Online Appendix for:
Learning About the Long Run

Leland E. Farmer, Emi Nakamura, and Jón Steinsson

December 2023

A Individual Forecaster Anomalies

Table A.1: Individual Forecast Anomalies

	Forecast Horizon			
	1	2	3	4
Panel A: Bias				
Consensus Forecast	-0.18***	-0.34***	-0.52***	-0.70***
	(0.05)	(0.09)	(0.14)	(0.19)
Pooled Individual	-0.19***	-0.34***	-0.50***	-0.68***
	(0.01)	(0.02)	(0.03)	(0.04)
Median Individual	-0.20	-0.34	-0.47	-0.62
Panel B: Autocorrelation				
Consensus Forecast	0.30*	0.27**	0.24*	0.13
	(0.14)	(0.12)	(0.12)	(0.13)
Pooled individual	0.34***	0.34***	0.30***	0.16***
	(0.04)	(0.026)	(0.024)	(0.025)
Median individual	0.39	0.37	0.26	0.09
Panel C: Mincer-Zarnowitz				
Consensus Forecast	0.97*	0.94**	0.90**	0.86**
	(0.02)	(0.02)	(0.04)	(0.05)
Pooled individual	0.95***	0.92***	0.88***	0.83***
	(0.006)	(0.008)	(0.009)	(0.012)
Median Individual	0.90	0.82	0.75	0.64
Panel D: Coibion-Gorodnichenko				
Consensus Forecast	0.23*	0.34*	0.62***	
	(0.12)	(0.16)	(0.16)	
Pooled individual	0.05	0.18***	0.29***	
	(0.030)	(0.031)	(0.034)	
Median Individual	0.13	0.19	0.28	

Note: The "Consensus Forecast" results are for regressions using the mean SPF forecast across forecasters – same as in Table 1. The "Pooled Individual" results are for regressions where the individual forecasts in the SPF are pooled together in single regression. For the "Median Individual" results we run separate regressions for each forecaster in the SPF and report the median across forecasters. The forecast horizons are quarters. Stars represent significance relative to the following hypotheses: $\alpha = 0$ for bias, $\beta = 0$ for autocorrelation, $\beta = 1$ for Mincer-Zarnowitz, $\beta = 0$ for Coibion-Gorodnichenko. Standard errors are reported in parentheses. P-values are computed using Newey-West Standard Errors. *p < 0.1, **p < 0.05, ***p < 0.01.

B Bayesian Updating about Parameters and States for Interest Rates

In this appendix, we describe in more detail the Gibbs Sampling algorithm we use to sample from the joint posterior of the model parameters and latent states in the UC model for the short rate. Define the vector of the parameters and latent states of the model through date t as $\boldsymbol{\theta} \equiv (\rho, \gamma, \sigma, \boldsymbol{\mu}_{1:t}, \boldsymbol{x}_{1:t})'$. Let $p(\boldsymbol{\theta})$ denote the joint prior over the parameter vector $\boldsymbol{\theta}$. Let $L(\boldsymbol{y}_{1:t}|\boldsymbol{\theta})$ denote the likelihood function of the data through time t, given a set of parameters $\boldsymbol{\theta}$. Our goal is to sample from the posterior of the parameters given the data, $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:t})$, where we know

$$p(\boldsymbol{\theta}|\boldsymbol{y}_{1:t}) \propto L(\boldsymbol{y}_{1:t}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

We assume functional forms for the initial beliefs as follows

$$\rho \sim N(\mu_\rho, \sigma_\rho^2)$$
$$\gamma \sim \mathcal{B}(\alpha_\gamma, \beta_\gamma)$$
$$\sigma^2 \sim \mathcal{IG}(\alpha_{\sigma^2}, \beta_{\sigma^2})$$

The initial beliefs for the states are given by

$$\mu_{1951Q2} \sim N(y_{1951Q2}, 1)$$
$$x_{1951Q2} \sim N(0, 1)$$

where y_{1952Q2} denotes the 3-month Treasury bill rate in 1952Q2.

We start with an initial guess of the parameters $\boldsymbol{\theta}^{(0)} = \left(\rho^{(0)}, \gamma^{(0)}, \sigma^{(0)}, \boldsymbol{\mu}_{1:t}^{(0)}, \boldsymbol{x}_{1:t}^{(0)}\right)'$. Given a draw of the parameters $\boldsymbol{\theta}^{(b)}$, we draw $\boldsymbol{\theta}^{(b+1)}$ as follows:

1. Draw $\rho^{(b+1)}|\gamma^{(b)}, \sigma^{(b)}, \boldsymbol{\mu}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}, \boldsymbol{y}_{1:t}$. Given the other parameters, beliefs about ρ can be updated from the autoregression

$$x_t^{(b)} = \rho x_{t-1}^{(b)} + \sqrt{1 - \gamma^{(b)}}\sigma^{(b)}\omega_t$$

Define

$$\tilde{\sigma}_\rho^2 \equiv \left[\sigma_\rho^{-2} + \frac{\sum_{s=2}^t \left(x_{s-1}^{(b)} \right)^2}{\left(1 - \gamma^{(b)} \right) \left(\sigma^{(b)} \right)^2} \right]^{-1}$$

$$\tilde{\mu}_\rho \equiv \tilde{\sigma}_\rho^2 \left[\frac{\mu_\rho}{\sigma_\rho^2} + \frac{\sum_{s=2}^t x_{s-1}^{(b)} x_s^{(b)}}{\left(1 - \gamma^{(b)} \right) \left(\sigma^{(b)} \right)^2} \right]$$

The posterior of ρ is $N(\tilde{\mu}_\rho, \tilde{\sigma}_\rho^2)$ and thus we draw $\rho^{(b+1)} \sim N(\tilde{\mu}_\rho, \tilde{\sigma}_\rho^2)$.

2. Draw $\gamma^{(b+1)}|\rho^{(b+1)}, \sigma^{(b)}, \boldsymbol{\mu}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}, \boldsymbol{y}_{1:t}$. There is no closed form expression for the posterior of γ. We therefore draw it using a random walk Metropolis-Hastings step. Specifically, we draw a proposal $\tilde{\gamma}^{(b+1)} \sim N(\gamma^{(b)}, \sigma_{\gamma,prop}^2)$ where $\sigma_{\gamma,prop}^2$ is a proposal variance chosen such that this step has between a 25 and 40% acceptance rate over the burn-in period. We then set $\gamma^{(b+1)} = \tilde{\gamma}^{(b+1)}$ with probability α_{b+1}, where

$$\alpha_{b+1} \equiv \frac{L\left(\boldsymbol{y}_{1:t}|\rho^{(b+1)}, \tilde{\gamma}^{(b+1)}, \sigma^{(b)}, \boldsymbol{\mu}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)} \right) p_\gamma \left(\tilde{\gamma}^{(b+1)} \right)}{L\left(\boldsymbol{y}_{1:t}|\rho^{(b+1)}, \gamma^{(b)}, \sigma^{(b)}, \boldsymbol{\mu}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)} \right) p_\gamma \left(\gamma^{(b)} \right)}$$

Otherwise we set $\gamma^{(b+1)} = \gamma^{(b)}$.

3. Draw $\sigma^{(b+1)}|\rho^{(b+1)}, \gamma^{(b+1)}, \boldsymbol{\mu}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}, \boldsymbol{y}_{1:t}$. Given the other parameters, beliefs about σ can be updated from the two equations

$$\mu_t^{(b)} = \mu_{t-1}^{(b)} + \sqrt{\gamma^{(b+1)}} \sigma \eta_t$$

$$x_t^{(b)} = \rho^{(b+1)} x_{t-1}^{(b)} + \sqrt{1 - \gamma^{(b+1)}} \sigma \omega_t$$

Since η_t and ω_t are independent, these regression equations can be treated as two independent sources of information for σ^2. It is as if beliefs about σ^2 are first updated using information about $\{\eta_s\}_{s=2}^t$ where $\sigma \eta_s = \frac{\mu_s - \mu_{s-1}}{\sqrt{\gamma}}$ and then updated using information about $\{\omega_s\}_{s=2}^t$ where $\sigma \omega_s = \frac{x_s - \rho x_{s-1}}{\sqrt{1-\gamma}}$. These are each samples of $t - 1$ observations which can be used to

learn about σ^2 using standard conjugate prior updating. Define

$$\tilde{\alpha}_{\sigma^2} \equiv \alpha_{\sigma^2} + (t-1)$$

$$\tilde{\beta}_{\sigma^2} \equiv \beta_{\sigma^2} + \frac{\sum_{s=2}^{t}\left(\mu_s^{(b)} - \mu_{s-1}^{(b)}\right)^2}{2\gamma^{(b+1)}} + \frac{\sum_{s=2}^{t}\left(x_s^{(b)} - \rho^{(b+1)}x_{s-1}^{(b)}\right)^2}{2\left(1 - \gamma^{(b+1)}\right)}$$

The posterior of σ^2 is $\mathcal{IG}(\tilde{\alpha}_{\sigma^2}, \tilde{\beta}_{\sigma^2})$ and thus we draw $\left(\sigma^{(b)}\right)^2 \sim \mathcal{IG}(\tilde{\alpha}_{\sigma^2}, \tilde{\beta}_{\sigma^2})$.

4. Draw $\boldsymbol{\mu}_{1:t}^{(b+1)}, \boldsymbol{x}_{1:t}^{(b+1)} | \rho^{(b+1)}, \gamma^{(b+1)}, \sigma^{(b+1)}, \boldsymbol{y}_{1:t}$. This can be done using the standard Kalman filter and simulation smoother.

This algorithm is repeated to produce B draws from the posterior distribution of the parameters and states at each time t.

C   Bayesian Forecasting of Interest Rates

The algorithm described in Appendix B yields B samples of the posterior of the states and parameters of our UC model at each point in time t. We index these samples by b as follows $\left\{\rho^{(b)}, \gamma^{(b)}, \sigma^{(b)}, \boldsymbol{\mu}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}\right\}_{b=1}^{B}$. Draws b for which $\rho^{(b)} > 1$ are not used for forecasting as they imply explosive dynamics in interest rates. We then use the following algorithm to produce a real-time forecast distribution for the yield curve at time t:

1. For each $b = 1, \ldots, B$

   (a) Simulate a path of shocks $\left\{\eta_{t+h}^{(b)}, \omega_{t+h}^{(b)}\right\}_{h=1}^{H}$ from the standard Normal distribution.

   (b) Starting from $h = 1$, construct a simulated path of the states over H subsequent periods using equations (7)-(8):

   $$\mu_{t+h}^{(b)} = \mu_{t+h-1|t}^{(b)} + \sqrt{\gamma^{(b)}}\sigma^{(b)}\eta_{t+h}^{(b)}$$
   $$x_{t+h}^{(b)} = \rho^{(b)}x_{t+h-1|t}^{(b)} + \sqrt{1-\gamma^{(b)}}\sigma^{(b)}\omega_{t+h}^{(b)}$$

   (c) Use the simulated states to construct the forecast distribution of the short rate $\left\{y_{t+h|t}^{(b)}\right\}_{h=1}^{H}$ where
   $$y_{t+h|t}^{(b)} = \mu_{t+h|t}^{(b)} + x_{t+h|t}^{(b)}$$

5

2. The forecast of y_{t+h} given time t information is computed as

$$F_t y_{t+h} = \frac{1}{B} \sum_{b=1}^{B} y_{t+h|t}^{(b)}$$

The implied yield of a bonds of maturity n is given by

$$y_t^{(n)} = c^{(n)} + \frac{1}{n} \sum_{h=0}^{n-1} F_t y_{t+h}$$

We estimate the constant $c^{(n)}$ as the average level of the corresponding n-period bond yield in the data since it is not identified from the expectations hypothesis alone. Note that this estimate of the constant does not affect the results of the expectations hypothesis regression tests we run since it only affects the level of the n-period yield.

At the end of the estimation we are left with a sequence of model-implied 1 to H-quarter ahead forecasts $\{F_t y_{t+h}\}_{h=1}^{H}$ and model-implied yields $\left\{ y_t^{(h)} \right\}_{h=1}^{H}$ for every quarter t from 1961Q3 to 2019Q4.

D   Search over Initial Beliefs for Nominal Short Rate

Let $\boldsymbol{\theta} = (\alpha_\rho, \beta_\rho, \alpha_\gamma, \beta_\gamma)'$. Let $\boldsymbol{\alpha} = \{\alpha_h\}_{h=1}^{H}$ and $\boldsymbol{\beta} = \{\beta_h\}_{h=1}^{H}$ denote vectors of estimated coefficients from the forecasting anomaly regressions for different horizons up through a maximum horizon of H using the SPF and yield curve data. Let $\hat{\boldsymbol{\alpha}} = \{\hat{\alpha}_h\}_{h=1}^{H}$ and $\hat{\boldsymbol{\beta}} = \left\{ \hat{\beta}_h \right\}_{h=1}^{H}$ denote those same quantities estimated on the model implied forecasts and yields for a particular value of $\boldsymbol{\theta}$. Define the moment function as

$$\hat{m}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\alpha}_{bias} - \hat{\boldsymbol{\alpha}}_{bias} \\ \boldsymbol{\beta}_{ar} - \hat{\boldsymbol{\beta}}_{ar} \\ \boldsymbol{\beta}_{mz} - \hat{\boldsymbol{\beta}}_{mz} \\ \boldsymbol{\beta}_{cg} - \hat{\boldsymbol{\beta}}_{cg} \\ \boldsymbol{\beta}_{sr} - \hat{\boldsymbol{\beta}}_{sr} \\ \boldsymbol{\beta}_{lr} - \hat{\boldsymbol{\beta}}_{lr} \end{bmatrix} \tag{17}$$

The parameters are then estimated via the simulated method of moments (SMM) with an iden-

6

tity weighting matrix

$$\hat{\boldsymbol{\theta}} = \text{argmax}_{\boldsymbol{\theta}} \|\hat{m}(\boldsymbol{\theta})\|_2 = \text{argmax}_{\boldsymbol{\theta}} \hat{m}(\boldsymbol{\theta})' \hat{m}(\boldsymbol{\theta})$$

Every evaluation of the moment function $\hat{m}(\boldsymbol{\theta})$ requires us to sample from the posterior of the UC model sequentially. Since this step is very computationally costly, we only re-estimate the model every 4 quarters rather than every quarter, and use a burn-in sample of 50,000 draws and keep the subsequent 25,000 draws rather than 75,000 for each of those quantities in our empirical specification. The global minimum is found using MATLAB's "particleswarm" optimization routine, subject to the constraint that the mean of ρ is larger than 0.5. As described in appendix D, the forecaster is assumed to discard posterior draws of ρ greater than 1 when computing forecasts.

E   Matching Forecasts and Yields Directly

Here, we report results for the T-Bill application where instead of targeting the regression statistics reported in section 3, we directly target the time series of consensus T-Bill forecasts from the SPF at horizons 1 to 4, and also the 5 and 10-year zero coupon nominal yields from the Liu and Wu (2020) data.
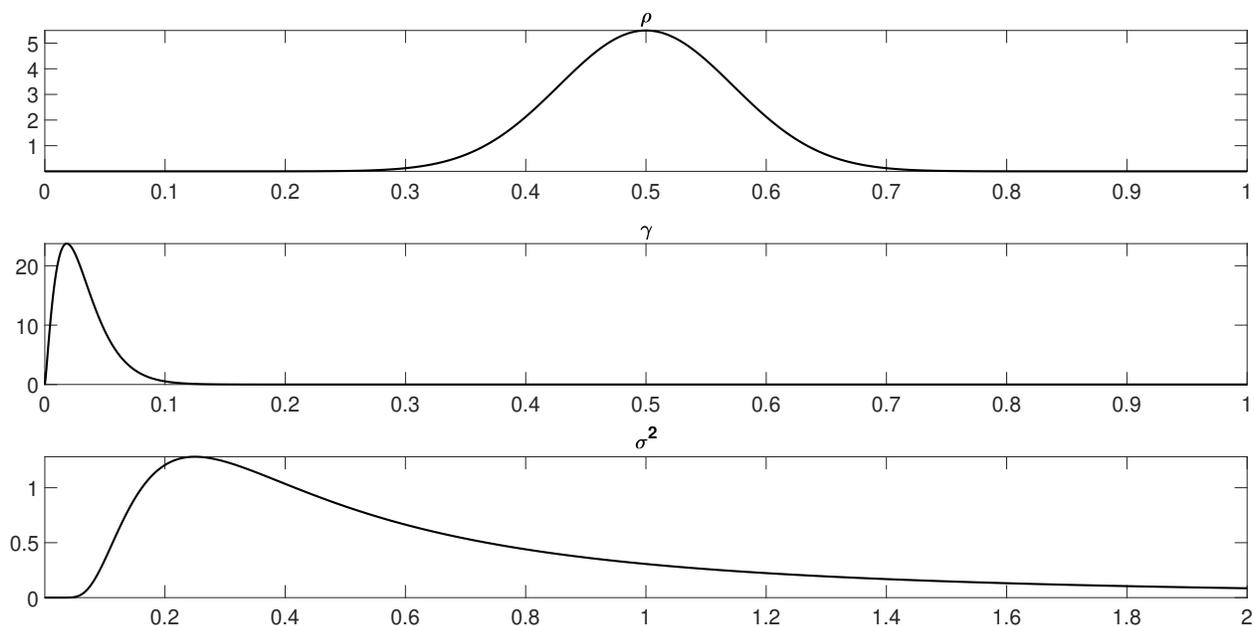
Figure E.1: Marginal Initial Beliefs Distributions: T-bill Rate Model

Note: Each panel plots the initial beliefs held in 1951Q2 by agents in our T-bill rate model for each of the three model parameters: ρ, γ, and σ^2 respectively.
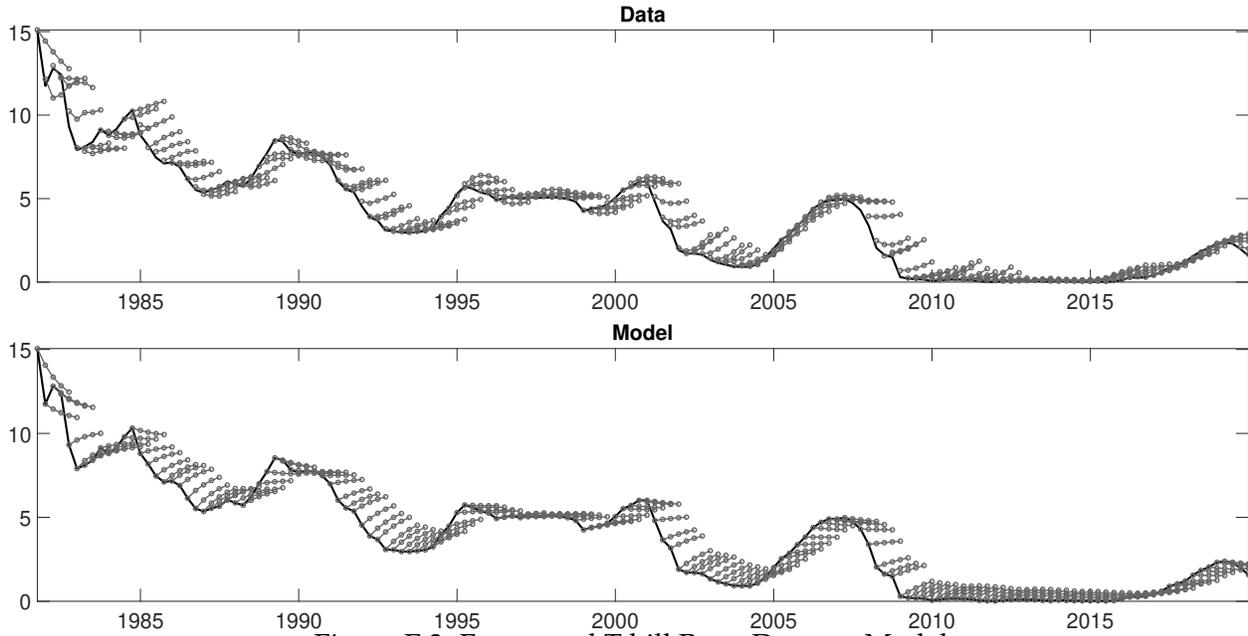
Data

Model

Figure E.2: Forecasted T-bill Rate: Data vs. Model

Note: The black solid line is the 3-month T-bill rate. Each short gray line with five circles represents forecasts made in a particular quarter about the then present quarter (first circle) and following four quarters (subsequent four circles). In the top panel, these forecasts are SPF forecasts. In the bottom panel, these forecasts are mean forecasts generated from the UC model estimated in real-time.

Model (direct targeting)

Model (baseline)
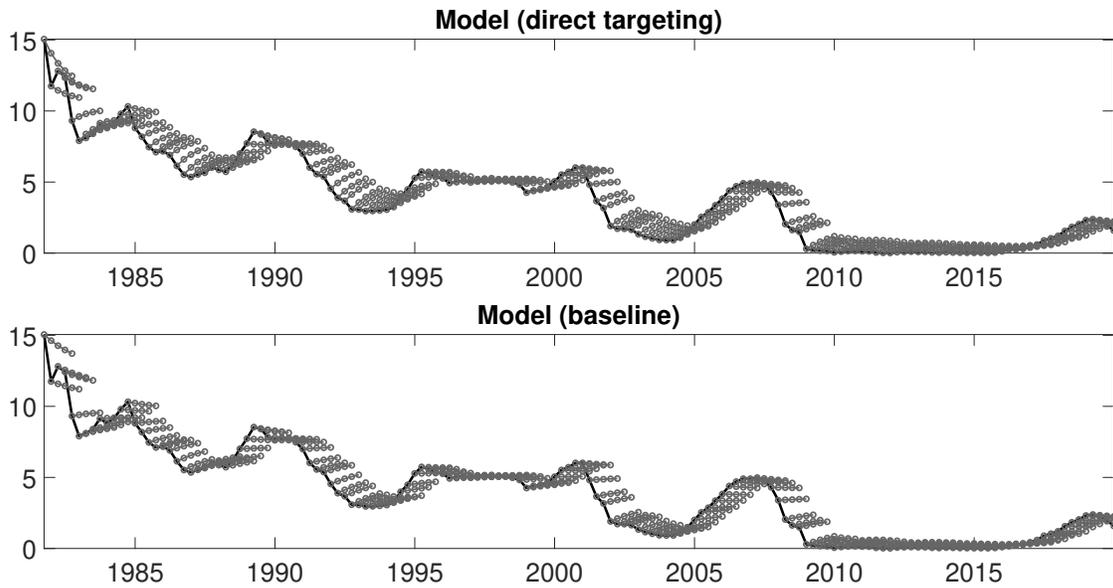
Figure E.3: Forecasted T-bill Rate: Direct targeting vs. baseline

Note: The black solid line is the 3-month T-bill rate. Each short gray line with five circles represents forecasts made in a particular quarter about the then present quarter (first circle) and following four quarters (subsequent four circles). All forecasts are mean forecasts generated from the UC model estimated in real-time. In the top panel, these forecasts come from the priors estimated by directly targeting the SPF forecasts and the 5 and 10-year nominal zero-coupon yields. The bottom panel plots the results for our baseline estimation.

Table E.1: T-Bill Rate Forecast Anomalies: Model vs. Data

	Forecast Horizon			
	1	2	3	4
Panel A: Bias				
SPF	-0.18***	-0.34***	-0.52***	-0.70***
	(0.06)	(0.11)	(0.16)	(0.20)
UC Model	-0.22***	-0.40***	-0.57***	-0.72***
	(0.06)	(0.12)	(0.17)	(0.21)
Panel B: Autocorrelation				
SPF	0.30*	0.27**	0.24**	0.13
	(0.15)	(0.11)	(0.11)	(0.12)
UC Model	0.44**	0.44**	0.41***	0.31**
	(0.16)	(0.15)	(0.11)	(0.10)
Panel C: Mincer-Zarnowitz				
SPF	0.97	0.94*	0.90**	0.86**
	(0.02)	(0.04)	(0.05)	(0.06)
UC Model	0.97*	0.94*	0.90**	0.86**
	(0.02)	(0.03)	(0.04)	(0.05)
Panel D: Coibion-Gorodnichenko				
SPF	0.23*	0.34**	0.62***	–
	(0.13)	(0.15)	(0.18)	
UC Model	0.53**	0.83*	1.28**	–
	(0.18)	(0.42)	(0.51)	

Note: The forecast horizons are quarters. Standard errors are reported in parentheses. Stars represent significance relative to the following hypotheses: $\alpha = 0$ for bias, $\beta = 0$ for autocorrelation, $\beta = 1$ for Mincer-Zarnowitz, $\beta = 0$ for Coibion-Gorodnichenko. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values, as proposed in Lazarus et al. (2018). This corresponds to a bandwidth of 17. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.2: Failures of the Expectations Hypothesis: Model vs. Data

	Long Horizon n						
	2	3	4	8	12	20	40
Panel A: Future Short Rates							
Data	-0.01***	0.11***	0.18***	0.39**	0.57	0.74	0.71
	(0.22)	(0.22)	(0.21)	(0.25)	(0.27)	(0.23)	(0.22)
UC Model	-0.01***	0.09***	0.12***	0.33**	0.51**	0.63*	0.71
	(0.14)	(0.15)	(0.16)	(0.21)	(0.22)	(0.19)	(0.23)
Panel B: Change in Long Rate							
Data	-1.02***	-0.91***	-1.03***	-1.29***	-1.61***	-2.04***	-2.75***
	(0.43)	(0.54)	(0.57)	(0.61)	(0.65)	(0.67)	(0.86)
UC Model	-1.02***	-1.01***	-1.02***	-1.10***	-1.25***	-1.61***	-2.52***
	(0.29)	(0.29)	(0.29)	(0.30)	(0.32)	(0.36)	(0.61)

Note: The sample period is from 1961Q3 to 2019Q4. The top panel reports estimates of β from regression (4). The bottom panel reports estimates of β from regression (5). In both cases, the horizon n is listed at the top of the table. Standard errors are reported in parentheses. Stars represent significance relative to the hypothesis that $\beta = 1$. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values, as proposed in Lazarus et al. (2018). This corresponds to a bandwidth of 19. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure E.4: Parameter Estimates: T-bill Rate Model

Note: Each panel plots the evolution of beliefs about one of the three UC model parameters: ρ, γ, and σ. The black solid line is the mean and the dotted black lines are the 5th and 95th percentiles of the posterior distribution for the parameter in question. Recall that we only update beliefs about these parameters every fourth quarter.
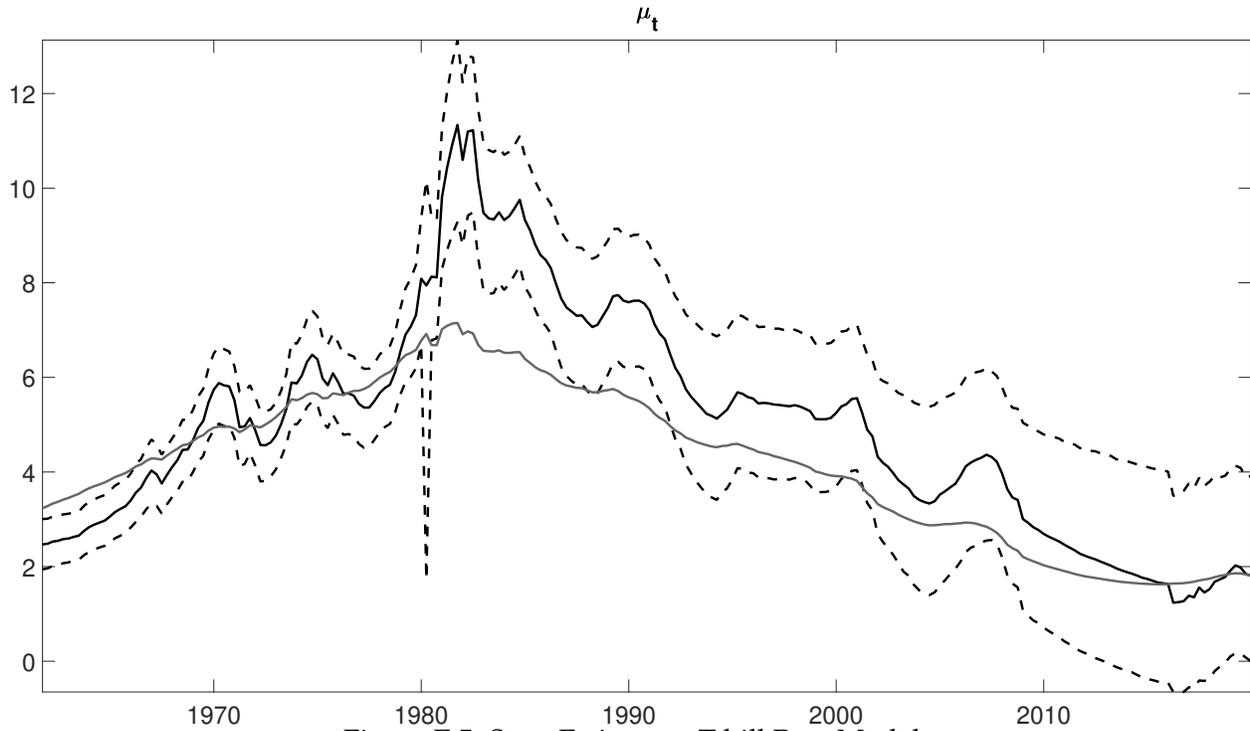
Figure E.5: State Estimates: T-bill Rate Model

Note: This figure plots the evolution of beliefs about the permanent component μ_t. The black solid line is the posterior mean of the real-time filtering distributions, the dotted black lines are the 5th and 95th percentiles of the posterior real-time filtering distributions, and the solid gray line is the posterior mean of the ex-post smoothing distributions.
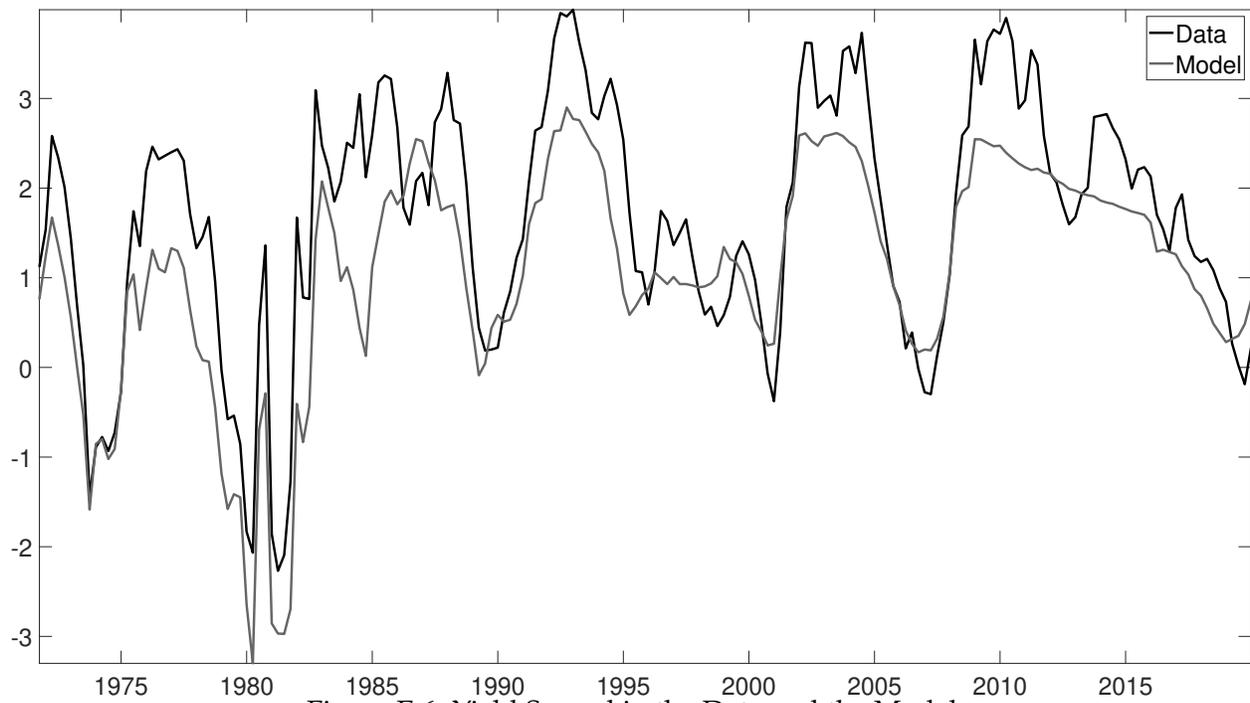
Figure E.6: Yield Spread in the Data and the Model

Note: The figure plots the spread between the yield on a 10-year zero coupon bond and the 3-month Treasury bill rate for the data and the model.

F   Alternative Initial Beliefs

F.1   Very Loose Initial Beliefs

Here we consider initial beliefs $\rho \sim N(0.5, 0.1)$ and $\gamma \sim \mathcal{B}(1.01, 1.01)$, which is close to uniform but not quite, since we need a little mass away from the boundaries for the model to be identified.
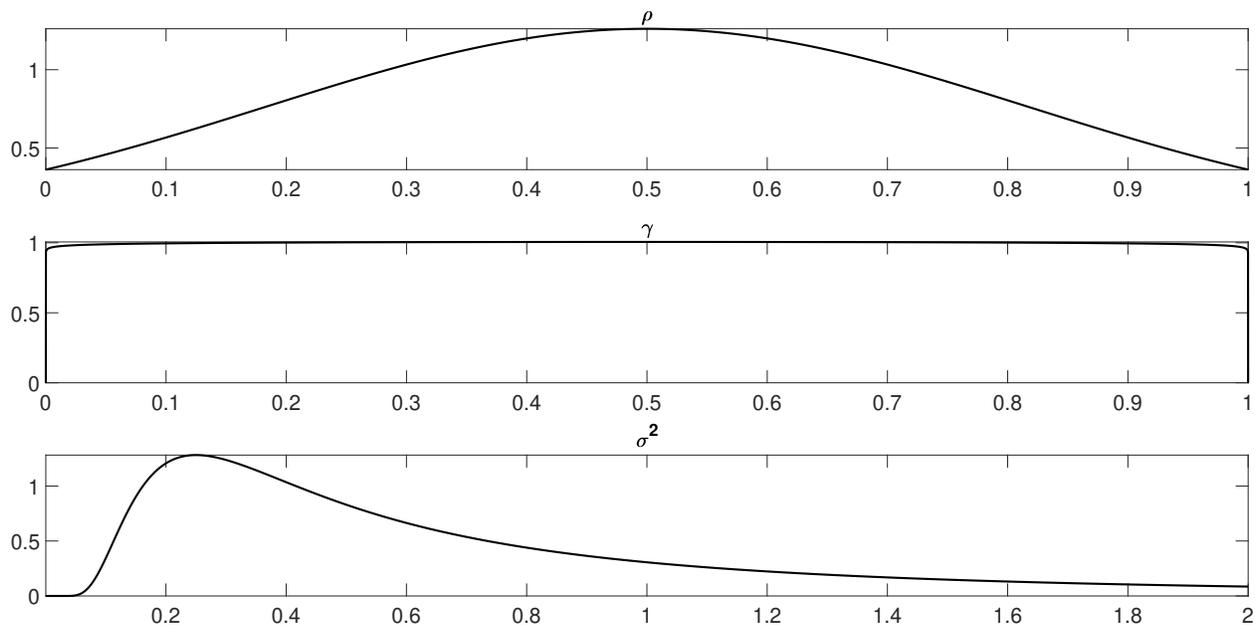
Figure F.1: Marginal Initial Beliefs Distributions: Loose Initial Beliefs Model

Note: Each panel plots the initial beliefs held in 1951Q2 by agents in our T-bill rate model for each of the three model parameters: ρ, γ, and σ^2 respectively.
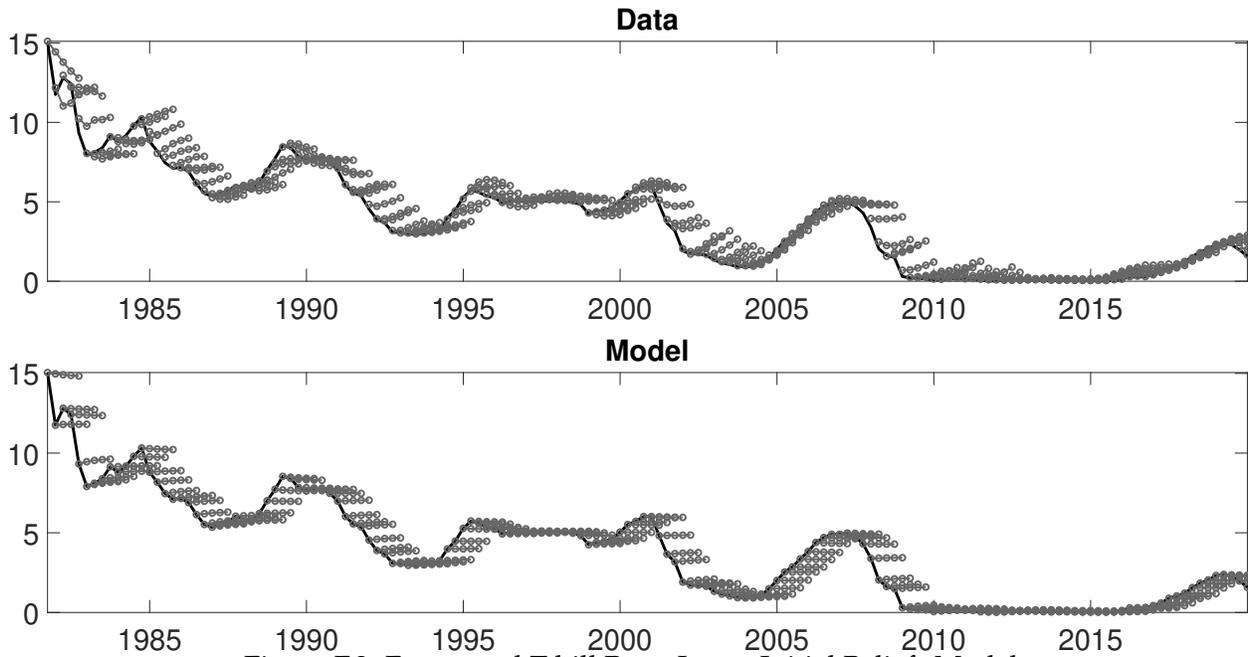
Figure F.2: Forecasted T-bill Rate: Loose Initial Beliefs Model

Note: The black solid line is the 3-month T-bill rate. Each short gray line with five circles represents forecasts made in a particular quarter about the then present quarter (first circle) and following four quarters (subsequent four circles). In the top panel, these forecasts are SPF forecasts. In the bottom panel, these forecasts are mean forecasts generated from the UC model estimated in real-time.
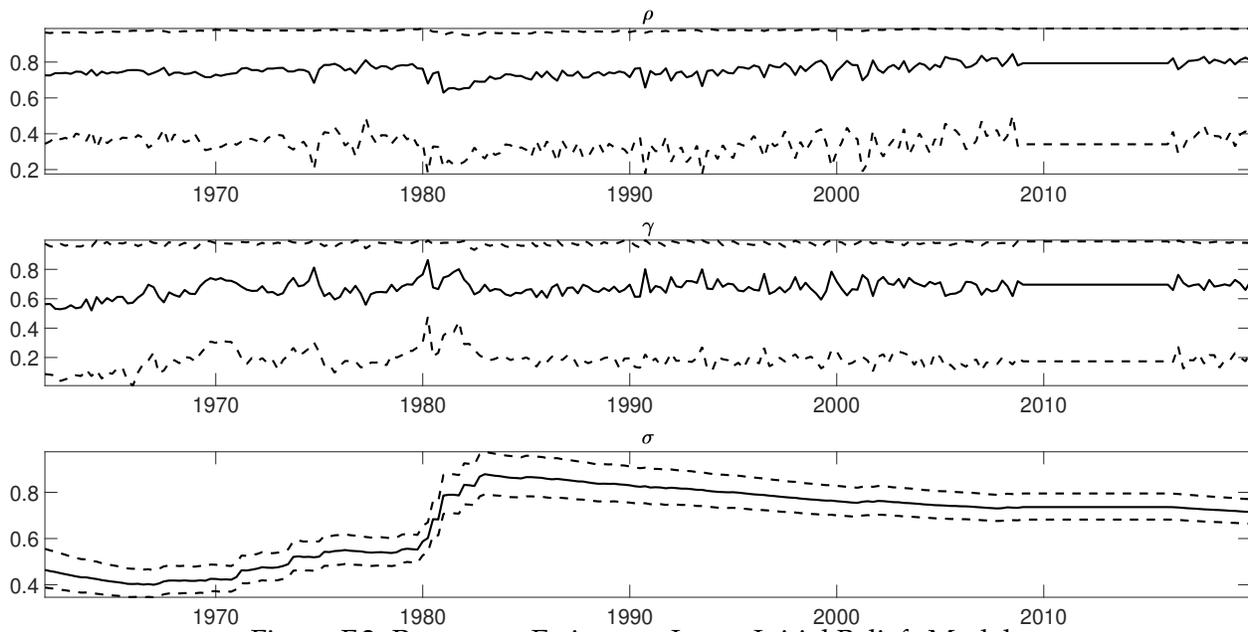
Figure F.3: Parameter Estimates: Loose Initial Beliefs Model

Note: Each panel plots the evolution of beliefs about one of the three UC model parameters: ρ, γ, and σ. The black solid line is the mean and the dotted black lines are the 5th and 95th percentiles of the posterior distribution for the parameter in question. Recall that we only update beliefs about these parameters every fourth quarter.

13

Table F.1: T-Bill Rate Forecast Anomalies: Loose Initial Beliefs Model

	Forecast Horizon			
	1	2	3	4
Panel A: Bias				
SPF	-0.18***	-0.34***	-0.52***	-0.70***
	(0.06)	(0.11)	(0.16)	(0.20)
UC Model	-0.10	-0.18	-0.27	-0.35
	(0.06)	(0.11)	(0.17)	(0.22)
Panel B: Autocorrelation				
SPF	0.30*	0.27**	0.24**	0.13
	(0.15)	(0.11)	(0.11)	(0.12)
UC Model	0.35*	0.39**	0.32**	0.18
	(0.15)	(0.12)	(0.11)	(0.12)
Panel C: Mincer-Zarnowitz				
SPF	0.97	0.94*	0.90**	0.86**
	(0.02)	(0.04)	(0.05)	(0.06)
UC Model	0.95**	0.91**	0.85***	0.79***
	(0.02)	(0.03)	(0.04)	(0.05)
Panel D: Coibion-Gorodnichenko				
SPF	0.23*	0.34**	0.62***	–
	(0.13)	(0.15)	(0.18)	
UC Model	0.36**	0.51	0.80**	–
	(0.16)	(0.30)	(0.32)	

Note: The forecast horizons are quarters. Standard errors are reported in parentheses. Stars represent significance relative to the following hypotheses: $\alpha = 0$ for bias, $\beta = 0$ for autocorrelation, $\beta = 1$ for Mincer-Zarnowitz, $\beta = 0$ for Coibion-Gorodnichenko. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values, as proposed in Lazarus et al. (2018). This corresponds to a bandwidth of 17. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table F.2: Failures of the Expectations Hypothesis: Loose Initial Beliefs Model

	Long Horizon n						
	2	3	4	8	12	20	40
Panel A: Future Short Rates							
Data	-0.01***	0.11***	0.18***	0.39**	0.57	0.74	0.71
	(0.22)	(0.22)	(0.21)	(0.25)	(0.27)	(0.23)	(0.22)
UC Model	-1.97***	-1.18**	-0.80*	1.03	2.35	3.19	3.53
	(0.84)	(0.90)	(0.94)	(1.25)	(1.39)	(1.24)	(1.41)
Panel B: Change in Long Rate							
Data	-1.02***	-0.91***	-1.03***	-1.29***	-1.61***	-2.04***	-2.75***
	(0.43)	(0.54)	(0.57)	(0.61)	(0.65)	(0.67)	(0.86)
UC Model	-4.95***	-5.45***	-5.88***	-7.35***	-8.55***	-10.85***	-16.71*
	(1.68)	(1.78)	(1.89)	(2.35)	(2.81)	(3.71)	(8.02)

Note: The sample period is from 1961Q3 to 2019Q4. The top panel reports estimates of β from regression (4). The bottom panel reports estimates of β from regression (5). In both cases, the horizon n is listed at the top of the table. Standard errors are reported in parentheses. Stars represent significance relative to the hypothesis that $\beta = 1$. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values, as proposed in Lazarus et al. (2018). This corresponds to a bandwidth of 19. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure F.4: State Estimates: Loose Initial Beliefs Model

Note: This figure plots the evolution of beliefs about the permanent component μ_t. The black solid line is the posterior mean of the real-time filtering distributions, the dotted black lines are the 5th and 95th percentiles of the posterior real-time filtering distributions, and the solid gray line is the posterior mean of the ex-post smoothing distributions.
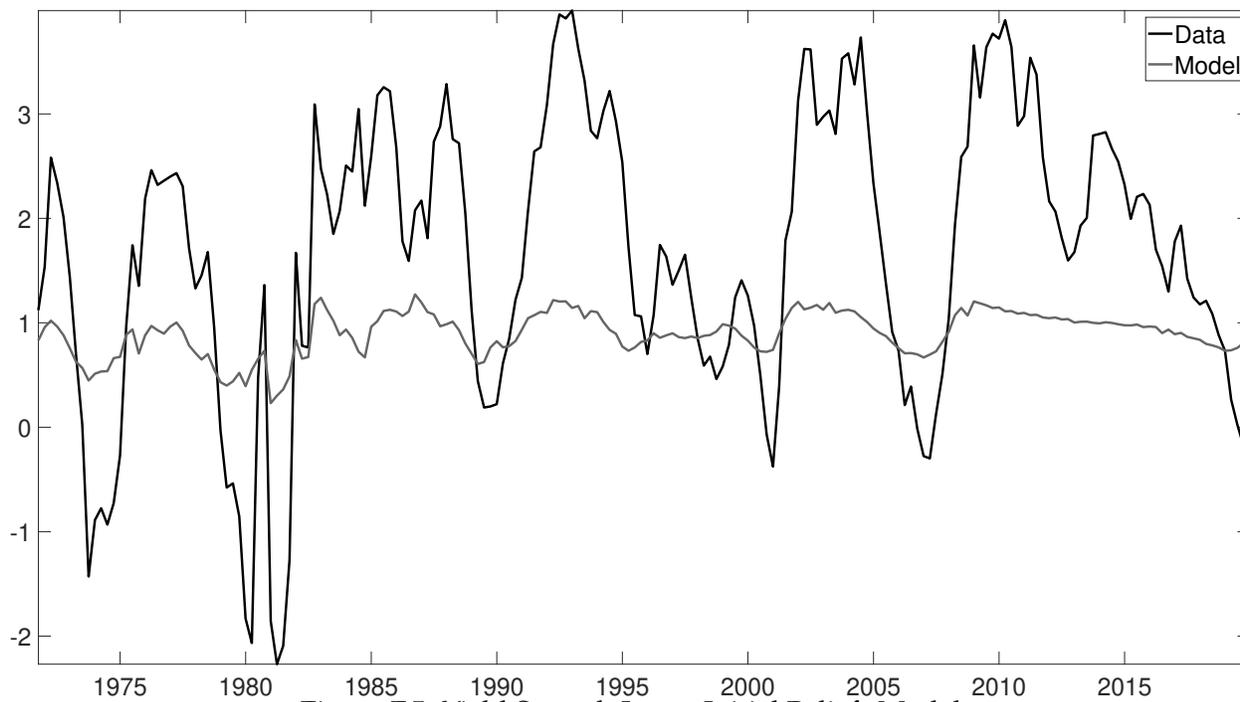
Figure F.5: Yield Spread: Loose Initial Beliefs Model

Note: The figure plots the spread between the yield on a 10-year zero coupon bond and the 3-month Treasury bill rate for the data and the model.

F.2   Look-Ahead Initial Beliefs

Next we consider initial beliefs $\rho \sim N(0.910, 0.00184)$ and $\gamma \sim \mathcal{B}(3.13, 11.04)$ which corresponds to a mode of 0.175 and a standard deviation of 0.107. These approximate the beliefs of agents in our model at the end of our sample in our baseline case.
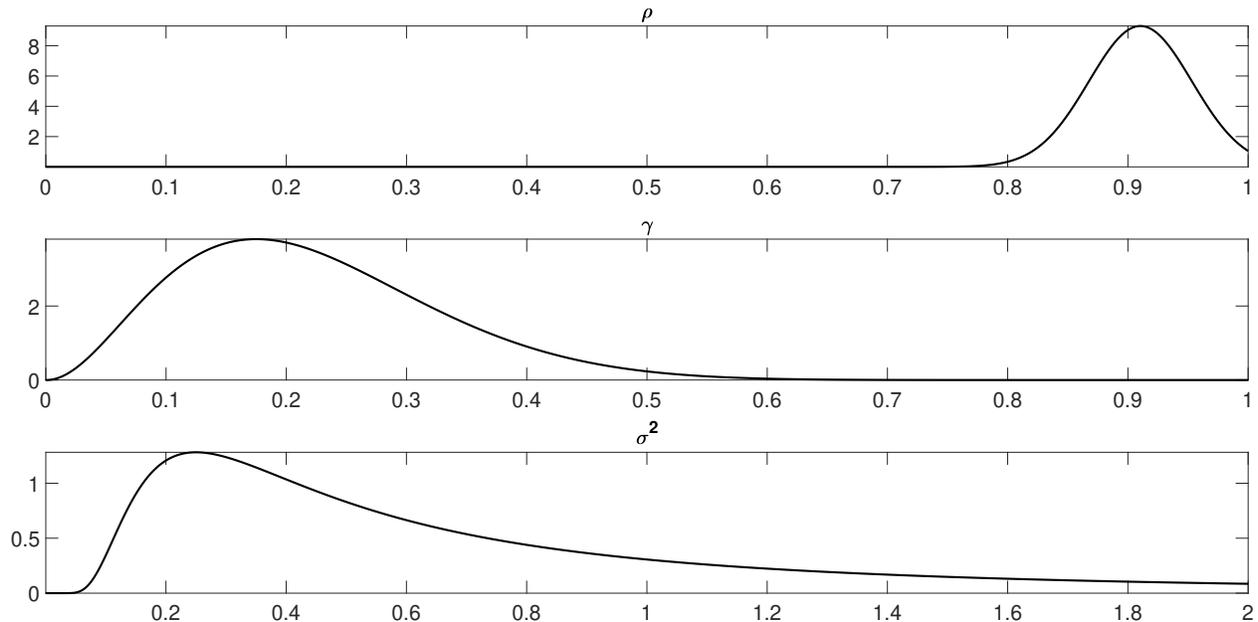
Figure F.6: Marginal Initial Beliefs Distributions: Look-Ahead Model

Note: Each panel plots the initial beliefs held in 1951Q2 by agents in our T-bill rate model for each of the three model parameters: ρ, γ, and σ^2 respectively.
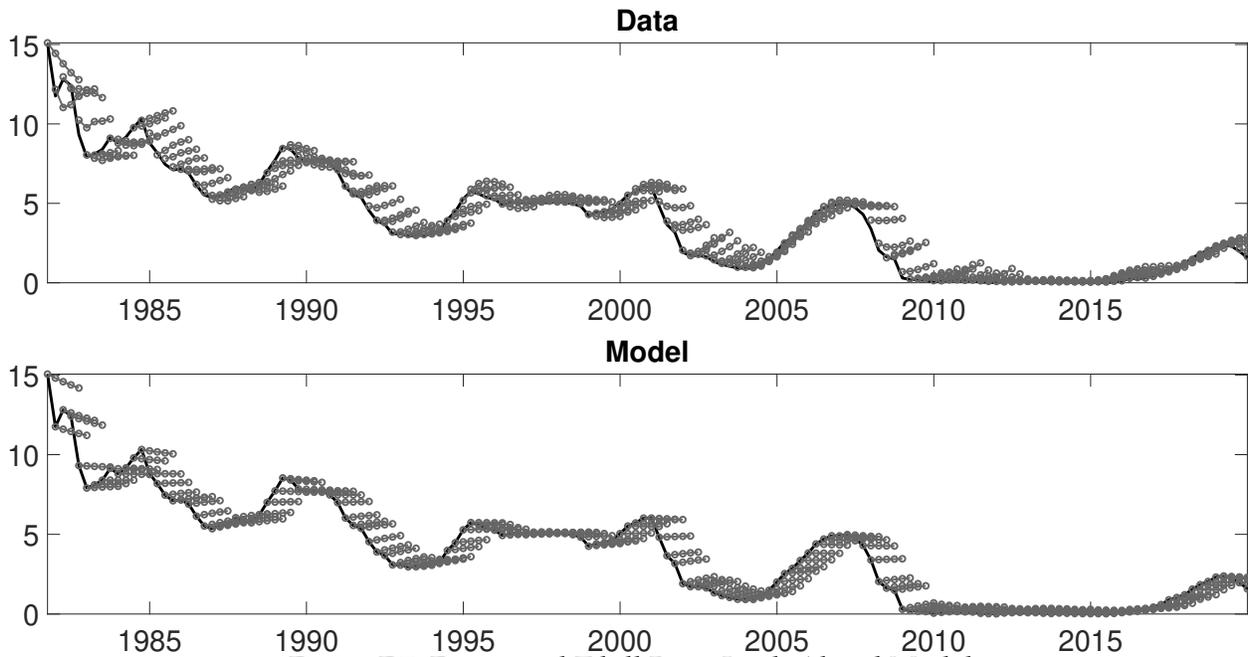
Figure F.7: Forecasted T-bill Rate: Look-Ahead Model

Note: The black solid line is the 3-month T-bill rate. Each short gray line with five circles represents forecasts made in a particular quarter about the then present quarter (first circle) and following four quarters (subsequent four circles). In the top panel, these forecasts are SPF forecasts. In the bottom panel, these forecasts are mean forecasts generated from the UC model estimated in real-time.
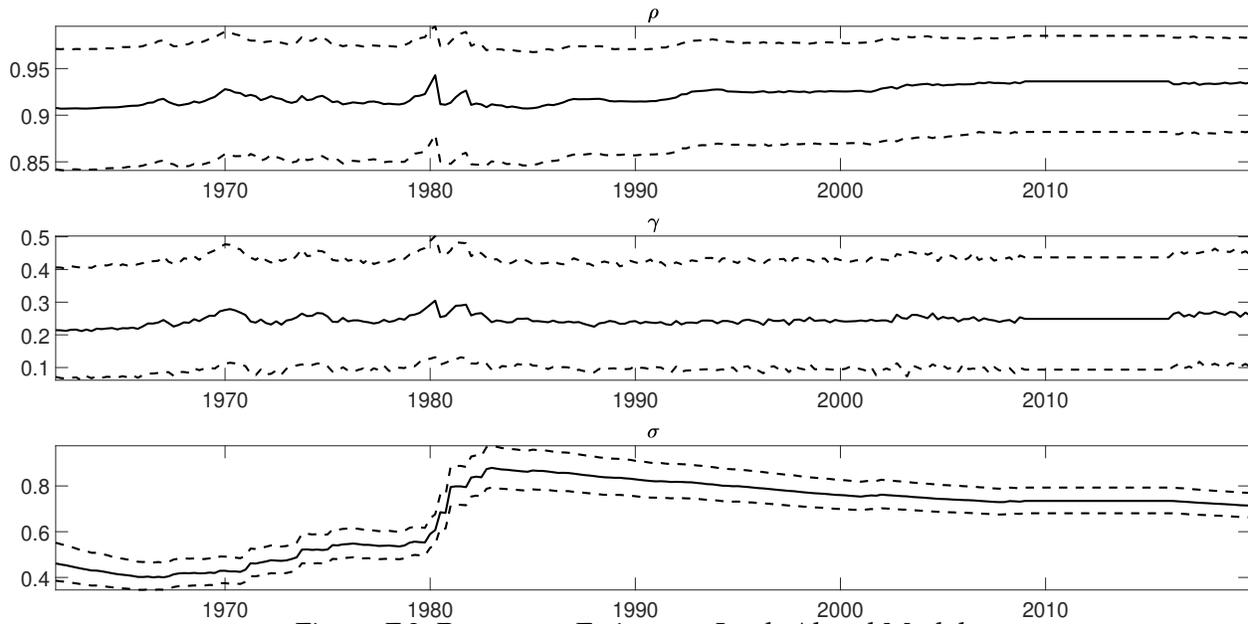
Figure F.8: Parameter Estimates: Look-Ahead Model

Note: Each panel plots the evolution of beliefs about one of the three UC model parameters: ρ, γ, and σ. The black solid line is the mean and the dotted black lines are the 5th and 95th percentiles of the posterior distribution for the parameter in question. Recall that we only update beliefs about these parameters every fourth quarter.

18

Table F.3: T-Bill Rate Forecast Anomalies: Look-Ahead Model

	Forecast Horizon			
	1	2	3	4
Panel A: Bias				
SPF	-0.18***	-0.34***	-0.52***	-0.70***
	(0.06)	(0.11)	(0.16)	(0.20)
UC Model	-0.13**	-0.23*	-0.33*	-0.43*
	(0.06)	(0.10)	(0.15)	(0.20)
Panel B: Autocorrelation				
SPF	0.30*	0.27**	0.24**	0.13
	(0.15)	(0.11)	(0.11)	(0.12)
UC Model	0.33*	0.36**	0.31**	0.18
	(0.17)	(0.14)	(0.12)	(0.13)
Panel C: Mincer-Zarnowitz				
SPF	0.97	0.94*	0.90**	0.86**
	(0.02)	(0.04)	(0.05)	(0.06)
UC Model	0.96*	0.93**	0.88**	0.83***
	(0.02)	(0.03)	(0.04)	(0.05)
Panel D: Coibion-Gorodnichenko				
SPF	0.23*	0.34**	0.62***	–
	(0.13)	(0.15)	(0.18)	
UC Model	0.34*	0.45	0.74*	–
	(0.18)	(0.34)	(0.38)	

Note: The forecast horizons are quarters. Standard errors are reported in parentheses. Stars represent significance relative to the following hypotheses: $\alpha = 0$ for bias, $\beta = 0$ for autocorrelation, $\beta = 1$ for Mincer-Zarnowitz, $\beta = 0$ for Coibion-Gorodnichenko. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values, as proposed in Lazarus et al. (2018). This corresponds to a bandwidth of 17. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table F.4: Failures of the Expectations Hypothesis: Look-Ahead Model

	Long Horizon n						
	2	3	4	8	12	20	40
Panel A: Future Short Rates							
Data	-0.01***	0.11***	0.18***	0.39**	0.57	0.74	0.71
	(0.22)	(0.22)	(0.21)	(0.25)	(0.27)	(0.23)	(0.22)
UC Model	0.06	0.33	0.47	0.98	1.24	1.25	1.13
	(0.62)	(0.63)	(0.63)	(0.63)	(0.56)	(0.45)	(0.48)
Panel B: Change in Long Rate							
Data	-1.02***	-0.91***	-1.03***	-1.29***	-1.61***	-2.04***	-2.75***
	(0.43)	(0.54)	(0.57)	(0.61)	(0.65)	(0.67)	(0.86)
UC Model	-0.89	-0.90	-0.91	-0.93	-0.97	-1.06	-1.09
	(1.24)	(1.25)	(1.26)	(1.33)	(1.41)	(1.56)	(2.38)

Note: The sample period is from 1961Q3 to 2019Q4. The top panel reports estimates of β from regression (4). The bottom panel reports estimates of β from regression (5). In both cases, the horizon n is listed at the top of the table. Standard errors are reported in parentheses. Stars represent significance relative to the hypothesis that $\beta = 1$. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values, as proposed in Lazarus et al. (2018). This corresponds to a bandwidth of 19. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure F.9: State Estimates: Look-Ahead Model

Note: This figure plots the evolution of beliefs about the permanent component μ_t. The black solid line is the posterior mean of the real-time filtering distributions, the dotted black lines are the 5th and 95th percentiles of the posterior real-time filtering distributions, and the solid gray line is the posterior mean of the ex-post smoothing distributions.
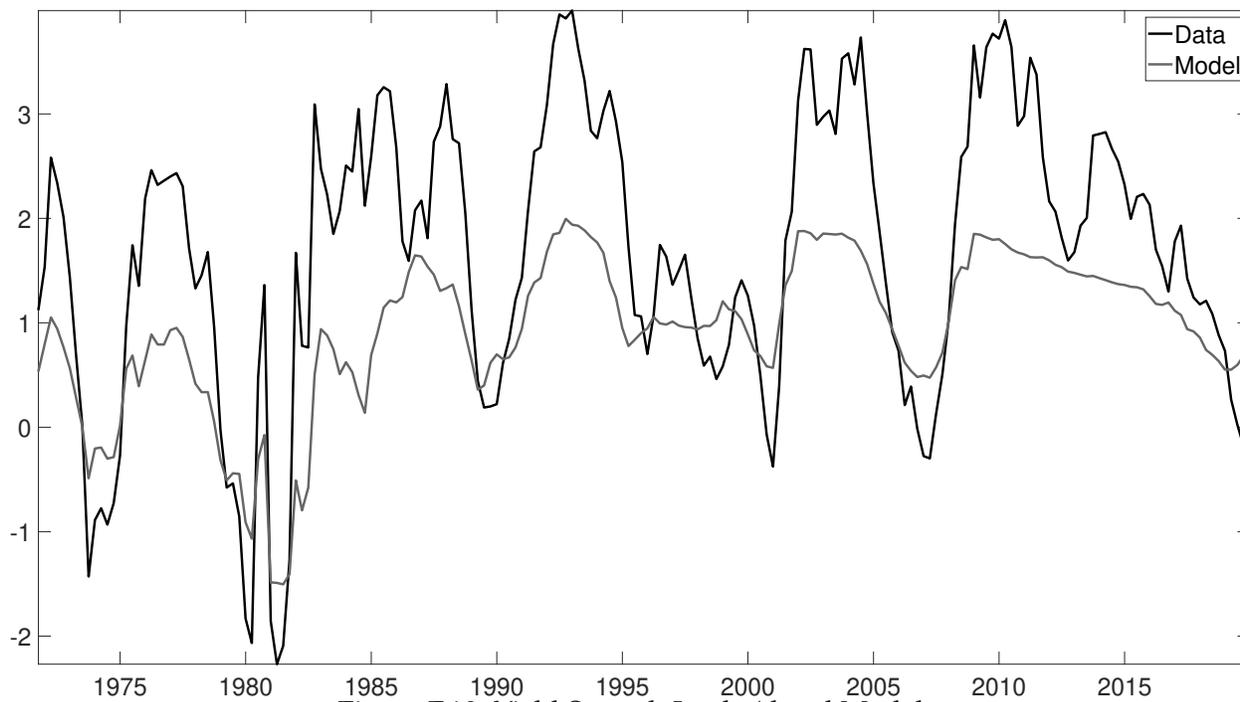
Figure F.10: Yield Spread: Look-Ahead Model

Note: The figure plots the spread between the yield on a 10-year zero coupon bond and the 3-month Treasury bill rate for the data and the model.

F.3 RMSE with Dispersed Initial Beliefs

Table F.5 reports results on the root-mean-squared error (RMSE) of the forecasts from our model relative to the RMSE of SFP forecasters. We do this for our baseline initial beliefs – $\rho \sim N(0.6, 0.12^2)$ and $\gamma \sim \mathcal{B}(2.3, 19.7)$ – and for a case with more highly dispersed initial beliefs – $\rho \sim N(0.6, 0.31^2)$ and $\gamma \sim \mathcal{B}(1.13, 2.9)$. The first four columns of Table F.5 report the ratio of the RMSE from our model relative to the RMSE of SFP forecasters for forecast horizons of one through four quarters. The last column reports the average ratio across the four horizons.

Overall, we conclude that a model with more dispersed initial beliefs generates forecasts of very similar quality as measured by RMSE. Using more highly dispersed initial beliefs leads to very slightly better forecasts from a RMSE perspective at short horizons, but slightly worse forecasts at longer horizons. Averaging across horizons, the difference is close to zero. Notice also that the forecasts from our model are slightly 'better' than the SPF forecasts for both initial beliefs: the ratios reported in the table are all smaller than 1.

Table F.5: Root-Mean-Squared Errors: Ratio of Model to SPF

	Forecast Horizon				
	1	2	3	4	Mean
Baseline	0.96	0.98	0.97	0.96	0.97
Highly Dispersed	0.94	0.97	0.96	0.97	0.97

Note: The table reports the ratio of the root-mean-squared error (RMSE) of forecasts from our model for interest rates with different initial beliefs to the RMSE of SPF forecasts. The baseline initial beliefs for ρ are $N(0.6, 0.12^2)$. The initial beliefs in the 'highly dispersed' case are $N(0.6, 0.31^2)$. The baseline initial beliefs for γ are $\mathcal{B}(2.3, 19.7)$. The initial beliefs in the 'highly dispersed' case are $\mathcal{B}(1.13, 2.9)$. The 'Mean' column reports the average ratio across the four horizons.

G Interest Rate Results with a Break in 1990

Here we present results for a case where we allow for a break in beliefs about γ in 1990. We redo our baseline short-rate analysis exactly as before except that we allow the agents in the model to "reset" their beliefs about γ in 1990. We assume that the new belief distribution of agents about γ in 1990 is $\gamma \sim \mathcal{B}(\alpha_{\gamma,2}, \beta_{\gamma,2})$ and we search over the values of $\alpha_{\gamma,2}$ and $\beta_{\gamma,2}$ as well as the hyperparameters in the baseline case to best match the forecast anomalies.
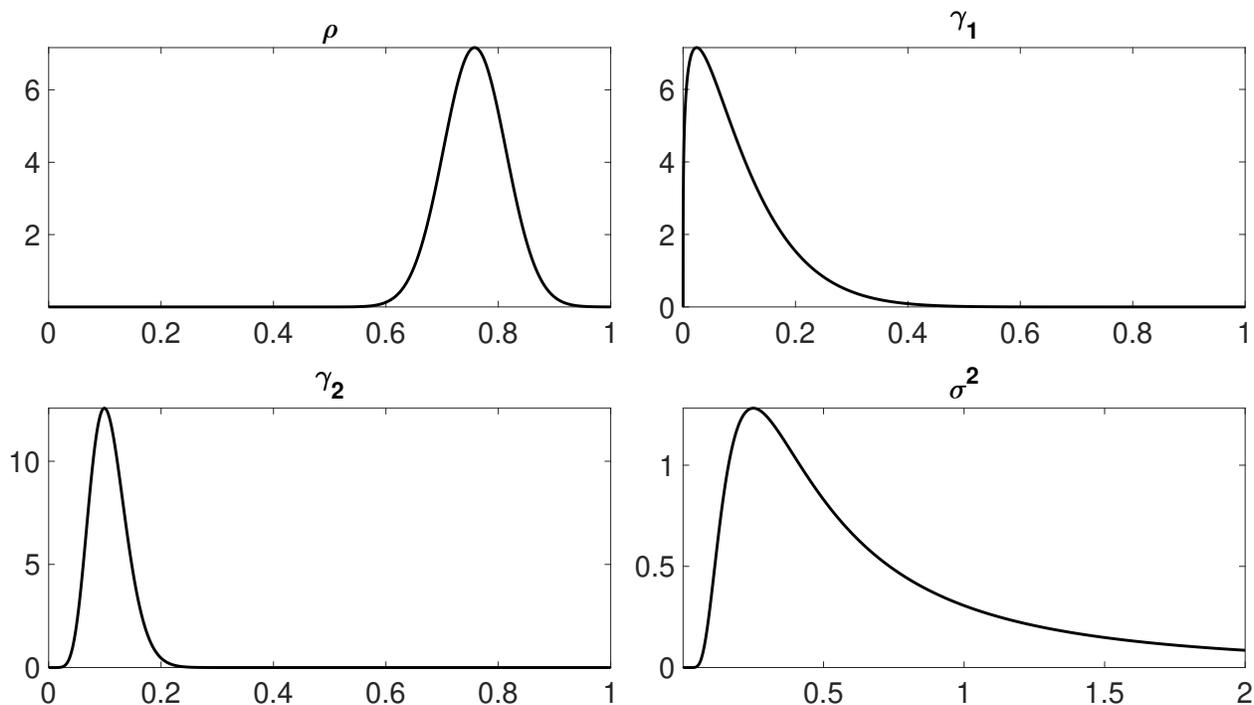
Figure G.1: Marginal Prior Distributions: 1990 Break Model

Note: These four panels plot the initial beliefs we estimate for the four parameters of the model. The panels labelled ρ, γ_1, and σ give the initial beliefs for ρ, γ, and σ, respectively, in 1951Q2. The panel labelled γ_2 give the belief distribution for γ in 1990Q1.

Results analogous to those presented for our baseline model in the main body are presented in Figures G.1-G.4 and Tables G.1-G.2. We estimate a substantial decrease in the mean of the distribution of beliefs about γ (from 0.19 to 0.11) and a sharp downward shift in the standard deviation (from 0.09 to 0.03) in 1990, leading to lower posterior mean estimates in the latter part of the sample as one would expect. Other results are quite similar to in our baseline case. The extra parameters allow up to improve the fit of the model to the data modestly.

Table G.1: T-Bill Rate Forecast Anomalies: 1990 Break Model

| | Forecast Horizon | | | |
	1	2	3	4
Panel A: Bias				
SPF	-0.18***	-0.34***	-0.52***	-0.70***
	(0.05)	(0.09)	(0.14)	(0.19)
UC Model	-0.18**	-0.32**	-0.47**	-0.60**
	(0.06)	(0.11)	(0.16)	(0.21)
Panel B: Autocorrelation				
SPF	0.30*	0.27**	0.24*	0.13
	(0.14)	(0.12)	(0.12)	(0.13)
UC Model	0.38*	0.41**	0.37**	0.25*
	(0.17)	(0.14)	(0.11)	(0.12)
Panel C: Mincer-Zarnowitz				
SPF	0.97*	0.94**	0.90**	0.86**
	(0.02)	(0.02)	(0.04)	(0.05)
UC Model	0.96*	0.93**	0.89**	0.84**
	(0.02)	(0.03)	(0.04)	(0.05)
Panel D: Coibion-Gorodnichenko				
SPF	0.23*	0.34*	0.62***	–
	(0.12)	(0.16)	(0.16)	
UC Model	0.41*	0.59	0.94*	–
	(0.19)	(0.39)	(0.45)	

Note: The forecast horizons are quarters. Standard errors are reported in parentheses. Stars represent significance relative to the following hypotheses: $\alpha = 0$ for bias, $\beta = 0$ for autocorrelation, $\beta = 1$ for Mincer-Zarnowitz, $\beta = 0$ for Coibion-Gorodnichenko. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values. This corresponds to a bandwidth of 17. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table G.2: Failures of the Expectations Hypothesis: 1990 Break Model

	Long Horizon n						
	2	3	4	8	12	20	40
Panel A: Future Short Rates							
Data	-0.01***	0.11***	0.18***	0.39**	0.57	0.74	0.71
	(0.23)	(0.23)	(0.23)	(0.23)	(0.26)	(0.23)	(0.20)
UC Model	-0.07***	0.09***	0.15***	0.48	0.72	0.84	0.92
	(0.25)	(0.26)	(0.27)	(0.33)	(0.33)	(0.28)	(0.33)
Panel B: Change in Long Rate							
Data	-1.02***	-0.91***	-1.03***	-1.29***	-1.61***	-2.04***	-2.75***
	(0.45)	(0.59)	(0.62)	(0.59)	(0.57)	(0.55)	(0.87)
UC Model	-1.14***	-1.17***	-1.19***	-1.31***	-1.45***	-1.79***	-2.61**
	(0.50)	(0.51)	(0.52)	(0.55)	(0.59)	(0.68)	(1.24)

Note: The sample period is from 1961Q3 to 2019Q4. The top panel reports estimates of β from regression (4). The bottom panel reports estimates of β from regression (5). In both cases, the horizon n is listed at the top of the table. Standard errors are reported in parentheses. Stars represent significance relative to the hypothesis that $\beta = 1$. P-values are computed using Newey-West standard errors with lag length selected as $L = \lceil 1.3 \times T^{1/2} \rceil$ and fixed-b critical values. This corresponds to a bandwidth of 19. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure G.2: Forecasted T-bill Rate: 1990 Break Model

Note: The black solid line is the 3-month T-bill rate. Each short gray line with five circles represents forecasts made in a particular quarter about the then present quarter (first circle) and following four quarters (subsequent four circles). In the top panel, these forecasts are SPF forecasts. In the bottom panel, these forecasts are mean forecasts generated from the UC model estimated in real-time.
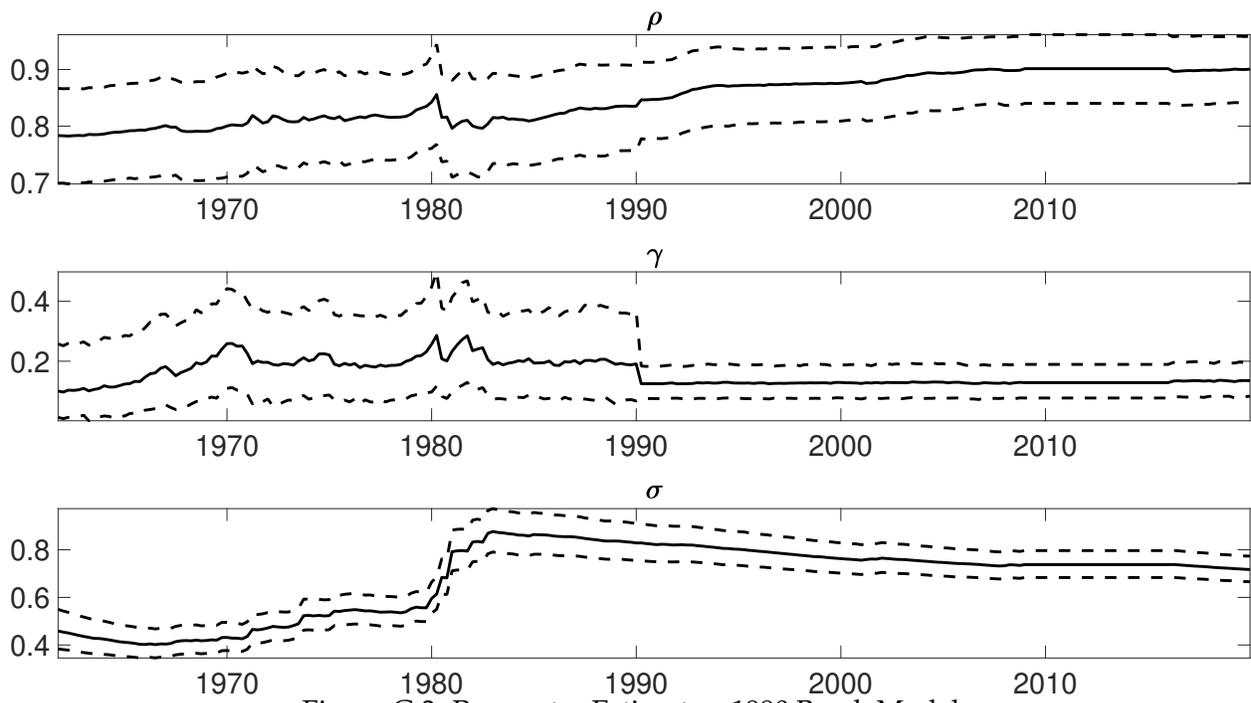
Figure G.3: Parameter Estimates: 1990 Break Model

Note: Each panel plots the evolution of beliefs about one of the three UC model parameters: ρ, γ, and σ. The black solid line is the mean and the dotted black lines are the 5th and 95th percentiles of the posterior distribution for the parameter in question. Recall that we only update beliefs about these parameters every fourth quarter.
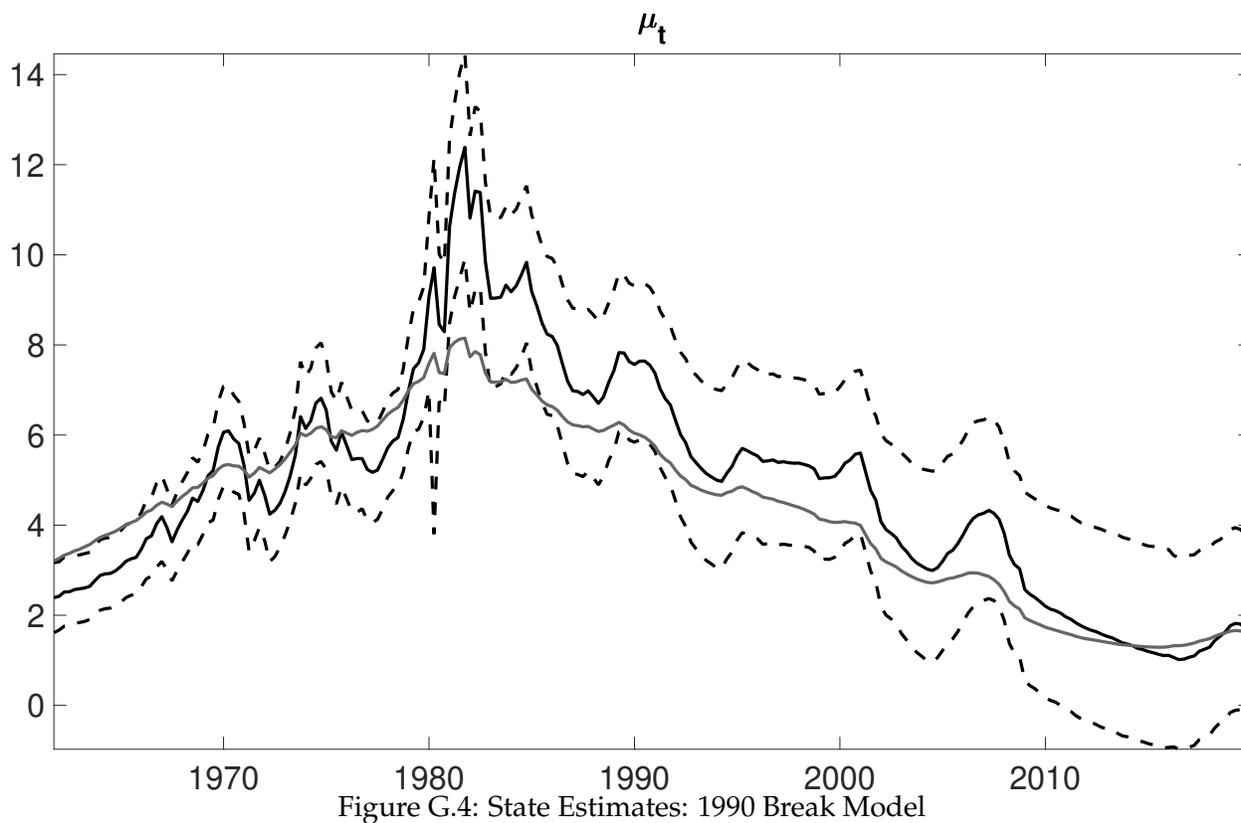
Figure G.4: State Estimates: 1990 Break Model

Note: Each panel corresponds to one of the two UC hidden state variables: μ_t and x_t respectively. The black solid line is the posterior mean of the real-time filtering distributions, the dotted black lines are the 5th and 95th percentiles of the posterior real-time filtering distributions, and the solid gray line is the posterior mean of the ex-post smoothing distributions for the corresponding parameter.

G.1 Cochrane-Piazzesi Regressions

Cochrane and Piazzesi (2005) show that a single factor predicts one-year excess returns on one- to five-year maturity bonds with an R^2 above 0.4. Here we show that our learning model can match this return predictability. Following Cochrane and Piazzesi (2005), we consider the period length to be measured in years in this section – i.e., n and t are measured in years in this section while it is measured in quarters elsewhere in the paper. Let $p_t^{(n)}$ denote the log price of an n-year zero coupon bond at time t. The relationship between the log yield and log price of an n-year zero coupon bond is $y_t^{(n)} = -p_t^{(n)}/n$. The forward rate at time t for a loan between time $t + n - 1$ and time $t + n$ is

$$f_t^{(n-1 \to n)} \equiv p_t^{(n)} - p_t^{(n-1)}.$$

We refer to this as the n-year forward rate. (It might alternatively be referred to as the n-year-ahead, one-year forward rate.) The log holding period return of buying an n-year bond at time t and selling it as an $n - 1$-year bond at time $t + 1$ is given by

$$r_{t+1}^{(n)} \equiv p_{t+1}^{(n-1)} - p_t^{(n)}.$$

Denote the log excess return on the bond as

$$rx_{t+1}^{(n)} \equiv r_{t+1}^{(n)} - y_t^{(1)}.$$

Cochrane and Piazzesi (2005) run two sets of return predictability regressions. First, they run a set of *unrestricted* regressions:

$$rx_{t+1}^{(n)} = \beta_0^{(n)} + \beta_1^{(n)} y_t^{(1)} + \beta_2^{(n)} f_t^{(1 \to 2)} + \ldots + \beta_5^{(n)} f_t^{(4 \to 5)} + \varepsilon_{t+1}^{(n)} \tag{18}$$

for $n = 2, \ldots, 5$. These regressions yield a similar pattern of coefficients across the four maturities. This motivates considering the notion that a single factor may forecast excess returns at all horizons as follows:

$$rx_{t+1}^{(n)} = b_n \left(\gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(1 \to 2)} + \ldots + \gamma_5 f_t^{(4 \to 5)} \right) + \varepsilon_{t+1}^{(n)}$$

for $n = 2, \ldots, 5$. Cochrane and Piazzesi normalize the loadings b_n so they have an average value

Table G.3: Restricted Regression Results for γ_n

			Data				
const.	$y^{(1)}$	$f^{(1\to2)}$	$f^{(2\to3)}$	$f^{(3\to4)}$	$f^{(4\to5)}$	$\overline{R}^2$
-0.19	-0.93	-1.25	1.41	1.45	-0.71	0.38
(0.39)	(0.45)	(0.85)	(0.82)	(1.00)	(0.73)	
			Model				
const.	$y^{(1)}$	$f^{(1\to2)}$	$f^{(2\to3)}$	$f^{(3\to4)}$	$f^{(4\to5)}$	$\overline{R}^2$
-2.67	-21.13	82.60	-112.64	53.76	-2.54	0.48
(1.82)	(7.31)	(34.02)	(51.08)	(37.28)	(17.71)	

of 1:

$$\frac{1}{4}\sum_{n=2}^{5} b_n = 1$$

and estimate the b_n and γ_n coefficients in two stages. First, they estimate the γ_n coefficients from

$$\frac{1}{4}\sum_{n=2}^{5} rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(1\to2)} + \ldots + \gamma_5 f_t^{(4\to5)} + \bar{\varepsilon}_{t+1}. \tag{19}$$

Then they estimate the b_n coefficients from

$$rx_{t+1}^{(n)} = b_n \left(\hat{\gamma}_0 + \hat{\gamma}_1 y_t^{(1)} + \hat{\gamma}_2 f_t^{(1\to2)} + \ldots + \hat{\gamma}_5 f_t^{(4\to5)} \right) + \varepsilon_{t+1}^{(n)}, \tag{20}$$

where $\hat{\gamma}_n$ are the fitted values from regression (19). We refer to these as the *restricted* regressions.

We perform this analysis on our quarterly zero-coupon bond data from Liu and Wu (2020) for the sample period 1961Q3-2019Q4. (Cochrane and Piazzesi (2005) use monthly data for the sample period 1964-2003.) We also perform this analysis on the bond yields implied by our model (allowing for a break in γ in 1990). Before running the regressions for the model-implied data, the sample mean of our model-implied yields is made the same as the sample mean of the yields in the real-world data over our sample period.

Tables G.3 and G.4 report the coefficients $[\gamma_1, \ldots \gamma_5]$ and $[b_2, \ldots b_5]$ from the restricted regressions (equations (19) and (20)) respectively, along with the adjusted R^2 for these regressions. We can match the high R^2 for one-year excess returns on 2- to 5-year zero coupon bonds in data from our model: the R^2 for these predictive regressions on data from our model are between 0.46 and 0.50. This finding of high predictability is quite robust across the different variants of our model we

Table G.4: Restricted Regression Results for b_n

	Data		
n	b_n	SE	$\overline{R}^2$
2	0.58	(0.07)	0.35
3	0.93	(0.08)	0.41
4	1.18	(0.09)	0.40
5	1.31	(0.10)	0.37
	Model		
n	b_n	SE	$\overline{R}^2$
2	0.60	(0.05)	0.47
3	0.91	(0.07)	0.50
4	1.16	(0.08)	0.50
5	1.33	(0.10)	0.49

have considered. The shape of the factor that predicts returns is much more sensitive to model specification, due to the high correlations between yields at various maturities generated by our model.

H Bayesian Updating about Parameters and States for GDP

Here we describe the initial beliefs and sampling algorithm for our GDP application. We assume that the initial belief of the CBO about the mean of the difference stationary component μ is Normal,

$$\mu \sim N(\mu_\mu, \sigma_\mu^2).$$

We assume that the CBO has independent Normal initial beliefs about the sum of the autoregressive parameters $\rho_1 + \rho_2$ and for the second autoregressive parameter ρ_2. We truncate these initial belief distributions in such a way as to put zero weight on parameter combinations that result in the x_t component being non-stationary. We can write these initial belief distributions as

$$\rho_1 + \rho_2 \sim N(\mu_\rho, \sigma_\rho^2)\mathcal{I}(\rho_1, \rho_2),$$

$$\rho_2 \sim N(\mu_{\rho_2}, \sigma_{\rho_2}^2)\mathcal{I}(\rho_1, \rho_2).$$

where $\mathcal{I}(\rho_1, \rho_2)$ is an indicator variable which is 1 for (ρ_1, ρ_2) combinations that result x_t being stationary and 0 otherwise.

This implies a joint initial belief distribution for ρ_1, ρ_2 the moments of which are

$$\mu_{\rho_1} = \mu_\rho - \mu_{\rho_2},$$

$$\sigma_\rho^2 = \sigma_{\rho_1}^2 + \sigma_{\rho_2}^2 + 2\sigma_{\rho_1,\rho_2},$$

$$\sigma_{\rho,\rho_2} = \sigma_{\rho_1,\rho_2} + \sigma_{\rho_2}^2 = 0,$$

$$\sigma_{\rho_1,\rho_2} = -\sigma_{\rho_2}^2,$$

$$\sigma_{\rho_1}^2 = \sigma_\rho^2 + \sigma_{\rho_2}^2.$$

In other words,

$$\begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_\rho - \mu_{\rho_2} \\ \mu_{\rho_2} \end{bmatrix}, \begin{bmatrix} \sigma_\rho^2 + \sigma_{\rho_2}^2 & -\sigma_{\rho_2}^2 \\ -\sigma_{\rho_2}^2 & \sigma_{\rho_2}^2 \end{bmatrix} \right) \mathcal{I}(\rho_1, \rho_2).$$

We assume that the CBOs initial belief distribution about the the variance share γ of shocks to the trend component is a Beta distribution,

$$\gamma \sim \mathcal{B}(\alpha_\gamma, \beta_\gamma).$$

We assume that the CBOs initial belief distribution about the conditional variance σ^2 is an Inverse Gamma distribution,

$$\sigma^2 \sim \mathcal{IG}(\alpha_{\sigma^2}, \beta_{\sigma^2}).$$

Lastly, we assume that agents' initial beliefs about z_t and x_t in 1959Q3 are $z_t \sim N(y_{1959Q3}, 0.01^2)$ and $x_t \sim N(0, 0.01^2)$.

We start with an initial guess of the unknown parameters

$$\boldsymbol{\theta}^{(0)} = \left(\mu^{(0)}, \rho_1^{(0)}, \rho_2^{(0)}, \gamma^{(0)}, \sigma^{(0)}, \boldsymbol{z}_{1:t}^{(0)}, \boldsymbol{x}_{1:t}^{(0)} \right)'.$$

Given a draw of the parameters $\boldsymbol{\theta}^{(b)}$, we draw $\boldsymbol{\theta}^{(b+1)}$ as follows:

1. Draw $\mu^{(b+1)}|\rho_1^{(b)}, \rho_2^{(b)}, \gamma^{(b)}, \sigma^{(b)}, \boldsymbol{z}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}, \boldsymbol{y}_{1:t}$. Given the other parameters, beliefs about μ can be updated from the equation for z_t:

$$\Delta z_t^{(b)} = \mu + \sqrt{\gamma^{(b)}} \sigma^{(b)} u_t.$$

Define

$$\tilde{\sigma}_\mu^2 \equiv \left[\sigma_\mu^{-2} + \frac{t-1}{\gamma^{(b)} \left(\sigma^{(b)} \right)^2} \right]^{-1},$$

$$\tilde{\mu}_\mu \equiv \tilde{\sigma}_\mu^2 \left[\frac{\mu_\mu}{\sigma_\mu^2} + \frac{\sum_{s=2}^{t} \Delta z_{s-1}^{(b)}}{\gamma^{(b)} \left(\sigma^{(b)} \right)^2} \right].$$

The posterior of μ is $N(\tilde{\mu}_\mu, \tilde{\sigma}_\mu^2)$ and thus we draw $\mu^{(b+1)} \sim N(\tilde{\mu}_\mu, \tilde{\sigma}_\mu^2)$.

2. Draw $\rho_1^{(b+1)}, \rho_2^{(b+1)} | \mu^{(b+1)}, \gamma^{(b)}, \sigma^{(b)}, \boldsymbol{z}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}, \boldsymbol{y}_{1:t}$. Given the other parameters, beliefs about ρ_1, ρ_2 can be updated from the equation for x_t:

$$x_t^{(b)} = \rho_1 x_{t-1}^{(b)} + \rho_2 x_{t-2}^{(b)} + \sqrt{(1 - \gamma^{(b)})} \sigma^{(b)} v_t.$$

Define

$$\tilde{\boldsymbol{\Sigma}}_\rho \equiv \left[\boldsymbol{\Sigma}_\rho^{-1} + \frac{\sum_{s=3}^{t} [x_{s-1}^{(b)}, x_{s-2}^{(b)}]'[x_{s-1}^{(b)}, x_{s-2}^{(b)}]}{(1 - \gamma^{(b)}) \left(\sigma^{(b)} \right)^2} \right]^{-1},$$

$$\tilde{\boldsymbol{\mu}}_\rho \equiv \tilde{\boldsymbol{\Sigma}}_\rho \left[\boldsymbol{\Sigma}_\rho^{-1} \boldsymbol{\mu}_\rho + \frac{\sum_{s=3}^{t} [x_{s-1}^{(b)}, x_{s-2}^{(b)}]' x_s^{(b)}}{(1 - \gamma^{(b)}) \left(\sigma^{(b)} \right)^2} \right].$$

The posterior of $(\rho_1, \rho_2)'$ is $N(\tilde{\boldsymbol{\mu}}_\rho, \tilde{\boldsymbol{\Sigma}}_\rho)$ and thus we draw $(\rho_1^{(b+1)}, \rho_2^{(b+1)})' \sim N(\tilde{\boldsymbol{\mu}}_\rho, \tilde{\boldsymbol{\Sigma}}_\rho)$.

3. Draw $\gamma^{(b+1)} | \mu^{(b+1)}, \rho_1^{(b+1)}, \rho_2^{(b+1)}, \sigma^{(b)}, \boldsymbol{z}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}, \boldsymbol{y}_{1:t}$. There is no closed form expression for the posterior of γ. We therefore draw it using a random walk Metropolis-Hastings step. Specifically, we draw a proposal $\tilde{\gamma}^{(b+1)} \sim N(\gamma^{(b)}, \sigma_{\gamma,prop}^2)$ where $\sigma_{\gamma,prop}^2$ is a proposal variance chosen such that this step has between a 25 and 40% acceptance rate over the burn-in period. We then set $\gamma^{(b+1)} = \tilde{\gamma}^{(b+1)}$ with probability α_{b+1}, where

$$\alpha_{b+1} \equiv \frac{L\left(\boldsymbol{y}_{1:t} | \mu^{(b+1)}, \rho_1^{(b+1)}, \rho_2^{(b+1)}, \tilde{\gamma}^{(b+1)}, \sigma^{(b)}, \boldsymbol{z}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)} \right) p_\gamma \left(\tilde{\gamma}^{(b+1)} \right)}{L\left(\boldsymbol{y}_{1:t} | \mu^{(b+1)}, \rho_1^{(b+1)}, \rho_2^{(b+1)}, \tilde{\gamma}^{(b+1)}, \sigma^{(b)}, \boldsymbol{z}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)} \right) p_\gamma \left(\gamma^{(b)} \right)}.$$

Otherwise we set $\gamma^{(b+1)} = \gamma^{(b)}$.

4. Draw $\sigma^{(b+1)} | \mu^{(b+1)}, \rho_1^{(b+1)}, \rho_2^{(b+1)}, \gamma^{(b+1)}, \boldsymbol{z}_{1:t}^{(b)}, \boldsymbol{x}_{1:t}^{(b)}, \boldsymbol{y}_{1:t}$. Given the other parameters, beliefs

about σ can be updated from the two equations

$$\Delta z_t^{(b)} = \mu^{(b+1)} + \sqrt{\gamma^{(b+1)}}\sigma u_t.$$

$$x_t^{(b)} = \rho_1^{(b+1)} x_{t-1}^{(b)} + \rho_2^{(b+1)} x_{t-2}^{(b)} + \sqrt{1 - \gamma^{(b+1)}}\sigma v_t.$$

Since u_t and v_t are independent, these regression equations can be treated as two independent sources of information for σ^2. It is as if beliefs about σ^2 are first updated using information about $\{u_s\}_{s=2}^t$ where $\sigma u_s = \frac{\Delta z_s - \mu}{\sqrt{\gamma}}$ and then updated using information about $\{v_s\}_{s=3}^t$ where $\sigma v_s = \frac{x_s - \rho_1 x_{s-1} - \rho_2 x_{s-2}}{\sqrt{1-\gamma}}$. These are samples of $t-1$ and $t-2$ observations respectively which can be used to learn about σ^2 using standard conjugate prior updating. Define

$$\tilde{\alpha}_{\sigma^2} \equiv \alpha_{\sigma^2} + (2t-3)/2,$$

$$\tilde{\beta}_{\sigma^2} \equiv \beta_{\sigma^2} + \frac{\sum_{s=2}^t \left(\Delta z_s^{(b)} - \mu^{(b+1)}\right)^2}{2\gamma^{(b+1)}} + \frac{\sum_{s=3}^t \left(x_s^{(b)} - \rho_1^{(b+1)} x_{s-1}^{(b)} - \rho_2^{(b+1)} x_{s-2}^{(b)}\right)^2}{2\left(1 - \gamma^{(b+1)}\right)}.$$

The posterior of σ^2 is $\mathcal{IG}(\tilde{\alpha}_{\sigma^2}, \tilde{\beta}_{\sigma^2})$ and thus we draw $\left(\sigma^{(b)}\right)^2 \sim \mathcal{IG}(\tilde{\alpha}_{\sigma^2}, \tilde{\beta}_{\sigma^2})$.

5. Draw $z_{1:t}^{(b+1)}, x_{1:t}^{(b+1)} | \mu^{(b+1)}, \rho_1^{(b+1)}, \rho_2^{(b+1)}, \gamma^{(b+1)}, \sigma^{(b+1)}, y_{1:t}$. This can be done using the standard Kalman filter and simulation smoother.

I   Bayesian Forecasting of GDP

The algorithm described in Appendix H yields B samples of the posterior of the states and parameters of our UC model for GDP at each point in time t. We index these samples by b as follows $\left\{\rho_1^{(b)}, \rho_2^{(b)}, \gamma^{(b)}, \mu^{(b)}, \sigma^{(b)}, z_{t|t}^{(b)}, x_{t|t}^{(b)}, x_{t-1|t}^{(b)}\right\}_{b=1}^B$. We then use the following algorithm to produce a real-time forecast distribution for the GDP at time t:

1. For each $b = 1, \ldots, B$

    (a) Simulate a path of shocks $\left\{u_{t+h}^{(b)}, v_{t+h}^{(b)}\right\}_{h=1}^H$ from the standard Normal distribution.

    (b) Starting from $h = 1$, construct a simulated path of the states over H subsequent periods

using equations

$$z_{t+h|t}^{(b)} = \mu^{(b)} + \sqrt{\gamma^{(b)}}\sigma^{(b)}u_{t+h}^{(b)},$$

$$x_{t+h|t}^{(b)} = \rho_1^{(b)}x_{t+h-1|t}^{(b)} + \rho_2^{(b)}x_{t+h-2|t}^{(b)} + \sqrt{1-\gamma^{(b)}}\sigma^{(b)}v_{t+h}^{(b)}.$$

(c) Use the simulated states to construct $\left\{y_{t+h|t}^{(b)}\right\}_{h=1}^{H}$ where

$$y_{t+h|t}^{(b)} = z_{t+h|t}^{(b)} + x_{t+h|t}^{(b)}.$$

2. The forecast of y_{t+h} given time t information is computed as

$$F_t y_{t+h} = \frac{1}{B}\sum_{b=1}^{B} y_{t+h|t}^{(b)}.$$

At the end of the estimation we are left with a sequence of model-implied 1 to H-quarter ahead forecasts $\{F_t y_{t+h}\}_{h=1}^{H}$ for every year t from 1976Q4 to 2019Q4.

We must perform a few additional steps to transform our forecasts to ones that are comparable to those produced by the CBO. The CBO publishes forecasts of growth in the average annual level of real output. We define the average annual level of real output over the year preceding quarter t as

$$\bar{Y}_t \equiv \frac{1}{4}\sum_{s=t-3}^{t} \exp(y_s).$$

As an example, in the CBO's economic outlook published in 1990, its 1-year ahead forecast of GDP growth is

$$100 \times \left(\frac{\bar{Y}_{1990Q4}}{\bar{Y}_{1989Q4}} - 1\right).$$

Thus to convert the model's forecasts of quarterly log real GDP to average annual h-year ahead level forecasts, we apply the following transformation to the simulated forecast distribution

$$F_t\bar{Y}_{t+h} \equiv \frac{1}{B}\sum_{b=1}^{B}\left[\frac{1}{4}\sum_{s=t+4h-3}^{t+4h}\exp\left(F_t y_{s|t}^{(b)}\right)\right].$$

The associated forecasts of growth in average annual levels between year $t+h-1$ and $t+h$ for $h = 1, \ldots, H$ are

$$100 \times \left(\frac{F_t\bar{Y}_{t+h}}{F_t\bar{Y}_{t+h-1}} - 1\right).$$

34

J   Search over Initial Beliefs for GDP Growth

We denote $\boldsymbol{\theta} = (\mu_\rho, \sigma_\rho, \mu_{\rho 2}, \sigma_{\rho 2}, \alpha_\gamma, \beta_\gamma)'$. Let $\boldsymbol{\alpha} = \{\alpha_h\}_{h=1}^H$ and $\boldsymbol{\beta} = \{\beta_h\}_{h=1}^H$ denote vectors of estimated coefficients from the forecasting anomaly regressions for different horizons up through a maximum horizon of H using the CBO data. Let $\hat{\boldsymbol{\alpha}} = \{\hat{\alpha}_h\}_{h=1}^H$ and $\hat{\boldsymbol{\beta}} = \{\hat{\beta}_h\}_{h=1}^H$ denote those same quantities estimated on the model implied forecasts and yields. Additionally, denote the $t-statistics$ associated with these coefficients as $\{\boldsymbol{t}_\alpha, \boldsymbol{t}_\beta\} = \{t_{\alpha,h}, t_{\beta,h}\}_{h=1}^H$ for the data and $\{\boldsymbol{t}_{\hat{\alpha}}, \boldsymbol{t}_{\hat{\beta}}\} = \{t_{\hat{\alpha},h}, t_{\hat{\beta},h}\}_{h=1}^H$ for the model. Define the moment function as

$$\hat{m}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\alpha}_{bias} - \hat{\boldsymbol{\alpha}}_{bias} \\ \boldsymbol{t}_{\alpha,bias} - \boldsymbol{t}_{\hat{\alpha},bias} \\ \boldsymbol{\beta}_{ar} - \hat{\boldsymbol{\beta}}_{ar} \\ \boldsymbol{t}_{\beta,ar} - \boldsymbol{t}_{\hat{\beta},ar} \\ \boldsymbol{\beta}_{mz} - \hat{\boldsymbol{\beta}}_{mz} \\ \boldsymbol{t}_{\beta,mz} - \boldsymbol{t}_{\hat{\beta},mz} \\ \boldsymbol{\beta}_{cg} - \hat{\boldsymbol{\beta}}_{cg} \\ \boldsymbol{t}_{\beta,cg} - \boldsymbol{t}_{\hat{\beta},cg} \end{bmatrix} \tag{21}$$

The parameters are then estimated via SMM with the following objective function

$$\hat{\boldsymbol{\theta}} = \text{argmax}_{\boldsymbol{\theta}} \hat{m}(\boldsymbol{\theta})' \boldsymbol{W} \hat{m}(\boldsymbol{\theta})$$

where the elements of the objective function associated with the Mincer-Zarnowitz and Coibion-Gorodnichenko coefficients are given 3 times the weight of all other elements in $\boldsymbol{W}$. We also place bounds on the estimated parameters as described in footnote 17 in the main text. The estimated initial belief distributions are plotted in Figure 10.

Every evaluation of the moment function $\hat{m}(\boldsymbol{\theta})$ requires us to sample from the posterior of the UC model sequentially. Since this step is very computationally costly, we only re-estimate the model every 4 quarters rather than every quarter, and use a burn-in sample of 15,000 draws and keep the subsequent 15,000 draws rather than 50,000 for each of those quantities in our empirical specification. The global minimum is found using MATLAB's "particleswarm" optimization routine.