Diagnostic Expectations and Stock Returns

by Pedro Bordalo, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer

Discussion

Pietro Veronesi

The University of Chicago Booth School of Business
Main Contribution and Outline of Discussion

- Main contribution of the paper:
  1. Provide numerous facts (some known, some new) about analysts expectations, their dynamics, and their relation to stock returns
  2. Put forward a behavioral learning model based on representative heuristics to make sense of these facts
Main Contribution and Outline of Discussion

- Main contribution of the paper:
  1. Provide numerous facts (some known, some new) about analysts expectations, their dynamics, and their relation to stock returns
  2. Put forward a behavioral learning model based on representative heuristics to make sense of these facts

- Outline of discussion
  1. What does rational filtering imply for growth?
  2. How do diagnostic expectations differ?
  3. An alternative learning story for expectations and returns
  4. Final comments
Kalman Filter in a Simple Setting

- $N$ firms, with realized growth rate

$$x_{it} = g_i + \varepsilon_{it}$$

- $g_i$ = unobservable long-term growth
Kalman Filter in a Simple Setting

- $N$ firms, with realized growth rate
  \[ x_{it} = g_i + \varepsilon_{it} \]
  \[- g_i = \text{unobservable long-term growth} \]

- Prior beliefs for each firm $i$ given information at $t-1$, $F_{t-1}$:
  \[ g_i \sim N \left( \hat{g}_{i,t-1}, \hat{\sigma}_{i,t-1}^2 \right) \]
Kalman Filter in a Simple Setting

- $N$ firms, with realized growth rate

$$x_{it} = g_i + \varepsilon_{it}$$

$- g_i =$ unobservable long-term growth

- Prior beliefs for each firm $i$ given information at $t - 1$, $F_{t-1}$:

$$g_i \sim N \left( \hat{g}_{i,t-1}, \hat{\sigma}_{i,t-1}^2 \right)$$

- Conditional on $F_{t-1}$:

$$\begin{pmatrix} x_{i,t} \\ g_i \end{pmatrix} \sim N \left( \begin{pmatrix} \hat{g}_{i,t-1} \\ \hat{g}_{i,t-1} \end{pmatrix}, \begin{pmatrix} \hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2 & \hat{\sigma}_{i,t-1}^2 \\ \hat{\sigma}_{i,t-1}^2 & \hat{\sigma}_{i,t-1}^2 \end{pmatrix} \right)$$
Kalman Filter in a Simple Setting

- $N$ firms, with realized growth rate
  \[ x_{it} = g_i + \varepsilon_{it} \]
  - $g_i = \text{unobservable long-term growth}$

- Prior beliefs for each firm $i$ given information at $t - 1$, $F_{t-1}$:
  \[ g_i \sim N(\hat{g}_{i,t-1}, \hat{\sigma}_{i,t-1}^2) \]

- Conditional on $F_{t-1}$:
  \[
  \begin{pmatrix}
  x_{i,t} \\
  g_i
  \end{pmatrix} \sim N
  \begin{pmatrix}
  \left( \hat{g}_{i,t-1} \right), \\
  \left( \frac{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2}, \hat{\sigma}_{i,t-1}^2 \right)
  \end{pmatrix}
  \]

- From the properties of conditional normal distribution
  \[ g_i | x_{i,t} \sim N \left( \hat{g}_{i,t-1} + \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2} (x_{i,t} - \hat{g}_{i,t-1}), \hat{\sigma}_{i,t-1}^2 - \frac{(\hat{\sigma}_{i,t-1}^2)^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2} \right) \]
Kalman Filter in a Simple Setting

- $N$ firms, with realized growth rate
  \[ x_{it} = g_i + \varepsilon_{it} \]
  \(- g_i = \text{unobservable long-term growth}

- Prior beliefs for each firm $i$ given information at $t - 1$, $F_{t-1}$:
  \[ g_i \sim N \left( \hat{g}_{i,t-1}, \hat{\sigma}_{i,t-1}^2 \right) \]

- Conditional on $F_{t-1}$:
  \[
  \begin{pmatrix}
  x_{i,t} \\
  g_i
  \end{pmatrix}
  \sim N \left( \begin{pmatrix}
  \hat{g}_{i,t-1} \\
  \hat{g}_{i,t-1}
  \end{pmatrix}, \begin{pmatrix}
  \hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2 & \hat{\sigma}_{i,t-1}^2 \\
  \hat{\sigma}_{i,t-1}^2 & \hat{\sigma}_{i,t-1}^2
  \end{pmatrix} \right)
  \]

- From the properties of conditional normal distribution
  \[
  g_{i|x_{i,t}} \sim N \left( \hat{g}_{i,t-1} + \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} (x_{i,t} - \hat{g}_{i,t-1}), \hat{\sigma}_{i,t-1}^2 \right)
  \]
Kalman Filter in a Simple Setting

- $N$ firms, with realized growth rate
  \[ x_{it} = g_i + \varepsilon_{it} \]
  
  $- g_i =$ unobservable long-term growth

- Prior beliefs for each firm $i$ given information at $t - 1$, $F_{t-1}$:
  \[ g_i \sim N (\hat{g}_{i,t-1}, \hat{\sigma}^2_{i,t-1}) \]

- Conditional on $F_{t-1}$:
  \[
  \begin{pmatrix}
  x_{i,t} \\
  g_i
  \end{pmatrix}
  \sim N
  \begin{pmatrix}
  \hat{g}_{i,t-1} \\
  \hat{g}_{i,t-1}
  \end{pmatrix}
  ,
  \begin{pmatrix}
  \hat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon & \hat{\sigma}^2_{i,t-1} \\
  \hat{\sigma}^2_{i,t-1} & \hat{\sigma}^2_{i,t-1}
  \end{pmatrix}
  \]

- From the properties of conditional normal distribution
  \[
  g_i \mid x_{i,t} \sim N
  \left(\hat{g}_{i,t-1} + \frac{\hat{\sigma}^2_{i,t-1}}{\hat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon} (x_{i,t} - \hat{g}_{i,t-1})
  ,
  \hat{\sigma}^2_{i,t-1} - \frac{(\hat{\sigma}^2_{i,t-1})^2}{\hat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}
  \right)
  \]
Kalman Filter in a Simple Setting

- $N$ firms, with realized growth rate
  \[ x_{it} = g_i + \varepsilon_{it} \]
  
  - $g_i$ = unobservable long-term growth

- Prior beliefs for each firm $i$ given information at $t - 1$, $F_{t-1}$:
  \[ g_i \sim N \left( \hat{g}_{i,t-1}, \hat{\sigma}_{i,t-1}^2 \right) \]

- Conditional on $F_{t-1}$:
  \[
  \begin{pmatrix}
  x_{i,t} \\
  g_i
  \end{pmatrix} \sim N \left( \begin{pmatrix}
  \hat{g}_{i,t-1} \\
  \hat{g}_{i,t-1}
  \end{pmatrix}, \begin{pmatrix}
  \hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2 & \hat{\sigma}_{i,t-1}^2 \\
  \hat{\sigma}_{i,t-1}^2 & \hat{\sigma}_{i,t-1}^2
  \end{pmatrix} \right)
  \]

- From the properties of conditional normal distribution
  \[
  g_i \mid x_{i,t} \sim N \left( \begin{pmatrix}
  \hat{g}_{i,t-1} + \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} (x_{i,t} - \hat{g}_{i,t-1}) \\
  \hat{g}_{i,t}
  \end{pmatrix}, \begin{pmatrix}
  \hat{\sigma}_{i,t-1}^2 - \frac{(\hat{\sigma}_{i,t-1})^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} & 0 \\
  0 & \hat{\sigma}_{i,t}^2
  \end{pmatrix} \right)
  \]
Sample Selection and Realized vs. Expected Growth

- I simulate 10,000 firms with (annualized) $g_i = 5\%, \hat{\sigma}_0 = 10\%, \sigma_\varepsilon = 15\%$
- At time $t^*$, sort firms on posterior mean $\hat{g}_{i,t}$ and form 10 decile portfolios.
- What growth rates should the top and bottom portfolios display?
Sample Selection and Realized vs. Expected Growth

- I simulate 10,000 firms with (annualized) $g_i = 5\%, \hat{\sigma}_0 = 10\%, \sigma_\varepsilon = 15\%$
- At time $t^*$, sort firms on posterior mean $\hat{g}_{i,t}$ and form 10 decile portfolios.
- What growth rates should the top and bottom portfolios display?
Sample Selection and Realized vs. Expected Growth

- I simulate 10,000 firms with (annualized) \( g_i = 5\% \), \( \hat{\sigma}_0 = 10\% \), \( \sigma_\varepsilon = 15\% \).
- At time \( t^* \), sort firms on posterior mean \( \hat{g}_{i,t} \) and form 10 decile portfolios.
- What growth rates should the top and bottom portfolios display?
Sample Selection and Realized vs. Expected Growth

- I simulate 10,000 firms with (annualized) $g_i = 5\%$, $\hat{\sigma}_0 = 10\%$, $\sigma_{\epsilon} = 15\%$
- At time $t^*$, sort firms on posterior mean $\hat{g}_{i,t}$ and form 10 decile portfolios.
- What growth rates should the top and bottom portfolios display?
Sample Selection and Realized vs. Expected Growth

- I simulate 10,000 firms with (annualized) $g_i = 5\%$, $\hat{\sigma}_0 = 10\%$, $\sigma_\varepsilon = 15\%$
- At time $t^*$, sort firms on posterior mean $\hat{g}_{i,t}$ and form 10 decile portfolios.
- What growth rates should the top and bottom portfolios display?
Sample Selection and Realized vs. Expected Growth

- I simulate 10,000 firms with (annualized) $g_i = 5\%$, $\hat{\sigma}_0 = 10\%$, $\sigma_\varepsilon = 15\%$
- At time $t^*$, sort firms on posterior mean $\hat{g}_{i,t}$ and form 10 decile portfolios.
- What growth rates should the top and bottom portfolios display?

![Graph showing realized and expected growth rates over time with a jump due to sample selection and rational expectations.]
Sample Selection and Realized vs. Expected Growth – 2

- That is, the High Long-Term-Growth (HLTG) portfolio displays:
That is, the High Long-Term-Growth (HLTG) portfolio displays:

1. Realized growth that:
   (a) is high before portfolio formation
   (b) is low after portfolio formation
   (c) jumps down after portfolio formation
   (d) equals expected future growth
That is, the High Long-Term-Growth (HLTG) portfolio displays:

1. Realized growth that:
   (a) is high before portfolio formation
   (b) is low after portfolio formation
   (c) jumps down after portfolio formation
   (d) equals expected future growth

2. Expected growth $\hat{g}_{it}$ that:
   (a) increases before portfolio formation
   (b) is flat after portfolio formation, as expectations are martingales after the conditioning event.
• That is, the High Long-Term-Growth (H_LTG) portfolio displays:

1. Realized growth that:
   (a) is high before portfolio formation
   (b) is low after portfolio formation
   (c) jumps down after portfolio formation
   (d) equals expected future growth

2. Expected growth $\hat{g}_{it}$ that:
   (a) increases before portfolio formation
   (b) is flat after portfolio formation, as expectations are martingales after the conditioning event.

• The Low Long-term-Growth (L_LTG) portfolio has symmetric properties.

• All these results can be proven formally, as BGLS in fact do in their paper.
Sample Selection and Realized vs. Expected Growth – 3

- Adding mean reversion generates more “curvy” results:

\[ x_{it} = \gamma_0 + g_{it} + \varepsilon_{it} \]
\[ g_{i,t+1} = \gamma_1 g_{it} + \eta_{i,t+1} \]
Adding mean reversion generates more “curvy” results:

\[ x_{it} = \gamma_0 + g_{it} + \varepsilon_{it} \]
\[ g_{i,t+1} = \gamma_1 g_{it} + \eta_{i,t+1} \]

Conditional on \( F_{t-1} \):

\[
\begin{pmatrix}
    x_{i,t} \\
    g_{i,t+1}
\end{pmatrix}
\sim N
\left( \begin{pmatrix}
    \gamma_0 + \hat{g}_{i,t-1} \\
    \gamma_1 \hat{g}_{i,t-1}
\end{pmatrix}
, \begin{pmatrix}
    \hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2 & \gamma_1 \hat{\sigma}_{i,t-1}^2 \\
    \gamma_1 \hat{\sigma}_{i,t-1}^2 & \gamma_1 \hat{\sigma}_{i,t-1}^2 + \sigma_\eta^2
\end{pmatrix} \right)
\]

After observing \( x_{it} \), the filter is then

\[
\hat{g}_{i,t} = \gamma_1 \hat{g}_{i,t-1} + \frac{\gamma_1 \hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} (x_{it} - \delta_0 - \hat{g}_{i,t-1})
\]
\[
\hat{\sigma}_{i,t}^2 = \gamma_1 \hat{\sigma}_{i,t-1}^2 + \sigma_\eta^2 - \frac{(\gamma_1 \hat{\sigma}_{i,t-1}^2)^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2}
\]
Sample Selection and Realized vs. Expected Growth – 3

- Adding mean reversion generates more “curvy” results:
  \[
  x_{it} = \gamma_0 + g_{it} + \varepsilon_{it}
  \]
  \[
  g_{i,t+1} = \gamma_1 g_{it} + \eta_{i,t+1}
  \]

- Conditional on \( F_{t-1} \):
  \[
  \begin{pmatrix}
  x_{i,t} \\
  g_{i,t+1}
  \end{pmatrix}
  \sim N
  \left(
  \begin{pmatrix}
  \gamma_0 + \hat{g}_{i,t-1} \\
  \gamma_1 \hat{g}_{i,t-1}
  \end{pmatrix},
  \begin{pmatrix}
  \tilde{\sigma}_i,t-1 + \sigma_\varepsilon^2 & \gamma_1 \tilde{\sigma}_i,t-1 \\
  \gamma_1 \tilde{\sigma}_i,t-1 & \gamma_1^2 \tilde{\sigma}_i,t-1 + \sigma_\eta^2
  \end{pmatrix}
  \right)
  \]

- After observing \( x_{it} \), the filter is then
  \[
  \hat{g}_{i,t} = \gamma_1 \hat{g}_{i,t-1} + \frac{\gamma_1 \tilde{\sigma}^2_{i,t-1}}{\tilde{\sigma}^2_{i,t-1} + \sigma_\varepsilon^2} (x_{it} - \delta_0 - \hat{g}_{i,t-1})
  \]
  \[
  \tilde{\sigma}_i,t = \gamma_1^2 \tilde{\sigma}_i,t-1 + \sigma_\eta^2 - \frac{(\gamma_1 \tilde{\sigma}_i,t-1)^2}{\tilde{\sigma}_i,t-1 + \sigma_\varepsilon^2}
  \]

- Simulation with 10,000 firms as before, now with (annualized)
  \( \gamma_0 = 5\%, \, \gamma_1 = 0.96, \, \sigma_\eta = 1\% \)
We observe that for HL TG portfolio, after portfolio formation:
1. Mean reversion in fundamentals ($g_{it}$) make expected growth decline.
2. Average realized growth is on top of expected (rational expectations).
We observe that for HLTG portfolio, after portfolio formation:

1. Mean reversion in fundamentals ($g_{it}$) make expected growth decline.
We observe that for HLTG portfolio, after portfolio formation:

1. Mean reversion in fundamentals \((g_{it})\) make expected growth decline.
2. Average realized growth is on top of expected (rational expectations)
Representative Heuristics

- Agents tend to give excessive probability to states in which a given heuristic characteristics is more prevalent.

- Examples:
  - “Florida” has more senior citizens $\implies$ overstate the frequency of senior citizens in Florida compared to reality.
  - A patient positive to a medical test may be “sick” $\implies$ doctors may overstate the likelihood of really being sick.
  - A firm with high past growth tend to be a “growth firm” $\implies$ analysts may overstate the probability that it will have high growth going forward compared to real probability (too many Googles).
Representative Heuristics

- Agents tend to give excessive probability to states in which a given heuristic characteristics is more prevalent.

- Examples:
  - “Florida” has more senior citizens $\implies$ overstate the frequency of senior citizens in Florida compared to reality.
  - A patient positive to a medical test may be “sick” $\implies$ doctors may overstate the likelihood of really being sick.
  - A firm with high past growth tend to be a “growth firm” $\implies$ analysts may overstate the probability that it will have high growth going forward compared to real probability (too many Googles).

- Formally:

  
  \[
  \text{Representativeness of } \tau \text{ in Group } G : R(\tau, G) = \frac{h(T = \tau|G)}{h(T = \tau| - G)}
  \]

- Agents use “distorted” probabilities:

  
  \[
  h^\theta(T = \tau|G) = h(T = \tau|G)R(\tau, G)^\theta Z
  \]
Representative Heuristics – Model

- Consider again the simpler case:

\[ x_{i,t} = g_i + \varepsilon_{it} \]
Representative Heuristics – Model

• Consider again the simpler case:

\[ x_{i,t} = g_i + \varepsilon_{it} \]

• If analyst observes a signal \( x_{i,t} > \hat{g}_{i,t-1} \) \( \implies \) Heuristics = “high growth firm” \( \implies \) increase likelihood of high future growth.
Representative Heuristics – Model

- Consider again the simpler case:

\[ x_{i,t} = g_i + \varepsilon_{it} \]

- If analyst observes a signal \( x_{i,t} > \hat{g}_{i,t-1} \Rightarrow \text{Heuristics} = \text{“high growth firm”} \Rightarrow \text{increase likelihood of high future growth.} \)

- If analysts observes a signal \( x_{i,t} < \hat{g}_{i,t-1} \Rightarrow \text{Heuristics} = \text{“low growth firm”} \Rightarrow \text{increase likelihood of low future growth.} \)
Representative Heuristics – Model

- Consider again the simpler case:

\[ x_{i,t} = g_i + \varepsilon_{it} \]

- If analyst observes a signal \( x_{i,t} > \hat{g}_{i,t-1} \implies \) Heuristics = “high growth firm” \( \implies \) increase likelihood of high future growth.

- If analysts observes a signal \( x_{i,t} < \hat{g}_{i,t-1} \implies \) Heuristics = “low growth firm” \( \implies \) increase likelihood of low future growth.

- Thanks to normal distribution, BGLS show:

\[
g_i \mid x_{i,t} \sim N \left( \hat{g}_{i,t-1} + (1 + \theta) \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma^2_\varepsilon} (x_{i,t} - \hat{g}_{i,t-1}), \ \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma^2_\varepsilon} \right)
\]
Representative Heuristics – Model

• Consider again the simpler case:

\[ x_{i,t} = g_i + \varepsilon_{it} \]

• If analyst observes a signal \( x_{i,t} > \hat{g}_{i,t-1} \) \( \implies \) Heuristics = “high growth firm” 
  \( \implies \) increase likelihood of high future growth.

• If analysts observes a signal \( x_{i,t} < \hat{g}_{i,t-1} \) \( \implies \) Heuristics = “low growth firm” 
  \( \implies \) increase likelihood of low future growth.

• Thanks to normal distribution, BGLS show:

\[
g_{i|x_{i,t}} \sim N \left( \hat{g}_{i,t-1} + (1 + \theta) \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2} (x_{i,t} - \hat{g}_{i,t-1}), \; \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2} \right)
\]

\[
\hat{g}_{i,t} \quad \hat{\sigma}_{i,t}^2
\]

\[
\hat{g}_{i,t-1} - \frac{(\hat{\sigma}_{i,t-1}^2)^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2}
\]
Representative Heuristics – Model

- Consider again the simpler case:
  \[ x_{i,t} = g_i + \varepsilon_{it} \]

- If analyst observes a signal \( x_{i,t} > \hat{g}_{i,t-1} \) \( \implies \) Heuristics = “high growth firm” \( \implies \) increase likelihood of high future growth.

- If analysts observes a signal \( x_{i,t} < \hat{g}_{i,t-1} \) \( \implies \) Heuristics = “low growth firm” \( \implies \) increase likelihood of low future growth.

- Thanks to normal distribution, BGLS show:
  \[
  g_i | x_{i,t} \sim N \left( \hat{g}_{i,t-1} + (1 + \theta) \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2} (x_{i,t} - \hat{g}_{i,t-1}), \hat{\sigma}_{i,t-1}^2 - \frac{(\hat{\sigma}_{i,t-1}^2)^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_{\varepsilon}^2} \right)
  \]

- \( \theta = 0 \) \( \implies \) rational expectations model (Kalman filter)
Representative Heuristics – Model

• Consider again the simpler case:

\[ x_{i,t} = g_i + \varepsilon_{it} \]

• If analyst observes a signal \( x_{i,t} > \hat{g}_{i,t-1} \) \( \implies \) Heuristics = “high growth firm” \( \implies \) increase likelihood of high future growth.

• If analysts observes a signal \( x_{i,t} < \hat{g}_{i,t-1} \) \( \implies \) Heuristics = “low growth firm” \( \implies \) increase likelihood of low future growth.

• Thanks to normal distribution, BGLS show:

\[
g_{i \mid x_{i,t}} \sim N \left( \hat{g}_{i,t-1} + (1 + \theta) \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} (x_{i,t} - \hat{g}_{i,t-1}), \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} \right)
\]

\[ \begin{aligned} \hat{g}_{i,t} &\quad \hat{\sigma}_{i,t}^2 \\ \theta &\quad \hat{\sigma}_{i,t-1}^2 - \frac{(\hat{\sigma}_{i,t-1}^2)^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} \end{aligned} \]

\(- \theta = 0 \implies \) rational expectations model (Kalman filter)

\(- \theta > 0 \implies \) representative heuristic model (Kahneman filter?)
Representative Heuristics – Model

• Consider again the simpler case:

\[ x_{i,t} = g_i + \varepsilon_{it} \]

• If analyst observes a signal \( x_{i,t} > \hat{g}_{i,t-1} \implies \text{Heuristics} = \text{“high growth firm”} \implies \text{increase likelihood of high future growth.} \)

• If analysts observes a signal \( x_{i,t} < \hat{g}_{i,t-1} \implies \text{Heuristics} = \text{“low growth firm”} \implies \text{increase likelihood of low future growth.} \)

• Thanks to normal distribution, BGLS show:

\[ g_{i | x_{i,t}} \sim N \left( \hat{g}_{i,t-1} + (1 + \theta) \frac{\hat{\sigma}_{i,t-1}^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} (x_{i,t} - \hat{g}_{i,t-1}), \hat{\sigma}_{i,t-1}^2 - \frac{(\hat{\sigma}_{i,t-1})^2}{\hat{\sigma}_{i,t-1}^2 + \sigma_\varepsilon^2} \right) \]

\[ \hat{g}_{i,t} \]
\[ \hat{\sigma}_{i,t}^2 \]

– \( \theta = 0 \implies \text{rational expectations model (Kalman filter)} \)

– \( \theta > 0 \implies \text{representative heuristic model (Kahneman filter?)} \)

• BGLS formally analyze the impact of \( \theta > 0 \) on expectation, stock returns, etc.
Diagnostic Kalman Filter, $\theta = 2$
Representative Heuristics – Simulations – 5

Diagnostic Kalman Filter, $\theta = 5$

Expected growth overshoots

- HLTG: Average Realized
- HLTG: Average Expected
- LLTG: Average Realized
- LLTG: Average Expected
Representative Heuristics -- Simulations -- 5

Expected growth does not return to normal

Diagnostic Kalman Filter, $\theta = 5$

Expected growth overshoots

Expected growth declines after formation even without mean reversion
Expected growth overshoots

Realized growth < Expected growth

Diagnostic Kalman Filter, $\theta = 5$

- HLTG: Average Realized
- HLTG: Average Expected
- LLTG: Average Realized
- LLTG: Average Expected

Expected growth declines after formation even without mean reversion.
Representative Heuristics – Implications

• For the HTLG portfolio:
  – Before formation:
    1. Expected growth increase even more rapidly
  – After formation:
    1. Expected growth declines even without mean reversion
    2. Realized growth is below expected growth
    3. Realized growth similar across $\theta$’s
Representative Heuristics – Implications

- For the HTLG portfolio:
  - Before formation:
    1. Expected growth increase even more rapidly
  - After formation:
    1. Expected growth declines even without mean reversion
    2. Realized growth is below expected growth
    3. Realized growth similar across $\theta$’s

- If stock price is $P_t = D_t F(\hat{g}_t^\theta)$ with $F' > 0 \implies$ similar implications must hold for realized stock returns:
  - For HTLG firms,
    1. Realized stock returns are higher than expected before formation
    2. Realized stock return are lower than expected after formation
Representative Heuristics – Implications

- For the HTLG portfolio:
  - Before formation:
    1. Expected growth increase even more rapidly
  - After formation:
    1. Expected growth declines even without mean reversion
    2. Realized growth is below expected growth
    3. Realized growth similar across θ’s

- If stock price is $P_t = D_t F(\hat{g}_t^\theta)$ with $F' > 0 \implies$ similar implications must hold for realized stock returns:
  - For HTLG firms,
    1. Realized stock returns are higher than expected before formation
    2. Realized stock return are lower than expected after formation

- Most important contribution of the model is about distorted beliefs.
  - The other implications (e.g. on stock returns) occur in other models.
Reverse Causality: Learning from Prices

• Analysts got to look at stock prices when they form their predictions
  – Stock price goes up $\implies$ “It got to be high growth stock!”
Reverse Causality: Learning from Prices

• Analysts got to look at stock prices when they form their predictions
  – Stock price goes up $\implies$ “It got to be high growth stock!”

• What if analysts learn from prices too?
Reverse Causality: Learning from Prices

- Analysts got to look at stock prices when they form their predictions
  - Stock price goes up $\Rightarrow$ “It got to be high growth stock!”

- What if analysts learn from prices too?

- Same model as before but with:
  - Unobservable mean reverting growth $g_{it}$ (as earlier)
Reverse Causality: Learning from Prices

• Analysts got to look at stock prices when they form their predictions
  – Stock price goes up $\implies$ “It got to be high growth stock!”

• What if analysts learn from prices too?

• Same model as before but with:
  – Unobservable mean reverting growth $g_{it}$ (as earlier)
  – Unobservable mean reverting expected return $R_{it} = \delta_0 + r_{it}$:

$$r_{i,t+1} = \delta_1 r_{it} + \eta_{i,t+1}^r$$
Reverse Causality: Learning from Prices

- Analysts got to look at stock prices when they form their predictions
  - Stock price goes up $\implies$ “It got to be high growth stock!”

- What if analysts learn from prices too?

- Same model as before but with:
  - Unobservable mean reverting growth $g_{it}$ (as earlier)
  - Unobservable mean reverting expected return $R_{it} = \delta_0 + r_{it}$:
    \[ r_{i,t+1} = \delta_1 r_{it} + \eta^r_{i,t+1} \]

- Analysts observe the (log) price-dividend ratio $pd_{i,t}$. From Campbell and Shiller:
  \[ pd_{it} = A + B_g g_{it} - B_r r_{it} + \varepsilon^{pd}_t \]
Reverse Causality: Learning from Prices

• Analysts got to look at stock prices when they form their predictions
  – Stock price goes up $\implies$ “It got to be high growth stock!”

• What if analysts learn from prices too?

• Same model as before but with:
  – Unobservable mean reverting growth $g_{it}$ (as earlier)
  – Unobservable mean reverting expected return $R_{it} = \delta_0 + r_{it}$:
    $$r_{i,t+1} = \delta_1 r_{it} + \eta_{i,t+1}^r$$

• Analysts observe the (log) price-dividend ratio $pd_{i,t}$. From Campbell and Shiller:
  $$pd_{it} = A + B_g g_{it} - B_r r_{it} + \varepsilon_{t}^{pd}$$

• where $\varepsilon_{t}^{pd}$ is an approximation error, $\rho = e^{pd}/(1 + e^{pd}); \kappa = \log(1 + e^{pd}) - \rho pd$:
  $$A = (\kappa + \delta_0 - \gamma_0)/(1 - \rho); \quad B_g = 1/(1 - \rho \gamma_1); \quad B_r = 1/(1 - \rho \delta_1);$$
Reverse Causality: Learning from Prices

- Analysts got to look at stock prices when they form their predictions
  - Stock price goes up $\implies$ “It got to be high growth stock!”

- What if analysts learn from prices too?
- Same model as before but with:
  - Unobservable mean reverting growth $g_{it}$ (as earlier)
  - Unobservable mean reverting expected return $R_{it} = \delta_0 + r_{it}$:
    \[
    r_{i,t+1} = \delta_1 r_{it} + \eta_{i,t+1}^r
    \]

- Analysts observe the (log) price-dividend ratio $pd_{i,t}$. From Campbell and Shiller:
  \[
  pd_{it} = A + B_g g_{it} - B_r R_{it} + \varepsilon_{td}^{pd}
  \]
  where $\varepsilon_{td}^{pd}$ is an approximation error, $\rho = e^{pd}/(1+e^{pd})$; $\kappa = \log(1+e^{pd}) - \rho pd$:
  \[
  A = (\kappa + \delta_0 - \gamma_0)/(1 - \rho) ; \quad B_g = 1/(1 - \rho \gamma_1) ; \quad B_r = 1/(1 - \rho \delta_1) ;
  \]

- **Key:** Higher stock price due to higher $g_{it}$ or lower $r_{it}$. 
Reverse Causality: Learning from Prices – Filtering

• Two state equations and two observation equations:

\[
\begin{align*}
\begin{pmatrix} g_{i,t+1} \\ r_{i,t+1} \end{pmatrix} &= \begin{pmatrix} \gamma_1 & 0 \\ 0 & \delta_1 \end{pmatrix} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma_g & 0 \\ 0 & \sigma_r \end{pmatrix} \begin{pmatrix} \eta^g_{i,t} \\ \eta^r_{i,t} \end{pmatrix} \\
\begin{pmatrix} x_{i,t} \\ p_{d,i,t} \end{pmatrix} &= \begin{pmatrix} \delta_0 \\ A \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ B_g & -B_r \end{pmatrix} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma^g_{\varepsilon} & 0 \\ 0 & \sigma^{pd}_{\varepsilon} \end{pmatrix} \begin{pmatrix} \varepsilon^g_{i,t} \\ \varepsilon^{pd}_{i,t} \end{pmatrix}
\end{align*}
\]
Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

\[
\begin{pmatrix}
    g_{i,t+1} \\
    r_{i,t+1}
\end{pmatrix} =
\begin{pmatrix}
    \gamma_1 & 0 \\
    0 & \delta_1
\end{pmatrix}
\begin{pmatrix}
    g_{i,t} \\
    r_{i,t}
\end{pmatrix} +
\begin{pmatrix}
    \sigma_g & 0 \\
    0 & \sigma_r
\end{pmatrix}
\begin{pmatrix}
    \eta_{i,t}^g \\
    \eta_{i,t}^r
\end{pmatrix}
\]

\[
\begin{pmatrix}
    x_{i,t} \\
    pd_{i,t}
\end{pmatrix} =
\begin{pmatrix}
    \delta_0 \\
    A
\end{pmatrix} +
\begin{pmatrix}
    1 & 0 \\
    B_g & -B_r
\end{pmatrix}
\begin{pmatrix}
    g_{i,t} \\
    r_{i,t}
\end{pmatrix} +
\begin{pmatrix}
    \sigma_{\varepsilon}^g & 0 \\
    0 & \sigma_{pd}^{pd}
\end{pmatrix}
\begin{pmatrix}
    \varepsilon_{i,t}^g \\
    \varepsilon_{i,t}^{pd}
\end{pmatrix}
\]

- Given the Kalman gain \( K_t = ZP_t T' (ZP_t Z' + \Sigma \Sigma')^{-1} \), the filter for \( g_{it} \) is:

\[
\hat{g}_{i,t} = \gamma_1 \hat{g}_{i,t-1} + K_{11} \left[ x_{i,t} - E_{t-1} (x_{i,t}) \right] + K_{12} \left[ pd_{i,t} - E_{t-1} (pd_{i,t}) \right]
\]
Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

\[
\begin{pmatrix}
g_{i,t+1} \\
r_{i,t+1}
\end{pmatrix}
= \begin{pmatrix}
\gamma_1 & 0 \\
0 & \delta_1
\end{pmatrix}
\begin{pmatrix}
g_{i,t} \\
r_{i,t}
\end{pmatrix}
+ \begin{pmatrix}
\sigma_g & 0 \\
0 & \sigma_r
\end{pmatrix}
\begin{pmatrix}
\eta_{i,t}^g \\
\eta_{i,t}^r
\end{pmatrix}
\]

\[
\begin{pmatrix}
x_{i,t} \\
pd_{i,t}
\end{pmatrix}
= \begin{pmatrix}
\delta_0 \\
A
\end{pmatrix}
+ \begin{pmatrix}
1 & 0 \\
B_g & -B_r
\end{pmatrix}
\begin{pmatrix}
g_{i,t} \\
r_{i,t}
\end{pmatrix}
+ \begin{pmatrix}
\sigma_{\varepsilon}^g & 0 \\
0 & \sigma_{\varepsilon}^{pd}
\end{pmatrix}
\begin{pmatrix}
\varepsilon_{i,t}^g \\
\varepsilon_{i,t}^{pd}
\end{pmatrix}
\]

- Given the Kalman gain \( K_t = ZP_t T'(ZP_t Z' + \Sigma \Sigma')^{-1} \), the filter for \( g_{it} \) is:

\[
\hat{g}_{i,t} = \gamma_1 \hat{g}_{i,t-1} + K_{11} [x_{i,t} - E_{t-1}(x_{i,t})] + K_{12} [pd_{i,t} - E_{t-1}(pd_{i,t})]
\]

- Higher \( \hat{g}_{i,t} \) occurs if:
Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

\[
\begin{pmatrix}
  g_{i,t+1} \\
r_{i,t+1}
\end{pmatrix} = \begin{pmatrix}
  \gamma_1 & 0 \\
  0 & \delta_1
\end{pmatrix} \begin{pmatrix}
  g_{i,t} \\
r_{i,t}
\end{pmatrix} + \begin{pmatrix}
  \sigma_g & 0 \\
  0 & \sigma_r
\end{pmatrix} \begin{pmatrix}
  \eta_{i,t}^g \\
  \eta_{i,t}^r
\end{pmatrix}
\]

\[
\begin{pmatrix}
  x_{i,t} \\
pd_{i,t}
\end{pmatrix} = \begin{pmatrix}
  \delta_0 \\
  A
\end{pmatrix} + \begin{pmatrix}
  1 & 0 \\
  B_g & -B_r
\end{pmatrix} \begin{pmatrix}
  g_{i,t} \\
r_{i,t}
\end{pmatrix} + \begin{pmatrix}
  \sigma_{g \varepsilon} & 0 \\
  0 & \sigma_{pd \varepsilon}
\end{pmatrix} \begin{pmatrix}
  \varepsilon_{i,t}^g \\
  \varepsilon_{i,t}^{pd}
\end{pmatrix}
\]

- Given the Kalman gain \( K_t = Z P_t T' (Z P_t Z' + \Sigma \Sigma')^{-1} \), the filter for \( g_{it} \) is:

\[
\hat{g}_{i,t} = \gamma_1 \hat{g}_{i,t-1} + K_{11} [x_{i,t} - E_{t-1}(x_{i,t})] + K_{12} [pd_{i,t} - E_{t-1}(pd_{i,t})]
\]

- Higher \( \hat{g}_{i,t} \) occurs if:

1. Realized growth is higher than expected: \( x_{i,t} > E_{t-1}(x_{i,t}) \)
Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

\[
\begin{pmatrix}
    g_{i,t+1} \\
    r_{i,t+1}
\end{pmatrix}
= 
\begin{pmatrix}
    \gamma_1 & 0 \\
    0 & \delta_1
\end{pmatrix}
\begin{pmatrix}
    g_{i,t} \\
    r_{i,t}
\end{pmatrix}
+ 
\begin{pmatrix}
    \sigma_g & 0 \\
    0 & \sigma_r
\end{pmatrix}
\begin{pmatrix}
    \eta_{i,t}^g \\
    \eta_{i,t}^r
\end{pmatrix}
\]

\[
\begin{pmatrix}
    x_{i,t} \\
    pd_{i,t}
\end{pmatrix}
= 
\begin{pmatrix}
    \delta_0 \\
    A
\end{pmatrix}
+ 
\begin{pmatrix}
    1 & 0 \\
    B_g & -B_r
\end{pmatrix}
\begin{pmatrix}
    g_{i,t} \\
    r_{i,t}
\end{pmatrix}
+ 
\begin{pmatrix}
    \sigma_{g} & 0 \\
    0 & \sigma_{pd}
\end{pmatrix}
\begin{pmatrix}
    \varepsilon_{i,t}^{g} \\
    \varepsilon_{i,t}^{pd}
\end{pmatrix}
\]

- Given the Kalman gain \( K_t = ZP_t T' (ZP_t Z' + \Sigma \Sigma')^{-1} \), the filter for \( g_{it} \) is:

\[
\hat{g}_{i,t} = \gamma_1 \hat{g}_{i,t-1} + K_{11} [x_{i,t} - E_{t-1}(x_{i,t})] + K_{12} [pd_{i,t} - E_{t-1}(pd_{i,t})]
\]

- Higher \( \hat{g}_{i,t} \) occurs if:

1. Realized growth is higher than expected: \( x_{i,t} > E_{t-1}(x_{i,t}) \)
2. Realized return is higher than expected: \( pd_{i,t} > E_{t-1}(pd_{i,t}) \)
Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

\[
\begin{pmatrix}
g_{i,t+1} \\
r_{i,t+1}
\end{pmatrix} = \underbrace{\begin{pmatrix}
\gamma_1 & 0 \\
0 & \delta_1
\end{pmatrix}}_{T} \begin{pmatrix}
g_{i,t} \\
r_{i,t}
\end{pmatrix} + \underbrace{\begin{pmatrix}
\sigma_g & 0 \\
0 & \sigma_r
\end{pmatrix}}_{Z} \begin{pmatrix}
\eta_{i,t}^g \\
\eta_{i,t}^r
\end{pmatrix}
\]

\[
\begin{pmatrix}
x_{i,t} \\
pd_{i,t}
\end{pmatrix} = \begin{pmatrix}
\delta_0 \\
A
\end{pmatrix} + \underbrace{\begin{pmatrix}
1 & 0 \\
B_g & -B_r
\end{pmatrix}}_{Z} \begin{pmatrix}
g_{i,t} \\
r_{i,t}
\end{pmatrix} + \underbrace{\begin{pmatrix}
\sigma_{\varepsilon}^g & 0 \\
0 & \sigma_{\varepsilon}^{pd}
\end{pmatrix}}_{\Sigma} \begin{pmatrix}
\varepsilon_{i,t}^g \\
\varepsilon_{i,t}^{pd}
\end{pmatrix}
\]

- Given the Kalman gain \( K_t = ZP_t T' (ZP_t Z' + \Sigma \Sigma')^{-1} \), the filter for \( g_{i,t} \) is:

\[
\hat{g}_{i,t} = \gamma_1 \hat{g}_{i,t-1} + K_{11} [x_{i,t} - E_{t-1}(x_{i,t})] + K_{12} [pd_{i,t} - E_{t-1}(pd_{i,t})]
\]

- Higher \( \hat{g}_{i,t} \) occurs if:

1. Realized growth is higher than expected: \( x_{i,t} > E_{t-1}(x_{i,t}) \)
2. Realized return is higher than expected: \( pd_{i,t} > E_{t-1}(pd_{i,t}) \)

- The second effect comes from a decrease in expected return \( r_{it} \).

  - High growth stocks \( \implies \) forecast lower future return
Same implications for realized and expected growth as in previous cases.
Reverse Causality: Learning from Prices – Returns

Realized and Expected Return

Rate of return

months

HLTG: Realized
HLTG: Expected
LLTG: Realized
LLTG: Expected
• HLTG portfolio:
  1. **Pre-ranking**: Higher realized returns and *declining* expected return
- HLTG portfolio:
  1. Pre-ranking: Higher realized returns and *declining* expected return
  2. **Post-ranking:** Very low realized returns, consistent with low expected return
• HLTG portfolio:

1. Pre-ranking: Higher realized returns and *declining* expected return
2. Post-ranking: Very low realized returns, consistent with low expected return
Reverse Causality: Learning from Prices – Returns

**Figure 5 BGLS:** Twelve-day Returns on Earnings Announcements for LTG Portfolios. (DATA)
Reverse Causality: Learning from Prices – Average Returns

Postformation Portfolio Average Returns

Annualized Returns

Deciles

1 (LLGT)
2
3
4
5
6
7
8
9
10 (HLGT)
Reverse Causality: Learning from Prices – Average Returns

Figure 1 BGLS: Annual Returns for Portfolios Formed on LTG (DATA)
What this Model Cannot Explain? Systematic Forecast Errors

Figure 4 BGLS: LTG Forecast Errors (DATA)

Negative Forecast Error
Final Comments

- Main evidence for the model is on beliefs: forecast errors and over-reaction.
  - What does this evidence imply for the magnitude of $\theta$?
  - Is $\theta$ calibrated from psychology able to quantitatively rationalize beliefs dynamics and asset prices?
    * The model is “portable”: Can we use estimates of $\theta$ from other studies?
    * In JF paper, $\theta$ is estimated to $\theta = 0.91$. What about here?
Final Comments

• Main evidence for the model is on beliefs: forecast errors and over-reaction.
  – What does this evidence imply for the magnitude of $\theta$?
  – Is $\theta$ calibrated from psychology able to quantitatively rationalize beliefs dynamics and asset prices?
    * The model is “portable”: Can we use estimates of $\theta$ from other studies?
    * In JF paper, $\theta$ is estimated to $\theta = 0.91$. What about here?

• I love the objective of the paper:

  “We stress what we see as the central point: the theory of asset pricing can incorporate fundamental psychological insights while retaining the rigor and the predictive discipline of rational expectations models” (page 38, Conclusions)

  – This is great. The “next step” is to quantitatively assess its properties.
  – All puzzles in asset pricing (“equity premium”, “excess volatility”, “value-spread”, etc) are quantitative puzzles.
  – How far does this theory go to explain the facts for plausible parameters?
Reverse Causality: Learning from Prices – Growth

Figure 3 BGLS: Evolution of LTG (DATA)