Learning About the Long Run

Leland E. Farmer[1] Emi Nakamura[2] Jón Steinsson[2]

[1]University of Virginia

[2]UC Berkeley

Oct 2022

Biased Forecast Errors

- Regression equation:

$$e_{t+h|t} = \alpha + u_{t+h}$$

where $e_{t+h|t}$ is the h period ahead forecast error ($y_{t+h} - F_t y_{t+h}$)

	$\hat{\alpha},\ H_0 : \alpha = 0$ 81Q3-19Q4 for SPF, 76-19 for CBO				
	Forecast Horizon				
	1	2	3	4	5
T-Bill	-0.18*** (0.05)	-0.34*** (0.09)	-0.52*** (0.14)	-0.70*** (0.19)	–
GDP Growth	0.27 (0.25)	-0.27 (0.35)	-0.54 (0.50)	-0.62 (0.53)	-0.52 (0.49)

Autocorrelated Forecast Errors

- Regression equation:

$$e_{t+h|t} = \alpha + \beta e_{t|t-h} + u_{t+h}$$

where $e_{t+h|t}$ is the h period ahead forecast error.

	$\hat{\beta},\ H_0 : \beta = 0$ 81Q3-19Q4 for SPF, 76-19 for CBO				
	Forecast Horizon				
	1	2	3	4	5
T-Bill	0.30*	0.27**	0.24*	0.13	–
	(0.14)	(0.12)	(0.12)	(0.13)	
GDP Growth	0.22	0.16	0.11	0.08	0.08
	(0.12)	(0.14)	(0.13)	(0.18)	(0.10)

MINCER-ZARNOWITZ REGRESSIONS

- Regression equation:

$$y_{t+h} = \alpha + \beta F_t y_{t+h} + u_{t+h}$$

where y_{t+h} is truth at $t+h$ and $F_t y_{t+h}$ is h period ahead forecast

	$\hat{\beta},\ H_0 : \beta = 1$ 81Q3-19Q4 for SPF, 76-19 for CBO				
	Forecast Horizon				
	1	2	3	4	5
T-Bill	0.97* (0.02)	0.94** (0.02)	0.90** (0.04)	0.86** (0.05)	–
GDP Growth	0.94 (0.10)	0.60 (0.38)	0.03** (0.27)	-0.42*** (0.18)	-0.43*** (0.29)

Forecasting Future Short Rates

$$\frac{1}{n}\sum_{i=0}^{n-1} y_{t+i}^{(1)} - y_t^{(1)} = \alpha + \beta(y_t^{(n)} - y_t^{(1)}) + u_t$$

	$\hat{\beta},\ H_0 : \beta = 1$ 61Q3-19Q4						
	Long Horizon *n*						
	2	3	4	8	12	20	40
Future Short Rates	-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)

Forecasting Changes in Long Rates

$$y_{t+1}^{(n-1)} - y_t^{(n)} = \alpha + \beta \left(\frac{1}{n-1} \right) (y_t^{(n)} - y_t^{(1)}) + u_t$$

	$\hat{\beta},\ H_0 : \beta = 1$ 61Q3-19Q4						
	Long Horizon n						
	2	3	4	8	12	20	40
Change in Long Rate	-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)

Term Structure Anomalies

- Time-varying risk premia

  (Wachter 06, Bansal-Shaliastovich 13, Vayanos-Vila 21)

- Forecasts deviate from full-information rational expectations

  (Froot 89, Piazzesi-Salomao-Schneider 15, Cieslak 18, Xu 19, Nagel-Xu 21)

Forecasting Anomalies in Macro

- Traditional reaction: Forecasters are irrational / inefficient
  (Mincer-Zarnowitz 69, Friedman 80, Maddala 91, Croushore 98, Schuh 01)
  - Behavioral explanation: Bordalo-Gennaioli-Ma-Shleifer 20

- Alternative reaction: Information rigidity/frictions
  (Mankiw-Reis-Wolfers 03, Coibion-Gorodnichenko 12, 15)

Types of Information Frictions

- Sticky Information (e.g., Mankiw-Reis 02)
  - **Key assumption**: Forecasters update information infrequently
  - But professional forecasters look at the newspaper every day!
- Noisy Information (e.g., Sims, 2003; Woodford, 2003)
  - **Key assumption**: Forecasters get a noisy signal of interest rates / GDP
  - But professional forecasters know the data exactly

Alternative: Not Knowing Model

- Assumption that forecasters know the model very strong assumption
- More realistic to assume that:
  - Forecasters are learning about the model that generates the data
- Parameter learning fundamentally changes dynamics
- Can lead standard rational expectations tests to fail

(Friedman 79, Lewis 89, Barsky-DeLong 93, Timmermann 93, Lewellen-Schanken 02, Brav-Heaton 02, Cogley-Sargent 05, Collin-Dufresne-Johannes-Lochstoer 16, Jahnnes-Lochstoer-Mou 16, Singleton 21)

CIESLAK (2008)

Ex post predictability of forecast errors does not imply that people make “obvious” mistakes that could be easily fixed in real time. Even when conducting a quasi-real-time estimation, an econometrician uses ex post knowledge of a statistical relationship that would have been much harder to uncover in real time.

Major Challenge

- Realistic models are hard to solve with parameter learning!
- Most earlier work on parameter learning used relatively simple models
  - In these models, Bayesian learning is fast
  - Can't explain persistent anomalies (i.e., over several decades)
- Informal discussion of parameter breaks that might sustain learning
- Not clear whether Bayesian (i.e. rational) learning can quantitatively explain forecasting anomalies over several decades

Unobserved Components Models

- Realistic to assume unobserved components in many settings
  - "Shifting end-points" model for term structure (Kozicki-Tinsley 01)
  - Difference and trend stationary components for GDP
- Such models can be very hard to learn!
- Different parameters can yield:
  - Similar fit to high frequency behavior
  - But very different implications for low frequency behavior
- Bayesian learning can be very slow

(Collin-Dufresne-Johannes-Lochstoer 16 and Kozlowski-Veldkamp-Venkataswaran 20 achieve slow learning in models with rare events)

What We Do

- Two applications:
  - Forecasting interest rates
  - Forecasting GDP growth
- Consider Bayesian forecasters
  - Endow them with unobserved components model
    + initial beliefs about parameters
  - Have them learn about model parameters in real time
  - Have them generate real-time forecasts
- Assess whether resulting forecasts are "anomalous"

Main Result

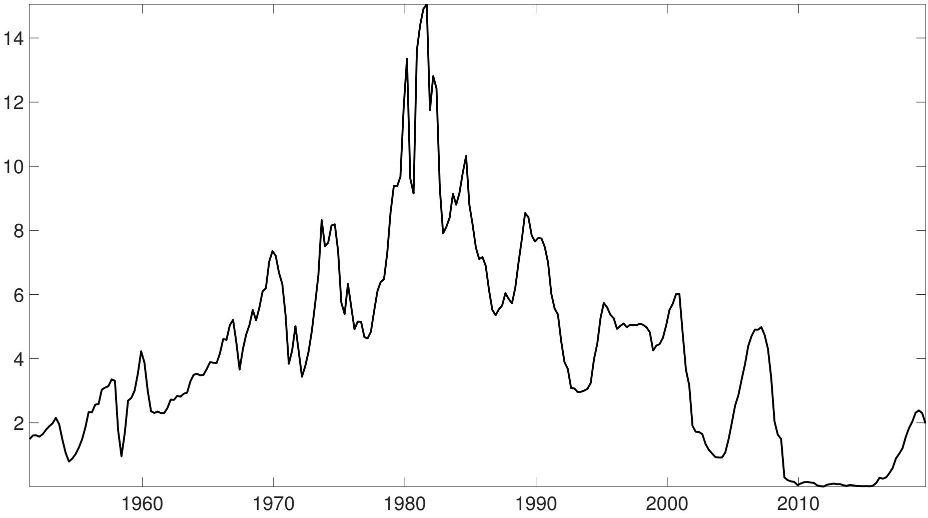
- Can match all the forecasting anomalies when forecasters are endowed with “reasonable” initial beliefs
- Interpretation:
  - Learning can generate very persistent forecasting anomalies
  - Low frequency phenomena very hard to learn about:
    - Long-run behavior of interest rates
    - Long-run changes in economic growth
  - Rational expectations tests can be very misleading (even over 50 years) when low frequency phenomena are important

Interest Rate Forecasting

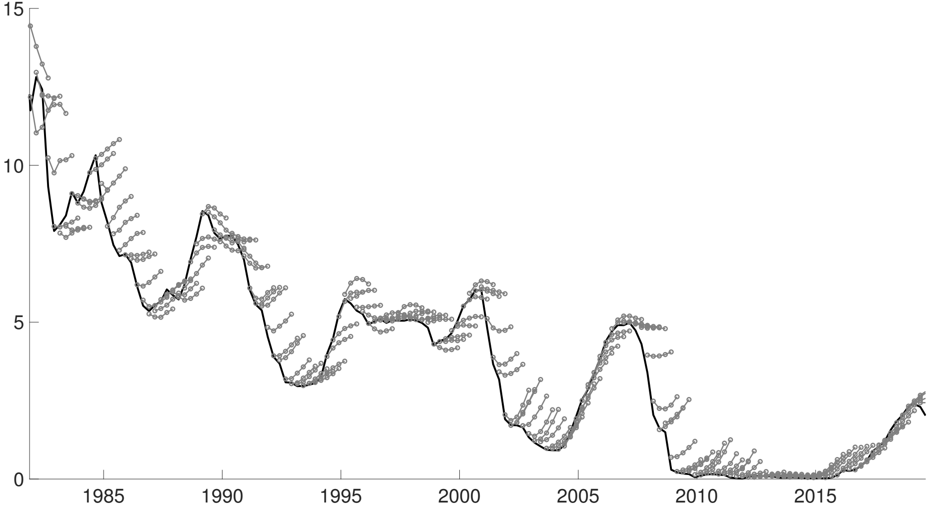
Data on Interest Rates

- Short-term interest rate: 3-month Treasury bill rate
  - Sample period: 1951Q2-2019Q4
  - Quarterly averages
- Zero-coupon yield curve: Liu and Wu (2020)
  - Sample period: 1961Q3-2019Q4
- Forecasts: Survey of Professional Forecasters
  - Sample period: 1981Q3-2019Q4
  - Participants surveyed near the middle of each quarter
  - Asked to produce nowcasts and forecasts up to 4 quarters in the future

3 Month Treasury Bill Yield



Professional Forecasts



- 3-month treasury bill yield: y_t

$$y_t = \mu_t + x_t$$
$$\mu_t = \mu_{t-1} + \sqrt{\gamma}\sigma\eta_t, \qquad \eta_t \sim N(0,1)$$
$$x_t = \rho x_{t-1} + \sqrt{(1-\gamma)}\sigma\omega_t, \qquad \omega_t \sim N(0,1)$$

- Parameters:
  - γ: variance share of permanent component
  - ρ: persistence of transitory component
  - σ: conditional volatility of short yield

Real-Time Learning and Forecasting

- "Online" estimation to mimic problem of forecasters
- Start with initial beliefs in 1951Q2
- Each quarter - new observation of short-term interest rate
- Re-estimate the parameters of UC model
- Use posteriors for (μ_t, x_t) and (ρ, γ, σ) to produce 1-40 quarter ahead forecasts
- Turn off parameter learning during ZLB period ZLB Learning
  (Interest rate process is censored.)

Initial Beliefs Will Matter

- Since learning will be slow, the initial beliefs of forecasters will matter
- Important question:
  - Can we match anomalies while assuming “reasonable” initial beliefs for forecasters?

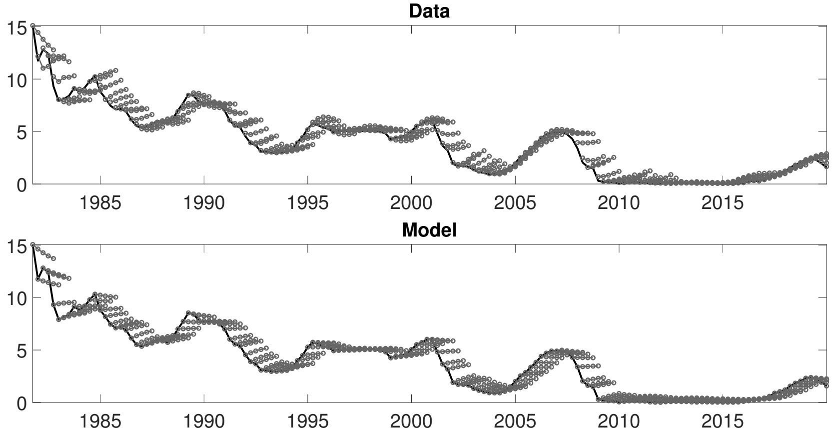
Initial Beliefs that Match Anomalies

- Search over space of initial beliefs:

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

- Four parameters: μ_ρ, σ_ρ^2, α_γ, β_γ
- Fix parameters associated with prior for $\sigma^2 \sim \mathcal{IG}(\alpha_\sigma, \beta_\sigma)$
  - α_σ and β_σ chosen such that prior has a mode of 0.25 and is highly dispersed
- Objective:
  - Minimize sum of squared deviations of model vs. data regression coefficients from anomaly regressions
  - 6 anomaly regressions estimated at different horizons (total of 29 targets)

RESULTS: MODEL-IMPLIED FORECASTS



Results: Bias

$$e_{t+h|t} = \alpha + u_{t+h}$$

	$\hat{\alpha},\ H_0 : \alpha = 0$			
	Forecast Horizon			
	1	2	3	4
SPF	-0.18***	-0.34***	-0.52***	-0.70***
	(0.05)	(0.09)	(0.14)	(0.19)
UC Model	-0.15**	-0.27**	-0.40**	-0.51**
	(0.06)	(0.11)	(0.16)	(0.21)

RESULTS: AUTOCORRELATION

$$e_{t+h|t} = \alpha + \beta e_{t|t-h} + u_{t+h}$$

	$\hat{\beta},\ H_0 : \beta = 0$			
	Forecast Horizon			
	1	2	3	4
SPF	0.30*	0.27**	0.24*	0.13
	(0.14)	(0.12)	(0.12)	(0.13)
UC Model	0.36*	0.39**	0.35**	0.23*
	(0.17)	(0.14)	(0.11)	(0.12)

RESULTS: MINCER-ZARNOWITZ

$$y_{t+h} = \alpha + \beta F_t y_{t+h} + u_{t+h}$$

	$\hat{\beta},\ H_0 : \beta = 1$			
	Forecast Horizon			
	1	2	3	4
SPF	0.97*	0.94**	0.90**	0.86**
	(0.02)	(0.02)	(0.04)	(0.05)
UC Model	0.96*	0.93**	0.88**	0.84***
	(0.02)	(0.03)	(0.04)	(0.05)

Results: Coibion-Gorodnichenko

$$e_{t+h|t} = \alpha + \beta \left(F_t y_{t+h} - F_{t-1} y_{t+h}\right) + u_{t+h}$$

	$\hat{\beta},\ H_0 : \beta = 0$			
	Forecast Horizon			
	1	2	3	4
SPF	0.23*	0.34*	0.62***	–
	(0.12)	(0.16)	(0.16)	
UC Model	0.39*	0.56	0.89*	–
	(0.18)	(0.37)	(0.42)	

Results: Forecasting Short Rates

$$\frac{1}{n}\sum_{i=0}^{n-1} y_{t+i}^{(1)} - y_t^{(1)} = \alpha + \beta(y_t^{(n)} - y_t^{(1)}) + u_t$$

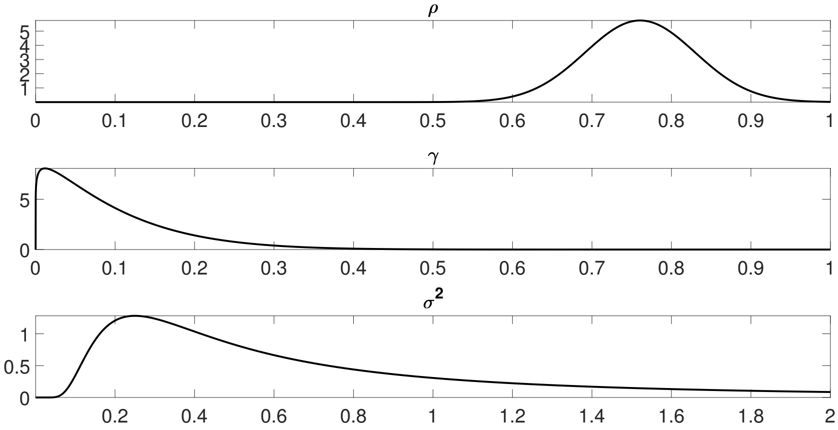
	$\hat{\beta},\ H_0 : \beta = 1$						
	Long Horizon n						
	2	3	4	8	12	20	40
Data	-0.01***	0.11***	0.18***	0.39**	0.57	0.74	0.71
	(0.23)	(0.23)	(0.23)	(0.23)	(0.26)	(0.23)	(0.20)
UC Model	-0.11***	0.08**	0.17**	0.56	0.81	0.93	0.99
	(0.32)	(0.32)	(0.33)	(0.38)	(0.37)	(0.31)	(0.36)

Results: Forecasting Changes in Long Rates

$$y_{t+1}^{(n-1)} - y_t^{(n)} = \alpha + \beta \left(\frac{1}{n-1} \right) (y_t^{(n)} - y_t^{(1)}) + u_t$$

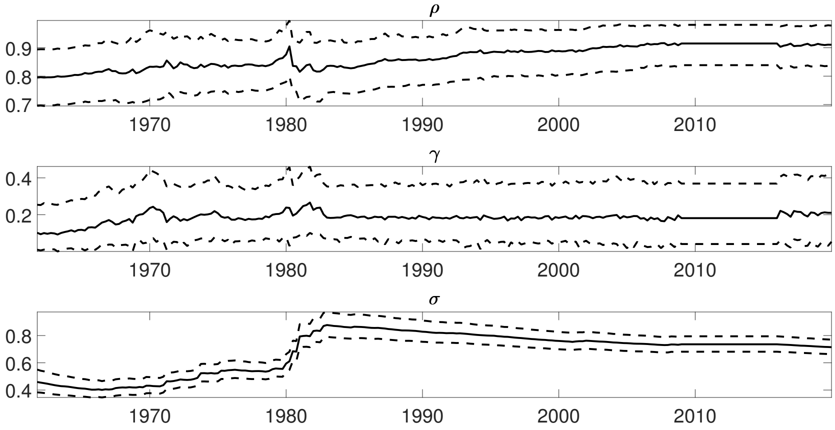
	$\hat{\beta},\ H_0 : \beta = 1$						
	Long Horizon n						
	2	3	4	8	12	20	40
Data	-1.02***	-0.91***	-1.03***	-1.29***	-1.61***	-2.04***	-2.75***
	(0.45)	(0.59)	(0.62)	(0.59)	(0.57)	(0.55)	(0.87)
UC Model	-1.21***	-1.25***	-1.28***	-1.40***	-1.54***	-1.84***	-2.55**
	(0.63)	(0.64)	(0.65)	(0.70)	(0.76)	(0.88)	(1.52)

Reasonable Initial Beliefs?



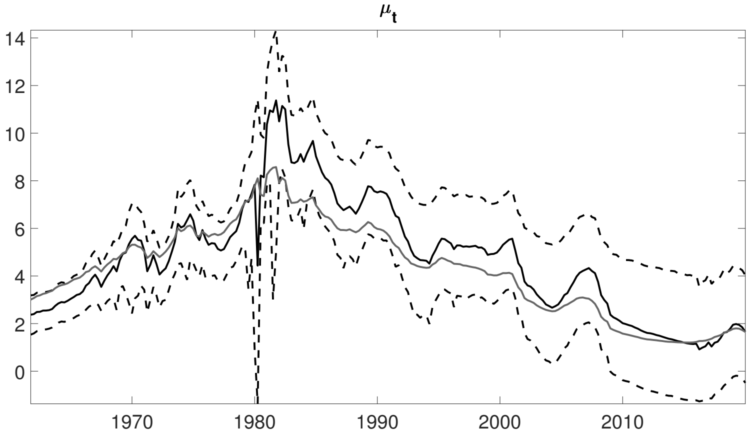
Model Families Estimation Table

Results: Parameter Estimates



Dist 1981-Q2 | Dist 1995-Q1 | Dist 2007-Q3 | Dist 2019-Q1

Results: State Estimates



▸ Yield Spread ▸ Long Rate

FAMA (2006)

There was little prior experience with a fiduciary currency when the right to exchange currency for gold was discontinued in 1971, and it is reasonable that the high inflation and interest rates that followed were a surprise...the experience led market participants to rationally predict that a fiduciary currency (a currency that is not backed by a commodity like gold) implied permanently higher expected inflation. In other words, the preceding positive shocks to expected inflation were judged to be permanent. It turns out, however, that the Federal Reserve...won...a long-odds game; they learned how to manage a fiduciary currency to bring about low inflation and interest rates. The result is a sequence of mostly negative permanent shocks to the spot rate. This story can explain why the spot rate appears to be slowly mean reverting...but the apparent mean reversion is missed by the forecasts....

Additional Results: Break in 1990

- Term structure researchers often start sample in 1990
  - Regime change at Fed after Volcker disinflation
  - Beliefs influenced by more than just past short rates

- We allow agents to "reset" their beliefs about γ in 1990

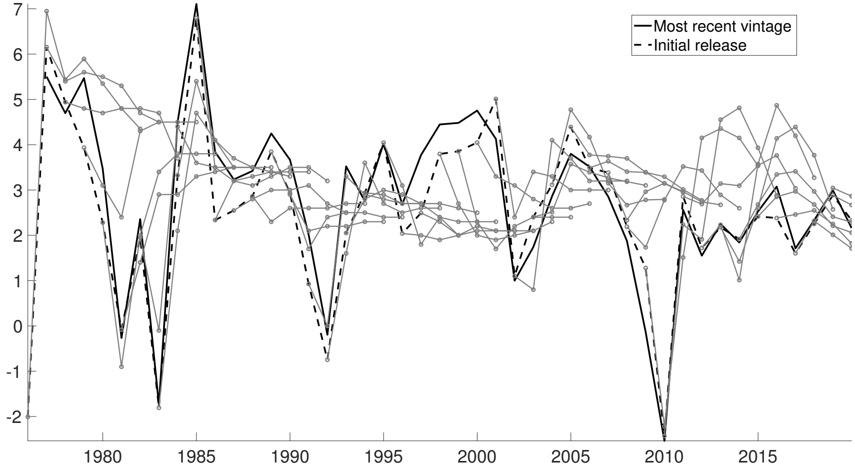
- Model can match additional facts about term structure:
  - Yield Spread ▸ Yield Spread
  - Cochrane-Piazzesi ▸ Cochrane-Piazzesi in Data ▸ Cochrane-Piazzesi in Model
  - Giglio-Kelly ▸ Giglio-Kelly

Real Output Growth Forecasting

Data for Real Output Growth Application

- Real output growth: Philadelphia Fed Real-Time Data
  - Sample period: 1959Q3 - 2019Q4
  - 1959Q3 is earliest date for which we have a full set of data vintages
- Forecasts: Congressional Budget Office
  - Sample period: 1976 - 2019
  - Forecasts of real output growth over 5 years
  - Survey typically produced with data available in early December
- In the model: Agents know the first release of Q4 when they forecast
- Data and forecasts are GNP until 1992 and GDP after that point

Professional Forecasts



Model for GDP Growth

- Quarterly Real Output Growth in logs: y_t

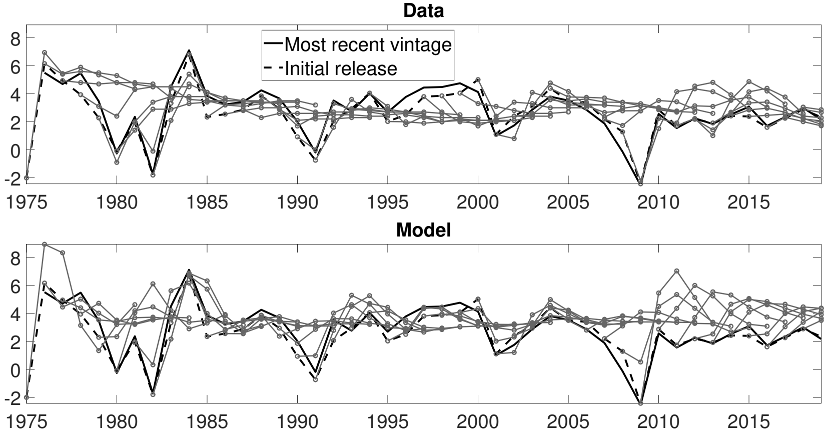
$$y_t = z_t + x_t$$

$$\Delta z_t = \mu + \sqrt{\gamma}\sigma u_t, \qquad u_t \sim N(0,1)$$

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \sqrt{1-\gamma}\sigma v_t, \qquad v_t \sim N(0,1)$$

- Parameters:
  - ρ_1, ρ_2: AR parameters of cyclical component
  - μ: mean of growth
  - γ: variance share of trend component
  - σ: conditional volatility of y_t

RESULTS: IMPLIED FORECASTS



Result: Mincer-Zarnowitz

$$y_{t+h} = \alpha + \beta F_t y_{t+h} + u_{t+h}$$

	$\hat{\beta},\ H_0 : \beta = 1$				
	Forecast Horizon				
	1	2	3	4	5
CBO	0.94	0.60	0.03**	-0.42***	-0.43***
	(0.10)	(0.38)	(0.27)	(0.18)	(0.29)
UC Model	0.84	0.35**	0.34*	-0.38***	-0.98**
	(0.11)	(0.17)	(0.31)	(0.19)	(0.53)

$$e_{t+h|t} = \alpha + \beta \left(F_t y_{t+h} - F_{t-1} y_{t+h} \right) + u_{t+h}$$

	$\hat{\beta}, H_0 : \beta = 0$				
	Forecast Horizon				
	1	2	3	4	5
CBO	0.08	0.00	0.50	-1.63**	-1.46**
	(0.08)	(0.28)	(0.58)	(0.36)	(0.40)
UC Model	0.06	-0.76	-0.11	-0.78	-1.22**
	(0.09)	(0.44)	(0.26)	(0.39)	(0.38)

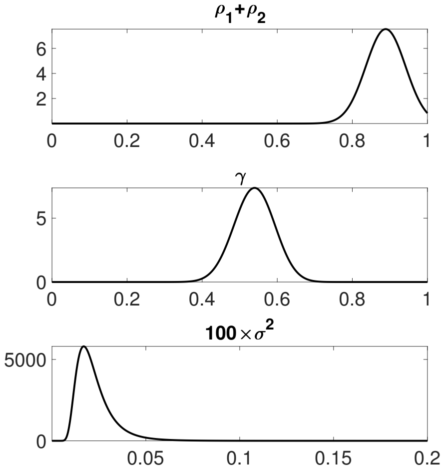
$$e_{t+h|t} = \alpha + u_{t+h}$$

	$\hat{\alpha},\ H_0 : \alpha = 0$				
	Forecast Horizon				
	1	2	3	4	5
CBO	0.27	-0.27	-0.54	-0.62	-0.52
	(0.25)	(0.35)	(0.50)	(0.53)	(0.49)
UC Model	-0.65	-1.65**	-1.36**	-0.85	-0.66
	(0.32)	(0.45)	(0.45)	(0.42)	(0.40)

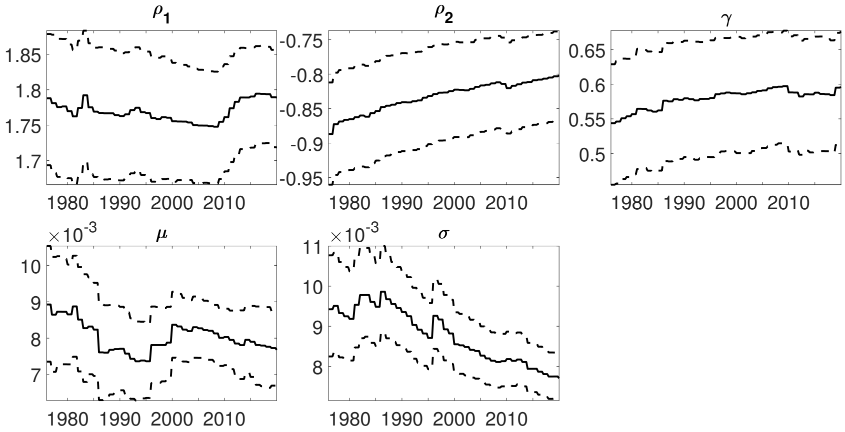
$$e_{t+h|t} = \alpha + \beta e_{t|t-h} + u_{t+h}$$

	$\hat{\beta},\ H_0 : \beta = 0$				
	Forecast Horizon				
	1	2	3	4	5
CBO	0.22	0.16	0.11	0.08	0.08
	(0.12)	(0.14)	(0.13)	(0.18)	(0.10)
UC Model	0.39*	0.31	0.23*	0.06	-0.05
	(0.17)	(0.16)	(0.10)	(0.10)	(0.05)

Reasonable Initial Beliefs?



Results: Parameter Estimates



Why Does It Work?

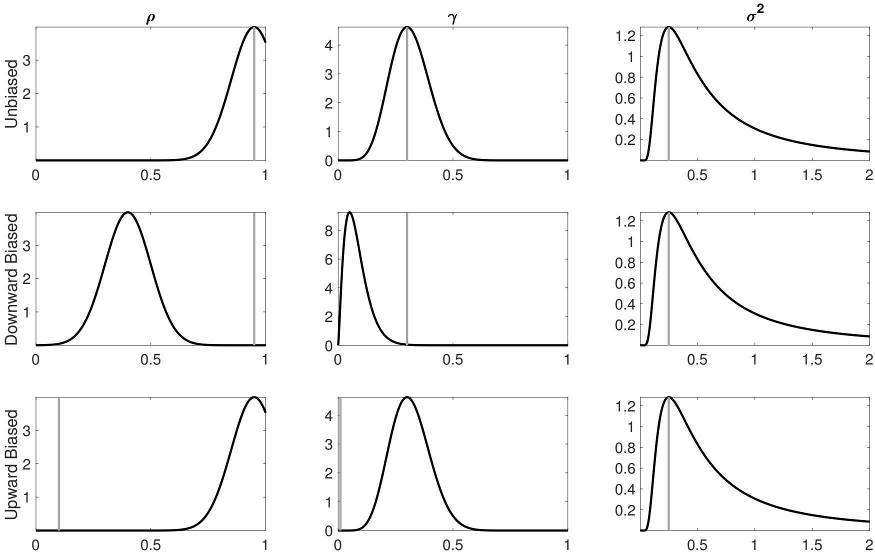
Why Does It Work?

- Interest rate model:

$$
\begin{aligned}
y_t &= \mu_t + x_t \\
\mu_t &= \mu_{t-1} + \sqrt{\gamma}\sigma\eta_t, && \eta_t \sim N(0,1) \\
x_t &= \rho x_{t-1} + \sqrt{(1-\gamma)}\sigma\omega_t, && \omega_t \sim N(0,1)
\end{aligned}
$$

- Consider three simulations:
  1. Truth: $\rho = 0.95$, $\gamma = 0.3$, $\sigma = 0.5$. Initial Beliefs: Unbiased
  2. Truth: $\rho = 0.95$, $\gamma = 0.3$, $\sigma = 0.5$. Initial Beliefs: Downward-biased
  3. Truth: $\rho = 0.1$, $\gamma = 0.01$, $\sigma = 0.5$. Initial Beliefs: Upward-biased

Prior Distributions



Results: Autocorrelation

$$e_{t+h|t} = \alpha + \beta e_{t|t-h} + u_{t+h}$$

	$\hat{\beta},\ H_0 : \beta = 0$			
	Forecast Horizon			
	1	2	3	4
Unbiased Initial Beliefs	0.01	0.00	-0.00	-0.01
	(0.08)	(0.09)	(0.11)	(0.13)
	1.00	1.00	0.99	0.84
Downward-Biased Initial Beliefs	0.16	0.19	0.19	0.18
	(0.09)	(0.10)	(0.12)	(0.14)
	0.93	0.78	0.61	0.33
Upward-Biased Initial Beliefs	-0.34	-0.32	-0.28	-0.26
	(0.06)	(0.07)	(0.08)	(0.08)
	1.00	1.00	1.00	1.00

Note: 1) Mean across simulations, 2) standard deviation across simulations,
3) Fraction of simulations that give a smaller estimate than the real-world data

RESULTS: COIBION-GORODNICHENKO

$$e_{t+h|t} = \alpha + \beta \left(F_t y_{t+h} - F_{t-1} y_{t+h} \right) + u_{t+h}$$

	$\hat{\beta},\ H_0 : \beta = 0$		
	Forecast Horizon		
	1	2	3
Unbiased Initial Beliefs	0.01 (0.09) 0.99	0.01 (0.12) 0.99	0.01 (0.15) 1.00
Downward-Biased Initial Beliefs	0.18 (0.11) 0.66	0.32 (0.19) 0.55	0.41 (0.25) 0.79
Upward-Biased Initial Beliefs	-0.52 (0.10) 1.00	-0.55 (0.13) 1.00	-0.53 (0.17) 1.00

Note: 1) Mean across simulations, 2) standard deviation across simulations, 3) Fraction of simulations that give a smaller estimate than the real-world data

Results: Forecasting Short Rates

$$\frac{1}{k}\sum_{i=0}^{k-1} y_{t+i}^{(1)} - y_t^{(1)} = \alpha + \beta(y_t^{(n)} - y_t^{(1)}) + u_t$$

	$\hat{\beta},\ H_0 : \beta = 1$						
	Long Horizon n						
	2	3	4	8	12	20	40
Unbiased Initial Beliefs	0.95	1.01	1.05	1.19	1.31	1.51	2.06
	(0.64)	(0.63)	(0.66)	(0.72)	(0.76)	(0.82)	(1.03)
	0.07	0.07	0.08	0.12	0.16	0.16	0.08
Downward-Biased Initial Beliefs	0.17	0.20	0.23	0.33	0.42	0.57	0.97
	(0.19)	(0.21)	(0.22)	(0.29)	(0.33)	(0.40)	(0.56)
	0.17	0.30	0.38	0.57	0.65	0.66	0.29
Upward-Biased Initial Beliefs	2.46	2.14	1.97	1.71	1.64	1.59	1.50
	(0.21)	(0.16)	(0.14)	(0.09)	(0.07)	(0.06)	(0.05)
	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: 1) Mean across simulations, 2) standard deviation across simulations,
3) Fraction of simulations that give a smaller estimate than the real-world data

Results: Forecasting Changes in Long Rates

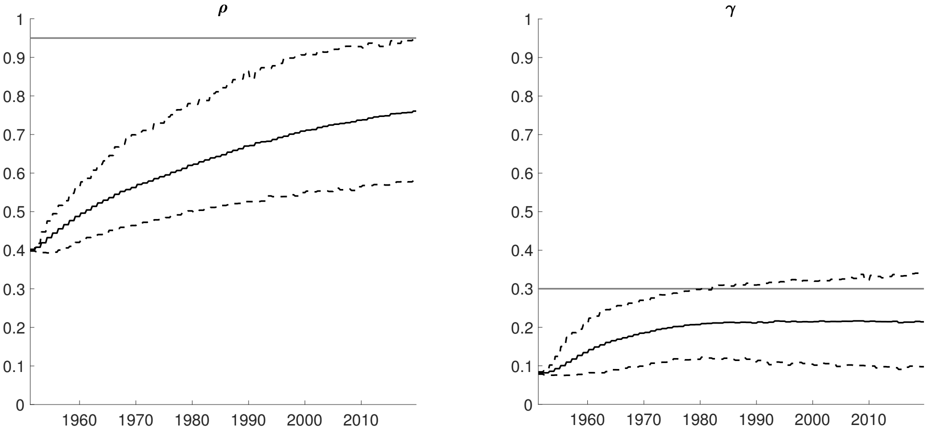
$$y_{t+1}^{(n-1)} - y_t^{(n)} = \alpha + \beta \left(\frac{1}{n-1} \right) (y_t^{(n)} - y_t^{(1)}) + u_t$$

	$\hat{\beta},\ H_0 : \beta = 1$						
	Long Horizon n						
	2	3	4	8	12	20	40
Unbiased Initial Beliefs	0.90 (1.27) 0.07	0.93 (1.32) 0.08	0.95 (1.36) 0.07	1.01 (1.50) 0.06	1.08 (1.63) 0.05	1.20 (1.92) 0.03	2.08 (3.00) 0.03
Downward-Biased Initial Beliefs	-0.66 (0.38) 0.17	-0.69 (0.40) 0.28	-0.74 (0.42) 0.24	-1.03 (0.52) 0.32	-1.39 (0.64) 0.34	-2.04 (0.91) 0.51	-3.62 (1.84) 0.66
Upward-Biased Initial Beliefs	3.91 (0.42) 0.00	4.13 (0.42) 0.00	4.38 (0.42) 0.00	5.59 (0.51) 0.00	6.90 (0.62) 0.00	9.56 (0.86) 0.00	13.77 (1.62) 0.00

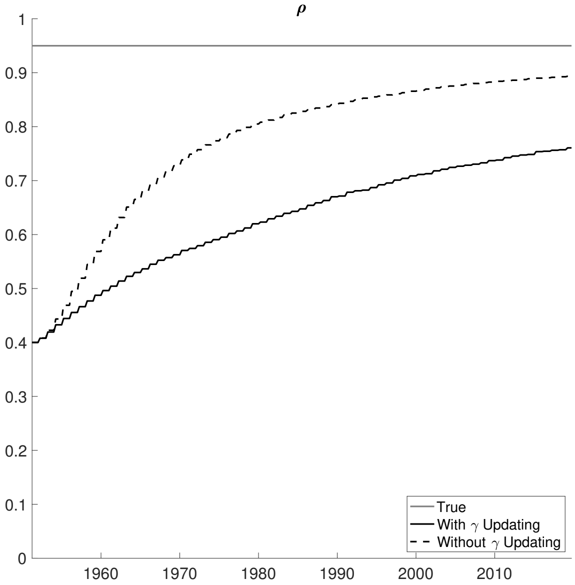
Note: 1) Mean across simulations, 2) standard deviation across simulations,

3) Fraction of simulations that give a smaller estimate than the real-world data

Very Slow Learning



Two Components Slows Learning



Conclusion

- Forecast anomalies can be explained with:
  - Bayesian learning about unobserved components models and forecasters that have “reasonable” initial beliefs
- Model explains deviations from expectations hypothesis of the term structure without reference to time-varying risk premia
- Lessons:
  - Low frequency behavior very hard to learn
  - Full information rational expectations analysis can be very misleading when low frequency behavior is important