Imperfect Macroeconomic Expectations: Evidence and Theory

George-Marios Angeletos, Zhen Huo, and Karthik A. Sastry
MIT and NBER, Yale, and MIT

University of Bocconi
June 1, 2020
State of The Art

Lots of lessons outside representative agent, rational expectations benchmark

But also a “wilderness” of alternatives

- Rational inattention, sticky info, etc. (Sims, Mankiw & Reis, Mackowiak & Wiederholt)
- Higher-order uncertainty (Morris & Shin, Woodford, Nimark, Angeletos & Lian)
- Level-K thinking (Garcia-Schmidt & Woodford, Farhi & Werning, Iovino & Sergeyev)
- Cognitive discounting (Gabaix)
- Over-extrapolation (Gennaioli, Ma & Shleifer, Fuster, Laibson & Mendel, Guo & Wachter)
- Over-confidence (Kohlhas & Broer, Scheinkman & Xiong)
- Representativeness (Bordalo, Gennaioli & Shleifer)
- Undue effect of historical experiences (Malmendier & Nagel)
- ...
Contributions:

- Use a parsimonious framework to organize existing evidence and various theories
- Provide new evidence
- Identify the “right” model of expectations for business cycle context

Main lessons:

- New fact: expectations under-react early but over-shoot later
- Best model: dispersed info + over-extrapolation
- Little support for FIRE, cognitive discounting, level-k thinking
This Paper

Contributions:

- Use a parsimonious framework to organize existing evidence and various theories
- Provide new evidence
- Identify the “right” model of expectations for business cycle context

Main lessons:

- New fact: expectations under-react early but over-shoot later
- Best model: dispersed info + over-extrapolation
- Little support for FIRE, cognitive discounting, level-k thinking
This Paper

Contributions:

• Use a parsimonious framework to organize existing evidence and various theories
• Provide new evidence
• Identify the “right” model of expectations for business cycle context

Main lessons:

• New fact: expectations under-react early but over-shoot later
• Best model: dispersed info + over-extrapolation
• Little support for FIRE, cognitive discounting, level-k thinking
This Paper

Contributions:

• Use a parsimonious framework to organize existing evidence and various theories
• Provide new evidence
• Identify the “right” model of expectations for business cycle context

Main lessons:

• New fact: expectations under-react early but over-shoot later
• Best model: dispersed info + over-extrapolation
• Little support for FIRE, cognitive discounting, level-k thinking
Outline

Three Existing Facts, with Conflicting Message

An “Umbrella Theory”

A New, Unifying Fact: Delayed Over-shooting in Aggregate Forecasts

Lessons for Theory

Going GE

Conclusion
Fact 1: Under-reaction in Aggregate Forecasts

Coibion and Gorodnichenko (2015)

\[(x_{t+k} - \bar{E}_t x_{t+k}) = a + K_{CG} \cdot (\bar{E}_t x_{t+k} - \bar{E}_{t-1} x_{t+k}) + u_t\]
Fact 1: **Under-reaction in Aggregate Forecasts**

Coibion and Gorodnichenko (2015)

\[
(x_{t+k} - \bar{E}_t x_{t+k}) = a + K_{CG} \cdot (\bar{E}_t x_{t+k} - \bar{E}_{t-1} x_{t+k}) + u_t
\]

<table>
<thead>
<tr>
<th>variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample</td>
<td>Unemployment</td>
<td>Inflation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revision(<em>t)(K(</em>{CG}))</td>
<td>0.741</td>
<td>0.809</td>
<td>1.528</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>(0.232)</td>
<td>(0.305)</td>
<td>(0.418)</td>
<td>(0.191)</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.111</td>
<td>0.159</td>
<td>0.278</td>
<td>0.016</td>
</tr>
<tr>
<td>Observations</td>
<td>191</td>
<td>136</td>
<td>190</td>
<td>135</td>
</tr>
</tbody>
</table>

Notes: The dataset is the Survey of Professional Forecasters and the observation is a quarter between Q4-1968 and Q4-2017. The forecast horizon is 3 quarters. Standard errors are HAC-robust, with a Bartlett (“hat”) kernel and lag length equal to 4 quarters. The data used for outcomes are first-release.
**Fact 1: Under-reaction in Aggregate Forecasts**

Coibion and Gorodnichenko (2015)

\[
(x_{t+k} - \bar{E}_t x_{t+k}) = a + K_{CG} \cdot (\bar{E}_t x_{t+k} - \bar{E}_{t-1} x_{t+k}) + u_t
\]

<table>
<thead>
<tr>
<th>variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample</td>
<td>Unemployment</td>
<td>Inflation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Revision$_t$(K$_{CG}$) | 0.741 (0.232) | 0.809 (0.305) | 1.528 (0.418) | 0.292 (0.191) |
| R$^2$            | 0.111 | 0.159 | 0.278 | 0.016 |
| Observations     | 191   | 136   | 190   | 135   |

*Notes: The dataset is the Survey of Professional Forecasters and the observation is a quarter between Q4-1968 and Q4-2017. The forecast horizon is 3 quarters. Standard errors are HAC-robust, with a Bartlett (“hat”) kernel and lag length equal to 4 quarters. The data used for outcomes are first-release.*

**Bad news for:** RE + *common* information

**Good news for:** (i) RE + *dispersed* noisy information

(ii) under-extrapolation, cognitive discounting, level-K
Fact 2: **Over-reaction in Individual Forecasts**

Bordalo, Gennaioli, Ma, and Shleifer (2018); Kohlhas and Broer (2018); Fuhrer (2018)

\[
(x_{t+k} - \mathbb{E}_{i,t} x_{t+k}) = a + K_{BGMS} \cdot (\mathbb{E}_{i,t} x_{t+k} - \mathbb{E}_{i,t-1} x_{t+k}) + u_t
\]
Fact 2: Over-reaction in Individual Forecasts
Bordalo, Gennaioli, Ma, and Shleifer (2018); Kohlhas and Broer (2018); Fuhrer (2018)

\[(x_{t+k} - \mathbb{E}_{i,t}x_{t+k}) = a + K_{BGMS} \cdot (\mathbb{E}_{i,t}x_{t+k} - \mathbb{E}_{i,t-1}x_{t+k}) + u_t\]

<table>
<thead>
<tr>
<th>variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment</td>
<td>0.321</td>
<td>0.398</td>
<td>0.143</td>
<td>-0.263</td>
</tr>
<tr>
<td>(0.107)</td>
<td>(0.149)</td>
<td>(0.123)</td>
<td>(0.054)</td>
<td></td>
</tr>
<tr>
<td>Inflation</td>
<td>0.028</td>
<td>0.052</td>
<td>0.005</td>
<td>0.025</td>
</tr>
<tr>
<td>(0.028)</td>
<td>(0.052)</td>
<td>(0.005)</td>
<td>(0.025)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5383</td>
<td>3769</td>
<td>5147</td>
<td>3643</td>
</tr>
</tbody>
</table>

Notes: The observation is a forecaster by quarter between Q4-1968 and Q4-2017. The forecast horizon is 3 quarters. Standard errors are clustered two-way by forecaster ID and time period. Both errors and revisions are winsorized over the sample to restrict to 4 times the inter-quartile range away from the median. The data used for outcomes are first-release.

BGMS argue that $K_{BGMS} < 0$ is more prevalent in other forecasts. If so, then:

**Bad news for:** under-extrapolation, cognitive discounting, and level-K thinking

**Good news for:** over-extrapolation and over-confidence (or “representativeness”)
Facts 1 + 2 ⇒ Dispersed Info

<table>
<thead>
<tr>
<th>variable sample</th>
<th>Unemployment</th>
<th>Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>K_{CG}</td>
<td>0.741</td>
<td>1.528</td>
</tr>
<tr>
<td>K_{BGMS}</td>
<td>0.321</td>
<td>0.143</td>
</tr>
</tbody>
</table>

Q: What does $K_{CG} > K_{BGMS}$ mean?

A: *My* forecast revision today predicts *your* forecast error tomorrow

Evidence of dispersed private information

combined regression
Fact 3: Over-reaction in Aggregate Forecasts
Kohlhas and Walther (2019)

\[
(x_{t+k} - \bar{E}_t x_{t+k}) = a + K_{KW} \cdot x_t + u_t
\]
**Fact 3: Over-reaction in Aggregate Forecasts**

Kohlhas and Walther (2019)

\[
(x_{t+k} - \hat{E}_t x_{t+k}) = a + K_{KW} \cdot x_t + u_t
\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t$ ($K_{KW}$)</td>
<td>$-0.061$</td>
<td>$-0.036$</td>
<td>$0.111$</td>
<td>$-0.068$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$(0.056)$</td>
<td>$(0.038)$</td>
<td>$(0.075)$</td>
<td>$(0.068)$</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.016</td>
<td>0.007</td>
<td>0.058</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>194</td>
<td>136</td>
<td>193</td>
<td>135</td>
<td></td>
</tr>
</tbody>
</table>

*Notes: The dataset is the Survey of Professional Forecasters and the observation is a quarter between Q4-1968 and Q4-2017. The forecast horizon is 3 quarters. Standard errors are HAC-robust, with a Bartlett (“hat”) kernel and lag length equal to 4 quarters. The data used for outcomes are first-release.*

Bad news for: noisy REE that generates sluggishness and inertia

Good news for: over-extrapolation
Fact 3: **Over-reaction in Aggregate Forecasts**

Kohlhas and Walther (2019)

\[
(x_{t+k} - \bar{x}_{t+k}) = a + K_{KW} \cdot x_t + u_t
\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>x_t (K_{KW})</td>
<td>-0.061 (0.056)</td>
<td>-0.036 (0.038)</td>
<td>0.111 (0.075)</td>
<td>-0.068 (0.068)</td>
</tr>
<tr>
<td>R^2</td>
<td>0.016</td>
<td>0.007</td>
<td>0.058</td>
<td>0.012</td>
</tr>
<tr>
<td>Observations</td>
<td>194</td>
<td>136</td>
<td>193</td>
<td>135</td>
</tr>
</tbody>
</table>

Notes: The dataset is the Survey of Professional Forecasters and the observation is a quarter between Q4-1968 and Q4-2017. The forecast horizon is 3 quarters. Standard errors are HAC-robust, with a Bartlett (“hat”) kernel and lag length equal to 4 quarters. The data used for outcomes are first-release.

Bad news for: noisy REE that generates sluggishness and inertia

**Good news for:** over-extrapolation

**But:** hard to reconcile with Fact 1
Outline

Three Existing Facts, with Conflicting Message

An “Umbrella Theory”

A New, Unifying Fact: Delayed Over-shooting in Aggregate Forecasts

Lessons for Theory

Going GE

Conclusion
An “Umbrella Theory”

Physical Environment

Noisy signal

\[ s_{i,t} = x_t + u_{i,t}/\sqrt{\tau} \]

Process for unemployment or inflation

\[ x_t = \rho x_{t-1} + \epsilon_t \]

Two non-rational Ingredients

Perception of signal

\[ s_{i,t} = x_t + u_{i,t}/\sqrt{\hat{\tau}} \]

Perception of process

\[ x_t = \hat{\rho} x_{t-1} + \eta_t \]
An “Umbrella Theory”

Physical Environment

Noisy signal

\[ s_{i,t} = x_t + u_{i,t}/\sqrt{\tau} \]

Process for unemployment or inflation

\[ x_t = \rho x_{t-1} + \epsilon_t \]

Two non-rational Ingredients

Perception of signal

\[ s_{i,t} = x_t + u_{i,t}/\sqrt{\hat{\tau}} \]

over- or under-confidence?
An “Umbrella Theory”

Physical Environment

Noisy signal

\[ s_{i,t} = x_t + u_{i,t}/\sqrt{\tau} \]

Process for unemployment or inflation

\[ x_t = \rho x_{t-1} + \epsilon_t \]

Two non-rational Ingredients

Perception of signal

\[ s_{i,t} = x_t + u_{i,t}/\sqrt{\hat{\tau}} \]

Perception of process

\[ x_t = \hat{\rho} x_{t-1} + \eta_t \]

later: \( \hat{\rho} < \rho \) in GE \( \approx \)

cognitive discounting,
level-K thinking

over- or under-confidence?

over- or under-extrapolation?
Facts 1-3 in the Model

**Proposition.** The theoretical counterparts of the regression coefficients are:

\[
K_{CG} = \kappa_1 \hat{\tau}^{-1} - \kappa_2 (\hat{\rho} - \rho) \quad \text{(Fact 1)}
\]

\[
K_{BGMS} = -\kappa_3 (\hat{\tau} - \tau) - \kappa_4 (\hat{\rho} - \rho) \quad \text{(Fact 2)}
\]

\[
K_{KW} = \kappa_5 \hat{\tau}^{-1} - \kappa_6 (\hat{\rho} - \rho) \quad \text{(Fact 3)}
\]

for some positive scalars \( \kappa_1, ..., \kappa_6 \) that depend on the deeper parameters.
Facts 1-3 in the Model

**Proposition.** The theoretical counterparts of the regression coefficients are:

\[
K_{CG} = \kappa_1 \hat{\tau}^{-1} - \kappa_2 (\hat{\rho} - \rho) \quad \text{(Fact 1)}
\]

\[
K_{BGMS} = -\kappa_3 (\hat{\tau} - \tau) - \kappa_4 (\hat{\rho} - \rho) \quad \text{(Fact 2)}
\]

\[
K_{KW} = \kappa_5 \hat{\tau}^{-1} - \kappa_6 (\hat{\rho} - \rho) \quad \text{(Fact 3)}
\]

for some positive scalars \(\kappa_1, ..., \kappa_6\) that depend on the deeper parameters.

Key lessons:

- **Moments of average forecasts depend on perceived, not actual, precision**
- **Actual level of noise matters only for moments of individual forecasts**
- **Fact 2 conflates over-confidence and over-extrapolation**
- **Facts 1 and 3 conflate noise and over-extrapolation (in different ways)**
Facts 1-3 in the Model

**Proposition.** The theoretical counterparts of the regression coefficients are:

\[
K_{CG} = \kappa_1 \hat{\tau}^{-1} - \kappa_2 (\hat{\rho} - \rho) \quad \text{(Fact 1)}
\]

\[
K_{BGMS} = -\kappa_3 (\hat{\tau} - \tau) - \kappa_4 (\hat{\rho} - \rho) \quad \text{(Fact 2)}
\]

\[
K_{KW} = \kappa_5 \hat{\tau}^{-1} - \kappa_6 (\hat{\rho} - \rho) \quad \text{(Fact 3)}
\]

for some positive scalars \(\kappa_1, ..., \kappa_6\) that depend on the deeper parameters.

Key lessons:

- Moments of *average* forecasts depend on *perceived*, not actual, precision
- **Actual level of noise matters only for moments of *individual* forecasts**
- Fact 2 conflates over-confidence and over-extrapolation
- Facts 1 and 3 conflate noise and over-extrapolation (in different ways)
Facts 1-3 in the Model

**Proposition.** The theoretical counterparts of the regression coefficients are:

\[
K_{CG} = \kappa_1 \hat{\tau}^{-1} - \kappa_2 (\hat{\rho} - \rho) \quad \text{(Fact 1)}
\]

\[
K_{BGMS} = -\kappa_3 (\hat{\tau} - \tau) - \kappa_4 (\hat{\rho} - \rho) \quad \text{(Fact 2)}
\]

\[
K_{KW} = \kappa_5 \hat{\tau}^{-1} - \kappa_6 (\hat{\rho} - \rho) \quad \text{(Fact 3)}
\]

for some positive scalars \( \kappa_1, \ldots, \kappa_6 \) that depend on the deeper parameters.

Key lessons:

- Moments of *average* forecasts depend on *perceived*, not actual, precision
- Actual level of noise matters only for moments of *individual* forecasts
- Fact 2 conflates over-confidence and over-extrapolation
- Facts 1 and 3 conflate noise and over-extrapolation (in different ways)
Facts 1-3 in the Model

**Proposition.** The theoretical counterparts of the regression coefficients are:

\[ K_{CG} = \kappa_1 \hat{\tau}^{-1} - \kappa_2 (\hat{\rho} - \rho) \]  
\[ K_{BGMS} = -\kappa_3 (\hat{\tau} - \tau) - \kappa_4 (\hat{\rho} - \rho) \]  
\[ K_{KW} = \kappa_5 \hat{\tau}^{-1} - \kappa_6 (\hat{\rho} - \rho) \]

for some positive scalars \( \kappa_1, ..., \kappa_6 \) that depend on the deeper parameters.

Key lessons:

- Moments of average forecasts depend on perceived, not actual, precision
- Actual level of noise matters only for moments of individual forecasts
- Fact 2 conflates over-confidence and over-extrapolation
- Facts 1 and 3 conflate noise and over-extrapolation (in different ways)
Facts 1-3 in the Model

**Proposition.** The theoretical counterparts of the regression coefficients are:

\[ K_{CG} = \kappa_1 \hat{\tau}^{-1} - \kappa_2 (\hat{\rho} - \rho) \]  
(Fact 1)

\[ K_{BGMS} = -\kappa_3 (\hat{\tau} - \tau) - \kappa_4 (\hat{\rho} - \rho) \]  
(Fact 2)

\[ K_{KW} = \kappa_5 \hat{\tau}^{-1} - \kappa_6 (\hat{\rho} - \rho) \]  
(Fact 3)

for some positive scalars \( \kappa_1, \ldots, \kappa_6 \) that depend on the deeper parameters.

Key lessons:

- Moments of *average* forecasts depend on *perceived*, not actual, precision
- Actual level of noise matters only for moments of *individual* forecasts
- Fact 2 conflates over-confidence and over-extrapolation
- Facts 1 and 3 conflate noise and over-extrapolation (in different ways)

Is there a better way to understand what’s going on both in the theory and in the data?
Proposition. Let $\{\zeta_k\}_{k=1}^{\infty}$ be the IRF of the average, one-step-ahead, forecast error.

(i) If $\hat{\rho} < \rho$, then $\zeta_k > 0 \ \forall k$.

(ii) If $\hat{\rho} > \rho$ and $\hat{\tau}$ large enough relative to $\hat{\rho} - \rho$, then $\zeta_k < 0 \ \forall k$.

(iii) If $\hat{\rho} > \rho$ and $\hat{\tau}$ small enough relative to $\hat{\rho} - \rho$, then $\zeta_k > 0 \ \forall k < k_{\text{IRF}}$ and $\zeta_k < 0$ for $\forall k > k_{\text{IRF}}$, for some $k_{\text{IRF}} \in (1, \infty)$.

That is, average forecasts under-react early and overshoot later if and only if there is both over-extrapolation and sufficiently slow learning.

Key idea:

- When shock hits: everything is noisy, forecasts under-react
- Many quarters after shock: noise is gone, tendency to over-extrapolate takes over
Visualizing the Theoretical Prediction

Without Over-Extrapolation, $\hat{\rho} = \rho$

Facts 1 and 3 ($K_{CG} > 0$ and $K_{KW} < 0$) consistent with noise and over-extrapolation and so is Fact 2 ($K_{BGMS} < 0$).
Visualizing the Theoretical Prediction

\[ K_{\text{CG}} \sim \text{Cov}(\text{errors}, \text{revisions}) \sim \text{IRF} \text{errors} \times \text{IRF} \text{revisions} \]

\[ K_{\text{KW}} \sim \text{Cov}(\text{errors}, \text{outcome}) \sim \text{IRF} \text{errors} \times \text{IRF} \text{outcome} \]

Facts 1 and 3 (\( K_{\text{CG}} > 0 \) and \( K_{\text{KW}} < 0 \)) consistent with noise and over-extrapolation and so is Fact 2 (\( K_{\text{BGMS}} < 0 \))
Visualizing the Theoretical Prediction

**Bonus:** regression coefficients deconstructed

\[ K_{CG} \sim \text{Cov}(\text{errors, revisions}) \sim IRF_{\text{errors}} \times IRF_{\text{revisions}} \]

\[ K_{KW} \sim \text{Cov}(\text{errors, outcome}) \sim IRF_{\text{errors}} \times IRF_{\text{outcome}} \]
Visualizing the Theoretical Prediction

Bonus: regression coefficients deconstructed

\[ K_{CG} \sim \text{Cov}(\text{errors}, \text{revisions}) \sim \text{IRF}_{\text{errors}} \times \text{IRF}_{\text{revisions}} \]

\[ K_{KW} \sim \text{Cov}(\text{errors}, \text{outcome}) \sim \text{IRF}_{\text{errors}} \times \text{IRF}_{\text{outcome}} \]

Facts 1 and 3 \((K_{CG} > 0 \text{ and } K_{KW} < 0)\) consistent with noise and over-extrapolation and so is Fact 2 \((K_{BGMS} < 0)\)
Outline

Three Existing Facts, with Conflicting Message

An “Umbrella Theory”

A New, Unifying Fact: Delayed Over-shooting in Aggregate Forecasts

Lessons for Theory

Going GE

Conclusion
Estimation Strategy

**Shocks:** usual suspects (e.g., Gali tech); or DSGE shocks (e.g., JPT inv); or “main business cycle shocks” (Angeletos, Collard & Dellas, 2020)

**Estimation method:** plain-vanilla linear projection; or big VARs; or ARMA-IV (novel approach)

**Moments of interest:**

\[
\left( \frac{\partial \text{ForecastError}_{t+k}}{\partial \text{BusinessCycleShock}_t} \right)_{k=0}^K = \text{Pattern of mistakes}
\]
Fact 4: Delayed Over-Shooting in Response to Main BC Shocks

Each "slice" compares 3-Q-ahead forecasts with outcome.
Fact 4: Delayed Over-Shooting in Response to Main BC Shocks

Slow recognition, big forecast errors

Shaded area = ± 1 SE
Fact 4: Delayed Over-Shooting in Response to Main BC Shocks

Delayed over-shooting, smaller but persistent forecast errors

Shaded area = ± 1 SE

Method
projection
ARMA-IV
Fact 4: Same Pattern with Other Identified Shocks

Gali (1999): Technology → Inflation

Fact 4: Same Pattern in Structural VARs

13-Variable Model: macro “usual suspects” + unemployment and inflation forecasts (SPF)

ACD, 2020 (max-share for BC)

Cholesky (one-step-ahead Error)
Corroborating Evidence: Over-extrapolation in the “Term Structure”

\[
\mathbb{E}_t[x_{t+k}] = \alpha_k + \beta_k^f \cdot \epsilon_t + \gamma' W_t + u_{t+k}
\]

\[
x_{t+k} = \alpha_k + \beta_k^o \cdot \epsilon_t + \gamma' W_t + u_{t+k}
\]

Expectation from \( t = 0 \)

Reality from \( t = 0 \)
Outline

Three Existing Facts, with Conflicting Message

An “Umbrella Theory”

A New, Unifying Fact: Delayed Over-shooting in Aggregate Forecasts

Lessons for Theory

Going GE

Conclusion
### Need to Combine Frictions to Explain Facts

<table>
<thead>
<tr>
<th></th>
<th>Models</th>
<th>Facts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Information</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noisy common info</td>
<td>No</td>
<td>No*</td>
</tr>
<tr>
<td>Noisy dispersed info</td>
<td>Yes</td>
<td>No*</td>
</tr>
<tr>
<td><strong>Confidence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over-confidence</td>
<td>No</td>
<td>Maybe</td>
</tr>
<tr>
<td>Representativeness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under-confidence</td>
<td>No</td>
<td>Maybe</td>
</tr>
<tr>
<td>Timidness</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Foresight</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over-extrapolation</td>
<td>No</td>
<td>Maybe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under-extrapolation</td>
<td>Yes</td>
<td>Maybe</td>
</tr>
<tr>
<td>or cognitive discounting or level-K</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Need to Combine Frictions to Explain Facts: A Winning Combination

<table>
<thead>
<tr>
<th>Information</th>
<th>Models</th>
<th>Facts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy common information</td>
<td>No</td>
<td>No*</td>
</tr>
<tr>
<td>Noisy dispersed information</td>
<td>Yes</td>
<td>No*</td>
</tr>
</tbody>
</table>

### Confidence

| Over-confidence or representativeness heuristic | No | Maybe | No | No |
| Under-confidence or “timidness”               | No | Maybe | No | No |

### Foresight

| Over-extrapolation                  | No | Maybe | Yes | Yes |
| Under-extrapolation or cognitive discounting or level-K | Yes | Maybe | No | No |
Outline

Three Existing Facts, with Conflicting Message

An “Umbrella Theory”

A New, Unifying Fact: Delayed Over-shooting in Aggregate Forecasts

Lessons for Theory

Going GE

Conclusion
Tractable NK Model with Imperfect Expectations

**Familiar Ingredients**

Euler equation/DIS

\[ c_t = \mathbb{E}_t^*[c_{t+1}] - \varsigma r_t + \epsilon_t \]

Market clearing

\[ c_t = y_t \]

Demand shock

\[ \xi_t \equiv -\varsigma r_t + \epsilon_t = \rho \xi_t + \epsilon_t \]

Prices fully rigid (relax later on)
Tractable NK Model with Imperfect Expectations

**Familiar Ingredients**

- Euler equation/DIS
  \[ c_t = E^*[c_{t+1}] - \varsigma r_t + \epsilon_t \]

- Market clearing
  \[ c_t = y_t \]

- Demand shock
  \[ \xi_t \equiv -\varsigma r_t + \epsilon_t = \rho \xi_t + \epsilon_t \]

**New Ingredients**: noise + irrationality

Prices fully rigid (relax later on)
Tractable NK Model with Imperfect Expectations

**Familiar Ingredients**

Euler equation/DIS

\[ c_t = \mathbb{E}_t^* [c_{t+1}] - \varsigma r_t + \epsilon_t \]

Market clearing

\[ c_t = y_t \]

Demand shock

\[ \xi_t \equiv -\varsigma r_t + \epsilon_t = \rho \xi_t + \epsilon_t \]

Prices fully rigid (relax later on)

**New Ingredients:** noise + irrationality

Noisy signal

\[ s_{i,t} = \xi_t + u_{i,t} / \sqrt{\tau} \]

Over- or under-confidence?

\[ \hat{\rho} < \rho \] in GE

≈ cognitive discounting, level-K
Tractable NK Model with Imperfect Expectations

**Familiar Ingredients**

- Euler equation/DIS
  \[ c_t = \mathbb{E}_t^*[c_{t+1}] - \varsigma r_t + \epsilon_t \]
- Market clearing
  \[ c_t = y_t \]
- Demand shock
  \[ \xi_t \equiv -\varsigma r_t + \epsilon_t = \rho \xi_t + \epsilon_t \]
- Prices fully rigid (relax later on)

**New Ingredients:** noise + irrationality

- Noisy signal
  \[ s_{i,t} = \xi_t + u_{i,t}/\sqrt{\tau} \]
- Perception of signal
  \[ s_{i,t} = \xi_t + u_{i,t}/\sqrt{\hat{\tau}} \]

over- or under-confidence?
Tractable NK Model with Imperfect Expectations

**Familiar Ingredients**

- **Euler equation/DIS**
  \[ c_t = E_t^*[c_{t+1}] - \varsigma r_t + \epsilon_t \]

- **Market clearing**
  \[ c_t = y_t \]

- **Demand shock**
  \[ \xi_t \equiv -\varsigma r_t + \epsilon_t = \rho \xi_t + \epsilon_t \]

- **Prices fully rigid (relax later on)**

**New Ingredients: noise + irrationality**

- **Noisy signal**
  \[ s_{i,t} = \xi_t + u_{i,t}/\sqrt{\tau} \]

- **Perception of signal**
  \[ s_{i,t} = \xi_t + u_{i,t}/\sqrt{\hat{\tau}} \]

- **Perception of demand process**
  \[ \xi_t = \hat{\rho} \xi_{t-1} + \epsilon_t \]

- **over- or under-confidence?**
- **over- or under-extrapolation?**
Tractable NK Model with Imperfect Expectations

**Familiar Ingredients**

Euler equation/DIS

\[ c_t = \mathbb{E}_t^*[c_{t+1}] - \varsigma r_t + \epsilon_t \]

Market clearing

\[ c_t = y_t \]

Demand shock

\[ \xi_t \equiv -\varsigma r_t + \epsilon_t = \rho \xi_t + \epsilon_t \]

Prices fully rigid (relax later on)

**New Ingredients**: noise + irrationality

Noisy signal

\[ s_{i,t} = \xi_t + u_{i,t}/\sqrt{\tau} \]

Perception of signal

\[ s_{i,t} = \xi_t + u_{i,t}/\sqrt{\hat{\tau}} \]

Perception of demand process

\[ \xi_t = \hat{\rho} \xi_{t-1} + \epsilon_t \]

over- or under-confidence?

over- or under-extrapolation?

\( \hat{\rho} < \rho \) in GE \( \approx \) cognitive discounting, level-K
**Proposition: Mapping to Forecast Data**

Closed-form expressions:

<table>
<thead>
<tr>
<th>Expression</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1. $K_{CG} = K_{CG}(\hat{\tau}, \rho, \hat{\rho}; \text{mpc})$</td>
<td></td>
</tr>
<tr>
<td>F2. $K_{BGMS} = K_{BGMS}(\tau, \hat{\tau}, \rho, \hat{\rho}; \text{mpc})$</td>
<td></td>
</tr>
<tr>
<td>F3. $K_{KW} = K_{KW}(\hat{\tau}, \rho, \hat{\rho}; \text{mpc})$</td>
<td></td>
</tr>
<tr>
<td>F4. $\left{ \frac{\partial \text{Error}<em>{t+k}}{\partial \eta_t} \right}</em>{k \geq 1} = F(\hat{\tau}, \rho, \hat{\rho}; \text{mpc})$</td>
<td></td>
</tr>
</tbody>
</table>

**Proposition: Equilibrium Outcomes**

As-if representative, rational agent with

$$c_t = -r_t + \omega_f \mathbb{E}_t^*[c_{t+1}] + \omega_b c_{t-1}$$

$$\begin{align*}
(\omega_f, \omega_b) &= \Omega(\hat{\tau}, \rho, \hat{\rho}, \text{mpc})
\end{align*}$$
Transparent Mapping between Data and Theory

**Proposition: Mapping to Forecast Data**

Closed-form expressions:

F1. \( K_{CG} = K_{CG}(\hat{\tau}, \rho, \hat{\rho}; \text{mpc}) \)

F2. \( K_{BGMS} = K_{BGMS}(\tau, \hat{\tau}, \rho, \hat{\rho}; \text{mpc}) \)

F3. \( K_{KW} = K_{KW}(\hat{\tau}, \rho, \hat{\rho}; \text{mpc}) \)

F4. \( \left\{ \frac{\partial \text{Error}}{\partial \eta_t} \right\}_{k \geq 1} = F(\hat{\tau}, \rho, \hat{\rho}; \text{mpc}) \)

**Proposition: Equilibrium Outcomes**

As-if representative, rational agent with

\[
\begin{align*}
   c_t &= -r_t + \omega_f T^*_t[c_{t+1}] + \omega_b c_{t-1} \\
   (\omega_f, \omega_b) &= \Omega(\hat{\tau}, \rho, \hat{\rho}, \text{mpc})
\end{align*}
\]

- **General equilibrium** matters through \( \text{mpc} = \) slope of Keynesian cross
- Key behavior pinned down by \( (\hat{\tau}, \rho, \hat{\rho}) \)
  - Moments of average forecasts are key; moments of individual forecasts (BGMS) less so
  - Our evidence helps pin down \( \omega_b, \omega_f \) and resulting dynamics
Transparent Mapping between Data and Theory

**Proposition: Mapping to Forecast Data**

Closed-form expressions:

F1. \[ K_{CG} = K_{CG}(\hat{\tau}, \rho, \hat{\rho}; mpc) \]
F2. \[ K_{BGMS} = K_{BGMS}(\tau, \hat{\tau}, \rho, \hat{\rho}; mpc) \]
F3. \[ K_{KW} = K_{KW}(\hat{\tau}, \rho, \hat{\rho}; mpc) \]
F4. \[ \left\{ \frac{\partial \text{Error}_{t+k}}{\partial \eta_t} \right\}_{k \geq 1} = F(\hat{\tau}, \rho, \hat{\rho}; mpc) \]

**Proposition: Equilibrium Outcomes**

As-if representative, rational agent with

\[ c_t = -r_t + \omega_f E_t^*[c_{t+1}] + \omega_b c_{t-1} \]

\[ (\omega_f, \omega_b) = \Omega(\hat{\tau}, \rho, \hat{\rho}, mpc) \]

- General equilibrium matters through mpc = slope of Keynesian cross
- **Key behavior** pinned down by \((\hat{\tau}, \rho, \hat{\rho})\)
  - Moments of average forecasts are key; moments of individual forecasts (BGMS) less so
  - Our evidence helps pin down \(\omega_b, \omega_f\) and resulting dynamics
New Keynesian Model Calibrated to Expectations Evidence

Full model: add NKPC (with imperfect expectations) and Taylor rule

Good fit for demand shock, mediocre for supply shock

Right qualitative ingredients but no abundance of free parameters
Counterfactuals: Interaction of Forces Matters

Perfect Expectations

- Output gap (minus)
- Forecast

Only Noise

- Output gap (minus)
- Forecast

Noise and Over-Extrapolation

- Output gap (minus)
- Forecast

Noise smooths and dampens IRF ("stickiness/inertia and myopia"). Over-extrapolation increases present value and amplifies initial response ("amplification and momentum").
Counterfactuals: Interaction of Forces Matters

Noise smooths and dampens IRF ("stickiness/inertia and myopia")
Counterfactuals: Interaction of Forces Matters

Noise smooths and dampens IRF ("stickiness/inertia and myopia")

Over-extrapolation increases present value and amplifies initial response ("amplification and momentum")
Outline

Three Existing Facts, with Conflicting Message

An “Umbrella Theory”

A New, Unifying Fact: Delayed Over-shooting in Aggregate Forecasts

Lessons for Theory

Going GE

Conclusion
Conclusion

Contributions:

- Developed a simple framework to organize diverse theories and evidence
- Found little support for certain theories (FIRE, cognitive discounting, level-K)
- Argued that the “right” model combines dispersed info and over-extrapolation
- Clarified which moments of forecasts are most relevant in the theory
- Illustrated GE implications
Conclusion

Contributions:

• Developed a simple framework to organize diverse theories and evidence
• Found little support for certain theories (FIRE, cognitive discounting, level-K)
• Argued that the “right” model combines dispersed info and over-extrapolation
• Clarified which moments of forecasts are most relevant in the theory
• Illustrated GE implications

Limitations/Future Work:

• Context: “regular business cycles” vs. crises or specific policy experiments
• Forecast data: ideally we would like expectations of firms and consumers, and for the objects that matter the most for their choices
**Facts 1 + 2: Showing Under-reaction and Dispersion**

\[
\text{Error}_{i,t,k} = a - K_{\text{noise}} \cdot (\text{Revision}_{i,t,k} - \text{Revision}_{t,k}) + K_{\text{agg}} \cdot \text{Revision}_{t,k} + u_{i,t,k}
\]

<table>
<thead>
<tr>
<th>variable sample</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.166</td>
<td>0.162</td>
<td>-0.346</td>
<td>-0.410</td>
<td></td>
</tr>
<tr>
<td>(0.043)</td>
<td>(0.053)</td>
<td>(0.042)</td>
<td>(0.041)</td>
<td></td>
</tr>
<tr>
<td>0.745</td>
<td>0.841</td>
<td>1.550</td>
<td>0.412</td>
<td></td>
</tr>
<tr>
<td>(0.173)</td>
<td>(0.210)</td>
<td>(0.278)</td>
<td>(0.180)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.103</td>
<td>0.152</td>
<td>0.211</td>
<td>0.072</td>
</tr>
<tr>
<td>Observations</td>
<td>5383</td>
<td>3769</td>
<td>5147</td>
<td>3643</td>
</tr>
</tbody>
</table>

**Notes:** The observation is a forecaster by quarter between Q4-1968 and Q4-2017. The forecast horizon is 3 quarters. Standard errors are clustered two-way by forecaster ID and time period. Both errors and revisions are winsorized over the sample to restrict to 4 times the inter-quartile range away from the median. The data used for outcomes are first-release.
Estimation Strategy

Overall goal: allow flexibility for dynamics to be “shock-specific”

ARMA-IV: two-stage-least-squares estimate of

\[ x_t = \alpha + \sum_{p=1}^{P} \gamma_p \cdot x_{t-p}^{IV} + \sum_{k=1}^{K} \beta_k \cdot \epsilon_{t-k} + u_t \]

\[ X_{t-1} = \eta + \mathcal{E}_{t-1}' \Theta + e_t \]

where \( X_{t-1} \equiv (x_{t-p})_{p=1}^{P}, \mathcal{E}_{t-1} \equiv (\epsilon_{t-K-j})_{j=1}^{J} \) and \( J \geq P \). Main specification: \( P = 3, J = 6 \).

Projection: OLS estimation at each horizon \( h \) of

\[ x_{t+h} = \alpha_h + \beta_h \cdot \epsilon_t + \gamma' W_t + u_{t+h} \]

where the controls \( W_t \) are \( x_{t-1} \) and \( \mathbb{E}_{t-k-1}[x_{t-1}] \).
Estimation Strategy

Figure 1: *

Forecast error estimation with projection method (grey) and ARMA-OLS(1,1) (green).
Variable List for SVAR

10 usual suspects: real GDP, real investment, real consumption, labor hours, the labor share, the Federal Funds Rate, labor productivity, and utilization-adjusted TFP

3 forecast variables: three-period-ahead unemployment forecast, three-period annual inflation forecast, one-period-ahead quarter-to-quarter inflation forecast
The Role of Noise and HOB

As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f \mathbb{E}_t^* [c_{t+1}] + \omega_b c_{t-1} \]
The Role of Noise and HOB

As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f E_t^*[c_{t+1}] + \omega_b c_{t-1} \]

Only Dispersed Info \( \Rightarrow \omega_f < 1 \quad \omega_b > 0 \)

- \( \omega_f < 1 \): captures noise plus myopia due to HOB (Angeletos & Lian, 2018)
  \( \sim \) resolution to forward guidance puzzle etc
- \( \omega_b > 0 \): captures learning, or momentum in beliefs
  \( \sim \) resembles habit or adjustment costs
- both distortions disciplined by moments of average forecasts (CG or ours)
- both distortions increase with MPC, or Keynesian multiplier (HANK connection)
The Role of Noise and HOB

As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f E_t^*[c_{t+1}] + \omega_b c_{t-1} \]

Only Dispersed Info \( \Rightarrow \) \( \omega_f < 1 \) \( \omega_b > 0 \)

- \( \omega_f < 1 \) : captures noise plus myopia due to HOB (Angeletos & Lian, 2018)
  \( \leadsto \) resolution to forward guidance puzzle etc
- \( \omega_b > 0 \) : captures learning, or momentum in beliefs
  \( \leadsto \) resembles habit or adjustment costs

- both distortions disciplined by moments of average forecasts (CG or ours)
- both distortions increase with MPC, or Keynesian multiplier (HANK connection)
The Role of Noise and HOB

As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f \mathbb{E}_t^*[c_{t+1}] + \omega_b c_{t-1} \]

Only Dispersed Info \( \Rightarrow \omega_f < 1 \quad \omega_b > 0 \)

- \( \omega_f < 1 \) : captures noise plus myopia due to HOB (Angeletos & Lian, 2018)
  \( \sim \) resolution to forward guidance puzzle etc
- \( \omega_b > 0 \) : captures learning, or momentum in beliefs
  \( \sim \) resembles habit or adjustment costs
- both distortions disciplined by moments of average forecasts (CG or ours)
- both distortions increase with MPC, or Keynesian multiplier (HANK connection)
The Role of Under/Over-Extrapolation

As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f E^t[c_{t+1}] + \omega_b c_{t-1} \]

Only Under-extrapolation \( \Rightarrow \) \( \omega_f < 1 \) \( \omega_b = 0 \)

- myopia but not habit/momentum
- consistent with CG but rejected by BGMS and our fact
- same applies for cognitive-discounting and level-K thinking
The Role of Under/Over-Extrapolation

As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f \mathbb{E}_t^* [c_{t+1}] + \omega_b c_{t-1} \]

Only Under-extrapolation \( \Rightarrow \omega_f < 1 \quad \omega_b = 0 \)

• myopia but not habit/momentum
• consistent with CG but rejected by BGMS and our fact
• same applies for cognitive-discounting and level-K thinking
The Role of Under/Over-Extrapolation

As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f E^*_t[c_{t+1}] + \omega_b c_{t-1} \]

Only Under-extrapolation \( \Rightarrow \) \( \omega_f < 1 \) \( \omega_b = 0 \)

- myopia but not habit/momentum
- consistent with CG but rejected by BGMS and our fact
- same applies for cognitive-discounting and level-K thinking

Only Over-extrapolation \( \Rightarrow \) \( \omega_f > 1 \) \( \omega_b = 0 \)

- hyperopia but not habit/momentum
- consistent with BGMS but rejected by CG and our fact
As-if Representation (builds on Angeletos & Huo, 2018):

\[ c_t = -r_t + \omega_f \mathbb{E}_t^*[c_{t+1}] + \omega_b c_{t-1} \]

Over-extrapolation plus enough noise \( \Rightarrow \) \( \omega_f < 1 \) \( \omega_b > 0 \)

- matches all facts about expectations
- quantitative bite disciplined by our evidence
### Model Parameters

**Table 1: Exogenously Set Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>Calvo prob</td>
<td>0.6</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Slope of NKPC</td>
<td>0.02</td>
</tr>
<tr>
<td>$\chi$</td>
<td>Discount factor</td>
<td>0.99</td>
</tr>
<tr>
<td>mpc</td>
<td>MPC</td>
<td>0.3</td>
</tr>
<tr>
<td>$\varsigma$</td>
<td>IES</td>
<td>1.0</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Monetary policy</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Table 2: Calibrated Parameters**

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\rho}$</th>
<th>$\rho$</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand shock</td>
<td>0.94</td>
<td>0.80</td>
<td>0.38</td>
</tr>
<tr>
<td>Supply shock</td>
<td>0.82</td>
<td>0.57</td>
<td>0.15</td>
</tr>
</tbody>
</table>