2019Q3 and CRSP from 1952Q2 at quarterly frequency. Standard errors reported in parenthesis.

Fama-Bliss Switching Regression, Regime Specific Intercept										
	Maturity (yrs)									
	GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5
$\gamma^{(n)}_{0,s_t=0}$	0.50	0.91	-0.11	-0.20	-0.30	4.83	0.40	-0.50	-0.64	3.60
	(0.009)	(0.017)	(0.016)	(0.023)	(0.037)	(0.063)	(0.008)	(0.014)	(0.020)	(0.039)
$\gamma^{(n)}_{0,s_t=1}$	-0.18	-0.95	0.59	0.58	0.58	-4.37	-0.37	1.05	1.02	-2.83
	(0.005)	(0.009)	(0.036)	(0.046)	(0.059)	(0.036)	(0.004)	(0.022)	(0.031)	(0.025)
$\gamma^{(n)}_{1,s_t=1}$	1.15	1.72	1.43	1.57	1.71	2.33	1.60	1.54	1.81	1.77
	(0.009)	(0.008)	(0.020)	(0.022)	(0.025)	(0.016)	(0.005)	(0.016)	(0.018)	(0.014)
$\sigma^{(n)}_{s_t=0}$	1.80	3.24	1.71	2.20	2.76	7.57	1.75	1.44	2.03	5.29
	(0.006)	(0.012)	(0.011)	(0.014)	(0.020)	(0.039)	(0.006)	(0.007)	(0.011)	(0.024)
$\sigma^{(n)}_{s_t=1}$	0.25	0.55	4.53	5.65	6.71	3.23	0.29	2.87	4.01	2.45
	(0.002)	(0.004)	(0.020)	(0.025)	(0.032)	(0.017)	(0.002)	(0.014)	(0.020)	(0.012)
$p^{(n)}$	0.98	0.97	0.93	0.92	0.92	0.91	0.96	0.90	0.90	0.90
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)
$q^{(n)}$	0.92	0.90	0.95	0.95	0.95	0.92	0.89	0.89	0.90	0.91
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Obs	231	231	231	231	231	231	263	263	263	263

Table A.4: Parameter estimates for Fama-Bliss specification with regime-specific intercepts. Data is from GSW from 1961Q2-2019Q3 and CRSP from 1952Q2-2019Q3 at quarterly frequency. Standard errors reported in parenthesis.

B
Bayesian Approach

Another natural way to approach this is with a Bayesian estimation with a collapsed prior for the $\gamma_{1,S_t=0}^{(n)}$ coefficient on the EH-implied value. The main advantage to this approach is that we can obtain a posterior distribution for the probabilities of the EH state, as opposed to the MLE estimates. Figure B.1 plots the median values for the probabilities of the EH state using CRSP data for the CS specification with dashed lines indicating a 90% credible region. Not only do the median values corroborate with the previous results, but the credible region is also in line with the variation across data sets and maturities.

Figures B.2 and B.3 plot the posterior distributions for the model's parameters, as well as the priors assumed: independent Beta(1,1) for $p^{(n)}, q^{(n)}$, $(\gamma_{0,0}^{(n)}, \gamma_{1,0}^{(n)}, \gamma_{0,1}^{(n)}, \gamma_{1,1}^{(n)})' \sim \mathbf{N}(b_0, B_0)$ with $b_0 = (0, 1, 0, 0)'$ and $B_0 = diag(B, \epsilon, B, B)$ such that $\epsilon \to 0$ and $B = 5$. Finally, we assume independent inverse-gammas for $\left\{\sigma_S^{2,(n)}\right\}_{S=0}^{1} \sim \mathbf{IG}(\nu_0/2, \nu_0\tau_0^2/2)$ more concentrated on smaller values for the volatility of the EH state, i.e. $\nu_0 = 4$ and $\tau_0^2 = 1$ for the EH state while $\nu_0 = 6$ and $\tau_0^2 = 4$ for $S = 1$. Posterior draws come from a Gibbs sampler based on Kim2017 with 10^4 draws after burning the first 10^4 using GSW data.

It is worth noting that the posterior distribution for the intercepts, which are related to the term-premia of the two states, is a lot more concentrated for $\gamma_{0,S=0}$ than for $\gamma_{0,S=1}$, which is to be expected since our underlying assumption is that the EH term premium is constant through time. Also, both distributions have a lot of overlap, with close means, which gives us confidence that the simplifying assumption of one intercept for both regimes used in the main model isn't an overreach.

Fama-Bliss Results

Figures B.4, B.5 and B.6 plot the same results for the FB specification, only now using a prior $b_0 = (0, 0, 0, 0)'$. Results are again essentially the same.

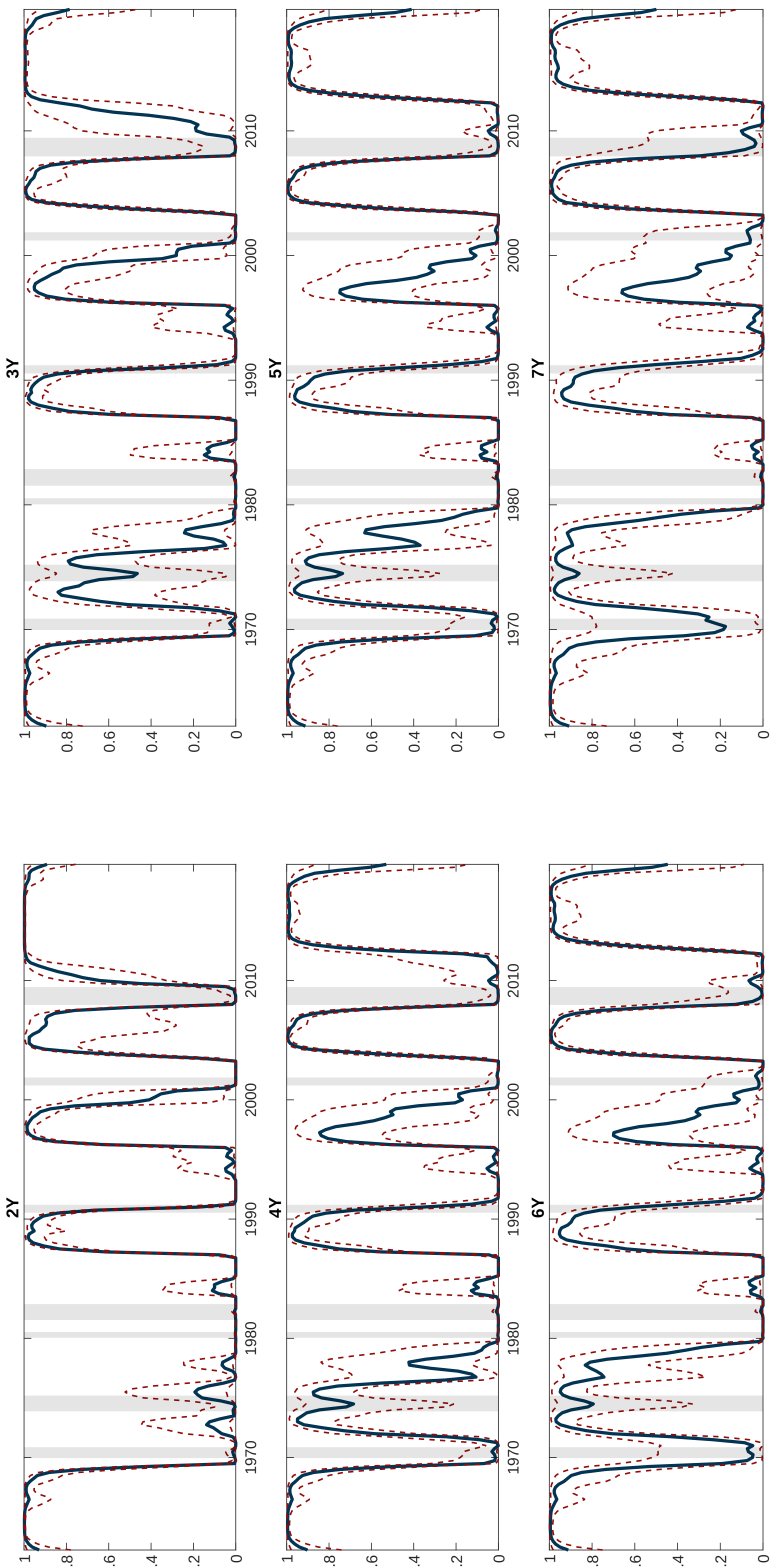


Figure B.1: Median and 90% credibility for the probabilities of the EH state based on Bayesian estimation of the CS specification with GSW data.

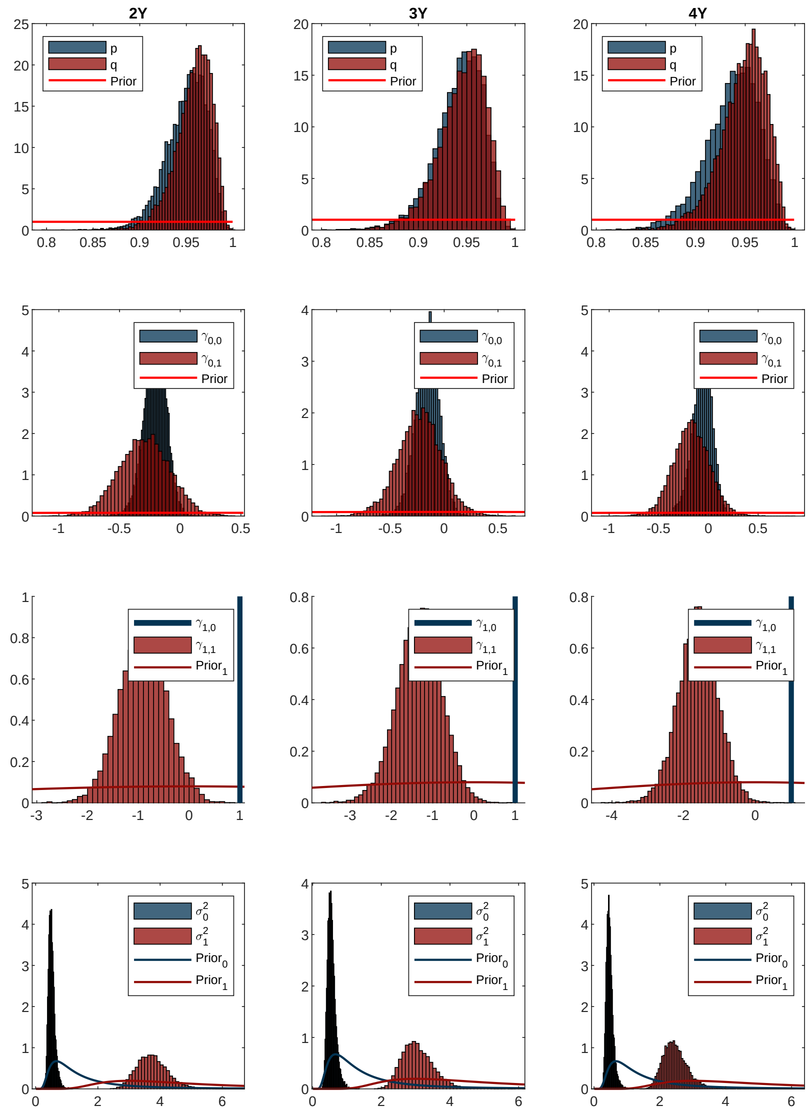


Figure B.2: Posterior draws for the model parameters of the CS specification, maturities 2-4Y, from a Gibbs Sampler with 10^4 draws after burning the first 10^4.

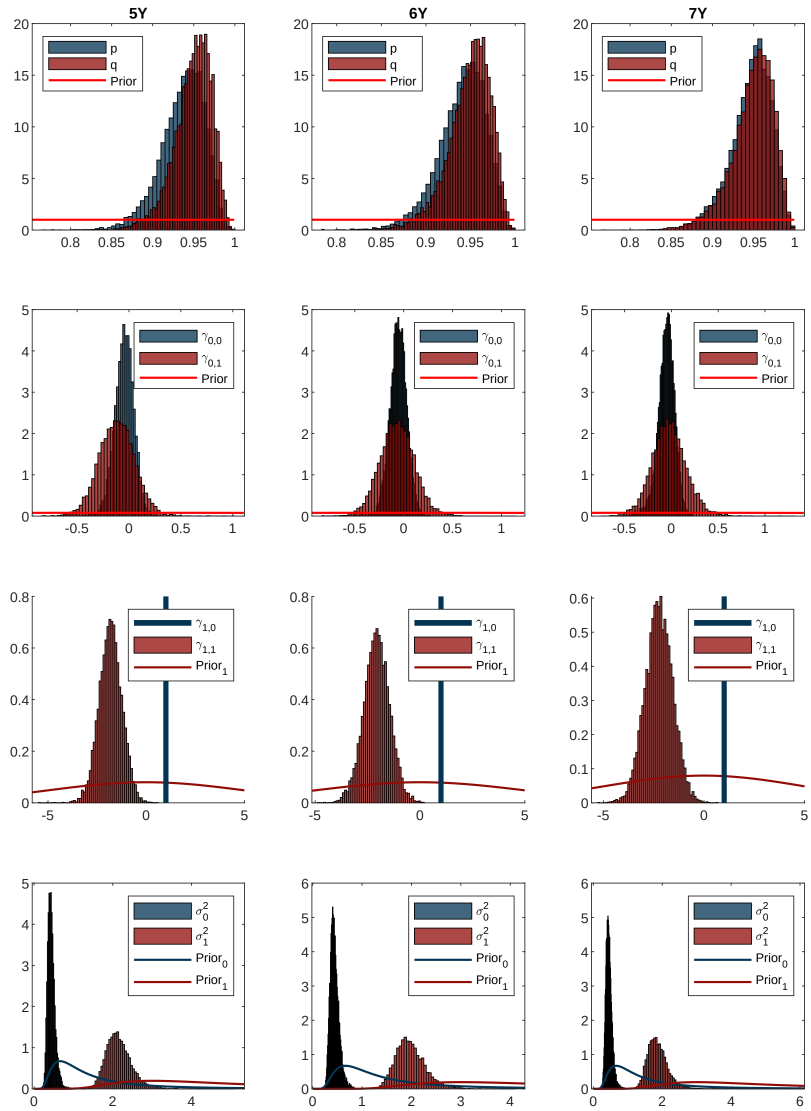


Figure B.3: Posterior draws for the model parameters of the CS specification, maturities 5-7Y, from a Gibbs Sampler with 10^4 draws after burning the first 10^4.

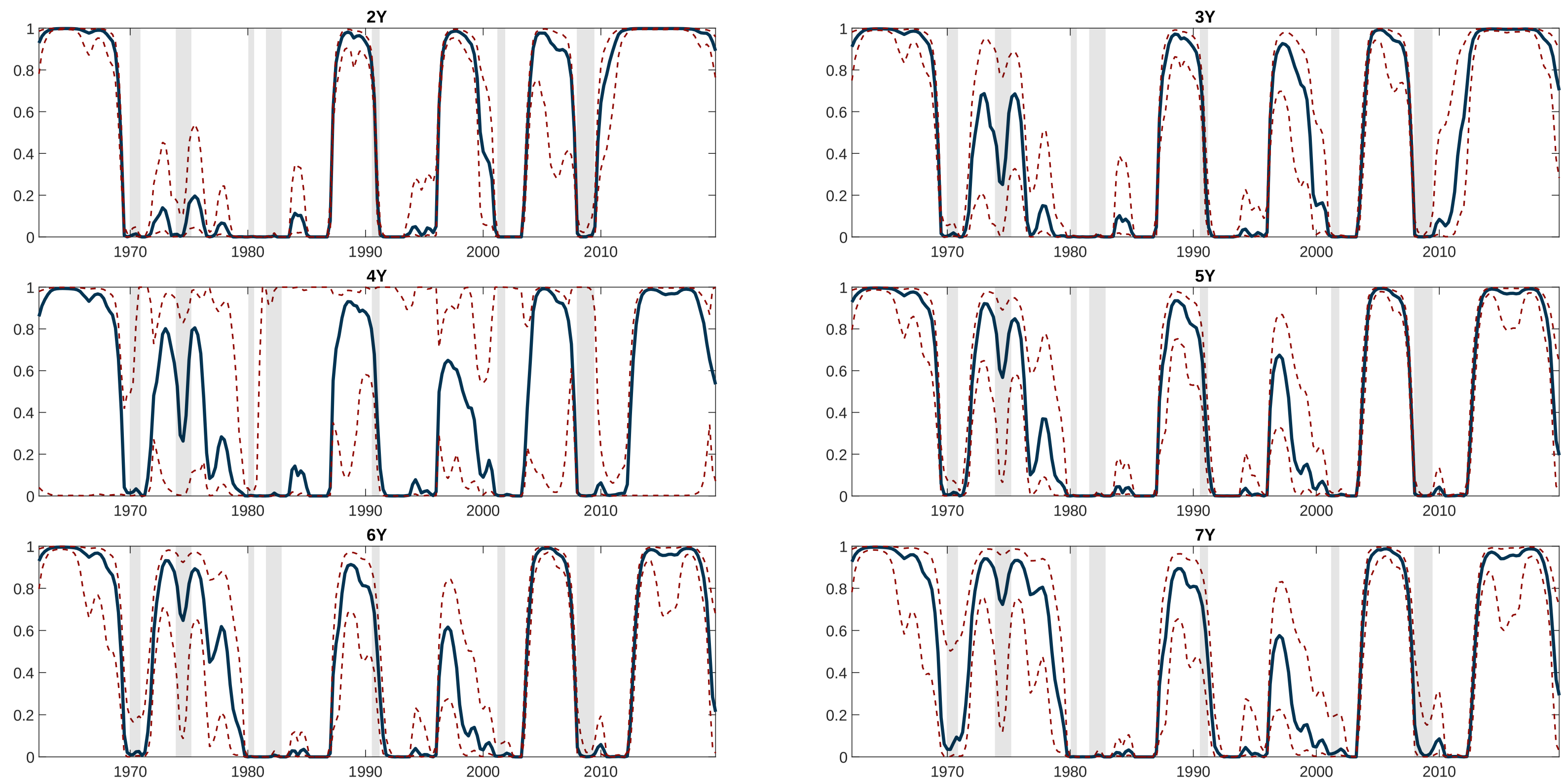


Figure B.4: Median and 90% credibility for the probabilities of the EH state based on Bayesian estimation of the FB specification with GSW data.

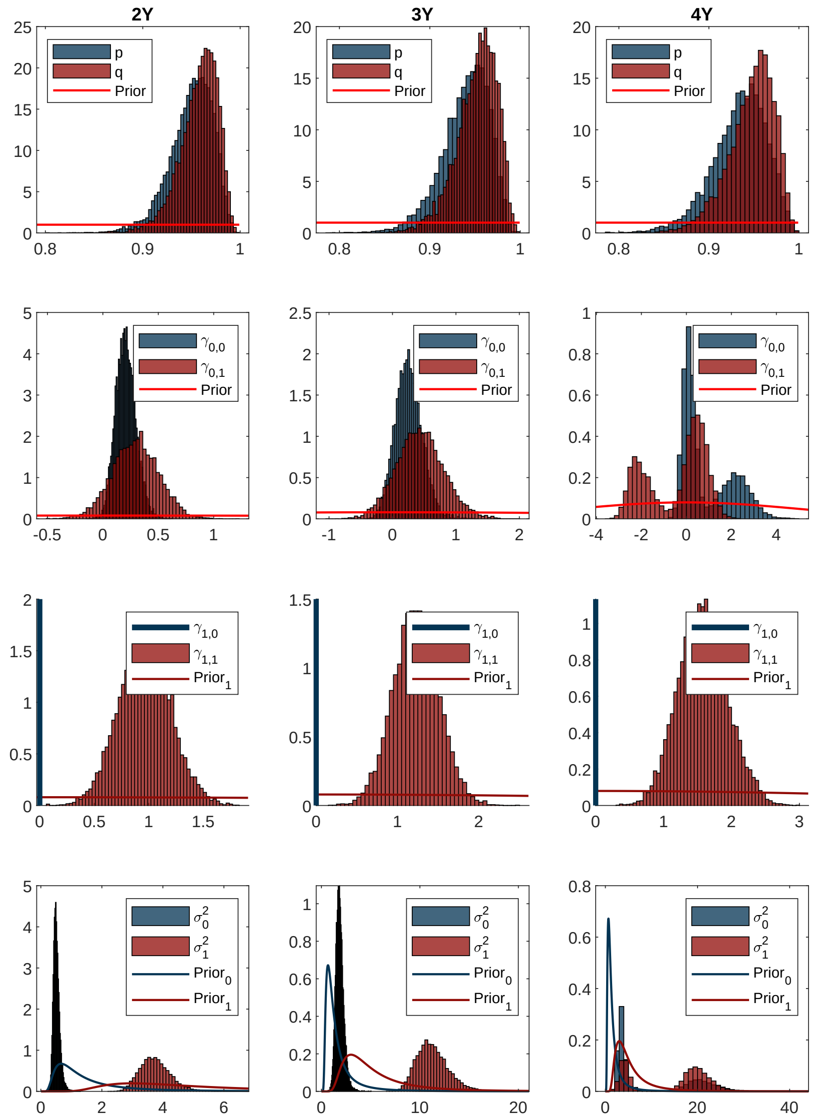


Figure B.5: Posterior draws for the model parameters of the FB specification, maturities 2-4Y, from a Gibbs Sampler with 10^4 draws after burning the first 10^4.

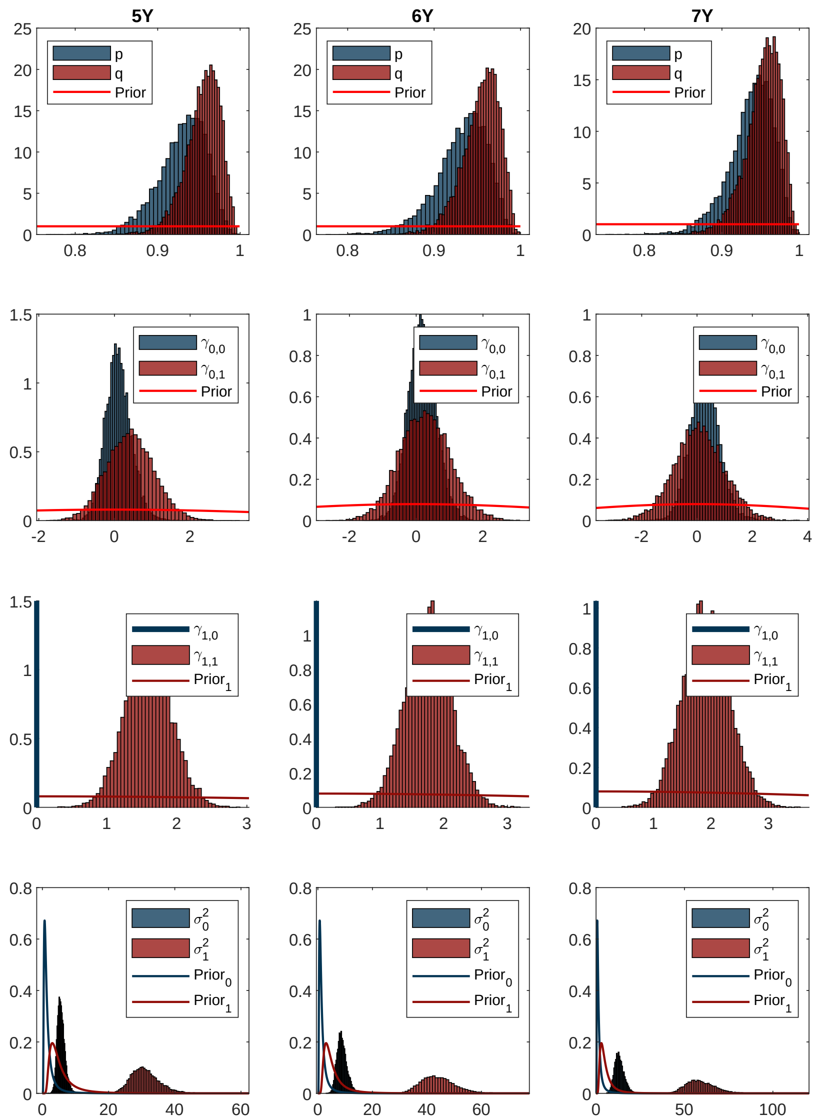


Figure B.6: Posterior draws for the model parameters of the FB specification, maturities 5-7Y, from a Gibbs Sampler with 10^4 draws after burning the first 10^4.

C
Time-Varying Parameters

Another natural way to allow for the variation implied by the rolling regressions is to use a time-varying coefficients model. One caveat here is that this may actually give the model too *much* freedom, which is aggravated by the fact that our sample is small and our variables persistent.

For those who are curious, we offer some insight into this by estimating a model that combines all the above approaches: a Bayesian time-varying coefficients model with Markov switching, in which the $S = 0$ regime implies that the CS intercept and slope are constant, with the intercept equal to 1, while in the second regime they have a VAR structure, i.e.

$$
\begin{aligned}
y_t &= X_t \gamma_t + \varepsilon_t \\
\gamma_t &= \begin{cases} \phi_0 & , s_t = 0 \\ \phi_1 + \Phi_1 \gamma_{t-1} + \epsilon_{1,t} & , s_t = 1 \end{cases} \\
\varepsilon_t &\sim \mathbf{N}\left(0, \sigma_S^2\right) \\
\epsilon_{1,t} &\sim \mathbf{N}\left(\mathbf{0}, \mathbf{T_1}\right) \\
\mathbb{P}(s_t = 0 | s_{t-1} = 0) &= p \\
\mathbb{P}(s_t = 1 | s_{t-1} = 1) &= q
\end{aligned}
$$

The prior for $\mathbb{P}(\phi_{0,1} = 1) = 1$ collapses to the EH value, with $\mathbb{P}(\phi_{0,0}) \sim \mathbf{N}(0, B)$, and the state VAR following a Normal-inverse Wishart[1]: $\mathbb{P}\left((\phi_{1,0}, \Phi_{1,00}, \Phi_{1,01}, \phi_{1,1}, \Phi_{1,10}, \Phi_{1,11})' | \mathbf{T_1}\right) \sim \mathbf{N}\left(\underline{b_1}, \mathbf{T_1} \otimes \underline{\mathbf{B_1}}\right), \mathbb{P}\left(\mathbf{T_1}\right) \sim iW\left(\underline{\mathbf{T_1}}, \underline{\nu_1^{(T)}}\right)$ with $\underline{b_1} = (0, 1, 0, -1, 0, 1)'$, $\underline{\mathbf{T_1}} = \mathbf{I_2}$,

$$
\underline{\mathbf{B_1}} = B \begin{bmatrix} 1 & 1/2 & 1/2 \\ & 1 & 1/2 \\ \cdot & & 1 \end{bmatrix}
$$

and $B = 10$. Priors for p, q, σ_S^2 follow the previous specifications.

Results for maturity $n = 2$ come from a Gibbs sampler and are still unstable and preliminary, but corroborate with the analysis above. The first

[1]See Kadiyala and Karlsson (1997) for details.

plot in Figure C.1 shows the posterior median for $-\frac{1}{2}\gamma_{0,t}$, which is equal to the term premium when $\gamma_{1,t} = 1$, with 90% credibility interval, and compares with the term premium from ACM. The second plot shows the posterior median path of $\gamma_{1,t}$ coefficient and 90% C.I., and the third plots posterior median and C.I. for the probability of the EH state.

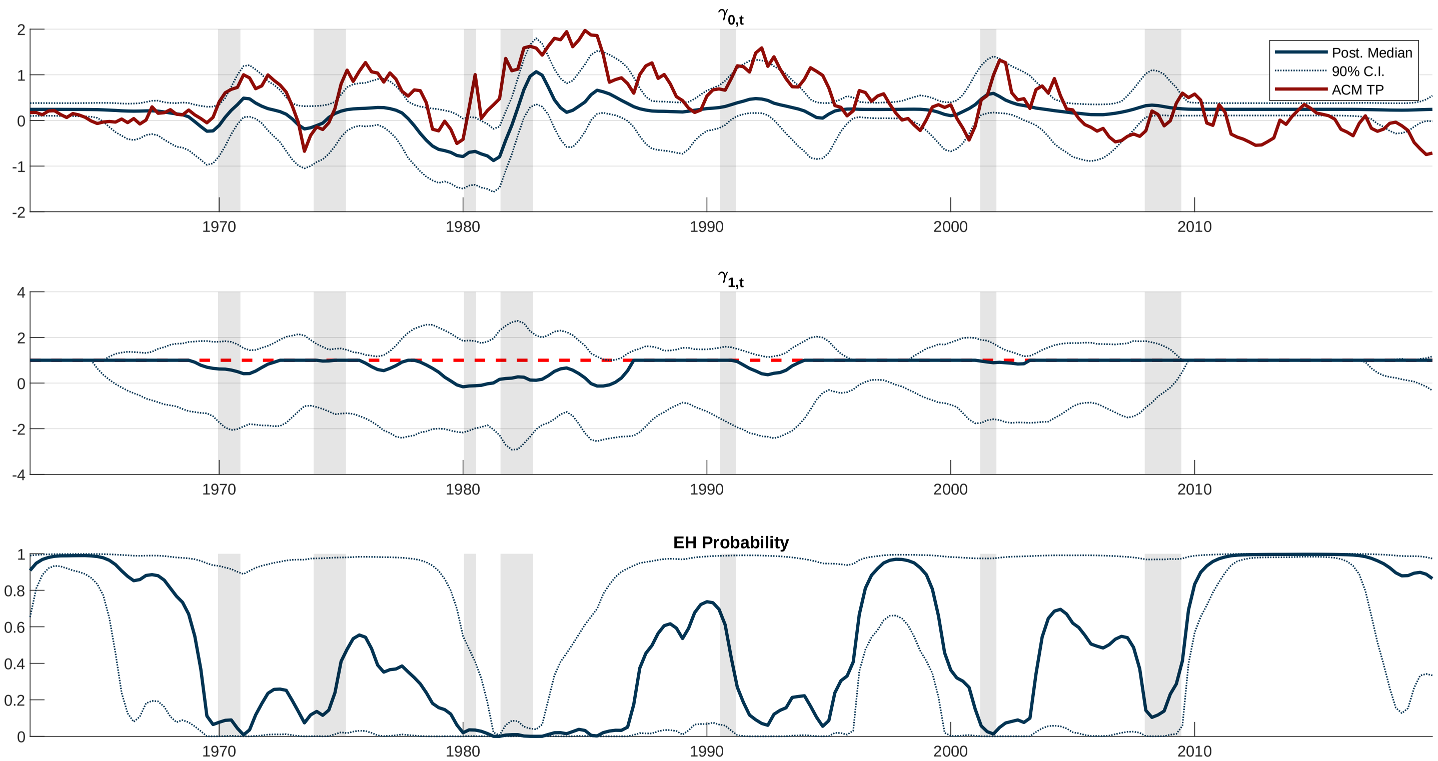
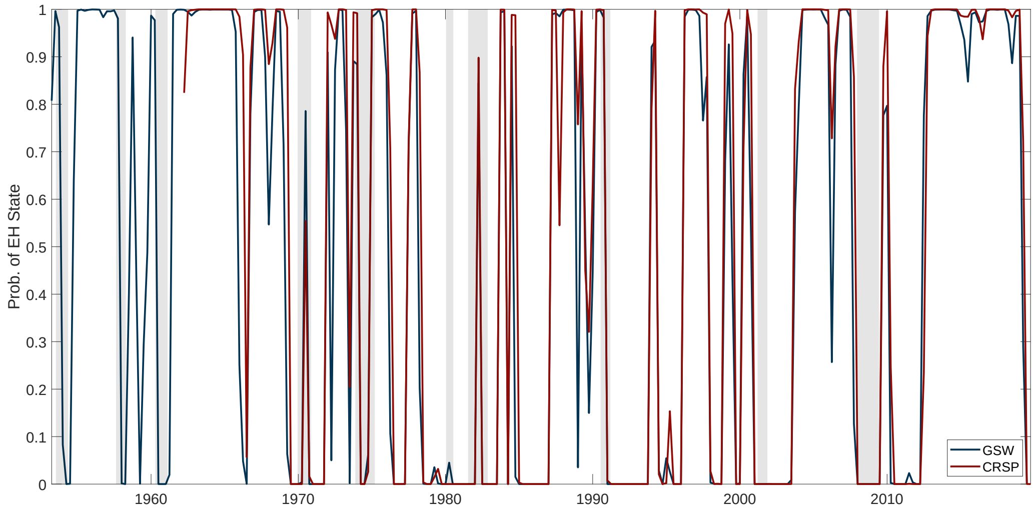


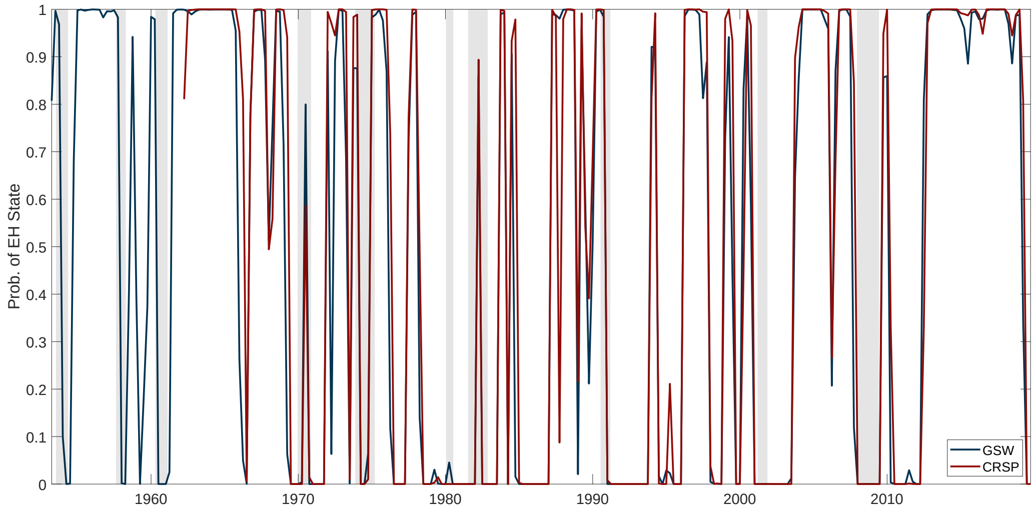
Figure C.1: Results for the time-varying parameter model with Markov switching for maturity $n = 2$. The first plot shows the posterior median for $-\frac{1}{2}\gamma_{0,t}$, which is equal to the term premium when $\gamma_{1,t} = 1$, with 90% credibility interval, and compares with the term premium from ACM. The second plot shows the posterior median path of the $\gamma_{1,t}$ coefficient and 90% C.I., and the third plots posterior median and 90% C.I. for the probability of the EH state.

D
Joint Estimation

The joint model is estimated by stacking and imposing the EH state across all available maturities. An important caveat is that the variance matrix in the estimation is has all off-diagonal elements equal to zero. A full variance matrix proved too unstable. The smoothed probabilities for both the CS and FB specification are reported in Figure D.1, for GSW and CRSP data.



(a)



(b)

Figure D.1: Figure D.1a plots smoothed probabilities of the EH state according to ML estimates of the joint model (3) for GSW data from 1961Q2-2019Q3 and CRSP data from 1952Q2-2019Q43 Figure D.1b is analogous, but using the FB specification. Shaded regions indicate US recessions.

E
Variable Selection with Lasso

Using the FRED-QD database discussed in Section 3, Lasso regressions can be used for variable selection. Figure E.1 plots the coefficients from Lasso logistic regression of the EH probabilities from joint model on FRED-QD variables, with selected variables' names displayed. A point worth noting is that the only variable selected by the 1SE criteria is "Net Worth of HH and Nonprofit Organizations Relative to Disposable Personal Income", although we offer no intuition for why this is. Figure E.2 plots the coefficients for each maturity's estimated EH probability. For parsimony, a complete list with variable descriptions is omitted, but available on request.

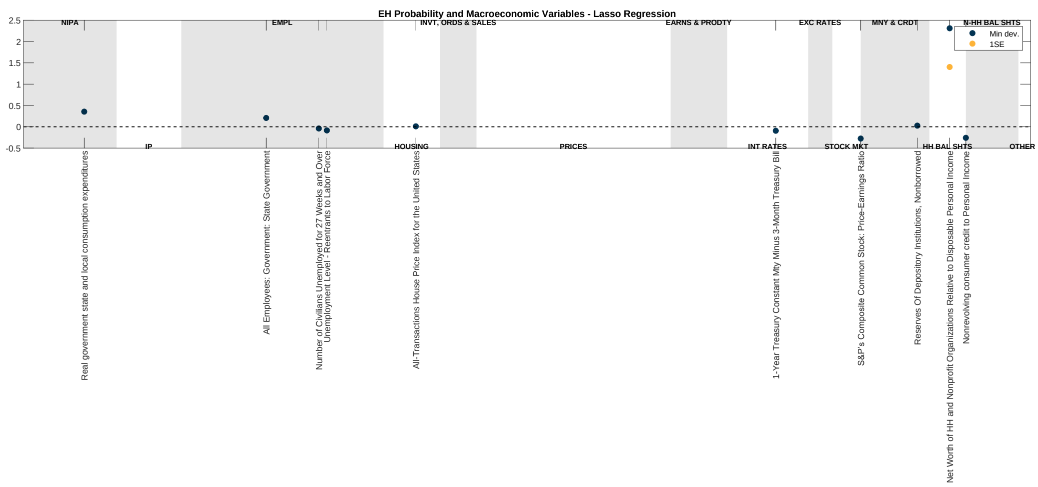


Figure E.1: Coefficients from Lasso logistic regression of EH probabilities from joint model on FRED-QD variables, using minimum deviance and 1 S.E. criteria. All but selected variables' names are omitted.

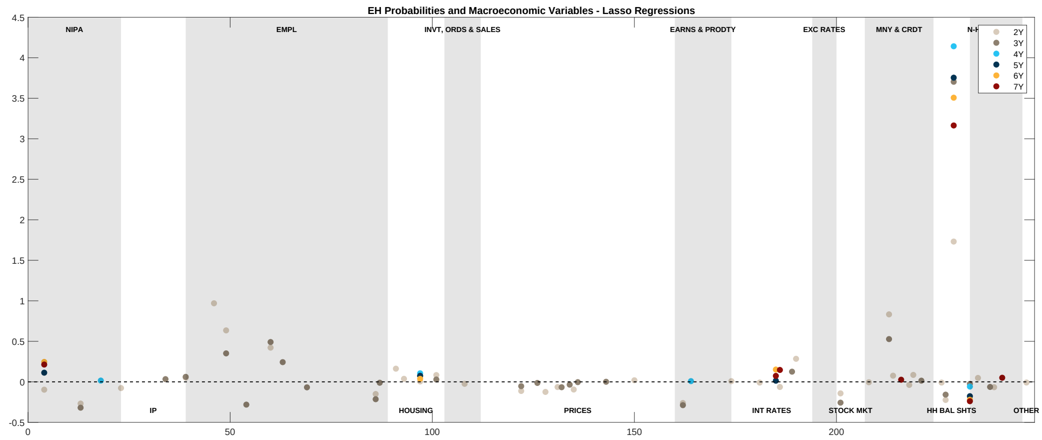


Figure E.2: Coefficients from Lasso logistic regression of EH probabilities from each maturity's estimate on FRED-QD variables, using 1 S.E. criteria. Variable's named omitted.

F
In-Sample Predictability

Predicting One Quarter-Ahead GDP Growth										
	Maturity (yrs)									
	GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5
Term Spread	8.44	4.59	3.11	2.30	1.79	1.45	7.78	4.33	2.96	2.22
	(1.72)	(1.03)	(0.79)	(0.67)	(0.60)	(0.55)	(1.57)	(0.97)	(0.73)	(0.64)
R^2	10%	8%	7%	5%	4%	3%	8%	6%	5%	4%
CS										
Term Spread	8.11	4.05	2.68	1.94	1.41	1.00	7.61	3.48	2.25	1.71
	(1.81)	(1.07)	(0.82)	(0.70)	(0.64)	(0.59)	(1.73)	(1.01)	(0.76)	(0.67)
Term Spread * $\mathbb{P}$(EH)	3.05	3.59	2.62	2.15	2.12	2.25	0.97	4.25	3.35	2.28
	(3.13)	(1.87)	(1.39)	(1.15)	(1.03)	(0.95)	(3.27)	(2.03)	(1.51)	(1.32)
R^2	10%	9%	8%	6%	5%	5%	8%	7%	6%	5%
CS-Joint										
Term Spread	6.42	3.53	2.40	1.76	1.35	1.07	6.16	3.36	2.35	1.80
	(1.87)	(1.11)	(0.85)	(0.72)	(0.65)	(0.60)	(1.78)	(1.07)	(0.79)	(0.71)
Term Spread * $\mathbb{P}$(EH)	7.12	3.94	2.75	2.13	1.75	1.51	5.15	3.86	2.56	1.74
	(3.09)	(1.74)	(1.27)	(1.03)	(0.88)	(0.78)	(3.15)	(1.85)	(1.32)	(1.11)
R^2	12%	10%	8%	7%	5%	4%	8%	8%	6%	4%
FB										
Term Spread	8.11	4.10	2.70	1.96	1.43	1.03	4.94	3.46	2.28	1.74
	(1.81)	(1.07)	(0.82)	(0.70)	(0.64)	(0.59)	(2.84)	(1.02)	(0.76)	(0.67)
Term Spread * $\mathbb{P}$(EH)	3.05	3.40	2.51	2.02	1.95	2.02	3.37	4.25	3.24	2.15
	(3.13)	(1.86)	(1.38)	(1.14)	(1.02)	(0.92)	(3.35)	(1.99)	(1.49)	(1.32)
R^2	10%	9%	7%	6%	5%	5%	8%	7%	6%	5%
FB-Joint										
Term Spread	6.47	3.57	2.42	1.78	1.37	1.08	6.19	3.38	2.37	1.82
	(1.85)	(1.10)	(0.85)	(0.72)	(0.65)	(0.60)	(1.78)	(1.08)	(0.80)	(0.71)
Term Spread * $\mathbb{P}$(EH)	7.16	3.91	2.69	2.07	1.69	1.45	5.03	3.73	2.45	1.66
	(3.08)	(1.73)	(1.26)	(1.02)	(0.87)	(0.78)	(3.12)	(1.83)	(1.30)	(1.09)
R^2	12%	10%	8%	7%	5%	4%	8%	8%	6%	4%
Obs	226	226	226	226	226	226	262	262	262	262

Table F.1: In-sample results for term spread predictability of GDP conditioned on the EH state probabilities. The first panel refers to the benchmark model while the second adds the interaction with the EH state maturity-specific probabilities from the Campbell-Shiller specification, third uses the jointly estimated probabilities and the fourth and fifth use the Fama-Bliss specification. Hodrick (1992) standard errors in parenthesis. Coefficients and standard errors $\times 10^{-3}$.

Predicting One Quarter-Ahead Excess Bond Returns										
	Maturity (yrs)									
	GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5
Term Spread	3.91	3.21	3.11	3.14	3.21	3.30	3.62	3.31	3.21	3.14
	(1.71)	(1.36)	(1.24)	(1.19)	(1.17)	(1.17)	(1.34)	(1.22)	(1.14)	(1.10)
R^2	6%	6%	6%	6%	6%	6%	6%	7%	7%	7%
CS										
Term Spread	4.59	3.95	3.78	3.76	3.83	3.92	4.59	4.09	3.93	3.84
	(1.78)	(1.45)	(1.34)	(1.28)	(1.27)	(1.28)	(1.51)	(1.38)	(1.28)	(1.25)
Term Spread * $\mathbb{P}$(EH)	-1.35	-0.87	-0.19	0.19	0.48	0.83	-1.27	-0.05	0.38	0.60
	(1.98)	(1.58)	(1.43)	(1.41)	(1.45)	(1.48)	(1.52)	(1.35)	(1.26)	(1.25)
R^2	9%	9%	8%	8%	8%	8%	9%	9%	9%	8%
CS-Joint										
Term Spread	5.39	4.43	4.22	4.18	4.20	4.25	5.00	4.28	4.12	4.07
	(2.00)	(1.55)	(1.40)	(1.34)	(1.30)	(1.29)	(1.69)	(1.42)	(1.30)	(1.25)
Term Spread * $\mathbb{P}$(EH)	0.19	-0.09	-0.02	0.14	0.33	0.52	0.48	0.39	0.30	0.24
	(1.71)	(1.42)	(1.32)	(1.30)	(1.31)	(1.34)	(1.27)	(1.23)	(1.21)	(1.21)
R^2	10%	10%	10%	10%	10%	9%	9%	10%	10%	10%
FB										
Term Spread	4.59	3.92	3.76	3.75	3.82	3.91	2.64	4.12	3.91	3.80
	(1.78)	(1.45)	(1.34)	(1.28)	(1.27)	(1.28)	(1.48)	(1.38)	(1.28)	(1.25)
Term Spread * $\mathbb{P}$(EH)	-1.35	-0.84	-0.16	0.26	0.61	1.00	3.81	-0.06	0.43	0.75
	(1.98)	(1.58)	(1.44)	(1.41)	(1.42)	(1.44)	(1.48)	(1.35)	(1.26)	(1.25)
R^2	9%	9%	8%	8%	8%	8%	7%	9%	9%	8%
FB-Joint										
Term Spread	5.35	4.41	4.21	4.18	4.21	4.26	5.01	4.29	4.13	4.08
	(1.98)	(1.54)	(1.40)	(1.33)	(1.30)	(1.29)	(1.69)	(1.42)	(1.30)	(1.25)
Term Spread * $\mathbb{P}$(EH)	0.17	-0.12	-0.05	0.11	0.30	0.50	0.47	0.41	0.33	0.28
	(1.71)	(1.41)	(1.32)	(1.29)	(1.30)	(1.33)	(1.27)	(1.21)	(1.20)	(1.20)
R^2	10%	10%	10%	10%	10%	10%	9%	10%	10%	10%
Obs	230	230	230	230	230	230	232	232	232	232

Table F.2: In-sample results for excess bond return predictability conditioned on the EH state probabilities. The first panel refers to the benchmark model while the second adds the interaction with the EH state maturity-specific probabilities from the Campbell-Shiller specification, third uses the jointly estimated probabilities and the fourth and fifth use the Fama-Bliss specification. Hodrick (1992) standard errors in parenthesis.

G
Out-of-Sample Predictability

Tables G.1 and G.2 gather out-of-sample results for CS, FB specifications using GSW and CRSP data for predicting GDP and bond excess returns. Table G.3 gathers results comparing with the affine model decomposition.

Pseudo Out-of-Sample Forecasting GDP Growth

Horizon (Q)	Campbell-Shiller Maturity (Y)										Fama-Bliss Maturity (Y)									
	GSW						CRSP				GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5	2	3	4	5	6	7	2	3	4	5
RMPE Ratio																				
1	0.983	0.969	0.970	0.969	0.968	0.963	0.995	0.981	0.989	0.984	1.005	0.990	0.974	0.977	0.977	0.944	1.010	0.983	0.984	0.980
2	0.985	0.971	0.973	0.968	0.964	0.959	0.997	0.977	0.985	0.983	1.013	0.996	0.977	0.977	0.975	0.937	1.025	0.980	0.981	0.976
3	0.987	0.980	0.982	0.978	0.975	0.972	1.001	0.982	0.992	0.990	1.019	1.014	0.986	0.988	0.988	0.947	1.035	0.986	0.989	0.980
4	0.988	0.988	0.991	0.988	0.985	0.984	1.010	0.986	0.999	0.997	1.019	1.019	0.996	0.999	1.000	0.953	1.036	0.990	0.996	0.986
CW																				
1	2.043	1.954	1.833	1.833	1.837	1.998	1.110	1.529	1.315	1.326	0.049	0.879	1.716	1.619	1.638	2.604	0.357	1.499	1.282	1.449
2	1.423	1.486	1.435	1.568	1.757	2.001	0.743	1.290	1.161	1.171	-0.077	0.626	1.365	1.407	1.568	2.804	0.227	1.255	1.221	1.338
3	1.141	1.244	1.244	1.392	1.588	1.779	0.552	1.134	1.036	1.011	-0.120	0.220	1.168	1.238	1.384	2.403	0.049	1.044	1.001	1.252
4	1.043	1.030	1.026	1.158	1.333	1.477	0.310	0.974	0.847	0.821	-0.067	0.090	0.965	0.999	1.124	2.200	0.021	0.876	0.781	1.115
Joint Models RMPE Ratio																				
1	0.982	0.985	0.989	0.993	0.995	0.997	0.991	0.992	0.998	0.996	0.984	0.988	0.992	0.994	0.996	0.996	0.991	0.993	0.999	0.996
2	0.978	0.981	0.986	0.991	0.994	0.997	0.998	0.994	0.998	1.001	0.983	0.987	0.992	0.995	0.998	1.000	0.997	0.994	0.999	1.000
3	0.982	0.985	0.991	0.996	1.000	1.003	1.010	1.004	1.007	1.010	0.988	0.993	0.998	1.002	1.005	1.006	1.009	1.004	1.007	1.010
4	0.983	0.987	0.994	0.999	1.003	1.006	1.016	1.008	1.011	1.014	0.989	0.994	1.000	1.004	1.006	1.008	1.016	1.009	1.012	1.014
CW																				
1	1.894	1.862	1.802	1.758	1.733	1.722	1.462	1.489	1.321	1.264	1.687	1.630	1.597	1.587	1.590	1.597	1.407	1.417	1.259	1.215
2	1.767	1.686	1.612	1.571	1.553	1.546	0.634	1.023	0.855	0.645	1.440	1.362	1.331	1.335	1.351	1.368	0.747	1.017	0.849	0.664
3	1.444	1.415	1.372	1.354	1.348	1.346	-0.881	0.035	0.084	-0.379	1.133	1.089	1.080	1.102	1.131	1.154	-0.833	0.044	0.061	-0.396
4	1.304	1.276	1.232	1.212	1.203	1.195	-1.145	-0.390	-0.265	-0.693	1.036	0.985	0.968	0.982	1.004	1.021	-1.129	-0.415	-0.310	-0.733
Obs	115	115	115	115	115	115	133	133	133	133	115	115	115	115	115	115	133	133	133	133

Table G.1: Out-of-sample results for term spread predictability of GDP conditioned on the EH state probabilities. Root mean prediction error ratios and Clark and West (2007) statistics are relative to term spread alone.

Pseudo Out-of-Sample Forecasting Excess Bond Returns

Horizon (Q)	Campbell-Shiller Maturity (Y)										Fama-Bliss Maturity (Y)									
	GSW						CRSP				GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5	2	3	4	5	6	7	2	3	4	5
RMPE Ratio																				
1	0.942	0.956	0.976	0.992	1.008	1.017	0.971	0.976	0.981	1.007	0.983	0.984	0.983	1.000	1.013	1.009	1.037	0.971	0.990	1.000
2	0.927	0.948	0.971	0.989	1.006	1.016	0.965	0.971	0.977	1.001	0.997	0.990	0.978	0.995	1.010	1.013	1.059	0.968	0.983	0.999
3	0.922	0.956	0.975	0.990	1.006	1.016	0.964	0.976	0.976	1.003	0.984	0.997	0.983	0.994	1.010	1.017	1.051	0.973	0.987	0.998
4	0.920	0.961	0.976	0.990	1.007	1.017	0.961	0.979	0.974	0.998	0.972	0.991	0.985	0.991	1.011	1.023	1.031	0.972	0.989	0.992
CW																				
1	3.133	3.705	2.885	2.072	0.649	-0.371	2.202	3.166	2.769	0.259	1.678	2.003	2.409	1.542	0.374	-0.133	-1.035	2.926	1.881	1.422
2	2.858	3.643	3.209	2.343	1.047	-0.045	2.003	3.078	3.228	0.823	1.035	1.422	2.820	2.019	0.838	-0.390	-2.499	2.628	2.681	1.518
3	2.515	2.904	2.513	1.904	0.936	0.011	1.820	2.089	2.731	0.445	1.559	0.798	2.066	1.675	0.792	-0.470	-2.185	2.060	1.780	1.263
4	2.333	2.367	2.013	1.573	0.726	0.008	1.869	1.654	2.603	0.786	2.229	0.863	1.637	1.485	0.623	-0.512	-1.111	1.961	1.336	1.475
Joint Models RMPE Ratio																				
1	0.971	0.978	0.987	0.995	1.003	1.010	0.957	0.967	0.976	0.981	0.968	0.976	0.986	0.994	1.002	1.009	0.955	0.964	0.972	0.976
2	0.972	0.972	0.979	0.988	0.997	1.004	0.940	0.953	0.962	0.972	0.968	0.970	0.979	0.989	0.998	1.006	0.937	0.950	0.958	0.969
3	0.976	0.975	0.979	0.984	0.990	0.995	0.956	0.961	0.963	0.972	0.972	0.973	0.978	0.985	0.991	0.997	0.952	0.957	0.960	0.970
4	0.979	0.979	0.981	0.984	0.988	0.991	0.967	0.967	0.967	0.973	0.976	0.977	0.981	0.984	0.988	0.992	0.963	0.963	0.963	0.969
CW																				
1	3.339	3.018	2.516	2.033	1.580	1.159	3.150	3.110	2.864	2.647	3.701	3.262	2.749	2.250	1.790	1.356	3.210	3.211	3.030	2.851
2	3.053	3.226	2.805	2.250	1.706	1.199	3.901	3.805	3.475	3.024	3.265	3.348	2.879	2.298	1.739	1.225	3.887	3.802	3.548	3.150
3	2.545	2.836	2.688	2.334	1.911	1.460	3.433	3.315	3.268	2.838	2.808	2.942	2.705	2.305	1.845	1.372	3.424	3.291	3.276	2.861
4	2.243	2.316	2.282	2.112	1.884	1.620	3.029	3.217	3.339	3.345	2.486	2.523	2.475	2.289	2.032	1.733	3.092	3.223	3.404	3.440
Obs	115	115	115	115	115	115	133	133	133	133	115	115	115	115	115	115	133	133	133	133

Table G.2: Out-of-sample results for term spread predictability of excess bond returns conditioned on the EH state probabilities. Root mean prediction error ratios and Clark and West (2007) statistics are relative to term spread alone.

Pseudo Out-of-Sample Forecasting Excess Bond Returns EH Probabilities vs. ACM Decomposition																				
Horizon (Q)	Campbell-Shiller Maturity (Y)										Fama-Bliss Maturity (Y)									
	GSW						CRSP				GSW						CRSP			
	2	3	4	5	6	7	2	3	4	5	2	3	4	5	6	7	2	3	4	5
RMPE Ratio																				
1	0.990	0.989	0.996	1.003	1.015	1.033	1.044	1.016	1.030	1.050	1.006	1.024	1.018	1.020	1.029	1.025	1.074	1.032	1.030	1.030
2	0.978	0.992	1.012	1.028	1.042	1.058	1.045	1.025	1.051	1.065	1.018	1.034	1.035	1.042	1.057	1.059	1.088	1.043	1.053	1.050
3	0.928	0.976	1.016	1.044	1.064	1.083	1.007	1.004	1.050	1.068	0.990	1.019	1.038	1.058	1.083	1.092	1.051	1.023	1.051	1.053
4	0.877	0.946	0.998	1.039	1.069	1.092	0.957	0.962	1.023	1.045	0.938	0.970	1.013	1.050	1.088	1.106	0.999	0.974	1.023	1.038
DM																				
1	0.200	0.274	0.111	-0.096	-0.441	-1.064	-0.882	-0.508	-0.801	-1.225	-0.122	-0.564	-0.571	-0.665	-0.929	-0.781	-1.668	-0.878	-0.846	-0.701
2	0.277	0.132	-0.220	-0.499	-0.765	-1.047	-0.617	-0.439	-0.767	-0.958	-0.204	-0.455	-0.616	-0.745	-1.016	-1.044	-1.174	-0.638	-0.807	-0.779
3	0.677	0.268	-0.179	-0.503	-0.726	-0.929	-0.066	-0.047	-0.514	-0.688	0.081	-0.177	-0.424	-0.652	-0.922	-0.955	-0.456	-0.234	-0.526	-0.571
4	1.001	0.478	0.014	-0.342	-0.595	-0.790	0.369	0.335	-0.186	-0.356	0.461	0.255	-0.112	-0.431	-0.746	-0.799	0.010	0.225	-0.190	-0.313
Joint Models RMPE Ratio																				
1	0.994	1.011	1.020	1.026	1.031	1.033	1.013	1.006	1.010	1.009	0.993	1.011	1.020	1.026	1.030	1.032	1.010	1.002	1.007	1.006
2	0.971	1.011	1.037	1.054	1.066	1.073	0.996	1.011	1.025	1.033	0.966	1.009	1.035	1.052	1.063	1.070	0.995	1.011	1.026	1.035
3	0.940	0.991	1.027	1.053	1.070	1.081	0.973	0.996	1.019	1.034	0.936	0.990	1.027	1.053	1.070	1.081	0.969	0.994	1.019	1.037
4	0.911	0.962	1.003	1.034	1.056	1.072	0.943	0.969	0.998	1.022	0.909	0.963	1.005	1.036	1.059	1.074	0.938	0.965	0.995	1.020
DM																				
1	0.122	-0.313	-0.619	-0.805	-0.906	-0.955	-0.354	-0.168	-0.307	-0.266	0.147	-0.309	-0.609	-0.773	-0.852	-0.883	-0.259	-0.063	-0.204	-0.183
2	0.390	-0.177	-0.614	-0.899	-1.073	-1.178	0.058	-0.177	-0.390	-0.517	0.453	-0.139	-0.580	-0.855	-1.016	-1.108	0.083	-0.173	-0.409	-0.568
3	0.604	0.103	-0.322	-0.637	-0.854	-0.996	0.330	0.053	-0.228	-0.410	0.646	0.117	-0.325	-0.640	-0.851	-0.985	0.374	0.069	-0.230	-0.443
4	0.743	0.348	-0.027	-0.329	-0.553	-0.709	0.573	0.304	0.019	-0.206	0.758	0.342	-0.048	-0.353	-0.573	-0.722	0.621	0.338	0.047	-0.195
Obs	115	115	115	115	115	115	133	133	133	133	115	115	115	115	115	115	133	133	133	133

Table G.3: Out-of-sample results for predicting excess returns conditioned on the EH state probabilities vs. the ACM decomposition.