Expectations Formation with Fat-Tailed Processes: Evidence and Theory

Tim de Silva
Stanford GSB & SIEPR

Eugene Larsen-Hallock
CFM

Adam Rej
CFM

David Thesmar
MIT, NBER, CEPR

NBER Behavioral Finance Meeting

April 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)

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

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)
- 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

⇒ Recognizing **complexity of DGP** is important for understanding belief formation!

1. Empirical evidence on under and overreaction in expectations

Related Literature

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

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
    - Bayesian or non-parametric learning Kozlowski et al. 20, Singleton 21, Farmer 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

Data and Variables

- **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} is adjusted for firm-specific mean and SD
    - Accounts for heterogenous DGPs across firms Wyatt-Bouchaud 03, but not crucial

- **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} is adjusted for firm-specific mean and SD
    - Accounts for heterogenous DGPs across firms Wyatt-Bouchaud 03, but not crucial
- **Forecasts**:

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

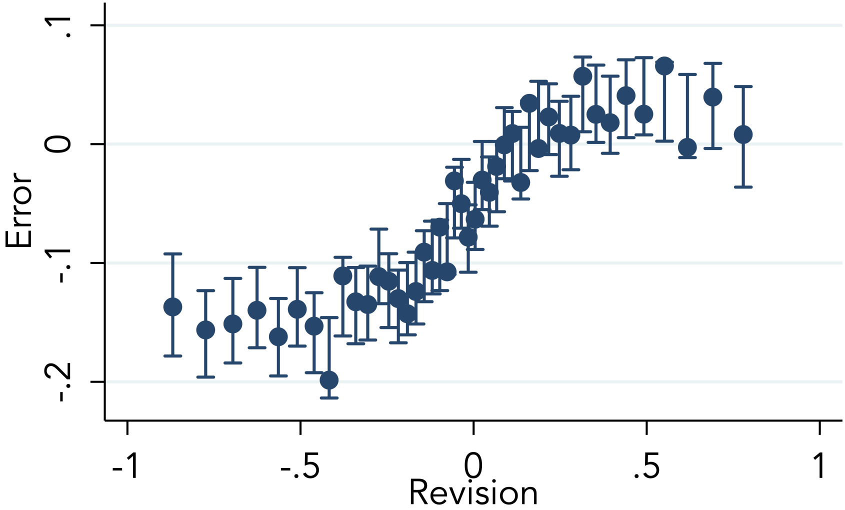
  - F_t = consensus analyst forecasts after year t FY-end announcement
  - $F_t g_{it+h}$ is adjusted using same firm-specific mean and SD as g_{it}
  - Note: (1) ignores a Jensen's adjustment, but not quantitatively important

Outline

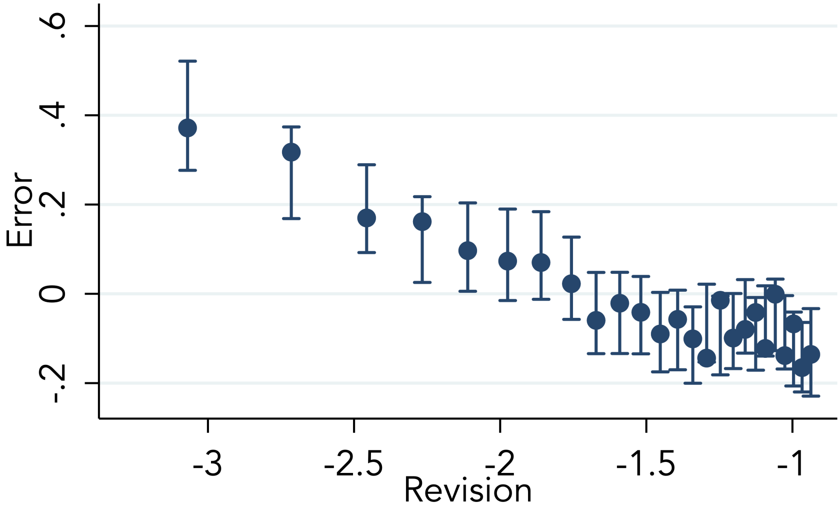
$$\underbrace{g_{it+1} - F_t g_{it+1}}_{\text{forecast error}} = \alpha + \beta \underbrace{\left(F_t g_{it+1} - F_{t-1} g_{it+1}\right)}_{\text{forecast revision}} + \epsilon_{it+1}$$

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

- $\beta \neq 0$ 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

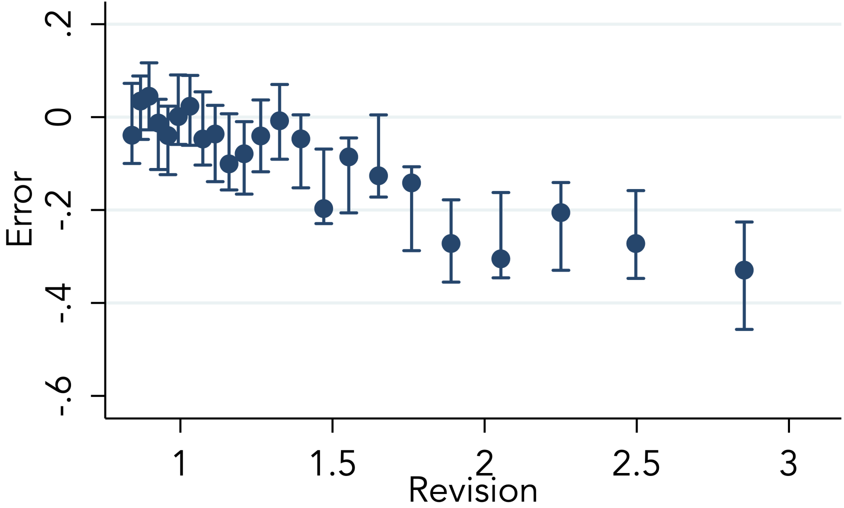


- Between **10-90%** of revisions, error-revision slope is **positive** Bouchaud et al. 19

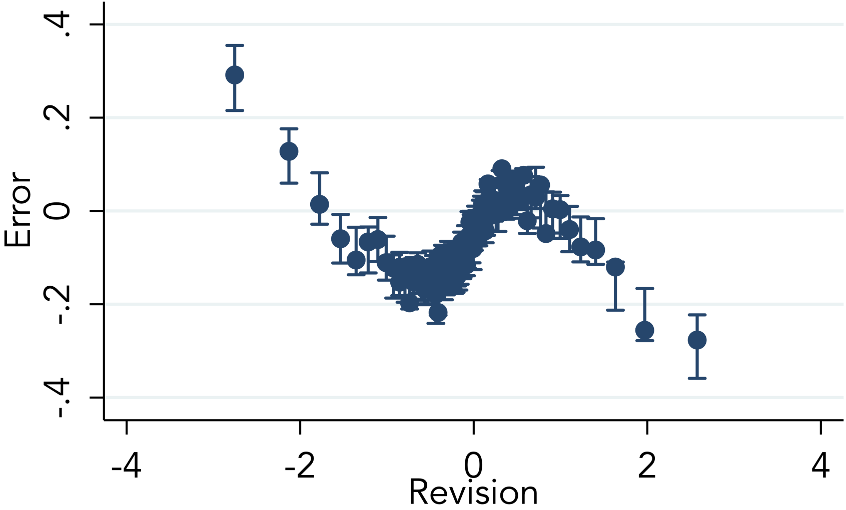


- Between **0-10%** of revisions, error-revision slope is **negative**

… But Overreaction in the Tails!



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

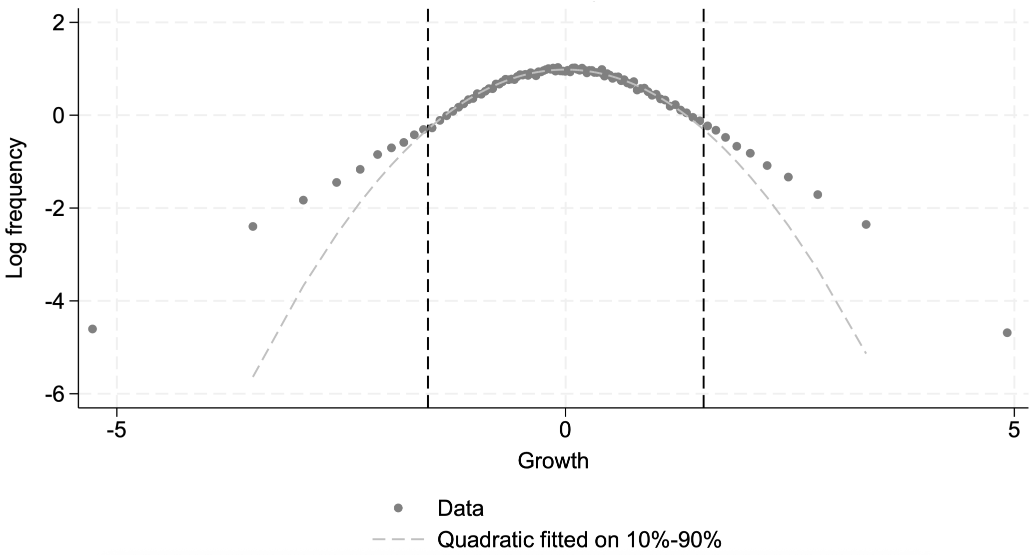


- Forecasts underreact **and** overreact within **same** variable and horizon

Robustness of Non-Linear Error-Revision Relationship

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 arise because of aggregate **time-varying volatility**
4. Not driven by aggregation: present in **individual** forecasts
5. Does not reflect sample: similar for both US and foreign firms

TAILS OF g_{it} ARE FATTER THAN GAUSSIAN

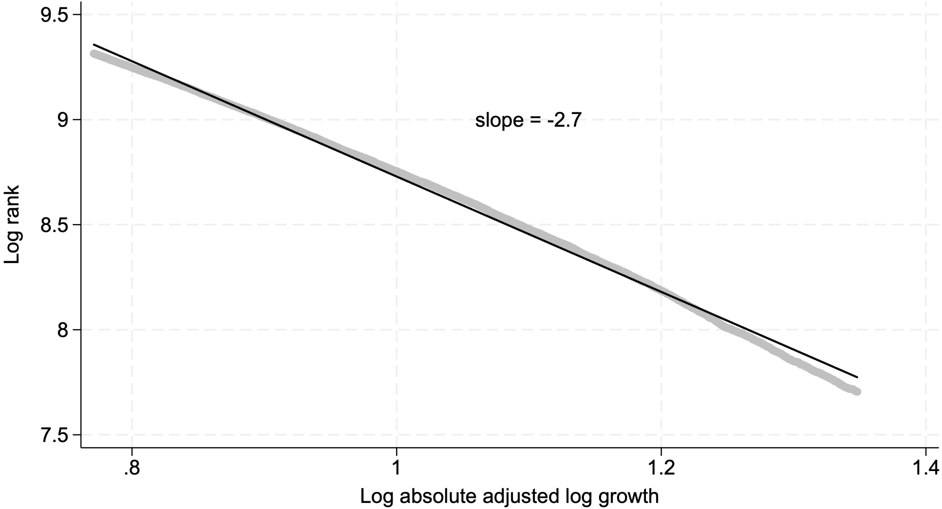


Tail Behavior in Top Deciles is Approximately a Power Law

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

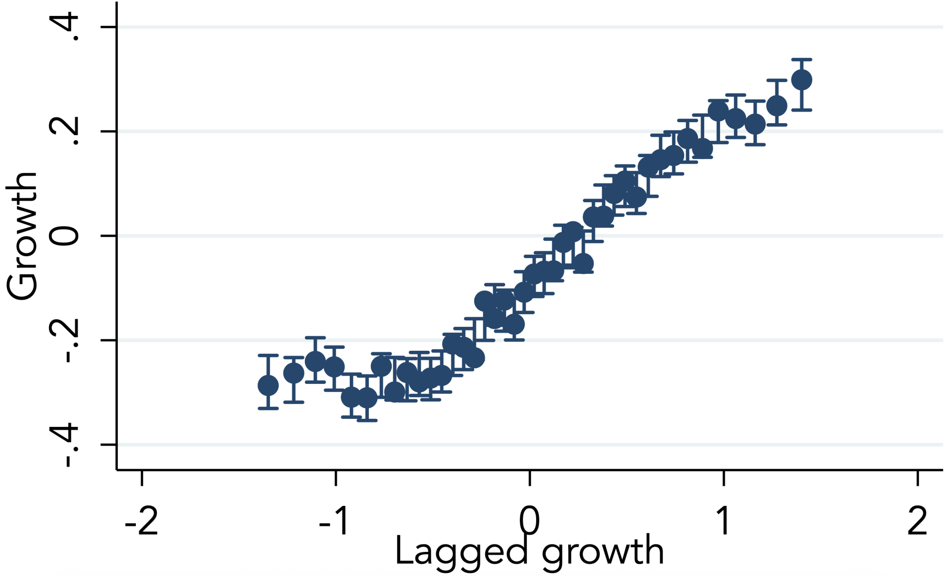
Tail Behavior in Top Deciles is Approximately a Power Law

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

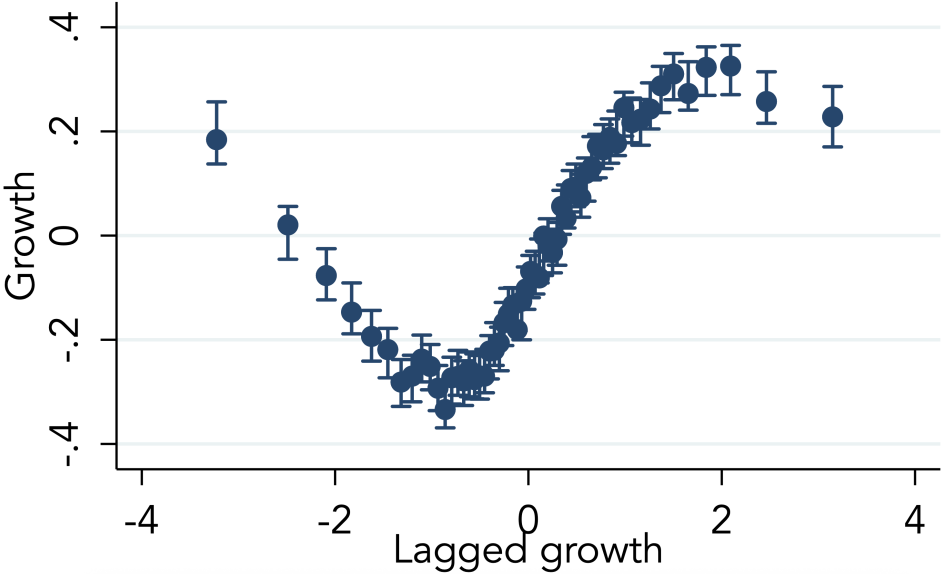


Outline

FACT 3: $\mathbf{E}(g_{it}|g_{it-1})$ IS NON-LINEAR: 10-90% OF g_{it-1}



FACT 3: $\mathbf{E}(g_{it}|g_{it-1})$ IS NON-LINEAR: FULL DISTRIBUTION OF g_{it-1}



Summarizing the Three Facts

1. Forecast errors of sales growth are **non-linear** in revisions
   - Underreaction in the bulk of the distribution, overreaction in the tails

2. Distribution of sales growth, g_{it}, follows a **power law**
   - In the tails, $\log P(|g_{it}| > x) \approx -2.7 \log x + \text{constant}$

3. $E(g_{it+1}|g_{it})$ is **non-linear**
   - Increasing function in the bulk, decreasing function in the tails

Summarizing the Three Facts

1. Forecast errors of sales growth are **non-linear** in revisions
   - Underreaction in the bulk of the distribution, overreaction in the tails

2. Distribution of sales growth, g_{it}, follows a **power law**
   - In the tails, $\log P(|g_{it}| > x) \approx -2.7 \log x + \text{constant}$

3. $E(g_{it+1}|g_{it})$ is **non-linear**
   - Increasing function in the bulk, decreasing function in the tails

Next: introduce a framework that connects these three facts

Outline

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

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

Data-Generating Process

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

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

- g_t is a combination of persistent and transitory processes 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):

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

- g_t is a combination of persistent and transitory processes 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
  - Pareto tail in u_t Guvenen et al. 14 instead of ϵ_t inconsistent with Fact 3

DGP Replicates Facts 2 and 3

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

DGP Replicates Facts 2 and 3

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

- 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

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

- 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

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

- 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

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

- 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

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

- 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

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

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

DGP Replicates Facts 2 and 3

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

- 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

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

- 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 ϵ_t $\Rightarrow$ likely transitory

- Consider a linear model of belief formation for intuition:

$$F_t g_{t+h} = \gamma g_t, \quad \gamma = \text{OLS coefficient of } g_{t+h} \text{ on } g_t$$

- Consider a linear model of belief formation for intuition:

$$F_t g_{t+h} = \gamma g_t, \quad \gamma = \text{OLS coefficient of } g_{t+h} \text{ on } g_t$$

- **Result**: Reaction of linear versus rational forecast to g_t:

$$\frac{d}{dg_t}\Big[F_t g_{t+h} - E(g_{t+1}|g_t)\Big] \geq 0 \iff \frac{d^2}{dg^2}\log h(g_t) \geq -\frac{1}{\sigma_g^2}$$

Replicating Fact 1 in a Simple Expectations Model

- Consider a linear model of belief formation for intuition:

$$F_t g_{t+h} = \gamma g_t, \quad \gamma = \text{OLS coefficient of } g_{t+h} \text{ on } g_t$$

- **Result**: Reaction of linear versus rational forecast to g_t:

$$\frac{d}{dg_t}\left[F_t g_{t+h} - E(g_{t+1}|g_t)\right] \geq 0 \iff \frac{d^2}{dg^2}\log h(g_t) \geq -\frac{1}{\sigma_g^2}$$

- If ϵ_t was Gaussian,

$$\frac{d^2}{dg^2}\log h(g_t) = -\frac{1}{\sigma_g^2} \Rightarrow \textbf{no} \text{ under or overreaction}$$

Replicating Fact 1 in a Simple Expectations Model

- Consider a linear model of belief formation for intuition:

$$F_t g_{t+h} = \gamma g_t, \quad \gamma = \text{OLS coefficient of } g_{t+h} \text{ on } g_t$$

- **Result**: Reaction of linear versus rational forecast to g_t:

$$\frac{d}{dg_t}\left[F_t g_{t+h} - E(g_{t+1}|g_t)\right] \geq 0 \iff \frac{d^2}{dg^2} \log h(g_t) \geq -\frac{1}{\sigma_g^2}$$

- If ϵ_t has Pareto tail, in the **bulk** of the distribution,

$$\frac{d^2}{dg^2} \log h(g_t) \approx -\frac{1}{\sigma_{g0}^2} < -\frac{1}{\sigma_g^2} \Rightarrow \textbf{under}\text{reaction}$$

Replicating Fact 1 in a Simple Expectations Model

- Consider a linear model of belief formation for intuition:

$$F_t g_{t+h} = \gamma g_t, \quad \gamma = \text{OLS coefficient of } g_{t+h} \text{ on } g_t$$

- **Result**: Reaction of linear versus rational forecast to g_t:

$$\frac{d}{dg_t}\left[F_t g_{t+h} - E(g_{t+1}|g_t)\right] \geq 0 \iff \frac{d^2}{dg^2}\log h(g_t) \geq -\frac{1}{\sigma_g^2}$$

- If ϵ_t has Pareto tail, in the **tails** of the distribution,

$$\frac{d^2}{dg^2}\log h(g_t) \approx \frac{\nu}{g_t^2} > -\frac{1}{\sigma_g^2} \Rightarrow \textbf{over}\text{reaction}$$

Replicating Fact 1 in a Simple Expectations Model

- Consider a linear model of belief formation for intuition:

$$F_t g_{t+h} = \gamma g_t, \quad \gamma = \text{OLS coefficient of } g_{t+h} \text{ on } g_t$$

- **Result**: Reaction of linear versus rational forecast to g_t:

$$\frac{d}{dg_t}\left[F_t g_{t+h} - E(g_{t+1}|g_t)\right] \geq 0 \iff \frac{d^2}{dg^2} \log h(g_t) \geq -\frac{1}{\sigma_g^2}$$

- **Intuition**:
  - Bulk: g_t is a stronger predictor of g_{t+1} than full-sample OLS predicts
  - Tails: g_t is a weak predictor of g_{t+1} because of transitory shocks

  $\Rightarrow$ Overreaction to weak signals + underreaction to strong signals Augenblick et al. 24

Replicating Fact 1 in a More Realistic Expectations Model

- Linear model of beliefs is tractable but unrealistic
- More realistic model: forecasts are **optimal** given full history of $\{g_s\}_{s=0}^{t}$
  - Add **one departure from RE**: forecasters think ϵ_t is **Gaussian** $\Rightarrow$ use Kalman filter

Replicating Fact 1 in a More Realistic Expectations Model

- Linear model of beliefs is tractable but unrealistic
- More realistic model: forecasts are **optimal** given full history of $\{g_s\}_{s=0}^{t}$
  - Add **one departure from RE**: forecasters think ϵ_t is **Gaussian** $\Rightarrow$ use Kalman filter
- **Result**: In the steady-state,

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

  - Proof uses that large revision$_t$ must reflect large current ϵ_t or past ϵ_{t-h}
  - Overreaction occurs because ϵ_t is transitory, which forecasters don't realize

Replicating Fact 1 in a More Realistic Expectations Model

- Linear model of beliefs is tractable but unrealistic
- More realistic model: forecasts are **optimal** given full history of $\{g_s\}_{s=0}^{t}$
  - Add **one departure from RE**: forecasters think ϵ_t is **Gaussian** $\Rightarrow$ use Kalman filter
- **Result**: In the steady-state, there exists an $R > 0$ such that:

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

  - Overreaction in tails + unbiased on average $\Rightarrow$ some underreaction in bulk
  - **Implication**: model generates Fact 1 qualitatively

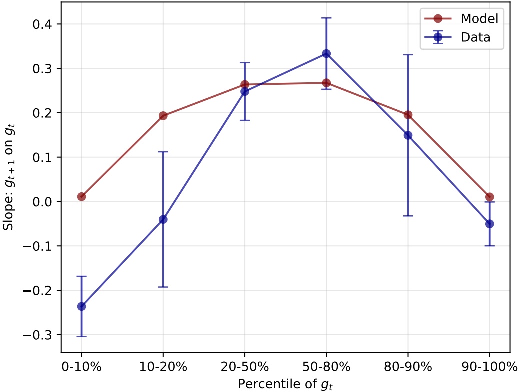
1. Three Key Facts
   - Fact 1: Non-Linear Error-Revision Relationship
   - Fact 2: Fat Tails in the Distribution of Growth
   - Fact 3: Expected Growth is Non-Linear in Past Growth
2. Model of Expectations Formation
3. **Additional Model Predictions**
   - Quantitative Fit
   - Forecasting Experiment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

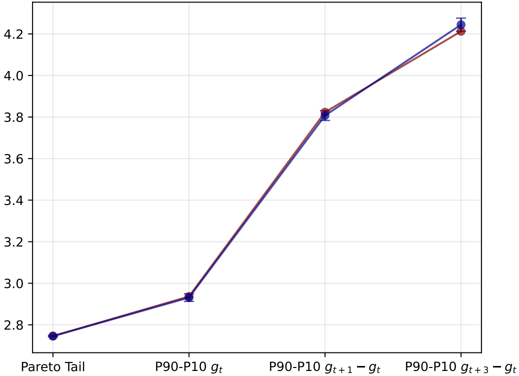
Outline

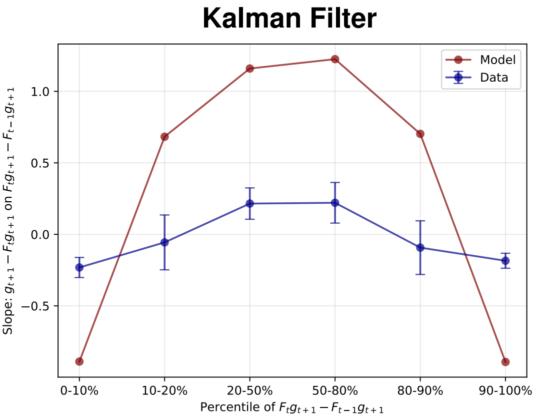
Model Fit: DGP

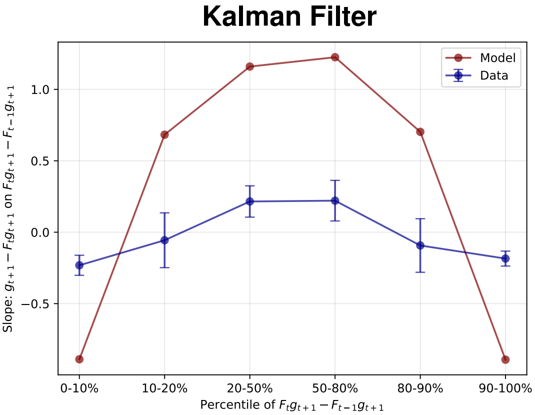
- Estimate DGP parameters using SMM by matching Facts 2 and 3
  - Assume $\epsilon \sim t$-distribution with ν degrees of freedom
  - Add additional moments to identify process scale and persistence

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

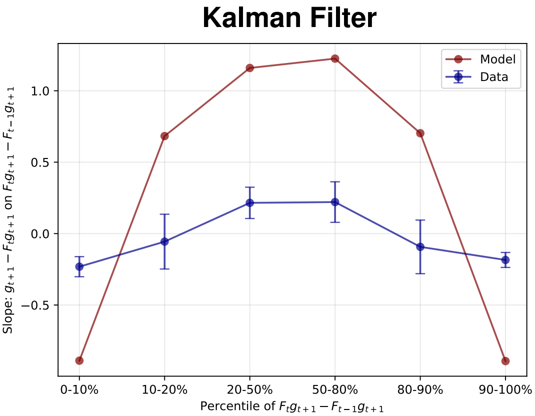




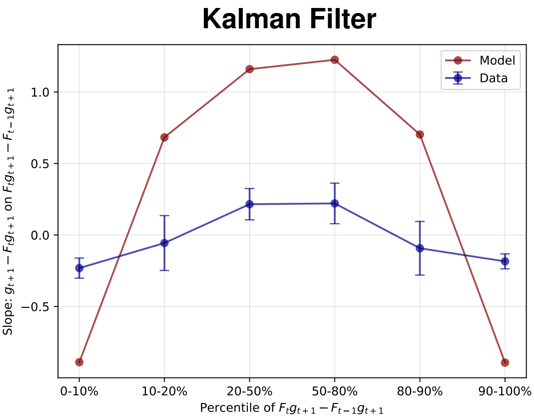

Kalman Filter
Model
Data
1.0
0.5
0.0
−0.5
Slope: $g_{t+1} - F_t g_{t+1}$ on $F_t g_{t+1} - F_{t-1} g_{t+1}$
0-10%
10-20%
20-50%
50-80%
80-90%
90-100%
Percentile of $F_t g_{t+1} - F_{t-1} g_{t+1}$

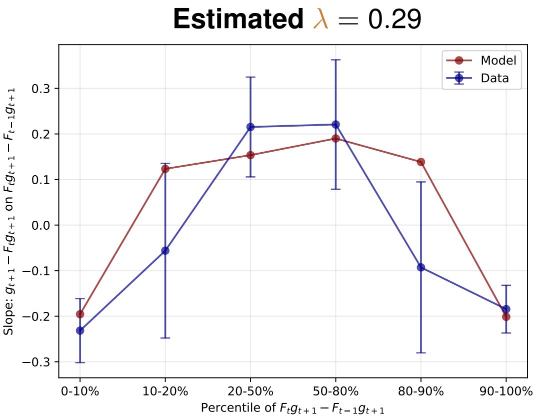


- Given DGP, Kalman filter generates Fact 1 **qualitatively**...
- ... but overdoes it **quantitatively**: too much predictability

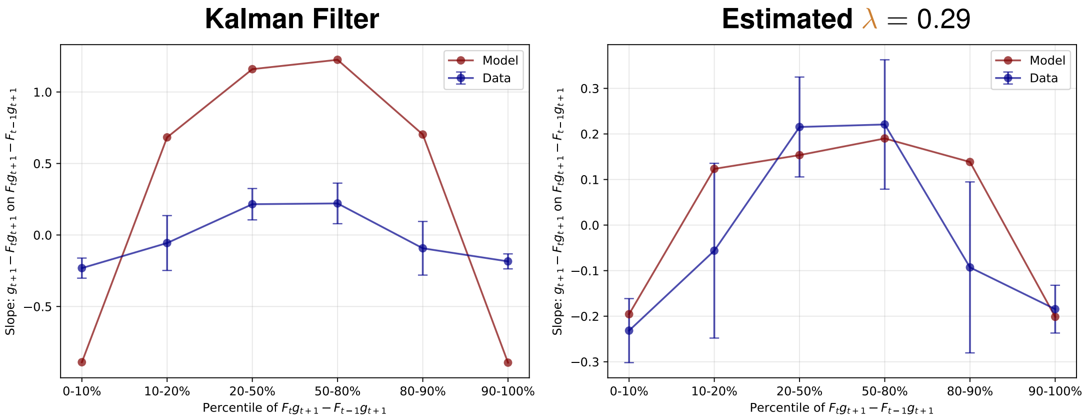


- Allow shrinkage to RE: $F_t^{\lambda} g_{t+h} = \lambda F_t g_{t+h} + \underbrace{(1-\lambda) E_t g_{t+h}}_{\text{particle filtering}}$ Fuster et al. 10, Gabaix 19





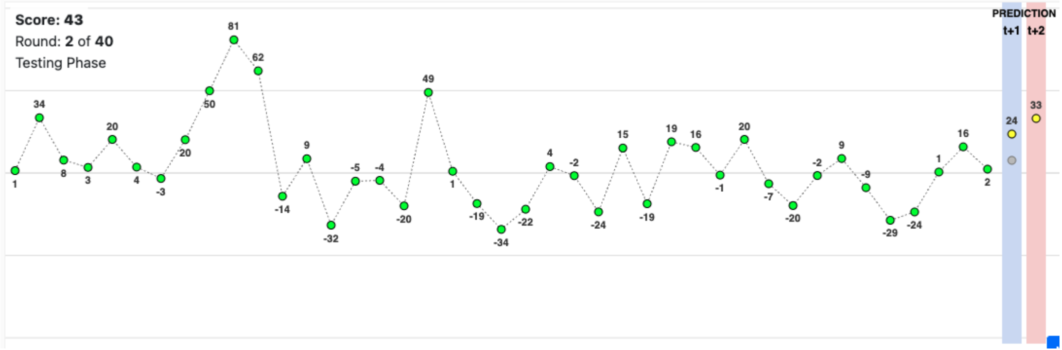
- Allow shrinkage to RE: $F_t^{\lambda} g_{t+h} = \lambda F_t g_{t+h} + (1 - \lambda) E_t g_{t+h}$ Fuster et al. 10, Gabaix 19
- $\lambda = 0.29 \Rightarrow$ replicates error predictability Fact 1



- Allow shrinkage to RE: $F_t^{\lambda} g_{t+h} = \lambda F_t g_{t+h} + (1-\lambda) E_t g_{t+h}$ Fuster et al. 10, Gabaix 19
- $\lambda = 0.29 \Rightarrow$ replicates error predictability Fact 1, but only lose **0.1%** of MSE

Outline

- Design follows Afrouzi et al. 23: participants make one and two-period forecasts
- 201 participants make 40 forecasts ⇒ 8K observations possibly scale up?
- DGP is a scaled version of the one estimated in data



Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error	
	Non-Gaussian DGP with Fat Tails (1)	Gaussian AR1
Revision	-0.40*** (0.02)	
Top 20%		
Bottom 20%		
Revision × Bottom 20%		
Revision × Top 20%		
Revision × Top & Bottom 20%		
Constant	✓	
Clustering by Participant	✓	
N	7839	

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error	
	Non-Gaussian DGP with Fat Tails	Gaussian AR1
	(1)	(4)
Revision	-0.40***	-0.44***
	(0.02)	(0.02)
Top 20%		
Bottom 20%		
Revision × Bottom 20%		
Revision × Top 20%		
Revision × Top & Bottom 20%		
Constant	✓	✓
Clustering by Participant	✓	✓
N	7839	5421

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error		
	Non-Gaussian DGP with Fat Tails		Gaussian AR1
	(1)	(2)	(4)
Revision	-0.40***	-0.28***	-0.44***
	(0.02)	(0.06)	(0.02)
Top 20%		-1.46	
		(4.41)	
Bottom 20%		-11.04**	
		(4.55)	
Revision × Bottom 20%		-0.27***	
		(0.09)	
Revision × Top 20%		-0.11	
		(0.08)	
Revision × Top & Bottom 20%			
Constant	✓	✓	✓
Clustering by Participant	✓	✓	✓
N	7839	7839	5421

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error			
	Non-Gaussian DGP with Fat Tails		Gaussian AR1	
	(1)	(2)	(4)	(5)
Revision	-0.40***	-0.28***	-0.44***	-0.42***
	(0.02)	(0.06)	(0.02)	(0.06)
Top 20%		-1.46		-1.10
		(4.41)		(2.52)
Bottom 20%		-11.04**		-10.17***
		(4.55)		(3.38)
Revision $\times$ Bottom 20%		-0.27***		-0.12
		(0.09)		(0.09)
Revision $\times$ Top 20%		-0.11		-0.07
		(0.08)		(0.07)
Revision $\times$ Top & Bottom 20%				
Constant	✓	✓	✓	✓
Clustering by Participant	✓	✓	✓	✓
N	7839	7839	5421	5421

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error				
	Non-Gaussian DGP with Fat Tails			Gaussian AR1	
	(1)	(2)	(3)	(4)	(5)
Revision	-0.40***	-0.28***	-0.28***	-0.44***	-0.42***
	(0.02)	(0.06)	(0.06)	(0.02)	(0.06)
Top 20%		-1.46	4.19		-1.10
		(4.41)	(4.24)		(2.52)
Bottom 20%		-11.04**	-5.42		-10.17***
		(4.55)	(3.76)		(3.38)
Revision × Bottom 20%		-0.27***			-0.12
		(0.09)			(0.09)
Revision × Top 20%		-0.11			-0.07
		(0.08)			(0.07)
Revision × Top & Bottom 20%			-0.18**		
			(0.08)		
Constant	✓	✓	✓	✓	✓
Clustering by Participant	✓	✓	✓	✓	✓
N	7839	7839	7839	5421	5421

Fat Tails Create Non-Linear Error-Revision Relationship

	Dependent Variable: Error					
	Non-Gaussian DGP with Fat Tails			Gaussian AR1		
	(1)	(2)	(3)	(4)	(5)	(6)
Revision	-0.40***	-0.28***	-0.28***	-0.44***	-0.42***	-0.42***
	(0.02)	(0.06)	(0.06)	(0.02)	(0.06)	(0.06)
Top 20%		-1.46	4.19		-1.10	-0.00
		(4.41)	(4.24)		(2.52)	(2.53)
Bottom 20%		-11.04**	-5.42		-10.17***	-8.83***
		(4.55)	(3.76)		(3.38)	(2.78)
Revision $\times$ Bottom 20%		-0.27***			-0.12	
		(0.09)			(0.09)	
Revision $\times$ Top 20%		-0.11			-0.07	
		(0.08)			(0.07)	
Revision $\times$ Top & Bottom 20%			-0.18**			-0.09
			(0.08)			(0.07)
Constant	✓	✓	✓	✓	✓	✓
Clustering by Participant	✓	✓	✓	✓	✓	✓
N	7839	7839	7839	5421	5421	5421

Outline

Positive Momentum in Bulk + Mean-Reversion in Tails

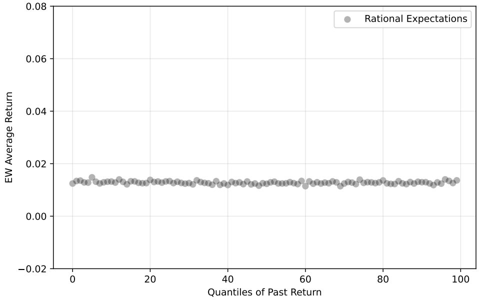
- Campbell 91 + constant $F_t(r_{t+k})$ + earnings growth$_t = \gamma \times g_t$

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

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

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

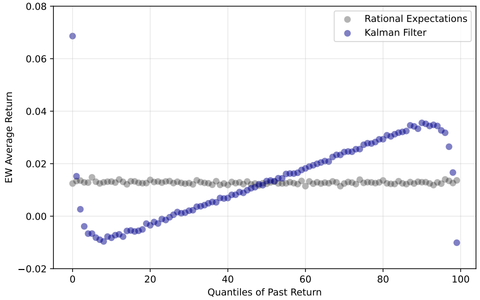
Model



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

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

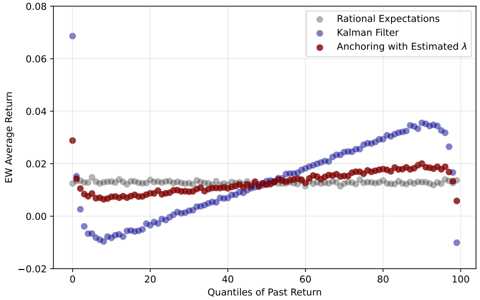
Model



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

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

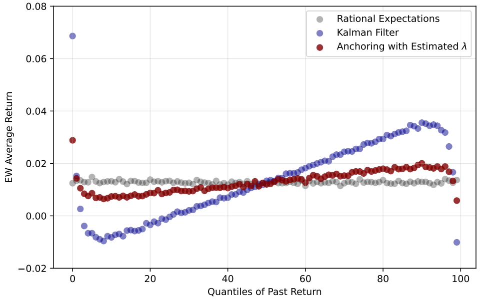
Model



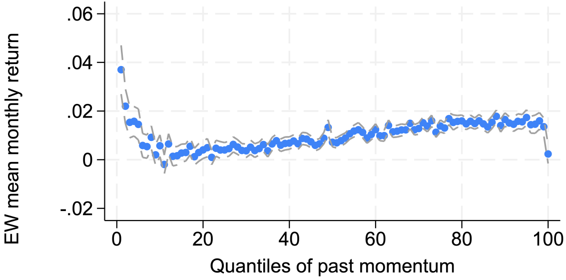
- Campbell 91 + constant $F_t(r_{t+k})$ + earnings growth$_t = \gamma \times g_t$

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

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

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. Recognizing **DGP complexity** important for understanding belief formation
  2. **Combining** experiments + surveys useful for assessing important features

THANK YOU!

tdesilva@stanford.edu thesmar@mit.edu