Time Varying Volatility as a Source of Error Predictability

Julio Ortiz
Boston University

March 2020

What Explains Error Predictability?

- Professional forecasts exhibit error predictability
  - Robust across level of aggregation, horizons, variables, sub-samples
  - Implies overreactions/underreactions
- So what?
  - Forecast performance, macro modeling
- Evidence against rational expectations (RE)?
  - Imperfect information
    - Rational inattention, noisy memory
  - Distorted incentives
    - Strategic interaction, smoothing motive, asymmetric loss
- Non-rational expectations
  - Overconfidence, misperceptions, diagnostic expectations

Matching the Survey Data

Existing models can replicate certain features of survey data

- Tend to restrict the sign of the bias
    - **Either** over-reaction or under-reaction

Survey data: forecasters **simultaneously** overreact and underreact

- I provide evidence of this from SPF
- Finding: Coinciding over- and underreactions are the norm

This Paper

Modifies standard model of expectations to explain the data

1. Assumes **nonlinear** unobserved state + **noisy information**
2. Forecasters **approximate** state using an approximation function
    - Choose from a set of costly alternatives (Branch 2004)
    - More sophisticated approximation methods are more expensive
3. Incorporates public signal observed with a lag
    - Macroeconomic variables as lagged public signals

Model can generate **both** individual over-reactions and under-reactions

Related Literature

Expectations Formation

- Azeredo da Silveira and Woodford (2019), Bordalo et. al. (2019), Branch (2004) , Brock and Hommes (1997), Broer and Kohlhas (2018), Coibion and Gorodnichenko (2012, 2015), Dovern et. al (2015), Fuhrer (2018), Fuster et. al (2012), Kucinskas and Peters (2019), Mankiw and Reis (2002), Messina et al (2015), Nordhaus (1985), Ottaviani and Sørensen (2007), Pesaran and Weale (2006), Rozsypal and Schlafmann (2019)

Stochastic Volatility

- Justiniano and Primiceri (2009), Kim and Nelson (1999), McConnell and Perez-Quiros (2000), Stock and Watson (2002, 2007)

Nonlinear Filtering

- Crisan and Doucet (2002), Doucet and Johansen (2009), Hu et. al (2011), Julier and Uhlman (2004)

Outline

Established Facts From Survey of Professional Forecasters

Notation: $x^i_{t+h|t}$ is forecaster i's conditional expectation of x devised at t for horizon $t+h$

1. Aggregate errors covary positively with revisions (Coibion and Gorodnichenko 2012, 2015)

$$x_t - x_{t|t} = \alpha_0 + \alpha_1 \big[x_{t|t} - x_{t|t-1} \big] + \epsilon_t$$

2. Individual errors covary negatively (and positively) with revisions (Bordalo et. al. 2018)

$$x_t - x^i_{t|t} = \beta_0 + \beta_1 \big[x^i_{t|t} - x^i_{t|t-1} \big] + \omega^i_t$$

Errors-on-Revisions Regressions

Consensus-Level

$$x_t - x_{t|t} = \alpha_0 + \alpha_1 \left[x_{t|t} - x_{t|t-1} \right] + \epsilon_t$$

Table 1: Consensus-Level Error Predictability Regressions: Real GDP

	Nowcast		One-Quarter Ahead		Two-Quarters Ahead	
	(1) Aggregate	(2) Individual	(3) Aggregate	(4) Individual	(5) Aggregate	(6) Individual
Revision	0.359**	-0.264***	0.675**	-0.227***	0.689*	-0.307***
	(0.178)	(0.077)	(0.314)	(0.072)	(0.374)	(0.056)
Observations	197	5739	196	4681	195	4527

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

- Data from Survey of Professional Forecasters (1968Q4-2018Q4)

Errors-on-Revisions Regressions

Forecaster-Level

$$x_t - x^i_{t|t} = \beta_0 + \beta_1 \left[x^i_{t|t} - x^i_{t|t-1}\right] + \omega^i_t$$

Table 2: Consensus-Level Error Predictability Regressions: Real GDP

	Nowcast		One-Quarter Ahead		Two-Quarters Ahead	
	(1) Aggregate	(2) Individual	(3) Aggregate	(4) Individual	(5) Aggregate	(6) Individual
Revision	0.359**	-0.264***	0.675**	-0.227***	0.689*	-0.307***
	(0.178)	(0.077)	(0.314)	(0.072)	(0.374)	(0.056)
Observations	197	5739	196	4681	195	4527

* p<0.10, ** p<0.05, *** p<0.01

- Data from Survey of Professional Forecasters (1968Q4-2018Q4)

Error Orthogonality Regressions

By Variable and Level of Aggregation

Variable	Mnemonic	β_1	α_1
Consumer price index	CPI	-0.085	0.868***
Employment	EMP	-0.123	0.564***
Housing starts	HOUSING	0.063	0.359***
Industrial production	IP	-0.147*	0.513***
Nominal GDP	NGDP	-0.310***	0.421**
GDP Deflator	PGDP	-0.363***	0.350**
Real consumption	RCONSUM	-0.401***	0.098
Real federal government spending	RFEDGOV	-0.483***	0.377
Real GDP	RGDP	-0.264***	0.359**
Real nonresidential investment	RNRESIN	-0.499**	0.362
Real residential investment	RRESINV	-0.234***	0.925***
Real state/local government spending	RSLGOV	-0.660***	-0.381
3-month Treasury bill	TBILL	0.010	0.178***
10-year Treasury bond	TBOND	0.020	0.154***
Unemployment rate	UNEMP	0.082**	0.247***

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 3: Pooled OLS Regressions at $h = 0$, by Variable

New Fact About Professional Forecasters

The same forecaster over- and underreacts to different variables

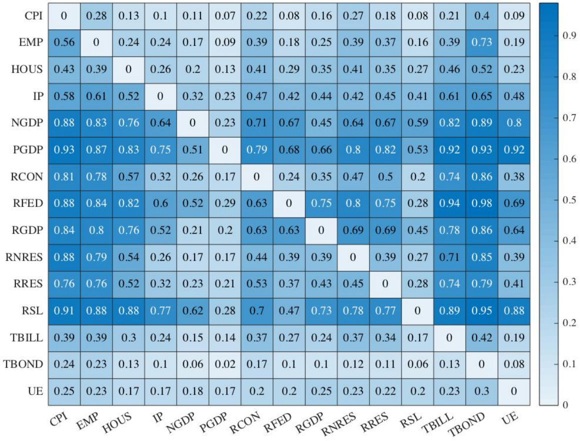
Not immediately obvious from pooled regression results

- Exit/entry among forecasters

To see this, consider forecaster i's predictions for variables j and k

- Compute share of forecasters for which $\widehat{\beta}_{1,ij} < 0$ and $\widehat{\beta}_{1,ik} > 0$

Figure 1: Frequency of Over- and Under-Reactions by the Same Forecaster



Note: Table displays share of forecasters who over-react to the row variable and under-react to the column variable.

Interpretation

≈ 84% of forecasters over-react to real GDP and simultaneously under-react to CPI!

	CPI	EMP	HOUS	IP	NGDP	PGDP	RCON	RFED	RGDP	RNRES	RRES	RSL	TBILL	TBOND	UE
CPI	0	0.28	0.13	0.1	0.11	0.07	0.22	0.08	0.16	0.27	0.18	0.08	0.21	0.4	0.09
EMP	0.56	0	0.24	0.24	0.17	0.09	0.39	0.18	0.25	0.39	0.37	0.16	0.39	0.73	0.19
HOUS	0.43	0.39	0	0.26	0.2	0.13	0.41	0.29	0.35	0.41	0.35	0.27	0.46	0.52	0.23
IP	0.58	0.61	0.52	0	0.32	0.23	0.47	0.42	0.44	0.42	0.45	0.41	0.61	0.65	0.48
NGDP	0.88	0.83	0.76	0.64	0	0.23	0.71	0.67	0.45	0.64	0.67	0.59	0.82	0.89	0.8
PGDP	0.93	0.87	0.83	0.75	0.51	0	0.79	0.68	0.66	0.8	0.82	0.53	0.92	0.93	0.92
RCON	0.81	0.78	0.57	0.32	0.26	0.17	0	0.24	0.35	0.47	0.5	0.2	0.74	0.86	0.38
RFED	0.88	0.84	0.82	0.6	0.52	0.29	0.63	0	0.75	0.8	0.75	0.28	0.94	0.98	0.69
RGDP	0.84	0.8	0.76	0.52	0.21	0.2	0.63	0.63	0	0.69	0.69	0.45	0.78	0.86	0.64
RNRES	0.88	0.79	0.54	0.26	0.17	0.17	0.44	0.39	0.39	0	0.39	0.27	0.71	0.85	0.39
RRES	0.76	0.76	0.52	0.32	0.23	0.21	0.53	0.37	0.43	0.45	0	0.28	0.74	0.79	0.41
RSL	0.91	0.88	0.88	0.77	0.62	0.28	0.7	0.47	0.73	0.78	0.77	0	0.89	0.95	0.88
TBILL	0.39	0.39	0.3	0.24	0.15	0.14	0.37	0.27	0.24	0.37	0.34	0.17	0	0.42	0.19
TBOND	0.24	0.23	0.13	0.1	0.06	0.02	0.17	0.1	0.1	0.12	0.11	0.06	0.13	0	0.08
UE	0.25	0.23	0.17	0.17	0.18	0.17	0.2	0.2	0.25	0.23	0.22	0.2	0.23	0.3	0

Outline

1. Facts About Professional Forecasters
2. **General Linear Noisy Information Model**
3. Nonlinear Noisy Information Model with Costly Approximation
4. Stylized Model: Stochastic Volatility
5. Cross-Sectional Evidence from the Survey of Professional Forecasters

General Linear Noisy Information Model

$$\text{State:} \quad \underset{n\times 1}{\mathbf{s_t}} = \underset{n\times n}{\mathbf{A}}\ \underset{n\times 1}{\mathbf{s_{t-1}}} + \underset{n\times 1}{\mathbf{w_t}}\,, \quad \mathbf{w_t} \sim \mathcal{N}(\mathbf{0}, \underset{n\times n}{\mathbf{B}})$$

$$\text{Signal:} \quad \underset{m\times 1}{\mathbf{z_t^i}} = \underset{m\times n}{\mathbf{C}}\ \underset{n\times 1}{\mathbf{s_t}} + \underset{m\times 1}{\mathbf{v_t^i}}\,, \ \mathbf{v_t^i} \sim \mathcal{N}(\mathbf{0}, \underset{m\times m}{\mathbf{D}})$$

- Forecaster *i* observes vector of signals $\mathbf{z_t^i}$
  - Signals contaminated with additive noise, $\mathbf{v_t^i}$
  - $\mathbf{z_t^i}$ can include private and public signals
- Can only observe measurement errors
  - More appropriate to model macroeconomic variables as public signals observed with a lag, $x_{t-1} \in \mathbf{z_t^i}$
- Agents devise forecasts of $\mathbf{s_t}$ and report predictions about $\mathbf{z_t^i}$

Forecasts, Errors, and Revisions

From the Kalman filter, one step ahead forecast is:

$$\mathbf{z^i_{t+1|t}} = \mathbf{z^i_{t+1|t-1}} + \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}})$$

where κ is the Kalman gain

Error: $\mathbf{z^i_{t+1}} - \mathbf{z^i_{t+1|t}} = (\mathbf{z^i_{t+1}} - \mathbf{z^i_{t+1|t-1}}) - \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}})$

Revision: $\mathbf{z^i_{t+1|t}} - \mathbf{z^i_{t+1|t-1}} = \mathbf{CA}\kappa(\mathbf{z^i_t} - \mathbf{z^i_{t|t-1}})$

Proposition 1

The general linear noisy information model implies:

1. $\beta_1 = 0$.
2. $\alpha_1 = \mathbf{CA}(\mathbf{I} - \mathbf{C}\kappa)(\kappa\mathbf{C})^{-1}(\mathbf{CA})^{-1} > 0$.

CG (2015) assume $A = \rho, B = \sigma_w, C = 1, D = \sigma_v$.

- $\beta_1 = 0$
- $\alpha_1 = \frac{1-\kappa}{\kappa}$, where κ is the Kalman gain

Error orthogonality (at the forecaster-level) always holds under linear rational expectations

Outline

Incorporating Nonlinearity: Volatility

Suppose that $\mathbf{B} \rightarrow \mathbf{B_t}$

- $\mathbf{s_t}$ exhibits time-varying volatility

Model becomes

$$\bar{\mathbf{s}}_\mathbf{t} = F(\bar{\mathbf{s}}_{\mathbf{t-1}}, \overline{\mathbf{w}}_\mathbf{t})$$
$$\mathbf{z_t^i} = \mathbf{C s_t} + \mathbf{D v_t^i}$$

- Forecasters do not observe $\mathbf{s_t}$
- In addition, variance of $\mathbf{s}_t$ is no longer constant
- State vector: $\bar{\mathbf{s}}_\mathbf{t} = [\mathbf{s_t} \quad diag(\mathbf{B_t})]^\top$
- $\overline{\mathbf{w}}_\mathbf{t}$ now includes $\mathbf{w}_t$ as well as shock to variance of $\bar{\mathbf{s}}_\mathbf{t}$

Error Predictability: Intuition

Kalman filter no longer delivers optimal prediction

- $\mathcal{Z}_t^i$ denotes history of signals observed by forecaster i through time t
- Recursive Bayes filter:

$$p(s_t|\mathcal{Z}_{t-1}^i) = \int p(s_t|s_{t-1})p(s_{t-1}|\mathcal{Z}_{t-1}^i)ds_{t-1} \quad \text{(Predict)}$$

$$p(s_t|\mathcal{Z}_t^i) = \frac{p(z_t^i|s_t)p(s_t|\mathcal{Z}_{t-1}^i)}{p(z_t|\mathcal{Z}_{t-1}^i)} \quad \text{(Update)}$$

- Outside of linear setting, $p(s_t|\mathcal{Z}_t^i)$ intractable
- **Note:** error predictability can arise with other nonlinearities
  - i.e. structural breaks

Costly Approximation

- Forecasters must approximate posterior → $\widehat{p}(s_t|\mathcal{Z}_t^i)$
- Choose approximation function $A \in \mathcal{A}$
- Each function subject to cost
  - $c_A^i \sim U[0, \overline{c}_A]$
  - Possible micro foundations: cognitive cost, computing time, etc.
- Assume $\overline{c}_A$ is increasing in the sophistication of A
  - Example: Particle filter more costly than naive approximation
- Loss function:

$$\mathcal{L} = \min_{A \in \mathcal{A}} \left\{ (\mathbf{z^i_{t+h}} - \widehat{\mathbf{z}}^{\mathbf{i}}_{\mathbf{t+h|t}})^\top (\mathbf{z^i_{t+h}} - \widehat{\mathbf{z}}^{\mathbf{i}}_{\mathbf{t+h|t}}) + c_A^i \right\}$$

Role of heterogeneous adoption costs

Approximate Predictions

After incurring cost and applying approximation, forecaster generates approximate prediction:

$$\hat{\bar{\mathbf{s}}}^{\mathbf{i}}_{\mathbf{t}|\mathbf{t}} = \int \bar{\mathbf{s}}_{\mathbf{t}} \, \widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^i_t) \, d\bar{\mathbf{s}}_{\mathbf{t}}$$

The approximate prediction can also be expressed as

$$\hat{\bar{\mathbf{s}}}^{i}_{t|t} = \underbrace{\mathbb{E}(\bar{\mathbf{s}}_t|\mathcal{Z}^i_t)}_{\text{Optimal}} + \underbrace{\int \bar{\mathbf{s}}_t \left[\widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^i_t) - p(\bar{\mathbf{s}}|\mathcal{Z}^i_t)\right] d\bar{\mathbf{s}}_t}_{\text{Approximation error}}$$

Signing β_1

From the nonlinear model subject to approximation errors, the covariance of errors and revision can be signed as follows:

$$\beta_1 \quad \propto \quad \underbrace{\mathbb{C}\left(\bar{\mathbf{s}}_\mathbf{t}, \int \bar{\mathbf{s}}_\mathbf{t}[\hat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t}) - \hat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t-1})]d\bar{\mathbf{s}}_\mathbf{t}\right)}_{\text{Under-reaction}}$$

$$- \underbrace{\mathbb{C}\left(\int \bar{\mathbf{s}}_\mathbf{t}\hat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t})d\bar{\mathbf{s}}_\mathbf{t}, \int \bar{\mathbf{s}}_\mathbf{t}[\hat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t}) - \hat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t-1})]\mathbf{d}\bar{\mathbf{s}}_\mathbf{t}\right)}_{\text{Over-reaction}}$$

The Relevance of Signal to Noise Ratio

High signal-to-noise ratio (SNR) $\implies$ revisions driven by movements in $\bar{\mathbf{s}}_\mathbf{t}$

$$\beta_1 \quad \propto \quad \underbrace{\mathbb{C}\left(\bar{\mathbf{s}}_\mathbf{t}, \int \bar{\mathbf{s}}_\mathbf{t}[\widehat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t}) - \widehat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t-1})]d\bar{\mathbf{s}}_\mathbf{t}\right)}_{\text{Under-reaction}}$$
$$- \underbrace{\mathbb{C}\left(\int \bar{\mathbf{s}}_\mathbf{t}\widehat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t})d\bar{\mathbf{s}}_\mathbf{t}, \int \bar{\mathbf{s}}_\mathbf{t}[\widehat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t}) - \widehat{p}(\bar{\mathbf{s}}_\mathbf{t}|\mathcal{Z}^\mathbf{i}_\mathbf{t-1})]\mathbf{d}\bar{\mathbf{s}}_\mathbf{t}\right)}_{\text{Over-reaction}}$$

The Relevance of Signal to Noise Ratio

Low SNR $\Longrightarrow$ revisions driven by noise in the system

$$\beta_1 \quad \propto \quad \underbrace{\mathbb{C}\left(\bar{\mathbf{s}}_{\mathbf{t}}, \int \bar{\mathbf{s}}_{\mathbf{t}}[\widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})]d\bar{\mathbf{s}}_{\mathbf{t}}\right)}_{\text{Under-reaction}}$$

$$- \underbrace{\mathbb{C}\left(\int \bar{\mathbf{s}}_{\mathbf{t}}\widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}})d\bar{\mathbf{s}}_{\mathbf{t}}, \int \bar{\mathbf{s}}_{\mathbf{t}}[\widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t}}) - \widehat{p}(\bar{\mathbf{s}}_{\mathbf{t}}|\mathcal{Z}^{\mathbf{i}}_{\mathbf{t-1}})]\mathbf{d}\bar{\mathbf{s}}_{\mathbf{t}}\right)}_{\text{Over-reaction}}$$

How Does Model Generate Over- and Underreactions?

Suppose that the state exhibits stochastic volatility according to

$$s_t = \rho s_{t-1} + \sigma_t w_t, \qquad w_t \sim \mathcal{N}(0,1)$$
$$\sigma_t = \chi\sigma_L + (1-\chi)\sigma_H$$

where $\chi = 1$ with probability q and $\chi = 0$ with probability $1-q$

- Forecasters receive private signal

  $$y_t^i = s_t + v_t^i$$

  each period, where $v_t^i \sim \mathcal{N}(0, \sigma_v^2)$
- Can only observe the sequence of private signals $\{y_k^i\}_{k=0}^t$.

Example (cont'd)

Suppose that forecasters approximate $\widehat{\sigma}_t = \frac{1}{2}(\sigma_L + \sigma_H)$

From Kalman filter:

$$\mathbb{C}(s_t - \widehat{s}^i_{t|t}, \widehat{s}^i_{t|t} - \widehat{s}^i_{t|t-1}) = \widehat{\kappa}\left[(1-\widehat{\kappa})\mathbb{V}(s_t - \widehat{s}^i_{t|t-1}) - \widehat{\kappa}\sigma_v^2\right]$$

where $\widehat{\kappa}$ is a distorted Kalman gain

- As q rises, state becomes less variable and SNR falls
  - Raises scope for over-reaction
  - Intuitively, forecasters believe the state to be more variable than it truly is, and therefore place a larger weight on new information
- As q falls, state becomes more variable and SNR rises
  - Underreactions arise

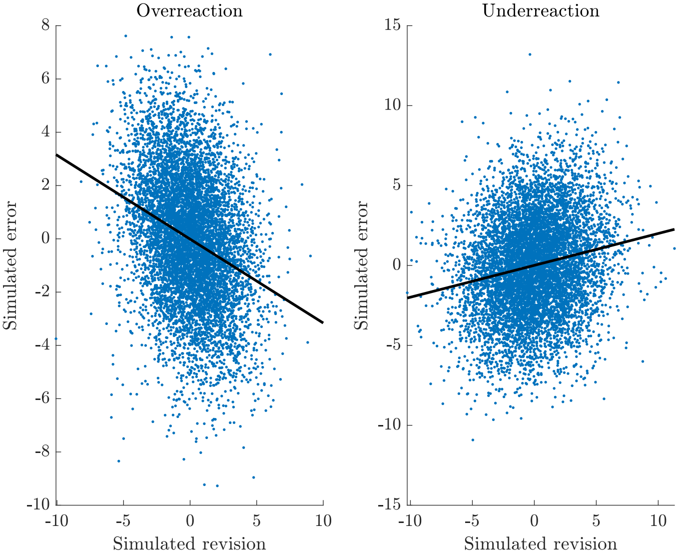


Figure 2: Simulated Examples

Proposition 2

The consensus errors on consensus revisions OLS coefficient is weakly greater than the individual errors on revisions coefficient:

$$\alpha_1 \geqslant \beta_1$$

Takeaway: model can explain simultaneous over- and under-reactions by same forecaster as well as Bordalo et. al. (2019) finding of over- and under- reactions across levels of aggregation

Alternate theories of expectation formation

Outline

1. Facts About Professional Forecasters
2. General Linear Noisy Information Model
3. Nonlinear Noisy Information Model with Costly Approximation
4. **Stylized Model: Stochastic Volatility**
5. Cross-Sectional Evidence from the Survey of Professional Forecasters

Stylized SV Model

State:

$$s_t = \rho s_{t-1} + e^{h_t/2} w_t, \qquad w_t \sim \mathcal{N}(0,1)$$
$$h_t = \phi_0 + \phi_1 h_{t-1} + \sigma_\eta \eta_t, \qquad \eta_t \sim \mathcal{N}(0,1)$$

Signals:

$$y_t^i = s_t + v_t^i \qquad v_t^i \sim \mathcal{N}(0, \sigma_v^2)$$
$$x_{t-1} = s_{t-1} + e_{t-1} \quad e_t \sim \mathcal{N}(0, \sigma_e^2)$$

- s_t exhibits stochastic volatility (nonlinear state space)
  - $\sigma_\eta = \sigma_e = 0 \implies$ CG (2015)
- Forecasters formulate expectations about s_t and report forecast for x_t
  - x_t is macroeconomic variable released with a lag

Stylized SV Model: Approximations

Forecasters know the underlying model and its parameters

- Assume that forecasters can make use of naive Kalman filter (KF) or particle filter (PF)
  - $\overline{c}_{PF} > \overline{c}_{KF}$
- Choose to use PF if and only if:

$$MSE_{PF} + c^{i}_{PF} \leqslant MSE_{KF} + c^{i}_{KF}$$

Details on PF

Simulation Result 1: $\beta_1 \neq 0$, $\alpha \neq 0$

Individual	Aggregate	SNR	β_1	α_1
Underreaction	Underreaction	1.43	0.12	0.46
Overreaction	Underreaction	0.43	-0.14	0.18

Table 4: Signal-to-Noise Ratio and Implied OLS Coefficients

Examples found in the survey data:

- First row: Unemployment
- Second row: Real GDP

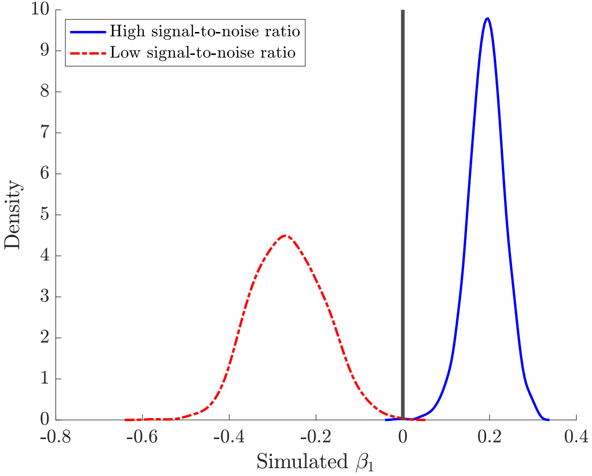
Covariance Between Errors and Revisions

$$\beta_1 \propto \underbrace{\mathbb{C}(s_t, \widehat{x}^i_{t|t} - \widehat{x}^i_{t|t-1})}_{\text{Under-reaction}} - \underbrace{\mathbb{C}(\widehat{x}^i_{t|t}, \widehat{x}^i_{t|t} - \widehat{x}^i_{t|t-1})}_{\text{Over-reaction}}$$

Covariance can take on either sign depending on the SNR

- When SNR is high, fluctuations in the underlying state drive forecast revisions
  - Result: underreactions
- When SNR is low, revisions are driven by the noise in the system
  - Upward revision in the reported forecast mechanically results in more negative forecast error
  - It is *as-if* forecasters report their predictions with measurement error
  - Result: overreaction

Simulation Result 2: Covariance of Errors and Revisions Depends on SNR



Outline

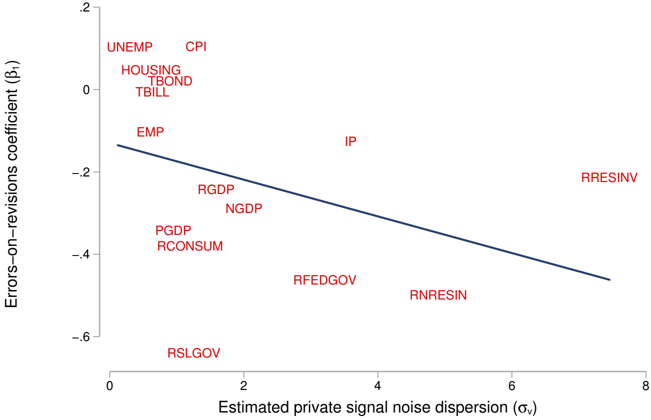
1. Facts About Professional Forecasters
2. General Linear Noisy Information Model
3. Nonlinear Noisy Information Model with Costly Approximation
4. Stylized Model: Stochastic Volatility
5. **Cross-Sectional Evidence from the Survey of Professional Forecasters**

Private Noise

- Model prediction: $\uparrow \sigma_v \implies \downarrow \beta_1$
- Cannot observe σ_v
- Proxy σ_v with dispersion in updates
  - Forecast revisions are a function of past forecast and news
  - News is obfuscated by noise
  - σ_v governs extent to which there is cross-sectional dispersion in revisions

Over-reaction and Private Noise

$$\uparrow \sigma_v \implies \downarrow \beta_1$$



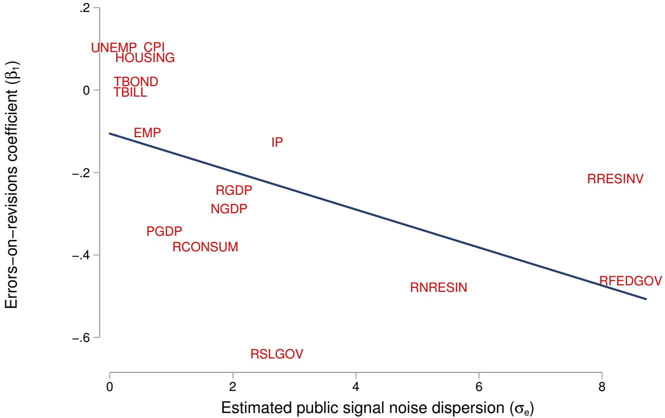
Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated private noise dispersion, proxied by the interquartile range of forecast revisions. Slope of fitted line is −0.045.

Public Noise

- Similarly, $\uparrow \sigma_e \implies \downarrow \beta_1$
- Again, σ_e unobservable
- Proxy with government revisions across vintages
  - SPF variables as public signals
  - Variables are revised ex-post
  - Extent of revision from initial to final release can inform "noisiness" of each of these public signals
  - Compute magnitude of revision across SPF variables

Over-reaction and Public Noise

$$\uparrow \sigma_e \implies \downarrow \beta_1$$



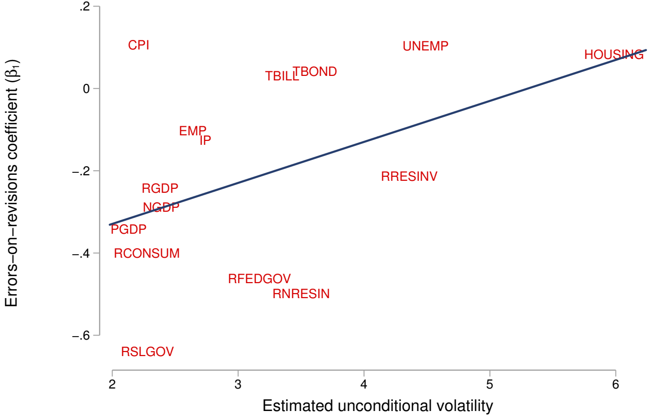
Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated public noise dispersion, proxied by the standard deviation of government revisions to real-time data. Slope of fitted line is -0.046.

State Variability

- More variable state $\implies \uparrow \beta_1$
- Estimate SV model for each SPF variable and and compute $e^{\frac{\phi_0}{1-\phi_1}}$
  - Estimation done via Bayesian methods as in Kastner and Fruhwirth-Schnatter (2014)
- Plot this against β_1
- Note that this also implies that β_1 is increasing in persistence of state
  - Consistent with Bordalo et. al. (2019)

Under-reaction and State Volatility

$\uparrow$ Variance $\implies \uparrow \beta_1$



Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against estimated unconditional volatility of the state, $\exp(\widehat{\phi}_0/2)$. Slope of fitted line is 0.100.

Release Frequency

- Forecasters report (quarterly) predictions about variables released at different frequencies
- Variables released at daily/monthly frequencies provide additional information relative to those released at quarterly frequency
- Variables released at higher frequencies $\implies \uparrow \beta_1$

Under-reaction and Release Frequency

$\uparrow$ Frequency $\Longrightarrow \uparrow \beta_1$

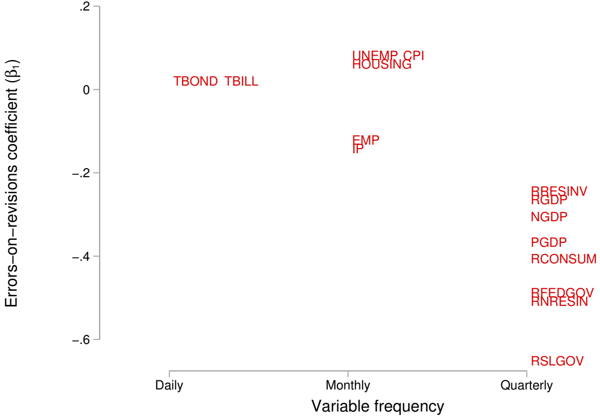


Figure 7: Error Predictability and Frequency

Note: For each SPF variable, the figure plots the estimated errors-on-revisions coefficient at the forecaster-level against variable's release frequency {Daily, Monthly, Quarterly}.

Conclusion

- Survey data inconsistent with **linear** RE models
- Can explain simultaneous over- and under-reactions by incorporating a relevant nonlinearity
  - Additional informational restriction: unobserved, time-varying state volatility
  - Modification to loss function: costly approximation functions
  - Law of iterated expectations breaks down
- This does not preclude non-rational behavior or alternative incentives
  - The right theory of expectation formation may depend on type economic agent in question
  - A model can incorporate nonlinearities as well as strategic interactions, misperceptions, over-confidence, etc.

Thank You!

Back-Up Slides

Constructing Over- and Under-reaction Shares

- Estimate $\widehat{\beta}_{1,ij}$ for each forecaster i forecasting variable j.
  - Delivers $N_i \times N_j$ matrix of estimates $\widehat{\beta}_{1,ij}$.
- Keep estimates that are significant at 5% level
- Fix a pair of variables j and k, and compute the number of forecasters such that $\widehat{\beta}_{1,ij} < 0$ and $\widehat{\beta}_{1,ik} > 0$
- Sum the number of forecasters for which this condition holds, and divide by the number of total forecasters reporting predictions about variables j and k.

More formally, I estimate a matrix P whose elements are p_{jk} with

$$p_{jk} = \frac{\sum_i \mathbb{1}\left(\widehat{\beta}_{1,ij} < 0 \text{ and } \widehat{\beta}_{1,ik} > 0\right)}{\min\{N_j, N_k\}}$$

where N_x denotes the number of forecasters providing predictions of variable x.

Back to Figure 1

Details on EKF

The EKF first linearizes the state around the most recent nowcast. The state is approximately

$$\bar{\mathbf{s}}_{\mathbf{t}} \approx F(\bar{\mathbf{s}}_{\mathbf{t-1}}, \mathbf{0}) + \mathbf{J}(\widehat{\bar{\mathbf{s}}}_{t-1|t-1})(\bar{\mathbf{s}}_{\mathbf{t-1}} - \bar{\mathbf{s}}^{\mathbf{i}}_{\mathbf{t-1|t-1}}) + h.o.t.$$

where J is the Jacobian of $F(\cdot)$ and $h.o.t.$ refers to higher order terms.

$$J = \begin{bmatrix} \frac{\partial F_1}{\partial \bar{\mathbf{s}}_1} & \frac{\partial F_1}{\partial \bar{\mathbf{s}}_2} \\ \frac{\partial F_2}{\partial \bar{\mathbf{s}}_1} & \frac{\partial F_2}{\partial \bar{\mathbf{s}}_2} \end{bmatrix} \Big|_{\bar{\mathbf{s}}=(\widehat{\bar{\mathbf{s}}}_{t-1|t-1})}$$

Given most recent nowcast, linearize state around this estimate. Then apply the prediction step from the Kalman filter to produce a one step-ahead forecast.

Upon observing the signals in the subsequent period, forecasters make use of the updating equations to generate a new nowcast.

Back to EKF

Details on PF

Boostrap filter of Gordon et. al (1993)

In principle, this approach makes use to mass points (particles) to construct $\widehat{p}(s_t|\mathcal{Z}_t^i)$.
Define set of particles and weights: $\chi = \{s^{(n)}, \omega^{(n)}\}_{n=1}^N$.
For each particle n, propagate estimate through system

$$s_t^{i,(n)} = F\left(s_{t-1}^{i(n)}, w_t\right)$$

Then update the weight: $\widetilde{\omega}_t^{i,(n)} = \omega_{t-1}^{i,(n)} \cdot p\left(z_t^i | s_t^{i,(n)}\right)$

Next normalize the weights $\omega_t^{i,(n)} = \frac{\widetilde{w}_t^{i,(n)}}{\sum_{n=1}^N \widetilde{\omega}_t^{i,(n)}}$

Nowcast: $\widehat{s}_{t|t}^i = \sum_{n=1}^N s_t^{i,(n)} \cdot \omega_t^{i,(n)}$.

Note: I assume that forecasters make use of the common sequential importance resampling scheme in which particles are resampled, each with a probability equal to its weight.

Back to PF

Role of Heterogeneous Adoption Costs

- Nonlinearity + costly approximation + homogeneous cost $\implies$ all forecasters select same A
- Assume c_A^i in order to generate heterogeneity in adoption of approximation functions
  - Heterogeneity in DGP alone is not enough to generate simultaneous over- and under-reactions
    - Forecasters would all over- and under-react to same variables
  - Heterogeneous costs alone not enough either
    - Would generate same shares of over- and under-reaction across variables
- **Note:** Could alternatively assume heterogeneous signal precision, or assume a fraction of forecasters use a given approximation function

Back to costly approx.

Alternate Theories of Expectation Formation

- Diagnostic expectations
  - Define over-reaction parameter, $\theta > 0$
  - $s^{i,\theta}_{t|t} = s^i_{t|t} + \theta(s^i_{t|t} - s^i_{t|t-1})$
  - Can only deliver individual over-reaction ($\beta_1 < 0$)
- Overconfidence
  - Suppose only signal is $z^i_t = s_t + v^i_t$ where $v^i_t \sim \mathcal{N}(0, \sigma^2_v)$
  - Overconfidence assumes that forecasters perceive $\widetilde{\sigma}^2_v < \sigma^2_v$
    - Believe own signal precision to be greater than it actually is
  - Can only deliver individual over-reaction ($\beta_1 < 0$)
  - Implies revision orthogonality (violated in the data)

Alternate Theories of Expectation Formation

- Strategic Interaction
  - Objective: minimize MSE but also consider how forecast relates to consensus forecast
  - Let γ be degree of strategic complementarity
    - $\gamma > 0 \implies \beta_1 > 0$
    - $\gamma < 0 \implies \beta_1 < 0$
  - Can explain both under-reaction and over-reaction, but doing so requires that forecasters place different values of γ on different variables
  - Also implies that revision orthogonality (violated in the data)
- Forecast Smoothing
  - Reputational considerations $\implies$ smoothing motive
  - Can only explain individual under-reactions ($\beta_1 > 0$)

Back to Proposition 2

Simultaneous Over- and Under-Reactions

State:

$$\begin{bmatrix} s_{1t} \\ s_{2t} \\ h_{1t} \\ h_{2t} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \phi_0 \\ \psi_0 \end{bmatrix} + \begin{bmatrix} \rho_1 \\ \rho_2 \\ \phi_1 \\ \psi_1 \end{bmatrix} \mathbf{I}_{4\times 4} \begin{bmatrix} s_{1,t-1} \\ s_{2,t-1} \\ h_{1,t-1} \\ h_{2,t-1} \end{bmatrix} + \begin{bmatrix} e^{h_{1t}/2} \\ e^{h_{2t}/2} \\ \sigma_{1\eta} \\ \sigma_{2\eta} \end{bmatrix} \mathbf{I}_{4\times 4} \begin{bmatrix} w_{1t} \\ w_{2t} \\ \eta_{1t} \\ \eta_{2t} \end{bmatrix}$$

Signals:

$$\begin{bmatrix} y^i_{1t} \\ y^i_{2t} \\ x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_{1t} \\ s_{2t} \\ h_{1t} \\ h_{2t} \end{bmatrix} + \begin{bmatrix} \sigma^i_{1v} \\ \sigma^i_{2v} \\ \sigma_{1e} \\ \sigma_{2e} \end{bmatrix} \mathbf{I}_{4\times 4} \begin{bmatrix} v^i_{1t} \\ v^i_{2t} \\ e_{1,t-1} \\ e_{2,t-1} \end{bmatrix}$$

Back to simulation results