Expectations Formation with Fat-Tailed Processes: Evidence and Theory

Tim de Silva	Eugene Larsen-Hallock	Adam Rej	David Thesmar
Stanford GSB & SIEPR	CFM	CFM	MIT, NBER, CEPR

NBER SI: Forecasting and Empirical Methods

July 2025

MOTIVATION

- Large recent literature documenting **predictability** in expectation errors
  - Underreaction: lab + field, often short-term or consensus forecasts
  - Overreaction: lab + field, often longer-term individual forecasts

Motivation

- Large recent literature documenting **predictability** in expectation errors
  - Underreaction: lab + field, often short-term or consensus forecasts
  - Overreaction: lab + field, often longer-term individual forecasts

- Models to explain this **predictability** typically have two ingredients
  1. Simple DGP for forecasting variable (e.g. AR1 with normal shocks)
  2. Biases given known DGP (e.g. diagnostic expectations)

- Large recent literature documenting **predictability** in expectation errors
  - Underreaction: lab + field, often short-term or consensus forecasts
  - Overreaction: lab + field, often longer-term individual forecasts

- Models to explain this **predictability** typically have two ingredients
  1. Simple DGP for forecasting variable (e.g. AR1 with normal shocks)
  2. Biases given known DGP (e.g. diagnostic expectations)

- However, many variables have non-Gaussian DGPs with Pareto tails Gabaix 09
  - **Challenge**: hard to study beliefs because rational expectations become intractable

- Large recent literature documenting **predictability** in expectation errors
  - Underreaction: lab + field, often short-term or consensus forecasts
  - Overreaction: lab + field, often longer-term individual forecasts

- Models to explain this **predictability** typically have two ingredients
  1. Simple DGP for forecasting variable (e.g. AR1 with normal shocks)
  2. Biases given known DGP (e.g. diagnostic expectations)

- However, many variables have non-Gaussian DGPs with Pareto tails Gabaix 09
  - **Challenge**: hard to study beliefs because rational expectations become intractable

- **This paper**: study expectations formation in the presence of “**fat**” **tails**
  - **Takeaway**: helps match data + parsimonious model of under & overreaction

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES
    1. Forecast errors are non-linear in revisions: underreaction in bulk, overreaction in tails

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES
    1. Forecast errors are non-linear in revisions: underreaction in bulk, overreaction in tails
    2. Distribution of growth has non-Gaussian “fat” tails
    3. Expected future growth is non-linear in current growth

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES
   1. Forecast errors are non-linear in revisions: underreaction in bulk, overreaction in tails
   2. Distribution of growth has non-Gaussian "fat" tails
   3. Expected future growth is non-linear in current growth

2. Show that a model of beliefs with two ingredients can explain these **three facts**
   1. DGP = persistent component + non-Gaussian shock $\Rightarrow$ Facts #2 and #3
   2. Forecasters use optimal expectations (partially) ignoring fat tails $\Rightarrow$ Fact #1

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES
    1. Forecast errors are non-linear in revisions: underreaction in bulk, overreaction in tails
    2. Distribution of growth has non-Gaussian "fat" tails
    3. Expected future growth is non-linear in current growth
2. Show that a model of beliefs with two ingredients can explain these **three facts**
    1. DGP = persistent component + non-Gaussian shock $\Rightarrow$ Facts #2 and #3
    2. Forecasters use optimal expectations (partially) ignoring fat tails $\Rightarrow$ Fact #1
3. Provide additional evidence in support of our theory
    1. Estimate the model and show it quantitatively **three facts**

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES
    1. Forecast errors are non-linear in revisions: underreaction in bulk, overreaction in tails
    2. Distribution of growth has non-Gaussian "fat" tails
    3. Expected future growth is non-linear in current growth
2. Show that a model of beliefs with two ingredients can explain these **three facts**
    1. DGP = persistent component + non-Gaussian shock $\Rightarrow$ Facts #2 and #3
    2. Forecasters use optimal expectations (partially) ignoring fat tails $\Rightarrow$ Fact #1
3. Provide additional evidence in support of our theory
    1. Estimate the model and show it quantitatively **three facts**
    2. Adding fat tails to DGP in experiment $\Rightarrow$ non-linear error-revision relationship

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES
    1. Forecast errors are non-linear in revisions: underreaction in bulk, overreaction in tails
    2. Distribution of growth has non-Gaussian “fat” tails
    3. Expected future growth is non-linear in current growth
2. Show that a model of beliefs with two ingredients can explain these **three facts**
    1. DGP = persistent component + non-Gaussian shock $\Rightarrow$ Facts #2 and #3
    2. Forecasters use optimal expectations (partially) ignoring fat tails $\Rightarrow$ Fact #1
3. Provide additional evidence in support of our theory
    1. Estimate the model and show it quantitatively **three facts**
    2. Adding fat tails to DGP in experiment $\Rightarrow$ non-linear error-revision relationship
    3. Returns to momentum are positive in bulk, but exhibit mean-reversion in tails

What We Do

1. Document **three facts** using analysts forecasts of sales growth from IBES
    1. Forecast errors are non-linear in revisions: underreaction in bulk, overreaction in tails
    2. Distribution of growth has non-Gaussian “fat” tails
    3. Expected future growth is non-linear in current growth

2. Show that a model of beliefs with two ingredients can explain these **three facts**
    1. DGP = persistent component + non-Gaussian shock ⇒ Facts #2 and #3
    2. Forecasters use optimal expectations (partially) ignoring fat tails ⇒ Fact #1

3. Provide additional evidence in support of our theory
    1. Estimate the model and show it quantitatively **three facts**
    2. Adding fat tails to DGP in experiment ⇒ non-linear error-revision relationship
    3. Returns to momentum are positive in bulk, but exhibit mean-reversion in tails

⇒ Allowing for **fat tails** is helpful for understanding belief formation!

1. Empirical evidence on under and overreaction in expectations

1. Empirical evidence on under and overreaction in expectations
   - Underreaction: lab Benjamin 19, ST earnings Bouchaud et al. 19, revenues Ma et al. 2024, ST rates Wang 21, macro (consensus) Coibion-Gorodnichenko 15
   - Overreaction: lab Afrouzi et al. 23, LT earnings growth Bordalo et al. 19, LT rates Giglio-Kelly 18, d'Arienzo 20, macro (individual) Bordalo et al. 20
   - **Contributions**:
     1. Field: evidence of both within **same** forecasting variable + horizon
     2. Lab: non-linearity in overreaction depends on the **Pareto tail** of DGP

Related Literature

1. Empirical evidence on under and overreaction in expectations
2. Models of under **or** overreaction in **individual** expectations
   - Underreaction: sticky expectations Bouchaud et al. 19, behavioral inattention Gabaix 19
   - Overreaction: diagnostic expectations Bordalo et al. 19, availability Afrouzi et al. 23

Related Literature

1. Empirical evidence on under and overreaction in expectations
2. Models of under **or** overreaction in **individual** expectations
3. Models of under **and** overreaction in **individual** expectations
   - Experience effects/constant-gain learning Malmendier-Nagel 16, Nagel-Xu 19
   - Selective recall with similarity and interference Bordalo et al. 22, 23
   - Shrinkage towards average persistence or precision Wang 21, Augenblick et al. 24
   - Overreaction to category-specific features Kwon-Tang 25
   - Within vs. across-category comparisons Graeber et al. 24
   - **Contribution**: model with under + overreaction **within category/DGP**

Related Literature

1. Empirical evidence on under and overreaction in expectations
2. Models of under **or** overreaction in **individual** expectations
3. Models of under **and** overreaction in **individual** expectations
4. Models of expectations with **unknown**/misspecified DGPs
   - Natural expectations Fuster et al. 10, 11
   - Learning Kozlowski et al. 20, Singleton 21, Farmer et al. 24, Dew-Becker et al. 24
   - No restrictions on the DGP de Silva-Thesmar 24
   - **Our focus**: misspecified model of distribution in the tails (could come from learning)

Related Literature

1. Empirical evidence on under and overreaction in expectations
2. Models of under **or** overreaction in **individual** expectations
3. Models of under **and** overreaction in **individual** expectations
4. Models of expectations with **unknown**/misspecified DGPs
5. Statistical models with **non-Gaussian** dynamics
   - Pareto tails, especially in firm growth Gabaix 09, Stanley et al. 96, Moran et al. 24
   - Skewness + kurtosis in income Guvenen et al. 14, 21, Braxton et al. 25
   - **Contribution**: connect with models of belief formation in a tractable way

Outline

Outline

- **Sample**: 122K observations from 2000-2023 of US and foreign firms in IBES
- **Forecasting variable**:

$$g_{it} \equiv \log \text{sales}_{it} - \log \text{sales}_{it-1 \text{ year}}$$

  - Advantages relative to EPS: larger sample + stationary
  - g_{it} standardized by firm: accounts for heterogenous DGPs Wyatt-Bouchaud 03

- **Sample**: 122K observations from 2000-2023 of US and foreign firms in IBES

- **Forecasting variable**:

$$g_{it} \equiv \log \text{sales}_{it} - \log \text{sales}_{it-1 \text{ year}}$$

  - Advantages relative to EPS: larger sample + stationary
  - g_{it} standardized by firm: accounts for heterogenous DGPs Wyatt-Bouchaud 03

- **Forecasts**:

$$F_t g_{it+h} \equiv \log F_t \text{sales}_{it+h \text{ years}} - \log F_t \text{sales}_{it+(h-1) \text{ years}}$$

  - F_t = consensus analyst forecasts after year t FY-end announcement
  - $F_t g_{it+h}$ standardized using same firm-specific mean and SD as g_{it}
  - Ignores a Jensen's term, but results similar with % growth

Outline

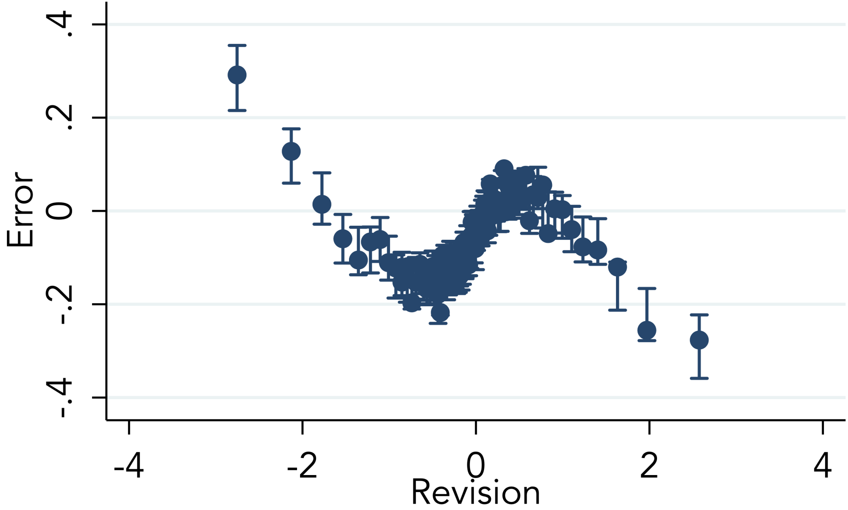
Coibion-Gorodnichenko Error-Revision Regressions

$$\underbrace{g_{it+1} - F_t g_{it+1}}_{\text{forecast error}} = \alpha + \beta \underbrace{\left(F_t g_{it+1} - F_{t-1} g_{it+1}\right)}_{\text{forecast revision}} + \epsilon_{it+1}$$

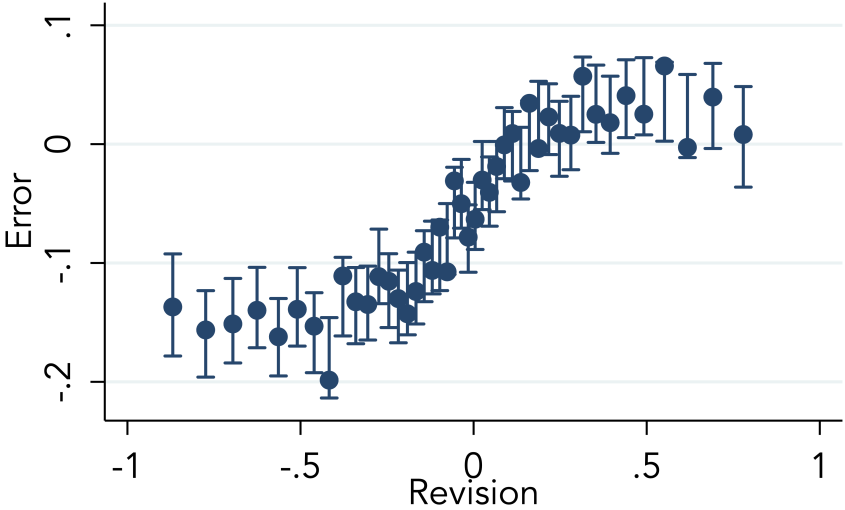
$$\underbrace{g_{it+1} - F_t g_{it+1}}_{\text{forecast error}} = \alpha + \beta \underbrace{\left(F_t g_{it+1} - F_{t-1} g_{it+1}\right)}_{\text{forecast revision}} + \epsilon_{it+1}$$

- $\beta \neq 0$ is inconsistent with rational expectations
  - Revisions are in forecasters' information set $\Rightarrow$ should not predict errors
- $\beta > 0 \Rightarrow$ revisions do not update "enough" $\Rightarrow$ **underreaction** Bouchaud et al. 19
- $\beta < 0 \Rightarrow$ revisions update "too much" $\Rightarrow$ **overreaction** Bordalo et al. 19
- Now a standard way of characterizing deviations from RE across datasets

Fact 1: Non-Linear Error-Revision Relationship

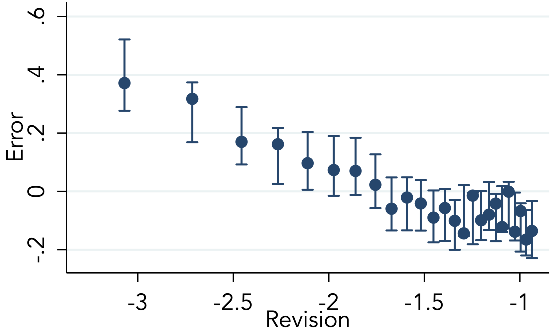


- Forecasts underreact **and** overreact within **same** variable and horizon



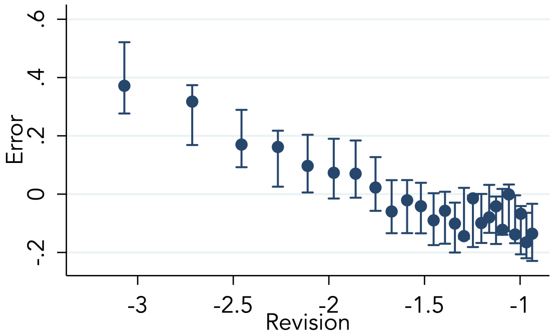
- Between **10-90%** of revisions, error-revision slope is **positive** Bouchaud et al. 19

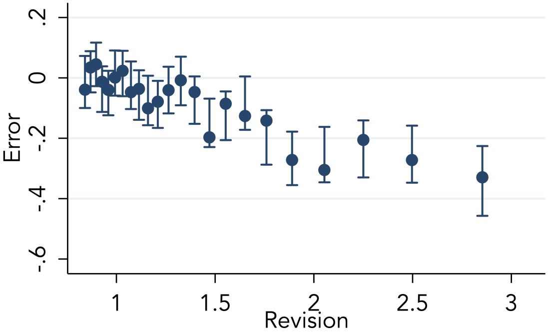
… But Overreaction in the Tails!



- Between **0-10%** of revisions, error-revision slope is **negative**

... But Overreaction in the Tails!





- Between **0-10%** and **90-100%** of revisions, error-revision slope is **negative**

Additional Results and Robustness

1. Not driven by within-firm adjustment: holds with **raw growth**
2. Does not reflect omitted Jensen's term: holds with **percent growth**
3. Does not reflect sample: similar for both **US and foreign** firms

Additional Results and Robustness

1. Not driven by within-firm adjustment: holds with **raw growth**
2. Does not reflect omitted Jensen's term: holds with **percent growth**
3. Does not reflect sample: similar for both **US and foreign** firms
4. Not specific to sales growth: present with **earnings-price** ratios

Additional Results and Robustness

1. Not driven by within-firm adjustment: holds with **raw growth**
2. Does not reflect omitted Jensen's term: holds with **percent growth**
3. Does not reflect sample: similar for both **US and foreign** firms
4. Not specific to sales growth: present with **earnings-price** ratios
5. Not driven by aggregation: present in **individual** forecasts

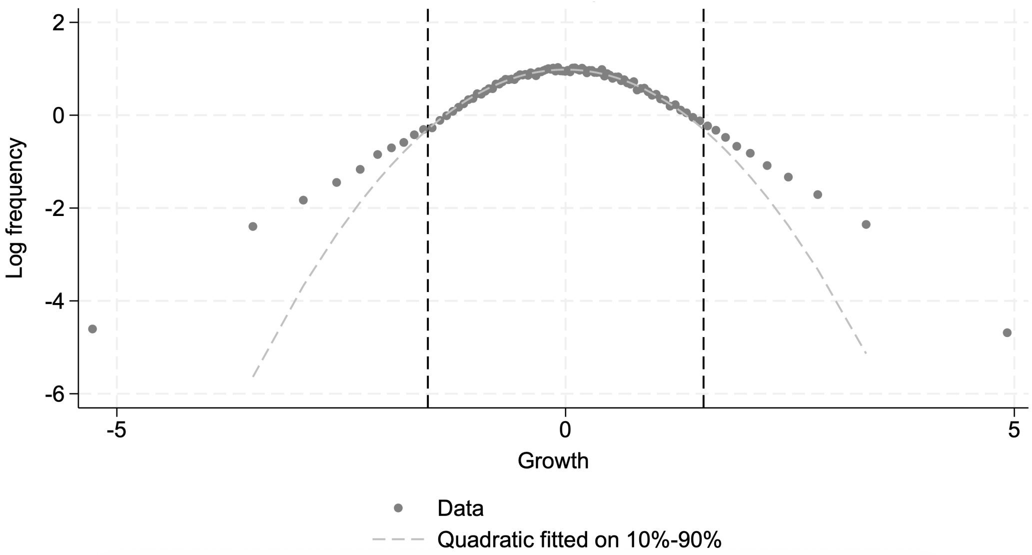
1. Not driven by within-firm adjustment: holds with **raw growth**
2. Does not reflect omitted Jensen's term: holds with **percent growth**
3. Does not reflect sample: similar for both **US and foreign** firms
4. Not specific to sales growth: present with **earnings-price** ratios
5. Not driven by aggregation: present in **individual** forecasts
6. Driven by cross-section not time-varying **aggregate** mean or volatility

Additional Results and Robustness

1. Not driven by within-firm adjustment: holds with **raw growth**
2. Does not reflect omitted Jensen's term: holds with **percent growth**
3. Does not reflect sample: similar for both **US and foreign** firms
4. Not specific to sales growth: present with **earnings-price** ratios
5. Not driven by aggregation: present in **individual** forecasts
6. Driven by cross-section not time-varying **aggregate** mean or volatility
7. Limited evidence of learning: does not vary with **analyst experience**

Outline

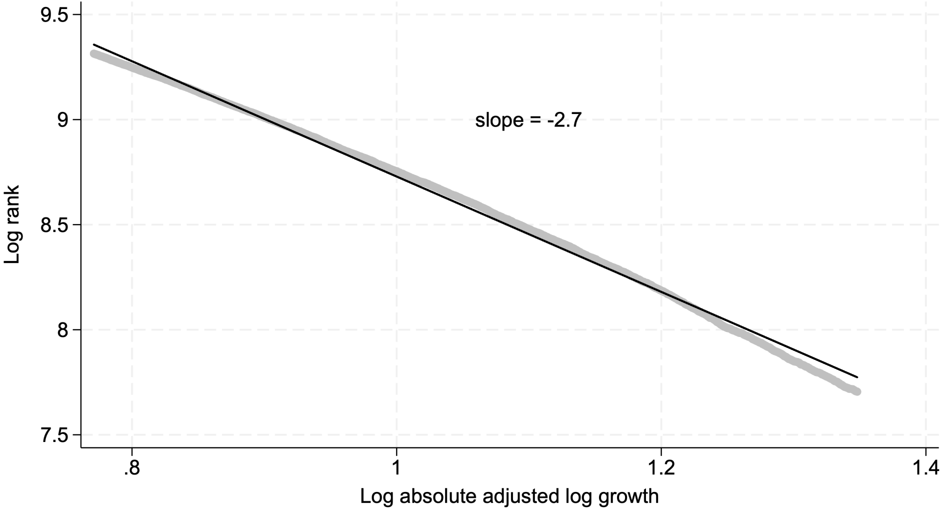
Tails of g_{it} are Fatter than Gaussian



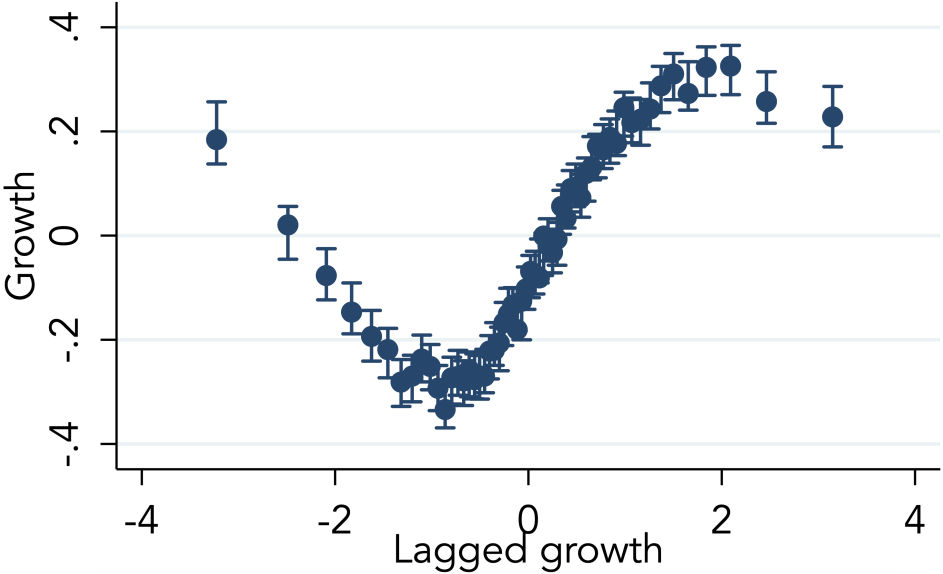
Tail Behavior in Top Deciles is Approximately a Power Law

$$\text{Power Law} : \log P(|g_{it}| > x) = -\nu \log x + \text{constant}$$

Power Law : $\log P(|g_{it}| > x) \approx -\mathbf{2.7} \log x + \text{constant}$



Outline


Growth
.4
.2
0
-.2
-.4
-4
-2
0
2
4
Lagged growth

Outline

Data-Generating Process

- DGP for sales growth (dropping *i* subscripts):

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0,1)$$

- g_t is a combination of persistent & transitory processes Bansal-Yaron 04, Lettau-Wachter 07
  - g^*_t = **unobservable** persistent latent state
  - ϵ_t = transitory shock with **Pareto tail**: $f(\epsilon) \propto \epsilon^{-\nu}$ as $|\epsilon| \longrightarrow \infty$, where $\nu > 2$

DATA-GENERATING PROCESS

- DGP for sales growth (dropping *i* subscripts):

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0,1)$$

- g_t is a combination of persistent & transitory processes Bansal-Yaron 04, Lettau-Wachter 07
  - g^*_t = **unobservable** persistent latent state
  - ϵ_t = transitory shock with **Pareto tail**: $f(\epsilon) \propto \epsilon^{-\nu}$ as $|\epsilon| \longrightarrow \infty$, where $\nu > 2$
- Remarks:
  - If ϵ_t was Gaussian, rational expectation would be the Kalman filter
  - **Key**: tail parameter in u_t larger than ϵ_t, otherwise inconsistent with Fact 3

DGP Replicates Facts 2 and 3

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$

$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

DGP Replicates Facts 2 and 3

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

- Replicates Fact 2: ϵ_t has Pareto tail ν and u_t is normal $\Rightarrow g_t$ has Pareto tail

DGP Replicates Facts 2 and 3

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

- Replicates Fact 2: ϵ_t has Pareto tail ν and u_t is normal $\Rightarrow$ g_t has Pareto tail
- What about Fact 3: $E(g_{t+1}|g_t)$ is non-linear in g_t?
    - Challenge: no closed-form expression $E(g_{t+1}|g_t)$ unless ϵ_t Gaussian

DGP Replicates Facts 2 and 3

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

- Replicates Fact 2: ϵ_t has Pareto tail ν and u_t is normal $\Rightarrow g_t$ has Pareto tail
- What about Fact 3: $E(g_{t+1}|g_t)$ is non-linear in g_t?
  - Challenge: no closed-form expression $E(g_{t+1}|g_t)$ unless ϵ_t Gaussian
- **Result** (Tweedie's formula): If $h(\cdot)$ is the observable marginal PDF of g_t, then

$$E(g_{t+1}|g_t) = -\rho\sigma^2_{g^*} \cdot \frac{d}{dg} \log h(g_t)$$

DGP Replicates Facts 2 and 3

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

- Replicates Fact 2: ϵ_t has Pareto tail ν and u_t is normal $\Rightarrow$ g_t has Pareto tail
- What about Fact 3: $E(g_{t+1}|g_t)$ is non-linear in g_t?
  - Challenge: no closed-form expression $E(g_{t+1}|g_t)$ unless ϵ_t Gaussian
- **Result** (Tweedie's formula): If $h(\cdot)$ is the observable marginal PDF of g_t, then

$$E(g_{t+1}|g_t) = -\rho\sigma^2_{g^*} \cdot \frac{d}{dg} \log h(g_t)$$

  - In bulk of distribution, $g_t \approx$ Gaussian $\Rightarrow$ $\log h(g_t) \approx -\frac{g_t^2}{2\sigma_g^2} + \text{constant}$

DGP Replicates Facts 2 and 3

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

- Replicates Fact 2: ϵ_t has Pareto tail ν and u_t is normal $\Rightarrow$ g_t has Pareto tail
- What about Fact 3: $E(g_{t+1}|g_t)$ is non-linear in g_t?
  - Challenge: no closed-form expression $E(g_{t+1}|g_t)$ unless ϵ_t Gaussian
- **Result** (Tweedie's formula): If $h(\cdot)$ is the observable marginal PDF of g_t, then

$$E(g_{t+1}|g_t) = \rho\sigma^2_{g^*} \cdot \frac{g_t}{\sigma^2_g} \qquad \nearrow g_t$$

  - In bulk of distribution, $g_t \approx$ Gaussian $\Rightarrow$ $\log h(g_t) \approx -\frac{g_t^2}{2\sigma_g^2} +$ constant
  - **Intution**: moderate values of g_t likely reflect g^*_t $\Rightarrow$ likely persistent

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

- Replicates Fact 2: ϵ_t has Pareto tail ν and u_t is normal $\Rightarrow g_t$ has Pareto tail
- What about Fact 3: $E(g_{t+1}|g_t)$ is non-linear in g_t?
  - Challenge: no closed-form expression $E(g_{t+1}|g_t)$ unless ϵ_t Gaussian
- **Result** (Tweedie's formula): If $h(\cdot)$ is the observable marginal PDF of g_t, then

$$E(g_{t+1}|g_t) = -\rho\sigma^2_{g^*} \cdot \frac{d}{dg} \log h(g_t)$$

  - In tails of distribution, $g_t \approx$ Pareto $\Rightarrow \log h(g_t) \approx -\nu \log(g_t)$

DGP Replicates Facts 2 and 3

$$g_{t+1} = g^*_{t+1} + \sigma_\epsilon \epsilon_{t+1} \quad \epsilon_t \sim f(\cdot) \quad var(\epsilon_t) = 1$$
$$g^*_{t+1} = \rho g^*_t + \sigma_u u_{t+1} \quad u_t \sim N(0, \sigma_u^2)$$

- Replicates Fact 2: ϵ_t has Pareto tail ν and u_t is normal $\Rightarrow g_t$ has Pareto tail
- What about Fact 3: $E(g_{t+1}|g_t)$ is non-linear in g_t?
  - Challenge: no closed-form expression $E(g_{t+1}|g_t)$ unless ϵ_t Gaussian
- **Result** (Tweedie's formula): If $h(\cdot)$ is the observable marginal PDF of g_t, then

$$E(g_{t+1}|g_t) = \rho\sigma^2_{g^*} \cdot \frac{\nu}{g_t} \qquad \searrow g_t$$

  - In tails of distribution, $g_t \approx$ Pareto $\Rightarrow \log h(g_t) \approx -\nu \log(g_t)$
  - **Intuition**: extreme values of g_t likely reflect $\epsilon_t \Rightarrow$ likely transitory

One Departure from Rational Expectations Replicates Fact 1

- **Assumption**: agents think $\epsilon_t \sim N(0, 1)$ but otherwise rational $\Rightarrow$ **Kalman filter**

One Departure from Rational Expectations Replicates Fact 1

- **Assumption**: agents think $\epsilon_t \sim N(0,1)$ but otherwise rational $\Rightarrow$ **Kalman filter**

- **Result**: In steady-state,

$$E(\text{error}_{t+1} \mid \text{revision}_t)$$

- **Assumption**: agents think $\epsilon_t \sim N(0,1)$ but otherwise rational $\Rightarrow$ **Kalman filter**
- **Result**: In steady-state,

$$E(\text{error}_{t+1} \mid \text{revision}_t) \xrightarrow{|\text{revision}_t| \to \infty}$$

One Departure from Rational Expectations Replicates Fact 1

- **Assumption**: agents think $\epsilon_t \sim N(0,1)$ but otherwise rational $\Rightarrow$ **Kalman filter**
- **Result**: In steady-state,

$$E(\text{error}_{t+1} \mid \text{revision}_t) \xrightarrow{|\text{revision}_t| \to \infty} -C \times \text{revision}_t$$

- **Assumption**: agents think $\epsilon_t \sim N(0,1)$ but otherwise rational $\Rightarrow$ **Kalman filter**
- **Result**: In steady-state,

$$E(\text{error}_{t+1} \mid \text{revision}_t) \xrightarrow{|\text{revision}_t| \to \infty} -C \times \text{revision}_t < 0$$

  - Large revision$_t$ reflects ϵ_t or ϵ_{t-h}, but forecasters overreact ignoring its **transitory**

One Departure from Rational Expectations Replicates Fact 1

- **Assumption**: agents think $\epsilon_t \sim N(0,1)$ but otherwise rational $\Rightarrow$ **Kalman filter**
- **Result**: In steady-state,

$$E(\text{error}_{t+1} \mid \text{revision}_t) \xrightarrow{|\text{revision}_t| \to \infty} -C \times \text{revision}_t < 0$$

  - Large revision_t reflects ϵ_t or ϵ_{t-h}, but forecasters overreact ignoring its **transitory**

- **Result**: In steady-state, there exists a $R > 0$ where:

$$E\left(\text{error}_{t+1} \times \text{revision}_t \middle| |\text{revision}_t| < R\right) > 0$$

  - Overreaction in tails + unbiased on average $\Rightarrow$ underreaction in bulk $\Rightarrow$ Fact 1

Connection to Insensitivity to Signal Strength

- Evidence of insensitivity to **signal strength** in lab inference problems
  - Overreaction to weak $+$ underreaction to strong signals Augenblick et al. 24, Ba et al. 24
  - Challenge: how to generalize this behavior to a time-series setting?

- Evidence of insensitivity to **signal strength** in lab inference problems
  - Overreaction to weak + underreaction to strong signals Augenblick et al. 24, Ba et al. 24
  - Challenge: how to generalize this behavior to a time-series setting?

- **Result**: Given our DGP,

$$\text{var}(g_{t+1}|g_t) = \sigma_g^2 + \rho^2\sigma_{g^*}^4 \frac{d^2}{dg^2} \log h(g_t)$$

- Evidence of insensitivity to **signal strength** in lab inference problems
  - Overreaction to weak + underreaction to strong signals Augenblick et al. 24, Ba et al. 24
  - Challenge: how to generalize this behavior to a time-series setting?
- **Result**: Given our DGP,

$$\text{var}(g_{t+1}|g_t) = \sigma_g^2 + \rho^2\sigma_{g^*}^4 \frac{d^2}{dg^2} \log h(g_t)$$

- Conditional variance is lower/higher when log density is **concave**/**convex**
  - Bulk: log density is concave $\Rightarrow$ strong signal, and have underreaction
  - Tails: log density is convex $\Rightarrow$ weak signal, and have overreaction

Connection to Insensitivity to Signal Strength

- Evidence of insensitivity to **signal strength** in lab inference problems
  - Overreaction to weak + underreaction to strong signals Augenblick et al. 24, Ba et al. 24
  - Challenge: how to generalize this behavior to a time-series setting?

- **Result**: Given our DGP,

$$\text{var}(g_{t+1}|g_t) = \sigma_g^2 + \rho^2 \sigma_{g^*}^4 \frac{d^2}{dg^2} \log h(g_t)$$

- Conditional variance is lower/higher when log density is **concave/convex**
  - Bulk: log density is concave $\Rightarrow$ strong signal, and have underreaction
  - Tails: log density is convex $\Rightarrow$ weak signal, and have overreaction

- Ignoring fat tails looks like ignoring signal strength in time-series setting!

Outline

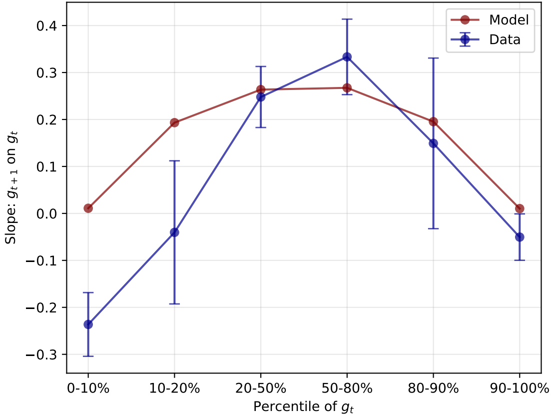
Outline

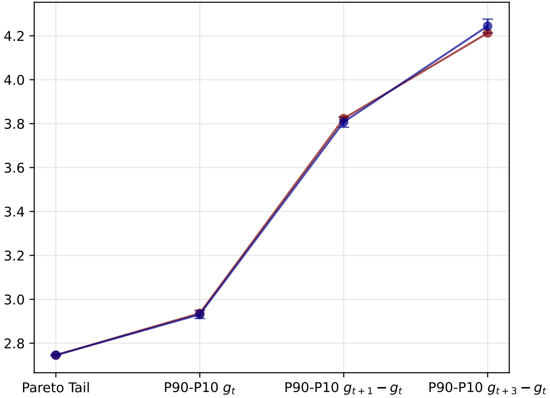
Model Fit: DGP

- Estimate DGP parameters using SMM by matching Facts 2 and 3
  - Assume $\epsilon \sim t$-distribution with ν degrees of freedom

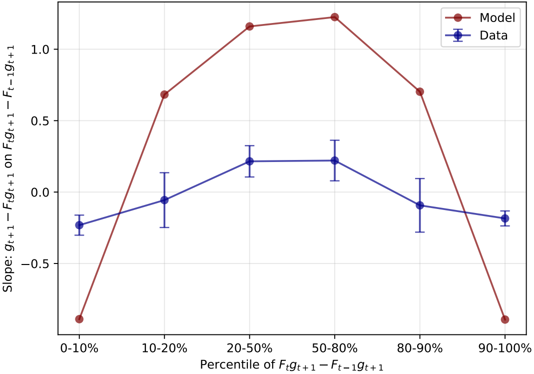
- Estimate DGP parameters using SMM by matching Facts 2 and 3
  - Assume $\epsilon \sim t$-distribution with ν degrees of freedom
- Parameter estimates: $\rho = 0.53$, $\nu = 2.53$, $\sigma_u = 0.63$, $\sigma_\epsilon = 1.33$

Standard Errors

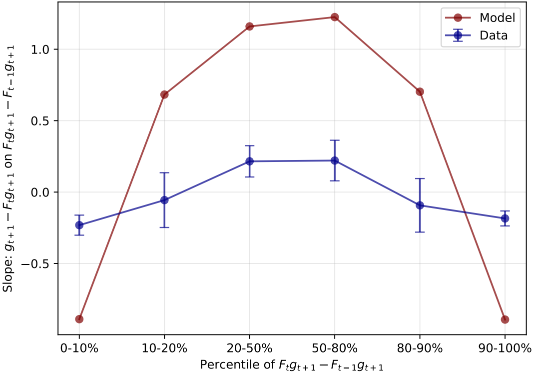




Model Fit: Beliefs

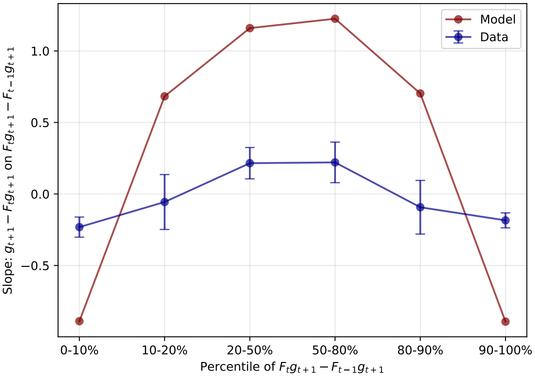


Model Fit: Beliefs



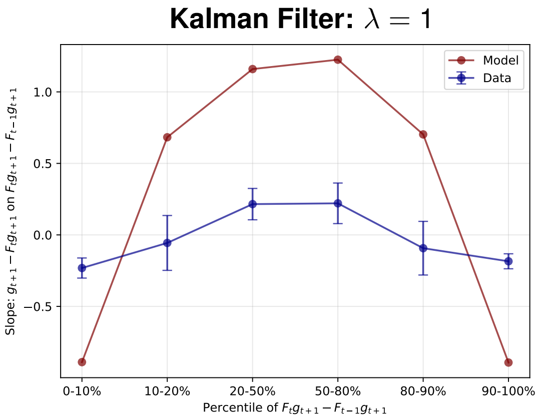
- Given DGP, model generates Fact 1 **qualitatively**, but not **quantitatively**

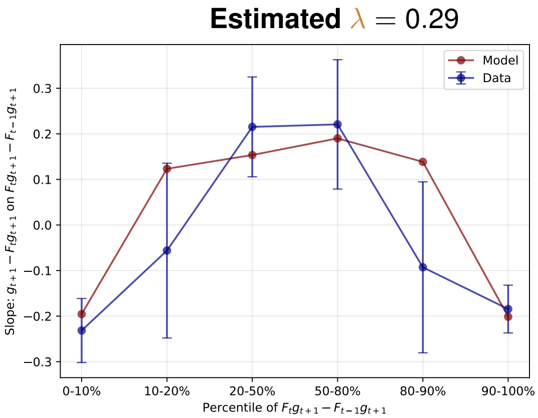
Model Fit: Beliefs



- Allow anchoring to RE: $F_t^{\lambda} g_{t+h} = \lambda F_t g_{t+h} + \underbrace{(1-\lambda) E_t g_{t+h}}_{\text{use particle filter}}$

Model Fit: Beliefs





- Allow anchoring to RE: $F_t^{\lambda} g_{t+h} = \lambda F_t g_{t+h} + (1 - \lambda) E_t g_{t+h}$
- Shrinkage between “default” and RE á la bounded rationality Fuster et al. 10, Gabaix 19

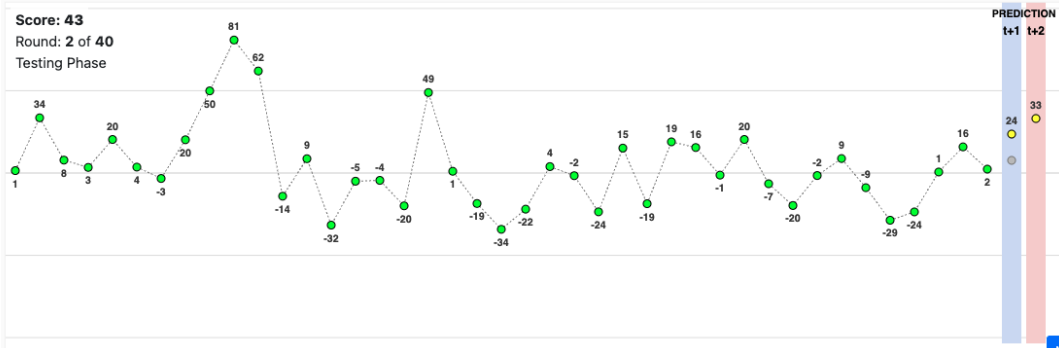
IS THE KALMAN FILTER A REASONABLE DEFAULT?

1. It is **simple**, like default models in Fuster et al. 10 and Gabaix 19
2. **Accuracy loss** relative to RE is small: **1.2%** reduction in MSE
   - With $\lambda = 0.29$, reduction in MSE is only **0.1%**
3. Kalman filter outperforms RE in small samples with **learning**
   - Need $T = 100$ for deviations from Kalman filter to outperform out-of-sample

Outline

Experimental Design

- Design follows Afrouzi et al. 23: participants make one and two-period forecasts
- 403 MTurk participants make 40 forecasts $\Rightarrow$ 16K observations
- DGPs: 1. rescaled estimated DGP $+$ 2. Gaussian AR1 with $\rho = 0.2$ Afrouzi et al. 23



	Dependent Variable: Error	
	Gaussian AR1	Estimated DGP
	(1)	
Revision	-0.44***	
	(0.02)	
Revision × Bottom 40%		
Revision × Top 40%		
Constant and Main Effects Included	✓	
SEs Clustered by Participant	✓	
Number of Observations	6,942	

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error	
	Gaussian AR1 (1)	Estimated DGP (3)
Revision	-0.44*** (0.02)	-0.41*** (0.01)
Revision × Bottom 40%		
Revision × Top 40%		
Constant and Main Effects Included	✓	✓
SEs Clustered by Participant	✓	✓
Number of Observations	6,942	15,717

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error		
	Gaussian AR1		Estimated DGP
	(1)	(2)	(3)
Revision	-0.44***	-0.39	-0.41***
	(0.02)	(0.41)	(0.01)
Revision × Bottom 40%		-0.03	
		(0.42)	
Revision × Top 40%		-0.11	
		(0.41)	
Constant and Main Effects Included	✓	✓	✓
SEs Clustered by Participant	✓	✓	✓
Number of Observations	6,942	6,942	15,717

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error			
	Gaussian AR1		Estimated DGP	
	(1)	(2)	(3)	(4)
Revision	-0.44***	-0.39	-0.41***	0.97**
	(0.02)	(0.41)	(0.01)	(0.39)
Revision × Bottom 40%		-0.03		-1.46***
		(0.42)		(0.39)
Revision × Top 40%		-0.11		-1.41***
		(0.41)		(0.40)
Constant and Main Effects Included	✓	✓	✓	✓
SEs Clustered by Participant	✓	✓	✓	✓
Number of Observations	6,942	6,942	15,717	15,717

Outline

1. Three Key Facts
   - Fact 1: Non-Linear Error-Revision Relationship
   - Fact 2: Fat Tails in the Distribution of Growth
   - Fact 3: Expected Growth is Non-Linear in Past Growth
2. Model of Expectations Formation
3. Additional Model Predictions
   - Quantitative Fit
   - Forecasting Experiment
   - Return Momentum
4. Conclusion

Positive Momentum in Bulk + Mean-Reversion in Tails

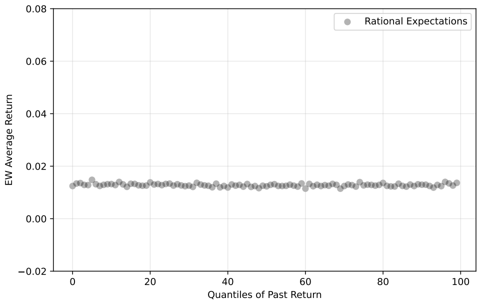
- Campbell 91 + assume constant $F_t(r_{t+k})$ & earnings growth$_t = \gamma \times g_t$

$$\Rightarrow r_{t+1} = \bar{r} + \gamma\left(F_{t+1} - F_t\right)\sum_{k=0}^{\infty} c^k g_{t+1+k}$$

- Campbell 91 + assume constant $F_t(r_{t+k})$ & earnings growth$_t = \gamma \times g_t$

$$\Rightarrow r_{t+1} = \bar{r} + \gamma \left(F_{t+1} - F_t\right) \sum_{k=0}^{\infty} c^k g_{t+1+k}$$

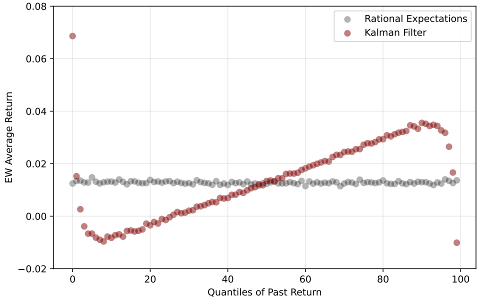
Model



- Campbell 91 + assume constant $F_t(r_{t+k})$ & earnings growth$_t = \gamma \times g_t$

$$\Rightarrow r_{t+1} = \bar{r} + \gamma \left(F_{t+1} - F_t\right) \sum_{k=0}^{\infty} c^k g_{t+1+k}$$

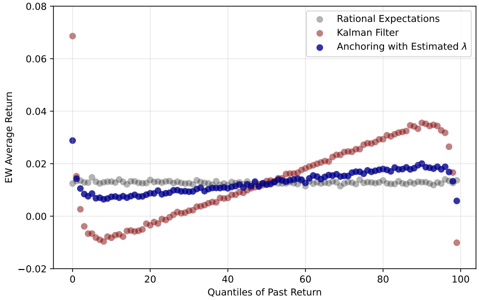
Model



- Campbell 91 + assume constant $F_t(r_{t+k})$ & earnings growth$_t = \gamma \times g_t$

$$\Rightarrow r_{t+1} = \bar{r} + \gamma \left(F_{t+1} - F_t\right) \sum_{k=0}^{\infty} c^k g_{t+1+k}$$

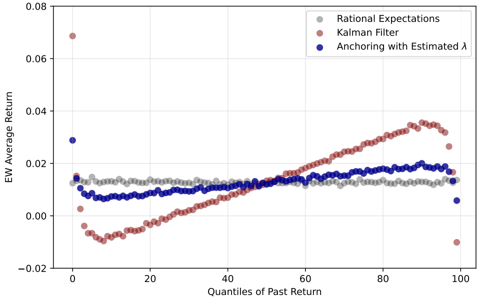
Model



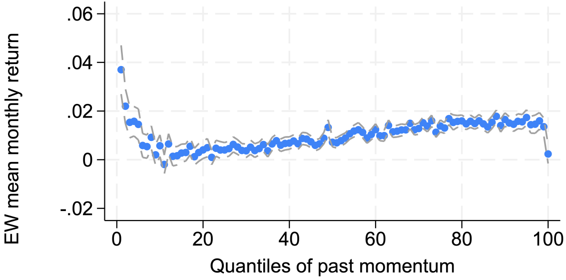
- Campbell 91 + assume constant $F_t(r_{t+k})$ & earnings growth$_t = \gamma \times g_t$

$$\Rightarrow r_{t+1} = \bar{r} + \gamma \left(F_{t+1} - F_t\right) \sum_{k=0}^{\infty} c^k g_{t+1+k}$$

Model



Data: Below Median Market Cap



Outline

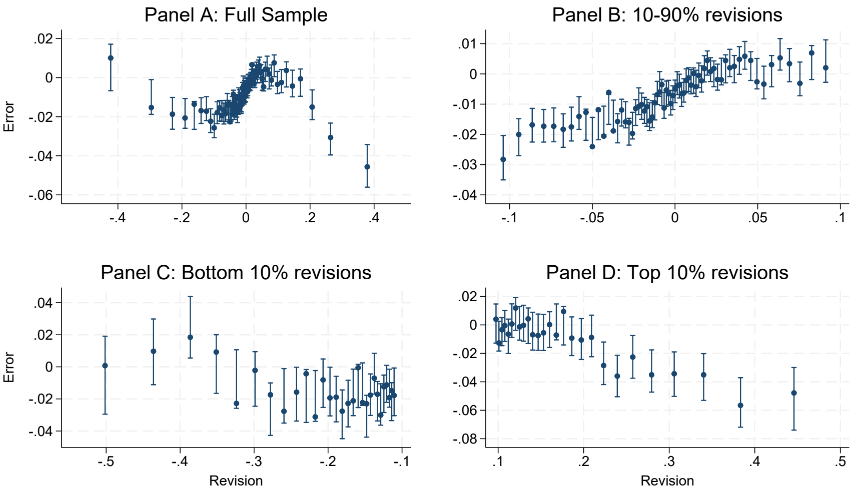
Conclusion

- Main fact: forecast errors are **non-linear** in forecast revisions
  - Underreaction in the bulk of the distribution, overreaction in the tails
- One deviation from RE can explain this: **ignoring fat tails**
  - **Intuition**: extreme realizations are less persistent than forecasters realize
  - Provides a parsimonious model of under **and** overreaction **within a DGP**
  - Also consistent with evidence from experiments and asset prices
- Broader takeaways:
  1. **Non-Gaussian models of DGP** are helpful for understanding belief formation
  2. **Combining** experiments + surveys useful for assessing important features

Thank You!

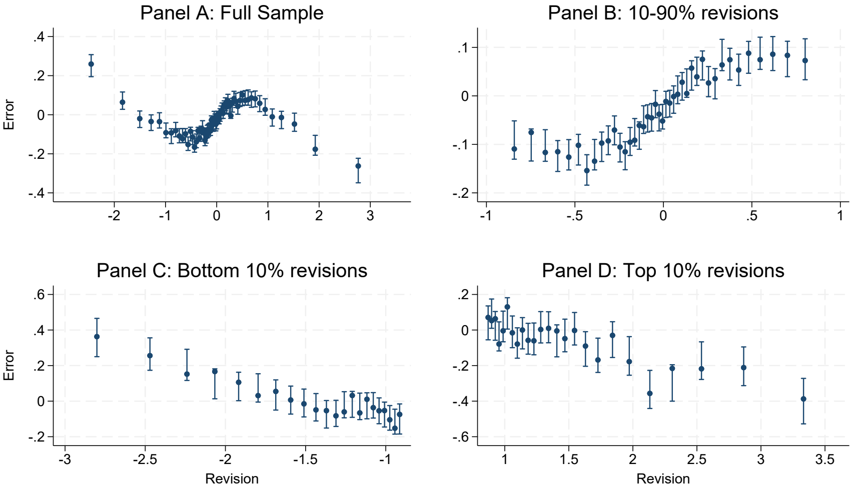
tdesilva@stanford.edu thesmar@mit.edu

Fact 1: Raw Growth

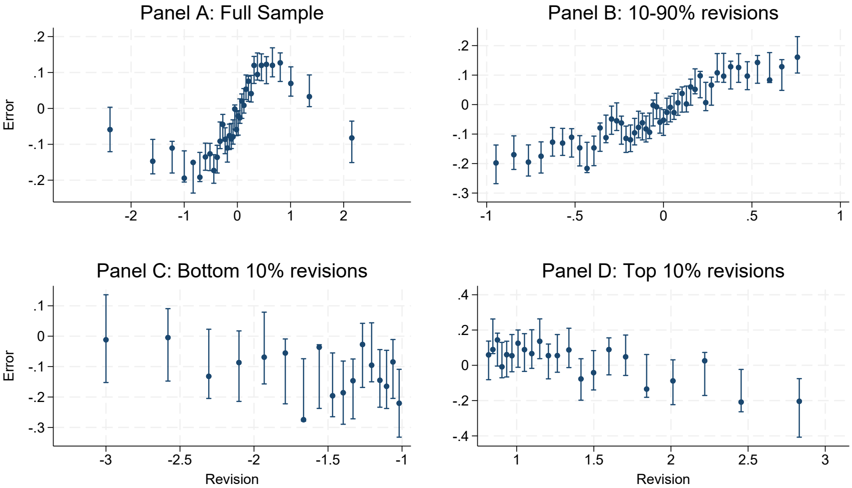


Back

FACT 1: PERCENT GROWTH

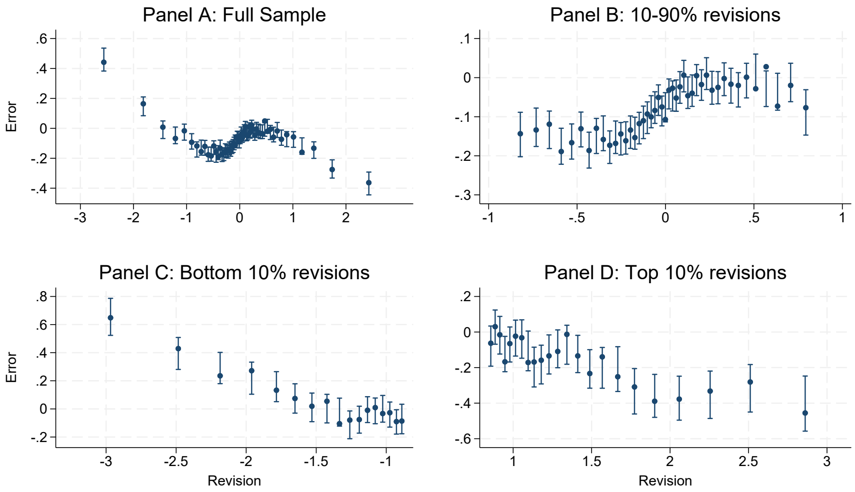


Back


Panel A: Full Sample
Panel B: 10-90% revisions
Panel C: Bottom 10% revisions
Panel D: Top 10% revisions
Error
Error
Revision
Revision
.2
.1
0
-.1
-.2
-2
-1
0
1
2
.2
.1
0
-.1
-.2
-.3
-1
-.5
0
.5
1
.1
0
-.1
-.2
-.3
-3
-2.5
-2
-1.5
-1
.4
.2
0
-.2
-.4
1
1.5
2
2.5
3

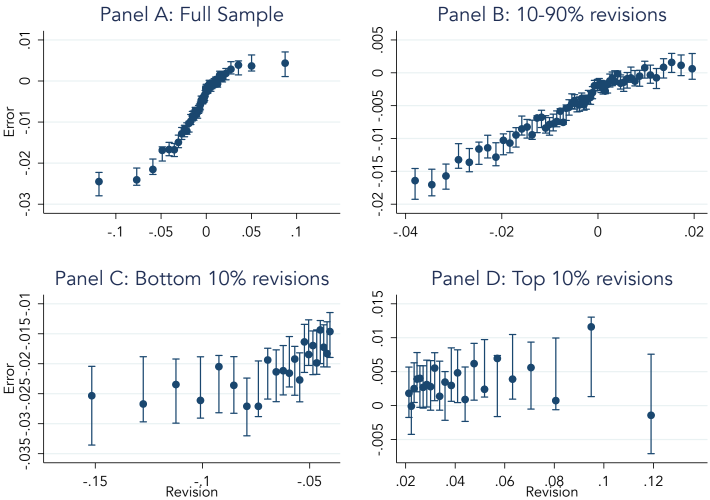
Back

Fact 1: Foreign Firms



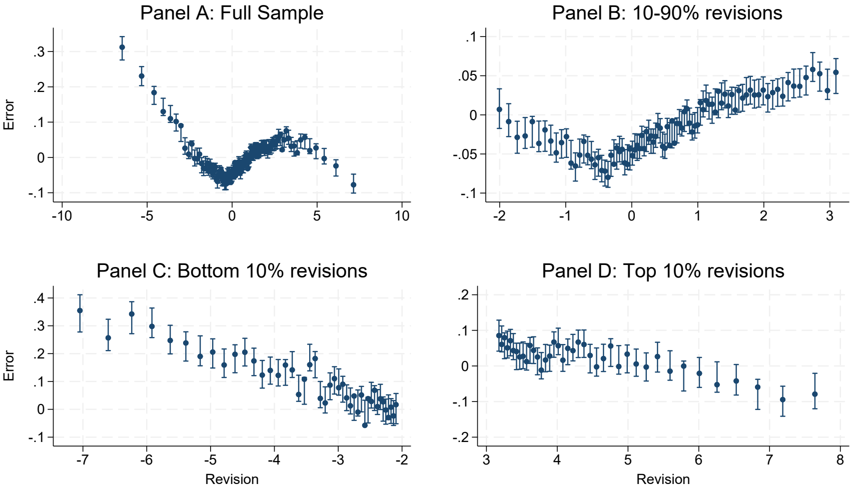
Back

FACT 1: EPS



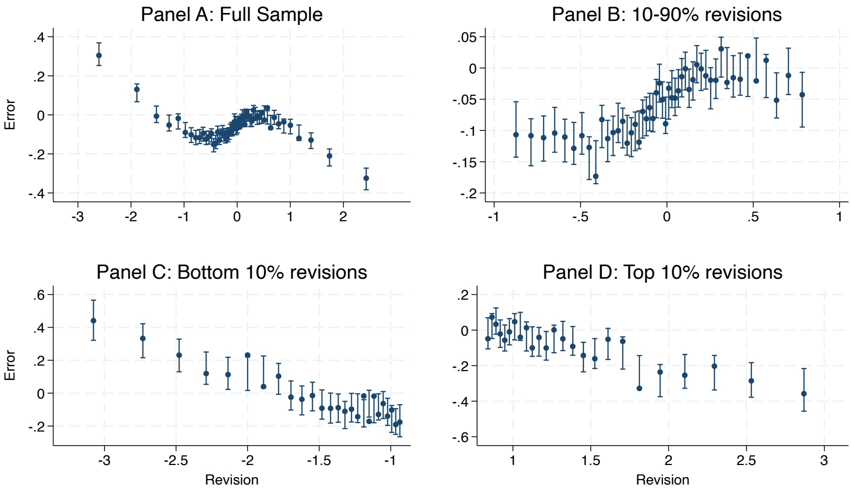
Back

Fact 1: Individual-Level Forecasts



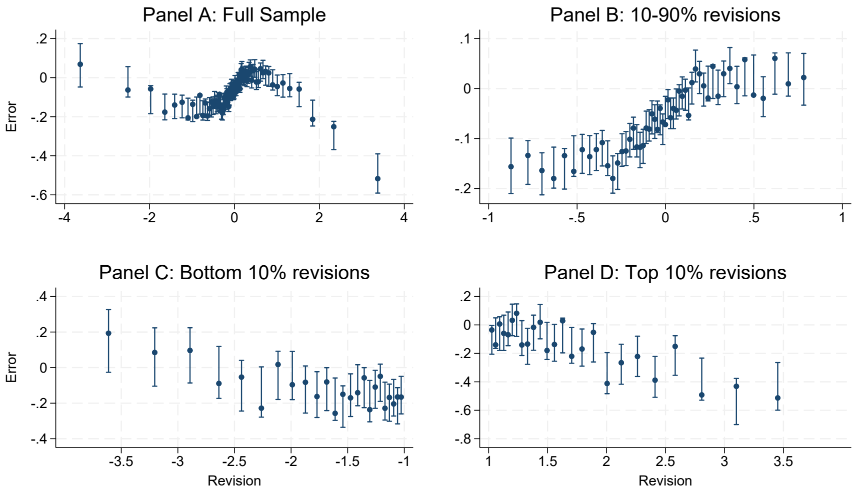
◂ Back

Fact 1: Removing Time FEs



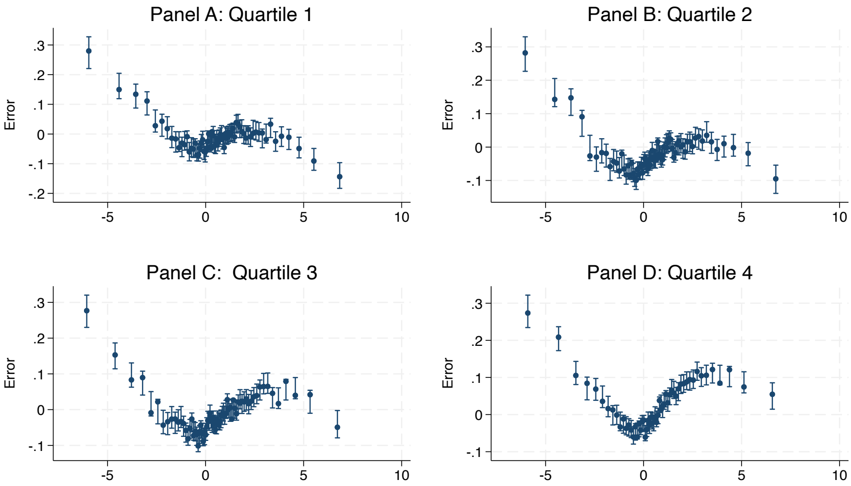
Back

Fact 1: Adjusting for Time-Varying Aggregate Volatility



Back

Fact 1: Variation with Analyst Experience



Back

Estimated Parameters with Standard Errors

	ρ	σ_u	σ_ϵ	ν	λ
Estimate	0.529	0.631	1.325	2.533	0.290
Std. Error	0.041	0.038	0.100	0.083	0.023

Back

Full Sample MSE Loss: Kalman Filter

	ν							
ρ	2.1	2.5	2.533	3.0	3.5	4.0	4.5	5.0
0.1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.2	0.1%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.0%
0.3	0.2%	0.4%	0.4%	0.3%	0.2%	0.2%	0.1%	0.1%
0.4	0.4%	0.7%	0.7%	0.6%	0.4%	0.3%	0.2%	0.2%
0.5	0.6%	1.1%	1.1%	0.9%	0.7%	0.5%	0.4%	0.3%
0.529	0.7%	1.2%	1.2%	1.0%	0.8%	0.6%	0.4%	0.3%
0.6	1.0%	1.6%	1.6%	1.3%	1.0%	0.7%	0.5%	0.4%
0.7	1.4%	2.2%	2.2%	1.8%	1.4%	1.0%	0.8%	0.6%
0.8	1.9%	3.0%	2.9%	2.4%	1.8%	1.4%	1.1%	0.8%
0.9	2.5%	3.6%	3.6%	3.0%	2.3%	1.8%	1.4%	1.1%

Back

Full Sample MSE Loss: Estimated λ

ρ	ν 2.1	2.5	2.533	3.0	3.5	4.0	4.5	5.0
0.1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.2	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.3	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.4	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
0.5	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%
0.529	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%
0.6	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%
0.7	0.1%	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%	0.0%
0.8	0.2%	0.3%	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%
0.9	0.2%	0.3%	0.3%	0.2%	0.2%	0.1%	0.1%	0.1%

Back

Approximating RE with State-Dependent Filter

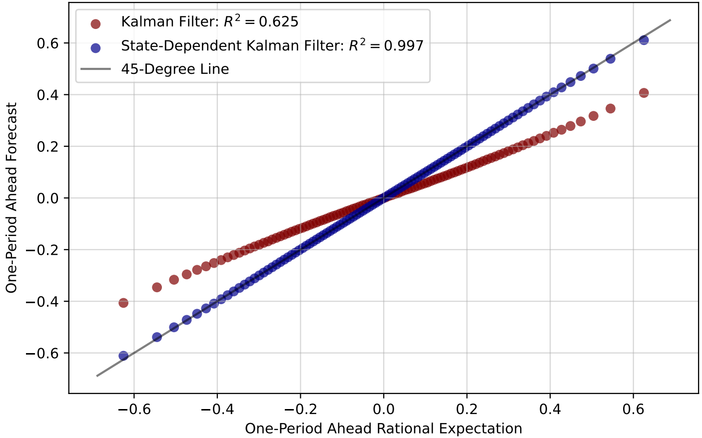
$$F_t g_{t+h} = \rho^h F_t g_t^* \qquad F_t g_t^* = (1 - K_t) F_{t-1} g_t^* + K_t g_t$$

$$K_t = \overline{K} \frac{\gamma_0}{1 + \exp(|g_t - F_{t-1} g_t| - \gamma_1)}$$

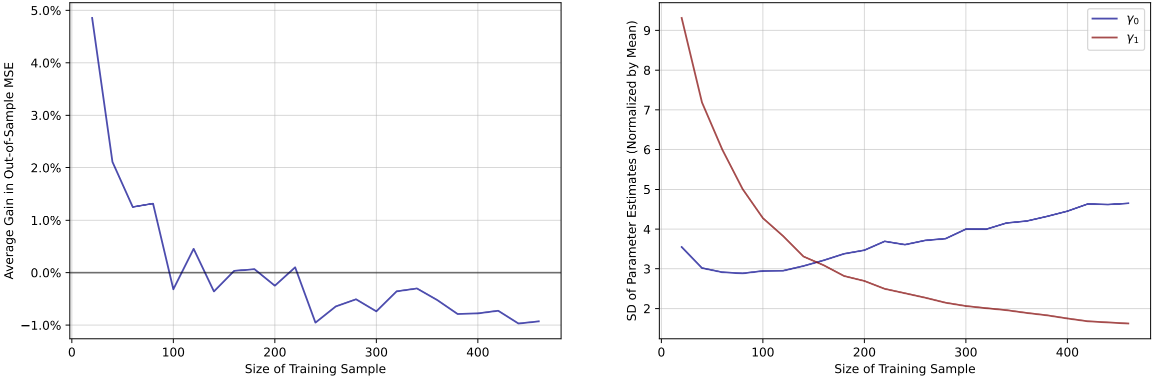
Approximating RE with State-Dependent Filter

$$F_t g_{t+h} = \rho^h F_t g_t^* \qquad F_t g_t^* = (1 - K_t) F_{t-1} g_t^* + K_t g_t$$

$$K_t = \overline{K} \frac{\gamma_0}{1 + \exp(|g_t - F_{t-1} g_t| - \gamma_1)}$$



Out-of-Sample Comparison with Kalman Filter



Note: γ_0, γ_1 estimated using past data only

Back