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 = unobservable long-term growth

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

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

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}:

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

- Conditional on F_{t-1}:

$$\begin{pmatrix} x_{i,t} \\ g_i \end{pmatrix} \sim N\left(\begin{pmatrix} \widehat{g}_{i,t-1} \\ \widehat{g}_{i,t-1} \end{pmatrix}, \begin{pmatrix} \widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon & , & \widehat{\sigma}^2_{i,t-1} \\ \widehat{\sigma}^2_{i,t-1} & , & \widehat{\sigma}^2_{i,t-1} \end{pmatrix}\right)$$

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}:

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

- Conditional on F_{t-1}:

$$\begin{pmatrix} x_{i,t} \\ g_i \end{pmatrix} \sim N\left(\begin{pmatrix} \widehat{g}_{i,t-1} \\ \widehat{g}_{i,t-1} \end{pmatrix}, \begin{pmatrix} \widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon} & , & \widehat{\sigma}^2_{i,t-1} \\ \widehat{\sigma}^2_{i,t-1} & , & \widehat{\sigma}^2_{i,t-1} \end{pmatrix}\right)$$

- From the properties of conditional normal distribution

$$g_i|_{x_{i,t}} \sim N\left(\widehat{g}_{i,t-1} + \frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}\left(x_{i,t} - \widehat{g}_{i,t-1}\right), \quad \widehat{\sigma}^2_{i,t-1} - \frac{\left(\widehat{\sigma}^2_{i,t-1}\right)^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}\right)$$

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}:

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

- Conditional on F_{t-1}:

$$\begin{pmatrix} x_{i,t} \\ g_i \end{pmatrix} \sim N\left(\begin{pmatrix} \widehat{g}_{i,t-1} \\ \widehat{g}_{i,t-1} \end{pmatrix}, \begin{pmatrix} \widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon, & \widehat{\sigma}^2_{i,t-1} \\ \widehat{\sigma}^2_{i,t-1}, & \widehat{\sigma}^2_{i,t-1} \end{pmatrix}\right)$$

- From the properties of conditional normal distribution

Covariance

$$g_i|_{x_{i,t}} \sim N\left(\widehat{g}_{i,t-1} + \frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}\left(x_{i,t} - \widehat{g}_{i,t-1}\right), \quad \widehat{\sigma}^2_{i,t-1} - \frac{\left(\widehat{\sigma}^2_{i,t-1}\right)^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}\right)$$

Variance

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}:

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

- Conditional on F_{t-1}:

$$\begin{pmatrix} \boxed{x_{i,t}} \\ g_i \end{pmatrix} \sim N\left(\begin{pmatrix} \boxed{\widehat{g}_{i,t-1}} \\ \widehat{g}_{i,t-1} \end{pmatrix}, \begin{pmatrix} \widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon & , & \widehat{\sigma}^2_{i,t-1} \\ \widehat{\sigma}^2_{i,t-1} & , & \widehat{\sigma}^2_{i,t-1} \end{pmatrix}\right)$$

- From the properties of conditional normal distribution

$$g_i|_{x_{i,t}} \sim N\left(\widehat{g}_{i,t-1} + \frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon} \underbrace{\boxed{\left(x_{i,t} - \widehat{g}_{i,t-1}\right)}}_{\text{News}}, \quad \widehat{\sigma}^2_{i,t-1} - \frac{\left(\widehat{\sigma}^2_{i,t-1}\right)^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}\right)$$

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}:

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

- Conditional on F_{t-1}:

$$\begin{pmatrix} x_{i,t} \\ g_i \end{pmatrix} \sim N\left(\begin{pmatrix} \widehat{g}_{i,t-1} \\ \widehat{g}_{i,t-1} \end{pmatrix}, \begin{pmatrix} \widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon & , & \widehat{\sigma}^2_{i,t-1} \\ \widehat{\sigma}^2_{i,t-1} & , & \widehat{\sigma}^2_{i,t-1} \end{pmatrix}\right)$$

- From the properties of conditional normal distribution

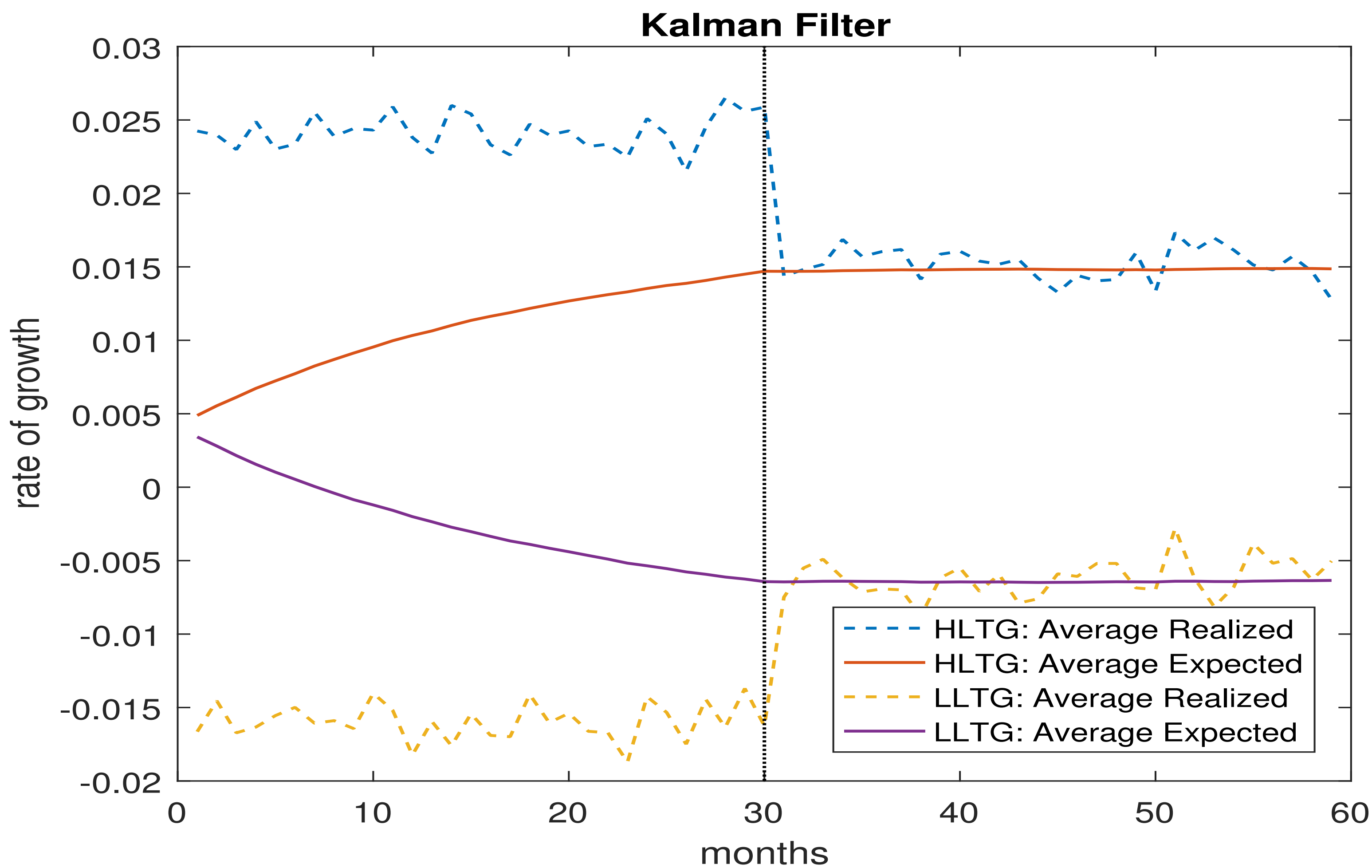
$$g_i|_{x_{i,t}} \sim N\left(\underbrace{\widehat{g}_{i,t-1} + \frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}\left(x_{i,t} - \widehat{g}_{i,t-1}\right)}_{\widehat{g}_{i,t}}, \quad \underbrace{\widehat{\sigma}^2_{i,t-1} - \frac{\left(\widehat{\sigma}^2_{i,t-1}\right)^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}}_{\widehat{\sigma}^2_{i,t}}\right)$$

Sample Selection and Realized vs. Expected Growth

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

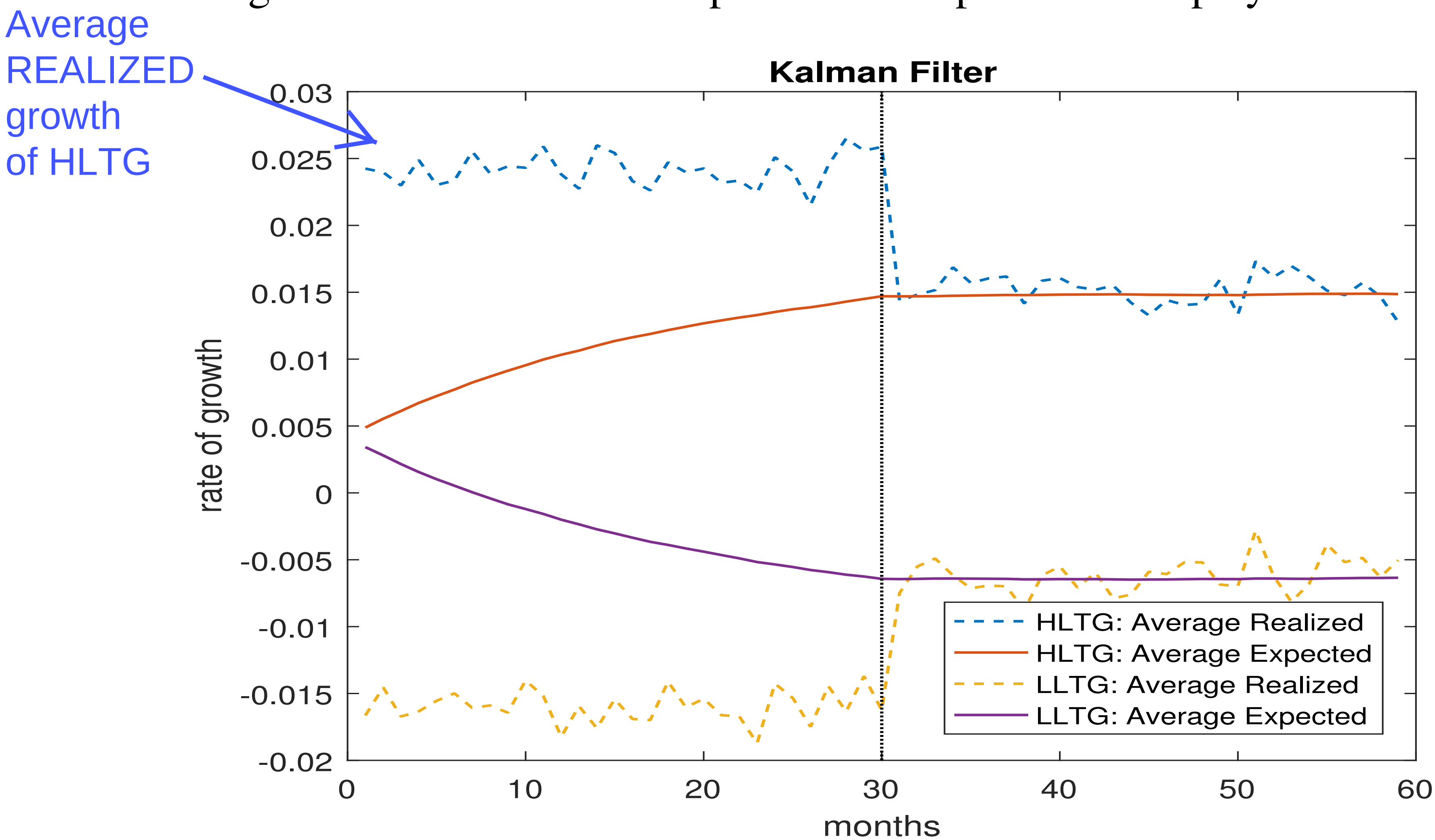
Sample Selection and Realized vs. Expected Growth

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



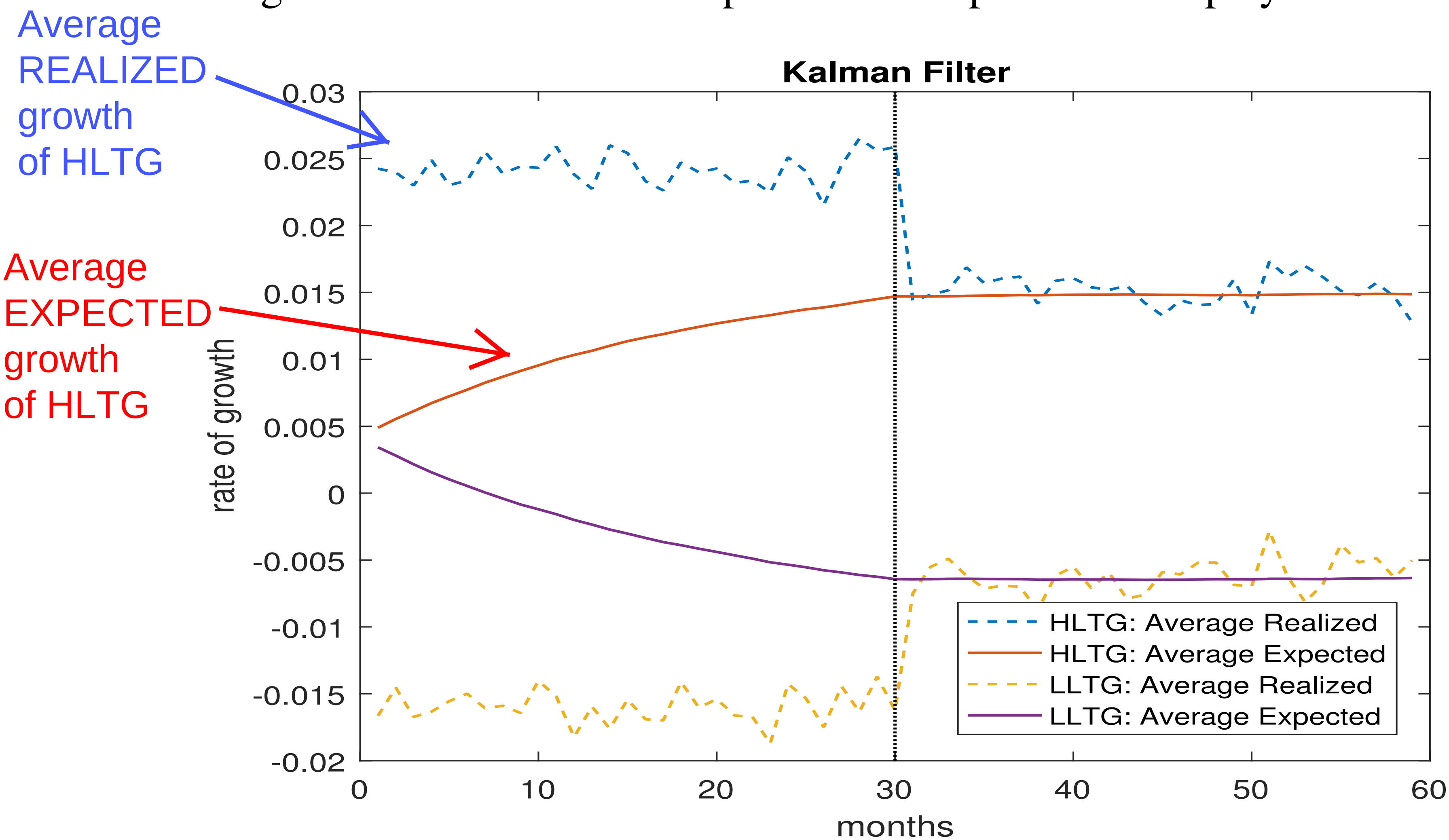
Sample Selection and Realized vs. Expected Growth

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



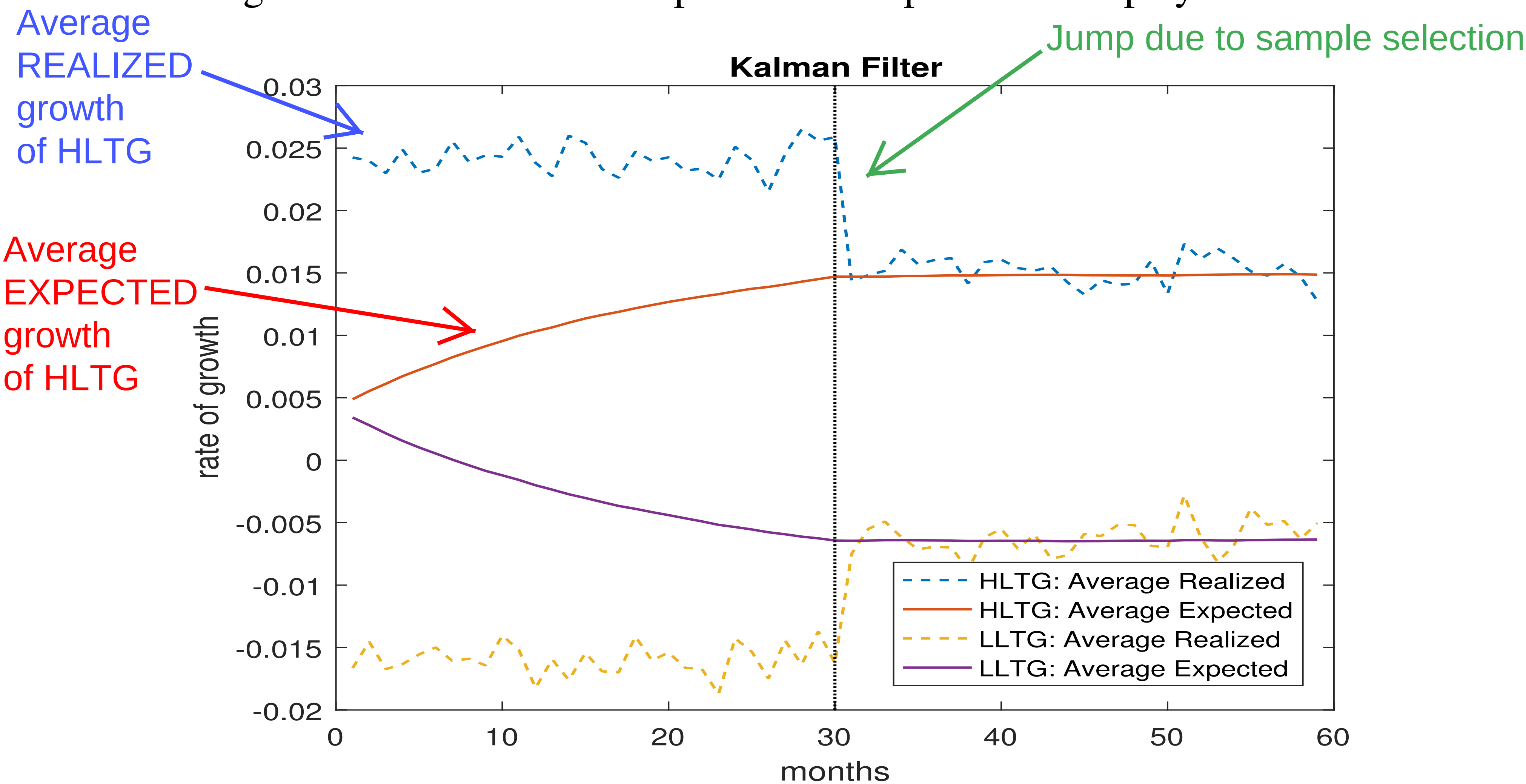
Sample Selection and Realized vs. Expected Growth

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



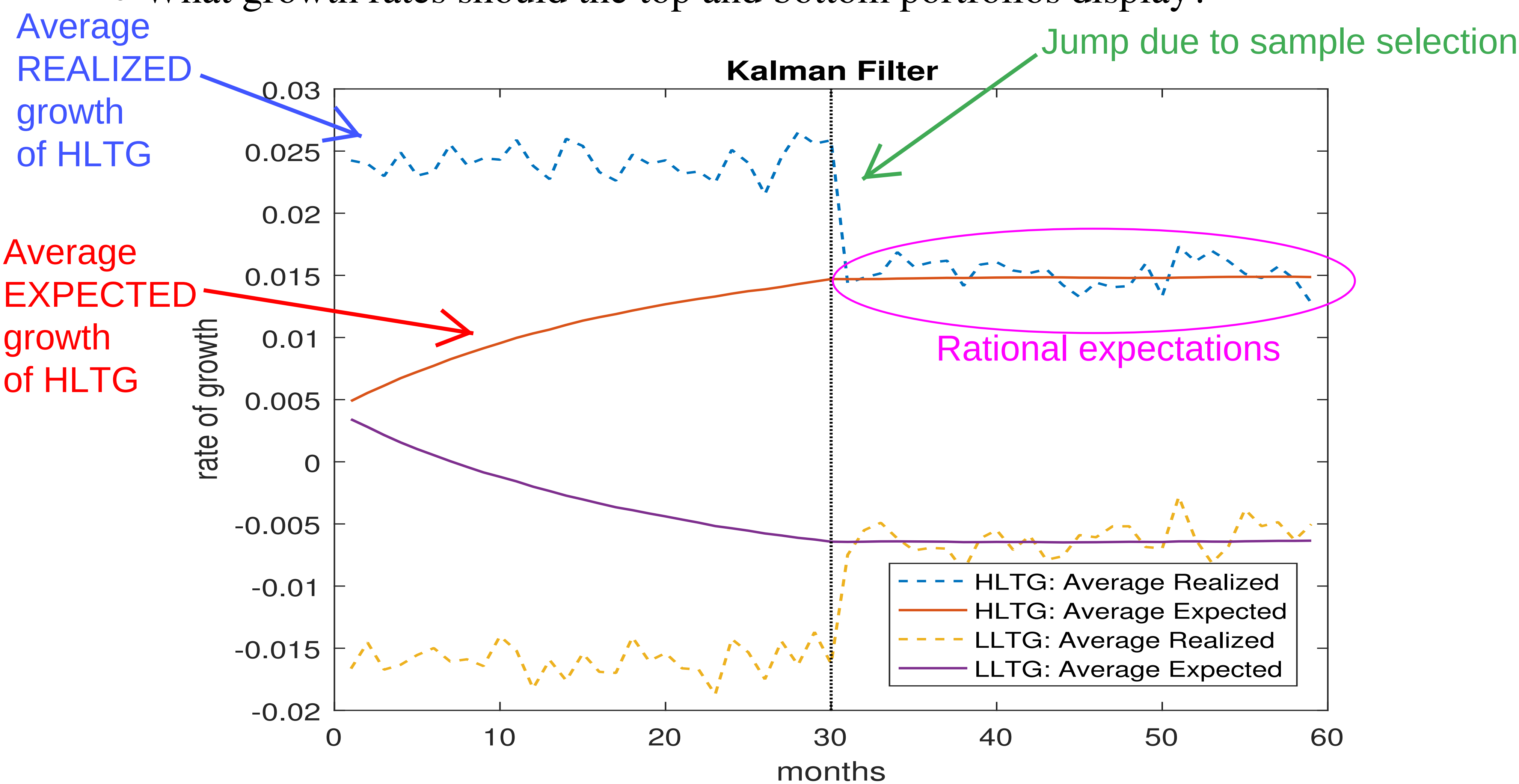
Sample Selection and Realized vs. Expected Growth

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



Sample Selection and Realized vs. Expected Growth – 2

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

Sample Selection and Realized vs. Expected Growth – 2

- 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

Sample Selection and Realized vs. Expected Growth – 2

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

Sample Selection and Realized vs. Expected Growth – 2

- 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 $\widehat{g}_{it}$ 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 (LLTG) porfolio 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:

$$
\begin{aligned}
x_{it} &= \gamma_0 + g_{it} + \varepsilon_{it} \\
g_{i,t+1} &= \gamma_1\, g_{it} + \eta_{i,t+1}
\end{aligned}
$$

Sample Selection and Realized vs. Expected Growth – 3

- Adding mean reversion generates more "curvy" results:

$$\begin{aligned} x_{it} &= \gamma_0 + g_{it} + \varepsilon_{it} \\ g_{i,t+1} &= \gamma_1 \, g_{it} + \eta_{i,t+1} \end{aligned}$$

- Conditional on F_{t-1}:

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

- After observing x_{it}, the filter is then

$$\begin{aligned} \widehat{g}_{i,t} &= \gamma_1 \, \widehat{g}_{i,t-1} + \frac{\gamma_1 \widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon} \left(x_{it} - \delta_0 - \widehat{g}_{i,t-1} \right) \\ \widehat{\sigma}^2_{i,t} &= \gamma_1^2 \, \widehat{\sigma}^2_{i,t-1} + \sigma^2_\eta - \frac{\left(\gamma_1 \widehat{\sigma}^2_{i,t-1}\right)^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon} \end{aligned}$$

Sample Selection and Realized vs. Expected Growth – 3

- Adding mean reversion generates more "curvy" results:

$$\begin{aligned} x_{it} &= \gamma_0 + g_{it} + \varepsilon_{it} \\ g_{i,t+1} &= \gamma_1\, g_{it} + \eta_{i,t+1} \end{aligned}$$

- Conditional on F_{t-1}:

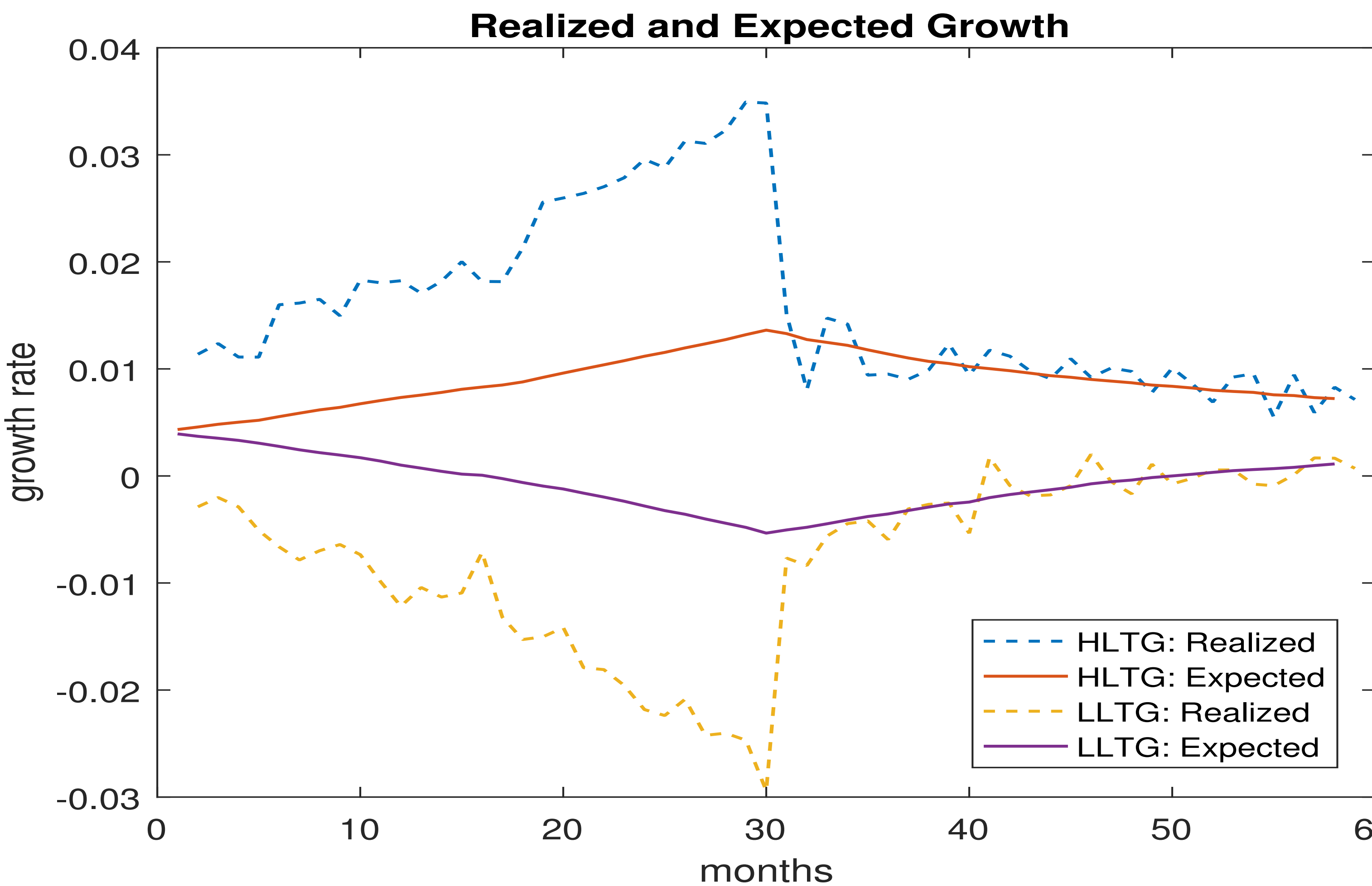
$$\begin{pmatrix} x_{i,t} \\ g_{i,t+1} \end{pmatrix} \sim N\left(\begin{pmatrix} \gamma_0 + \widehat{g}_{i,t-1} \\ \gamma_1\, \widehat{g}_{i,t-1} \end{pmatrix}, \begin{pmatrix} \widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon & , & \gamma_1\, \widehat{\sigma}^2_{i,t-1} \\ \gamma_1\, \widehat{\sigma}^2_{i,t-1}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

- After observing x_{it}, the filter is then

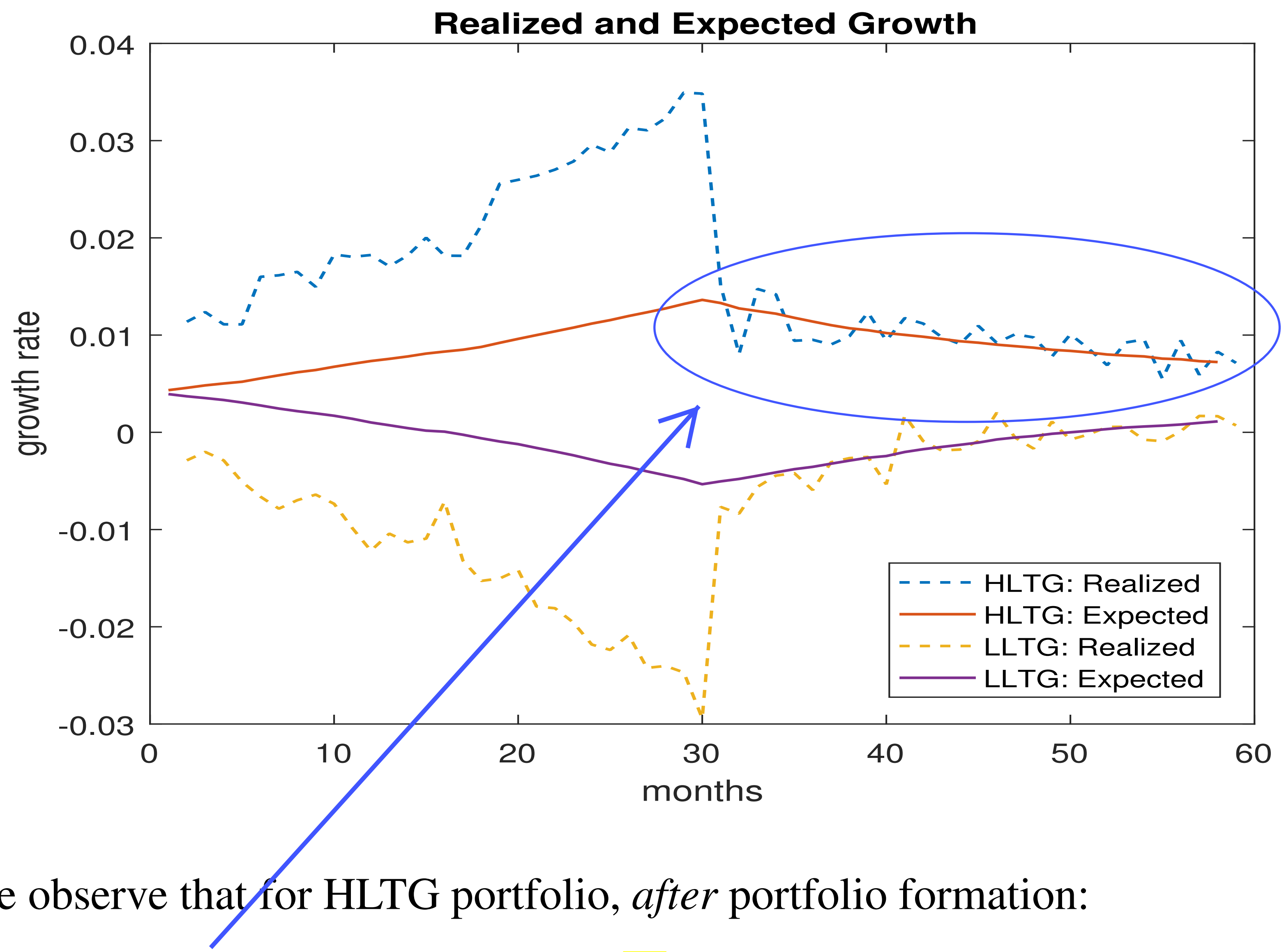
$$\begin{aligned} \widehat{g}_{i,t} &= \gamma_1\, \widehat{g}_{i,t-1} + \frac{\gamma_1 \widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon} \left(x_{it} - \delta_0 - \widehat{g}_{i,t-1} \right) \\ \widehat{\sigma}^2_{i,t} &= \gamma_1^2\, \widehat{\sigma}^2_{i,t-1} + \sigma^2_\eta - \frac{(\gamma_1 \widehat{\sigma}^2_{i,t-1})^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon} \end{aligned}$$

- Simulation with 10,000 firms as before, now with (annualized) $\gamma_0 = 5\%$, $\gamma_1 = 0.96$, $\sigma_\eta = 1\%$

Sample Selection and Realized vs. Expected Growth – 4

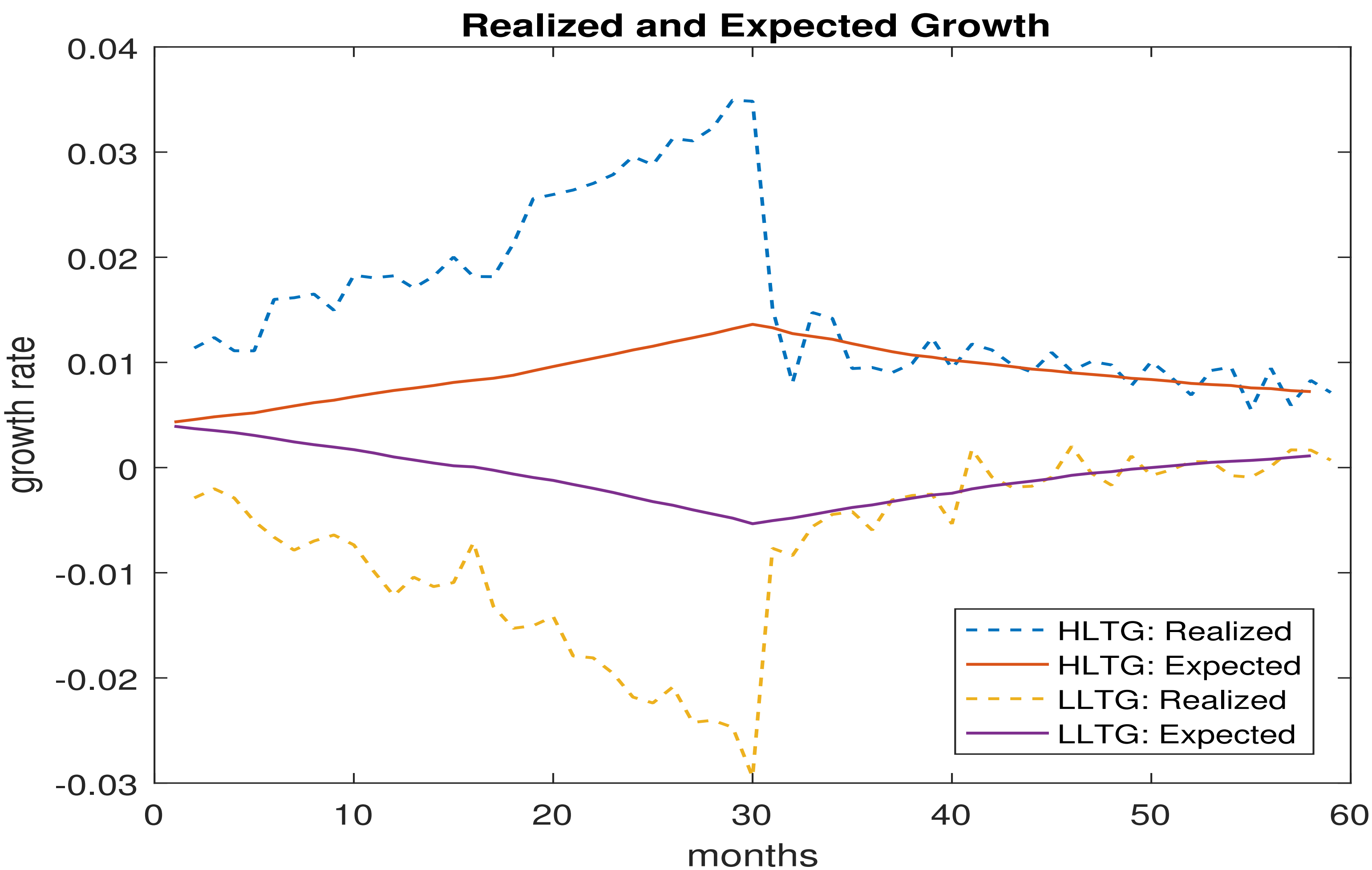


Sample Selection and Realized vs. Expected Growth – 4



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

Sample Selection and Realized vs. Expected Growth – 4



- 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 $\Longrightarrow$ overstate the frequency of senior citizens in Florida compared to reality.
  - A patient positive to a medical test may be "sick" $\Longrightarrow$ doctors may overstate the likelihood of really being sick.
  - A firm with high past growth tend to be a "growth firm" $\Longrightarrow$ 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 $\Longrightarrow$ overstate the frequency of senior citizens in Florida compared to reality.
  - A patient positive to a medical test may be "sick" $\Longrightarrow$ doctors may overstate the likelihood of really being sick.
  - A firm with high past growth tend to be a "growth firm" $\Longrightarrow$ 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 $x_{i,t} > \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics $=$ “high growth firm” $\Longrightarrow$ 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 $x_{i,t} > \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics $=$ "high growth firm" $\Longrightarrow$ increase likelihood of high future growth.
- If analysts observes a signal $x_{i,t} < \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics $=$ "low growth firm" $\Longrightarrow$ 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 $x_{i,t} > \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics $=$ "high growth firm" $\Longrightarrow$ increase likelihood of high future growth.
- If analysts observes a signal $x_{i,t} < \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics $=$ "low growth firm" $\Longrightarrow$ increase likelihood of low future growth.
- Thanks to normal distribution, BGLS show:

$$g_i|_{x_{i,t}} \sim N\left(\widehat{g}_{i,t-1} + (1+\theta)\frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}\left(x_{i,t} - \widehat{g}_{i,t-1}\right), \quad \widehat{\sigma}^2_{i,t-1} - \frac{(\widehat{\sigma}^2_{i,t-1})^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}\right)$$

Representative Heuristics – Model

- Consider again the simpler case:

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

- If analyst observes a signal $x_{i,t} > \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics = "high growth firm" $\Longrightarrow$ increase likelihood of high future growth.
- If analysts observes a signal $x_{i,t} < \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics = "low growth firm" $\Longrightarrow$ increase likelihood of low future growth.
- Thanks to normal distribution, BGLS show:

$$g_i|_{x_{i,t}} \sim N\left(\underbrace{\widehat{g}_{i,t-1} + (1+\theta)\frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}\left(x_{i,t} - \widehat{g}_{i,t-1}\right)}_{\widehat{g}^{\theta}_{i,t}},\quad \underbrace{\widehat{\sigma}^2_{i,t-1} - \frac{(\widehat{\sigma}^2_{i,t-1})^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}}_{\widehat{\sigma}^2_{i,t}}\right)$$

Representative Heuristics – Model

- Consider again the simpler case:

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

- If analyst observes a signal $x_{i,t} > \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics $=$ "high growth firm" $\Longrightarrow$ increase likelihood of high future growth.
- If analysts observes a signal $x_{i,t} < \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics $=$ "low growth firm" $\Longrightarrow$ increase likelihood of low future growth.
- Thanks to normal distribution, BGLS show:

$$g_i|_{x_{i,t}} \sim N\left(\underbrace{\widehat{g}_{i,t-1} + (1+\theta)\frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}\left(x_{i,t} - \widehat{g}_{i,t-1}\right)}_{\widehat{g}^{\theta}_{i,t}},\quad \underbrace{\widehat{\sigma}^2_{i,t-1} - \frac{(\widehat{\sigma}^2_{i,t-1})^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_{\varepsilon}}}_{\widehat{\sigma}^2_{i,t}}\right)$$

  - $\theta = 0 \Longrightarrow$ 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 $x_{i,t} > \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics = "high growth firm" $\Longrightarrow$ increase likelihood of high future growth.
- If analysts observes a signal $x_{i,t} < \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics = "low growth firm" $\Longrightarrow$ increase likelihood of low future growth.
- Thanks to normal distribution, BGLS show:

$$g_i|_{x_{i,t}} \sim N\left(\underbrace{\widehat{g}_{i,t-1} + (1+\theta)\frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1}+\sigma^2_\varepsilon}\left(x_{i,t} - \widehat{g}_{i,t-1}\right)}_{\widehat{g}^{\theta}_{i,t}},\quad \underbrace{\widehat{\sigma}^2_{i,t-1} - \frac{(\widehat{\sigma}^2_{i,t-1})^2}{\widehat{\sigma}^2_{i,t-1}+\sigma^2_\varepsilon}}_{\widehat{\sigma}^2_{i,t}}\right)$$

  - $\theta = 0 \Longrightarrow$ rational expectations model (Kalman filter)
  - $\theta > 0 \Longrightarrow$ representative heuristic model (Kahneman filter?)

Representative Heuristics – Model

- Consider again the simpler case:

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

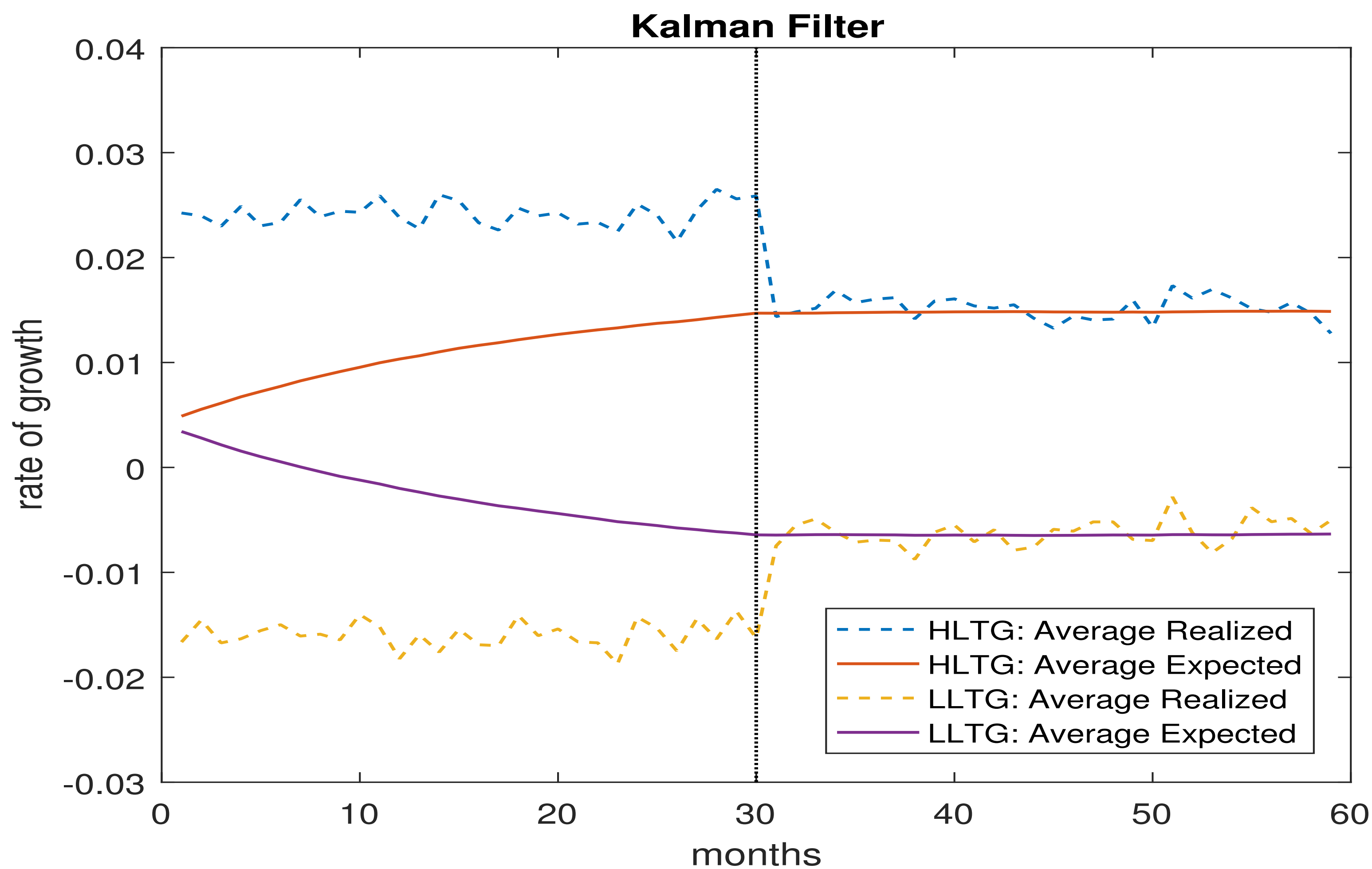
- If analyst observes a signal $x_{i,t} > \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics = "high growth firm" $\Longrightarrow$ increase likelihood of high future growth.
- If analysts observes a signal $x_{i,t} < \widehat{g}_{i,t-1} \Longrightarrow$ Heuristics = "low growth firm" $\Longrightarrow$ increase likelihood of low future growth.
- Thanks to normal distribution, BGLS show:

$$g_i|_{x_{i,t}} \sim N\left(\underbrace{\widehat{g}_{i,t-1} + (1+\theta)\frac{\widehat{\sigma}^2_{i,t-1}}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}\left(x_{i,t} - \widehat{g}_{i,t-1}\right)}_{\widehat{g}^{\theta}_{i,t}}, \quad \underbrace{\widehat{\sigma}^2_{i,t-1} - \frac{(\widehat{\sigma}^2_{i,t-1})^2}{\widehat{\sigma}^2_{i,t-1} + \sigma^2_\varepsilon}}_{\widehat{\sigma}^2_{i,t}}\right)$$

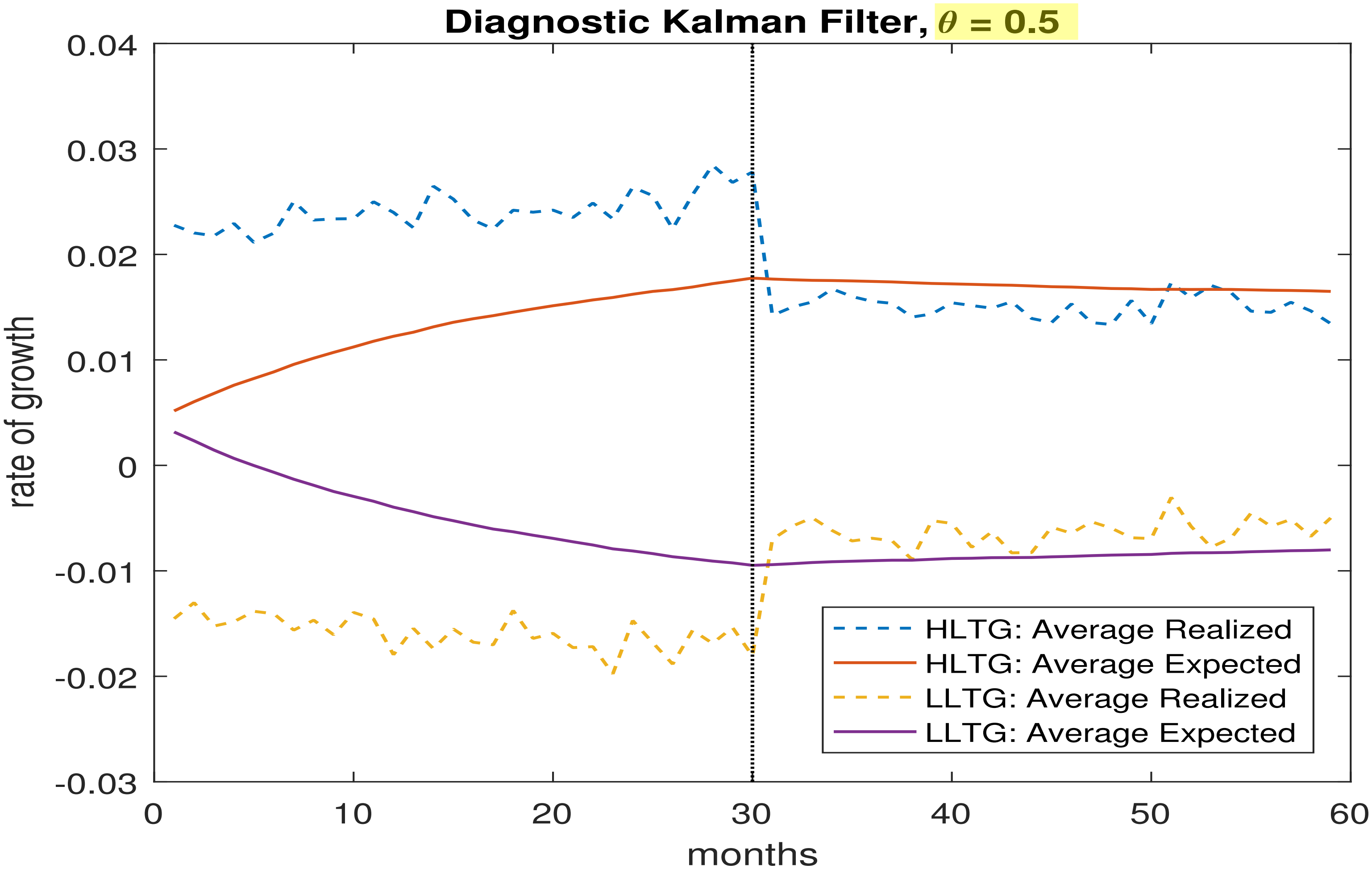
  - $\theta = 0 \Longrightarrow$ rational expectations model (Kalman filter)
  - $\theta > 0 \Longrightarrow$ representative heuristic model (Kahneman filter?)

- BGLS formally analyze the impact of $\theta > 0$ on expectation, stock returns, etc.

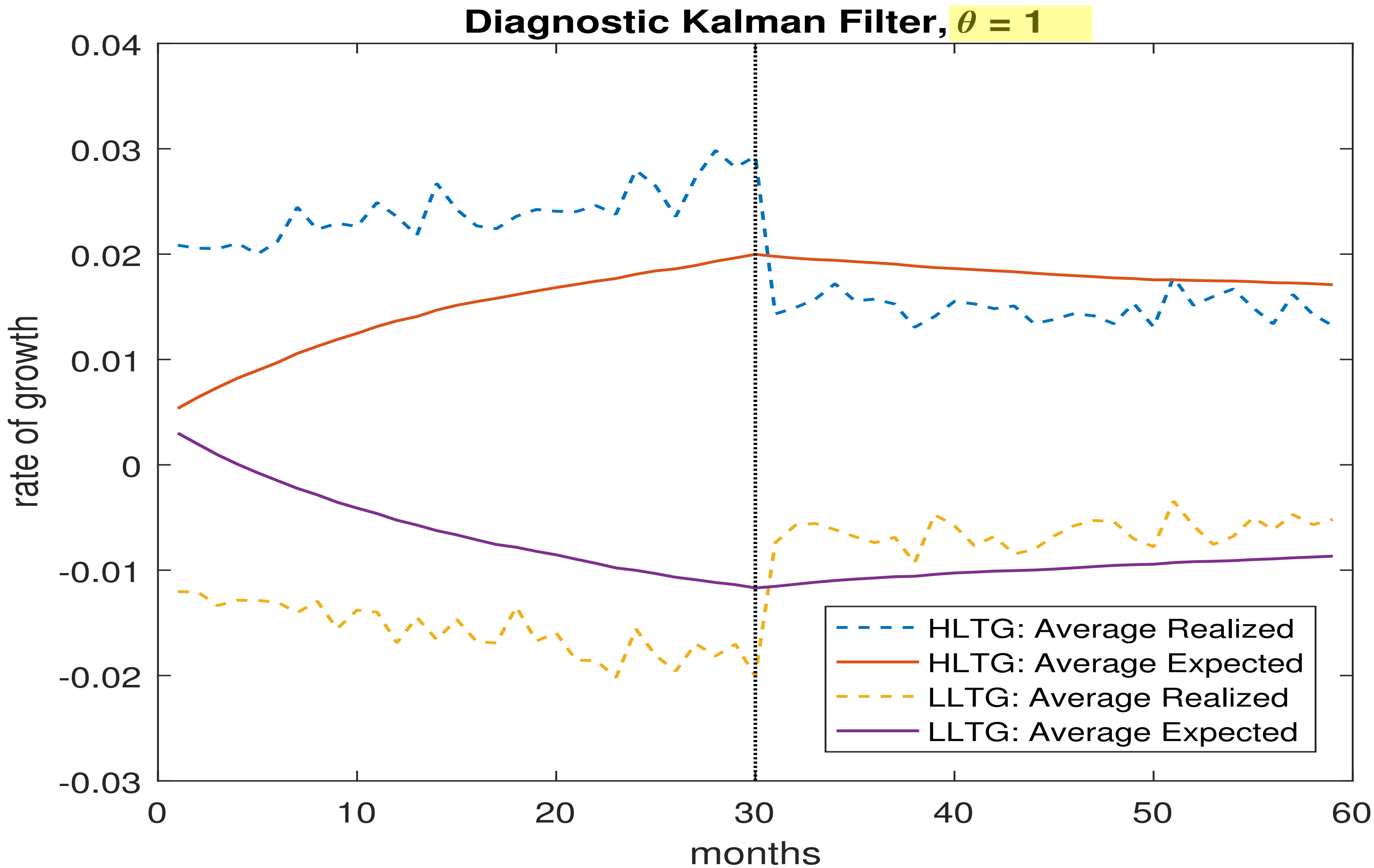
Representative Heuristics – Simulations – 1



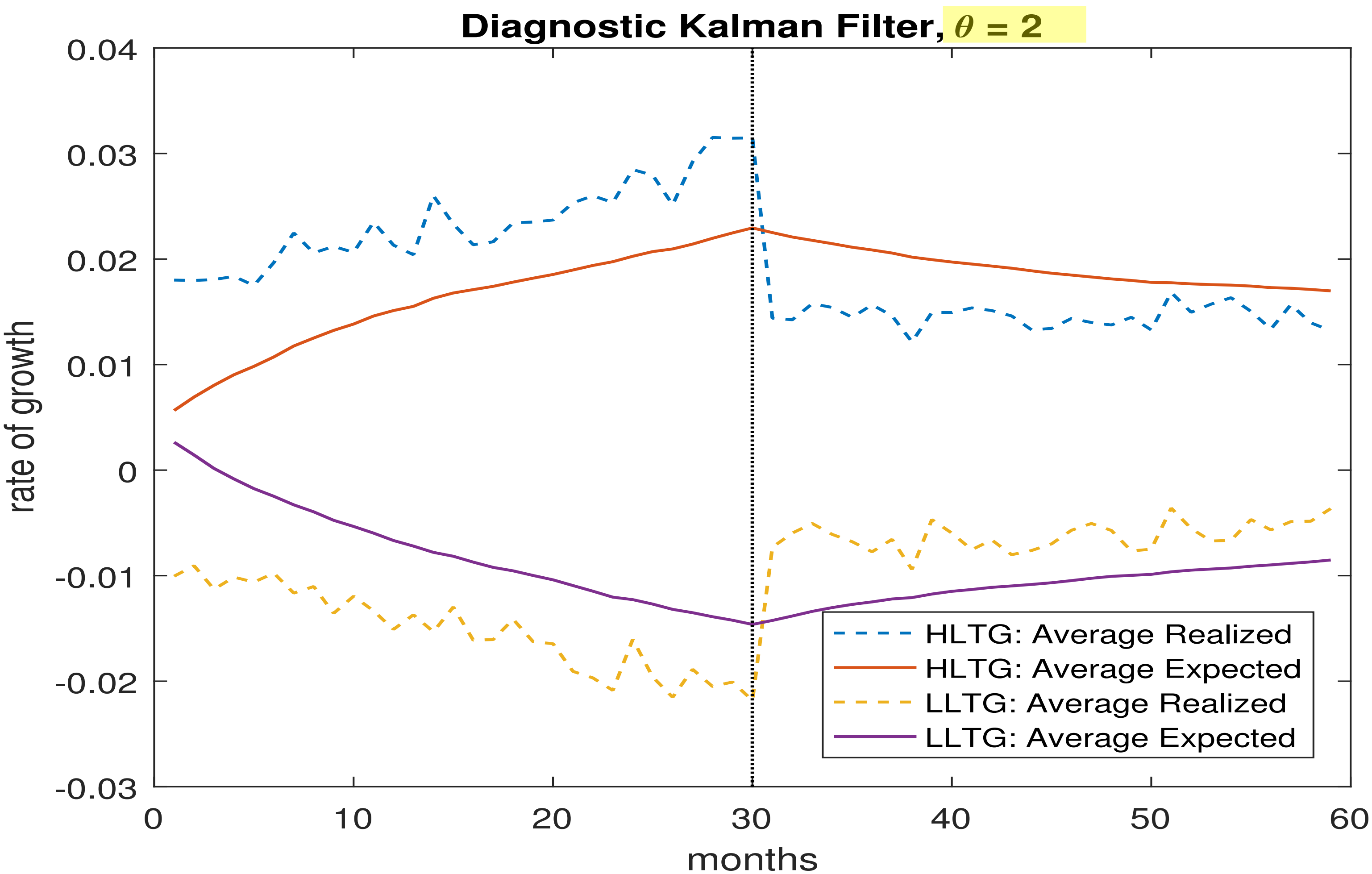
Representative Heuristics – – Simulations – 2



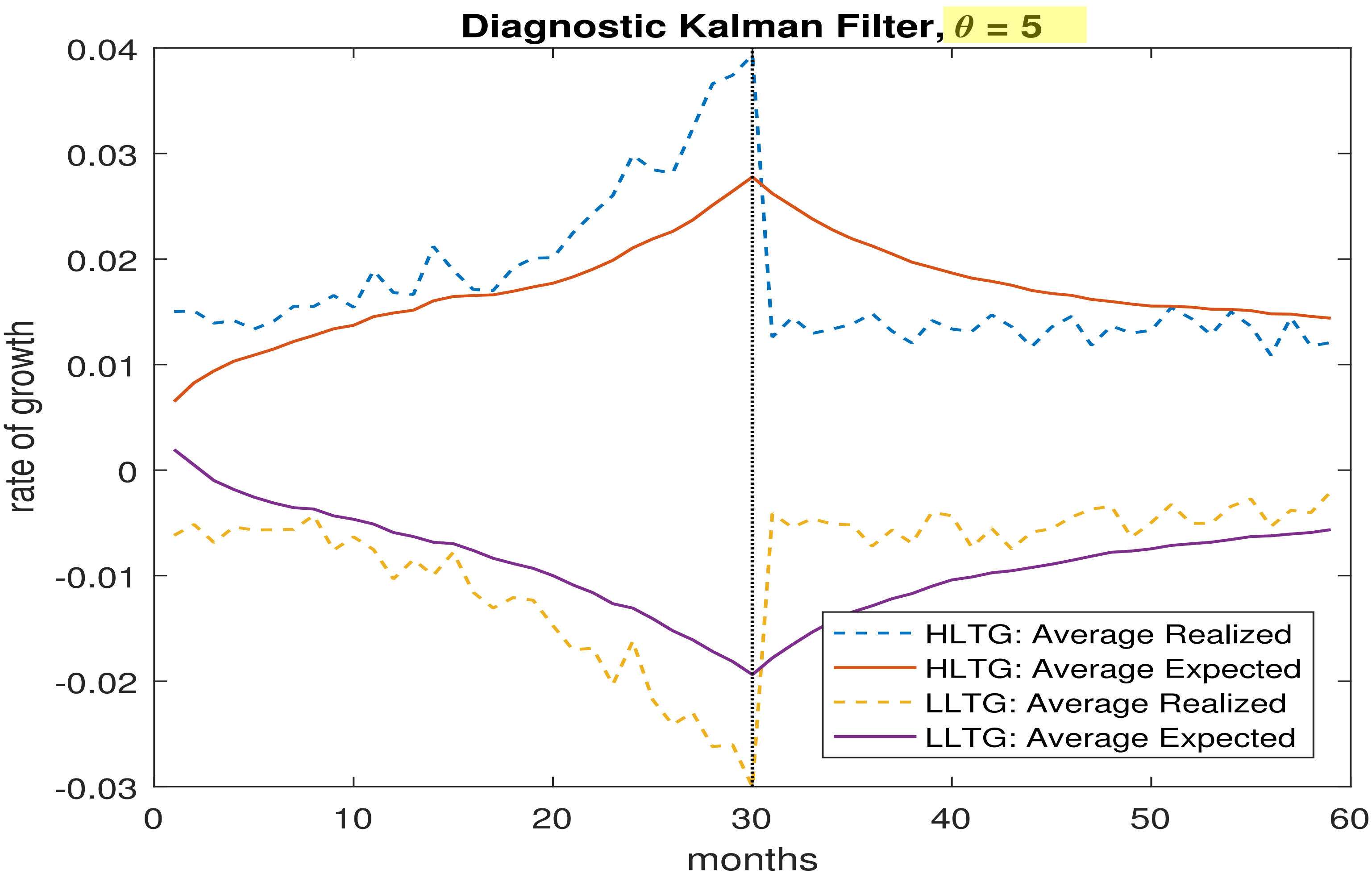
Representative Heuristics – – Simulations – 3



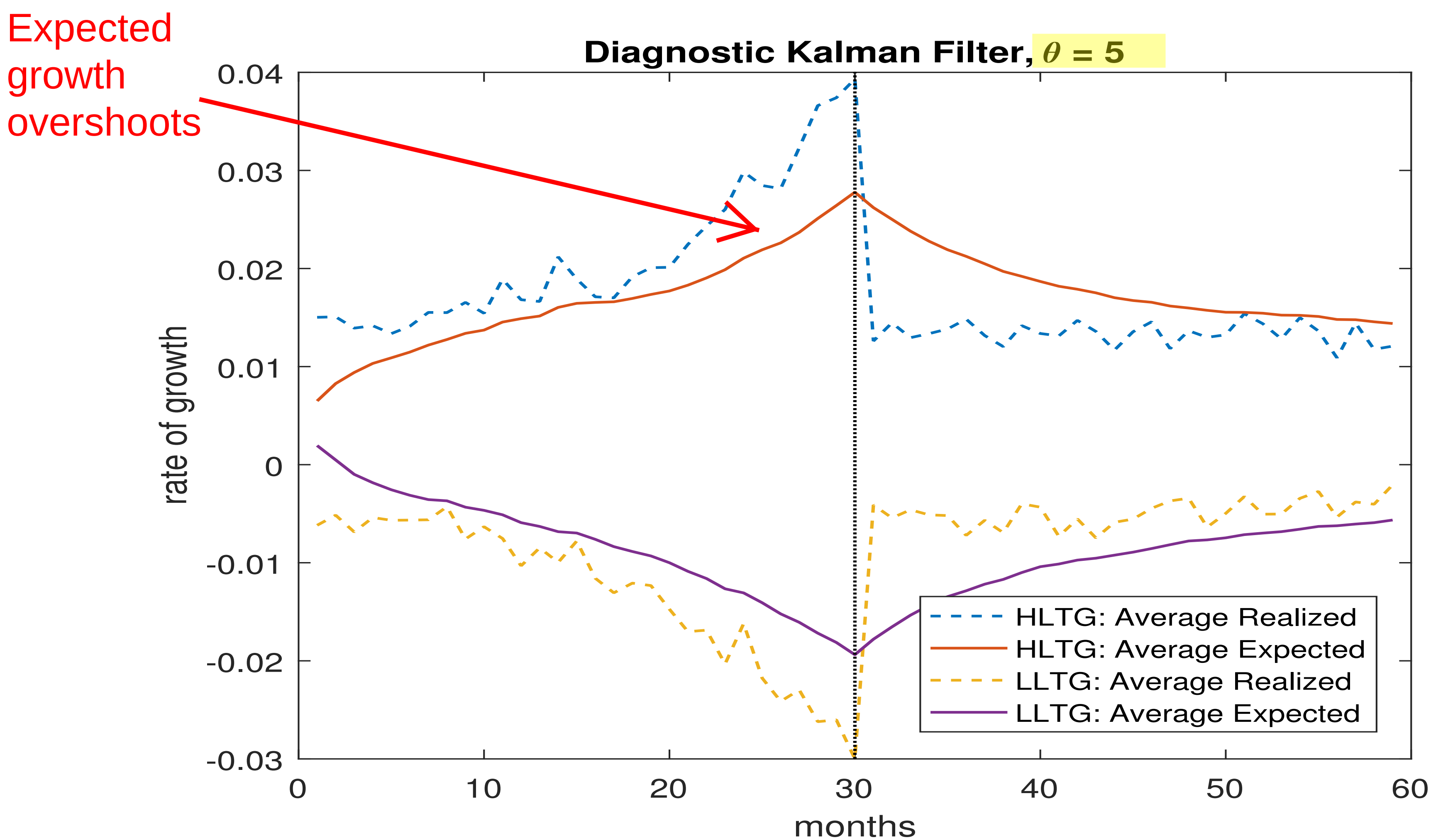
Representative Heuristics – – Simulations – 4



Representative Heuristics – – Simulations – 5



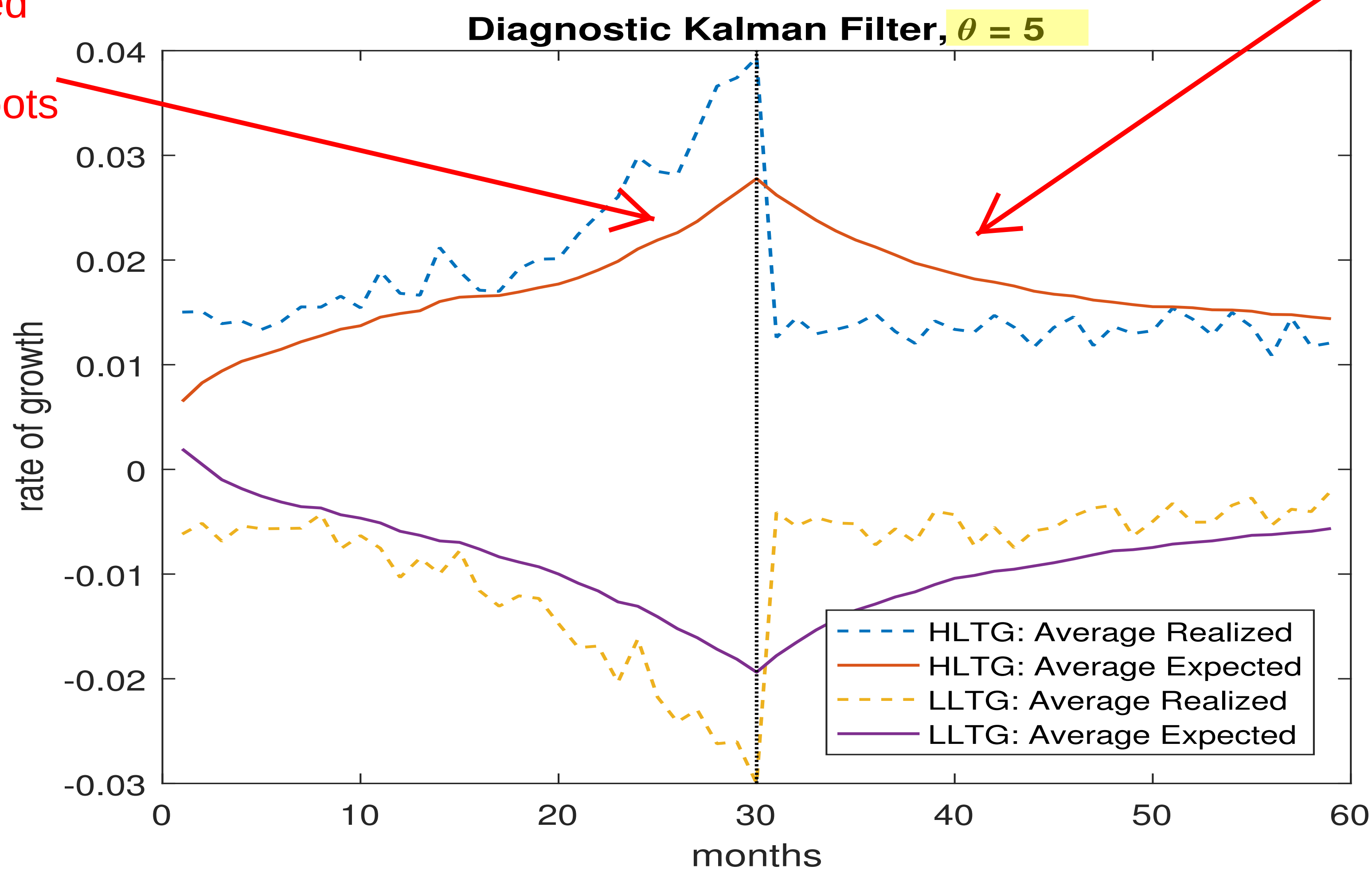
Representative Heuristics – – Simulations – 5



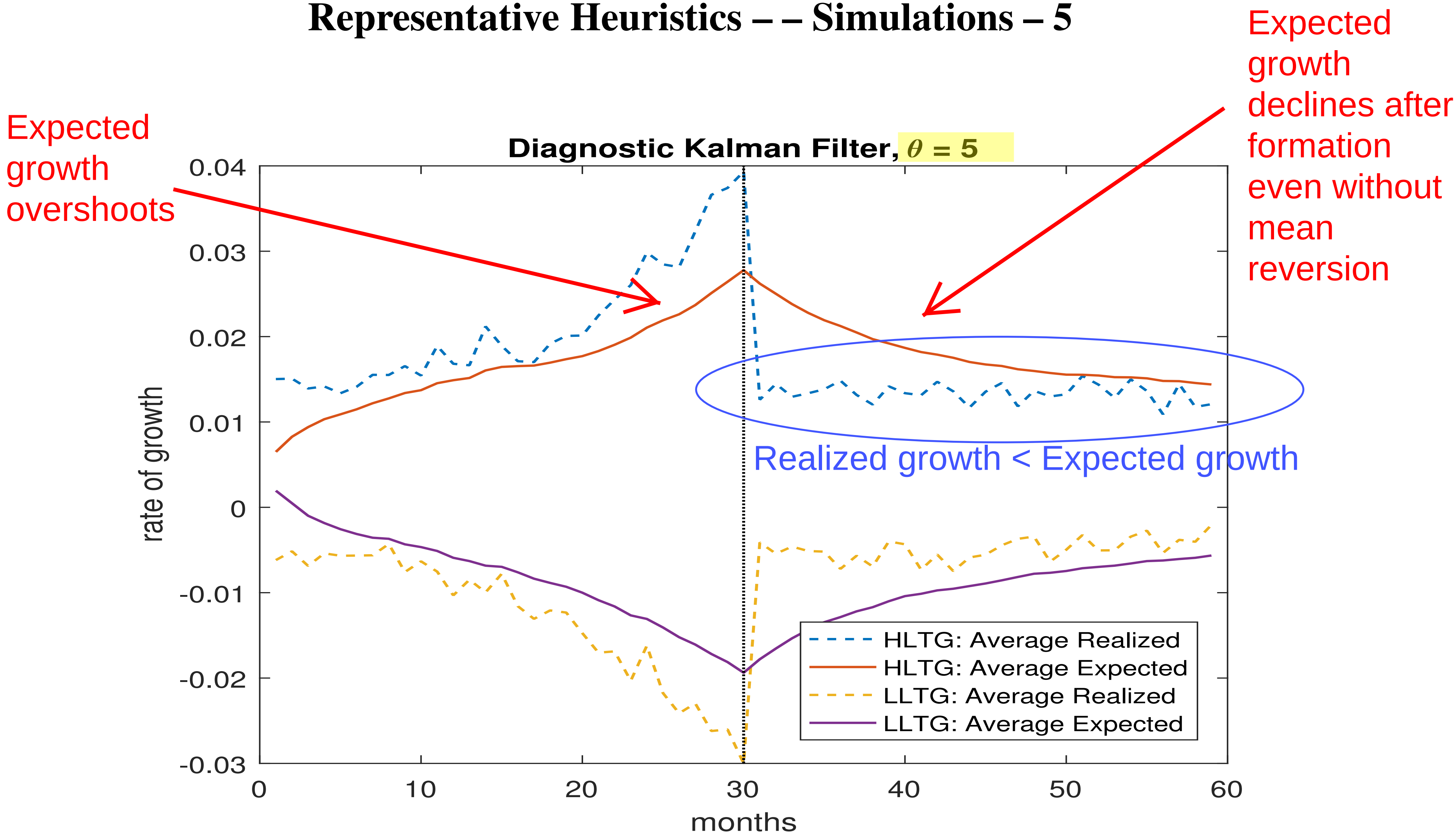
Representative Heuristics – – Simulations – 5

Expected growth overshoots

Expected growth declines after formation even without mean reversion



Representative Heuristics – – Simulations – 5



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

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 $P_t = D_t F\left(\widehat{g}_t^{\ \theta}\right)$ with $F' > 0 \Longrightarrow$ 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 $P_t = D_t F\left(\widehat{g}_t{}^{\theta}\right)$ with $F' > 0 \Longrightarrow$ 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 $\Longrightarrow$ "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 $\Longrightarrow$ "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 $\Longrightarrow$ "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 $\Longrightarrow$ "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 $R_{it} = \delta_0 + r_{it}$:

$$r_{i,t+1} \;=\; \delta_1 \, r_{it} + \eta^r_{i,t+1}$$

Reverse Causality: Learning from Prices

- Analysts got to look at stock prices when they form their predictions
  - Stock price goes up $\Longrightarrow$ "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 $\Longrightarrow$ "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:

$$pd_{it} = A + B_g \, g_{it} - B_r \, r_{it} + \varepsilon^{pd}_t$$

- where ε^{pd}_t is an approximation error, $\rho = e^{\overline{pd}}/(1 + e^{\overline{pd}})$; $\kappa = log(1 + e^{\overline{pd}}) - \rho\overline{pd}$:

$$A = (\kappa + \delta_0 - \gamma_0)/(1 - \rho); \quad B_g = 1/(1 - \rho\gamma_1); \quad B_r = 1/(1 - \rho\delta_1);$$

Reverse Causality: Learning from Prices

- Analysts got to look at stock prices when they form their predictions
  - Stock price goes up $\Longrightarrow$ "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:

$$pd_{it} = A + B_g \, g_{it} - B_r \, r_{it} + \varepsilon^{pd}_t$$

- where ε^{pd}_t is an approximation error, $\rho = e^{\overline{pd}}/(1 + e^{\overline{pd}})$; $\kappa = log(1 + e^{\overline{pd}}) - \rho\overline{pd}$:

$$A = (\kappa + \delta_0 - \gamma_0)/(1 - \rho); \quad B_g = 1/(1 - \rho\gamma_1); \quad B_r = 1/(1 - \rho\delta_1);$$

- **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:

$$\begin{pmatrix} g_{i,t+1} \\ r_{i,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma_1 & 0 \\ 0 & \delta_1 \end{pmatrix}}_{\mathbf{T}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma_g & 0 \\ 0 & \sigma_r \end{pmatrix} \begin{pmatrix} \eta^g_{i,t} \\ \eta^r_{i,t} \end{pmatrix}$$

$$\begin{pmatrix} x_{i,t} \\ pd_{i,t} \end{pmatrix} = \begin{pmatrix} \delta_0 \\ A \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ B_g & -B_r \end{pmatrix}}_{\mathbf{Z}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma^g_\varepsilon & 0 \\ 0 & \sigma^{pd}_\varepsilon \end{pmatrix}}_{\mathbf{\Sigma}} \begin{pmatrix} \varepsilon^g_{i,t} \\ \varepsilon^{pd}_{i,t} \end{pmatrix}$$

Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

$$\begin{pmatrix} g_{i,t+1} \\ r_{i,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma_1 & 0 \\ 0 & \delta_1 \end{pmatrix}}_{\mathbf{T}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma_g & 0 \\ 0 & \sigma_r \end{pmatrix} \begin{pmatrix} \eta^g_{i,t} \\ \eta^r_{i,t} \end{pmatrix}$$

$$\begin{pmatrix} x_{i,t} \\ pd_{i,t} \end{pmatrix} = \begin{pmatrix} \delta_0 \\ A \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ B_g & -B_r \end{pmatrix}}_{\mathbf{Z}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma^g_\varepsilon & 0 \\ 0 & \sigma^{pd}_\varepsilon \end{pmatrix}}_{\mathbf{\Sigma}} \begin{pmatrix} \varepsilon^g_{i,t} \\ \varepsilon^{pd}_{i,t} \end{pmatrix}$$

- Given the Kalman gain $\mathbf{K}_t = \mathbf{Z}\mathbf{P}_t\mathbf{T}'(\mathbf{Z}\mathbf{P}_t\mathbf{Z}' + \mathbf{\Sigma}\mathbf{\Sigma}')^{-1}$, the filter for g_{it} is:

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

Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

$$\begin{pmatrix} g_{i,t+1} \\ r_{i,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma_1 & 0 \\ 0 & \delta_1 \end{pmatrix}}_{\mathbf{T}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma_g & 0 \\ 0 & \sigma_r \end{pmatrix} \begin{pmatrix} \eta^g_{i,t} \\ \eta^r_{i,t} \end{pmatrix}$$

$$\begin{pmatrix} x_{i,t} \\ pd_{i,t} \end{pmatrix} = \begin{pmatrix} \delta_0 \\ A \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ B_g & -B_r \end{pmatrix}}_{\mathbf{Z}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma^g_\varepsilon & 0 \\ 0 & \sigma^{pd}_\varepsilon \end{pmatrix}}_{\mathbf{\Sigma}} \begin{pmatrix} \varepsilon^g_{i,t} \\ \varepsilon^{pd}_{i,t} \end{pmatrix}$$

- Given the Kalman gain $\mathbf{K}_t = \mathbf{Z}\mathbf{P}_t\mathbf{T}'(\mathbf{Z}\mathbf{P}_t\mathbf{Z}' + \mathbf{\Sigma}\mathbf{\Sigma}')^{-1}$, the filter for g_{it} is:

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

- Higher $\widehat{g}_{i,t}$ occurs if:

Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

$$\begin{pmatrix} g_{i,t+1} \\ r_{i,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma_1 & 0 \\ 0 & \delta_1 \end{pmatrix}}_{\mathbf{T}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma_g & 0 \\ 0 & \sigma_r \end{pmatrix} \begin{pmatrix} \eta^g_{i,t} \\ \eta^r_{i,t} \end{pmatrix}$$

$$\begin{pmatrix} x_{i,t} \\ pd_{i,t} \end{pmatrix} = \begin{pmatrix} \delta_0 \\ A \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ B_g & -B_r \end{pmatrix}}_{\mathbf{Z}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma^g_\varepsilon & 0 \\ 0 & \sigma^{pd}_\varepsilon \end{pmatrix}}_{\mathbf{\Sigma}} \begin{pmatrix} \varepsilon^g_{i,t} \\ \varepsilon^{pd}_{i,t} \end{pmatrix}$$

- Given the Kalman gain $\mathbf{K}_t = \mathbf{Z}\mathbf{P}_t\mathbf{T}'(\mathbf{Z}\mathbf{P}_t\mathbf{Z}' + \mathbf{\Sigma}\mathbf{\Sigma}')^{-1}$, the filter for g_{it} is:

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

- Higher $\widehat{g}_{i,t}$ 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:

$$\begin{pmatrix} g_{i,t+1} \\ r_{i,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma_1 & 0 \\ 0 & \delta_1 \end{pmatrix}}_{\mathbf{T}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma_g & 0 \\ 0 & \sigma_r \end{pmatrix} \begin{pmatrix} \eta^g_{i,t} \\ \eta^r_{i,t} \end{pmatrix}$$

$$\begin{pmatrix} x_{i,t} \\ pd_{i,t} \end{pmatrix} = \begin{pmatrix} \delta_0 \\ A \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ B_g & -B_r \end{pmatrix}}_{\mathbf{Z}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma^g_\varepsilon & 0 \\ 0 & \sigma^{pd}_\varepsilon \end{pmatrix}}_{\mathbf{\Sigma}} \begin{pmatrix} \varepsilon^g_{i,t} \\ \varepsilon^{pd}_{i,t} \end{pmatrix}$$

- Given the Kalman gain $\mathbf{K}_t = \mathbf{Z}\mathbf{P}_t\mathbf{T}'(\mathbf{Z}\mathbf{P}_t\mathbf{Z}' + \mathbf{\Sigma}\mathbf{\Sigma}')^{-1}$, the filter for g_{it} is:

$$\widehat{g}_{i,t} = \gamma_1\,\widehat{g}_{i,t-1} + K_{11}\left[x_{i,t} - E_{t-1}\left(x_{i,t}\right)\right] + K_{12}\boxed{\left[pd_{i,t} - E_{t-1}\left(pd_{i,t}\right)\right]}$$

- Higher $\widehat{g}_{i,t}$ 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: $\boxed{pd_{i,t} > E_{t-1}(pd_{i,t})}$

Reverse Causality: Learning from Prices – Filtering

- Two state equations and two observation equations:

$$\begin{pmatrix} g_{i,t+1} \\ r_{i,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma_1 & 0 \\ 0 & \delta_1 \end{pmatrix}}_{\mathbf{T}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \begin{pmatrix} \sigma_g & 0 \\ 0 & \sigma_r \end{pmatrix} \begin{pmatrix} \eta^g_{i,t} \\ \eta^r_{i,t} \end{pmatrix}$$

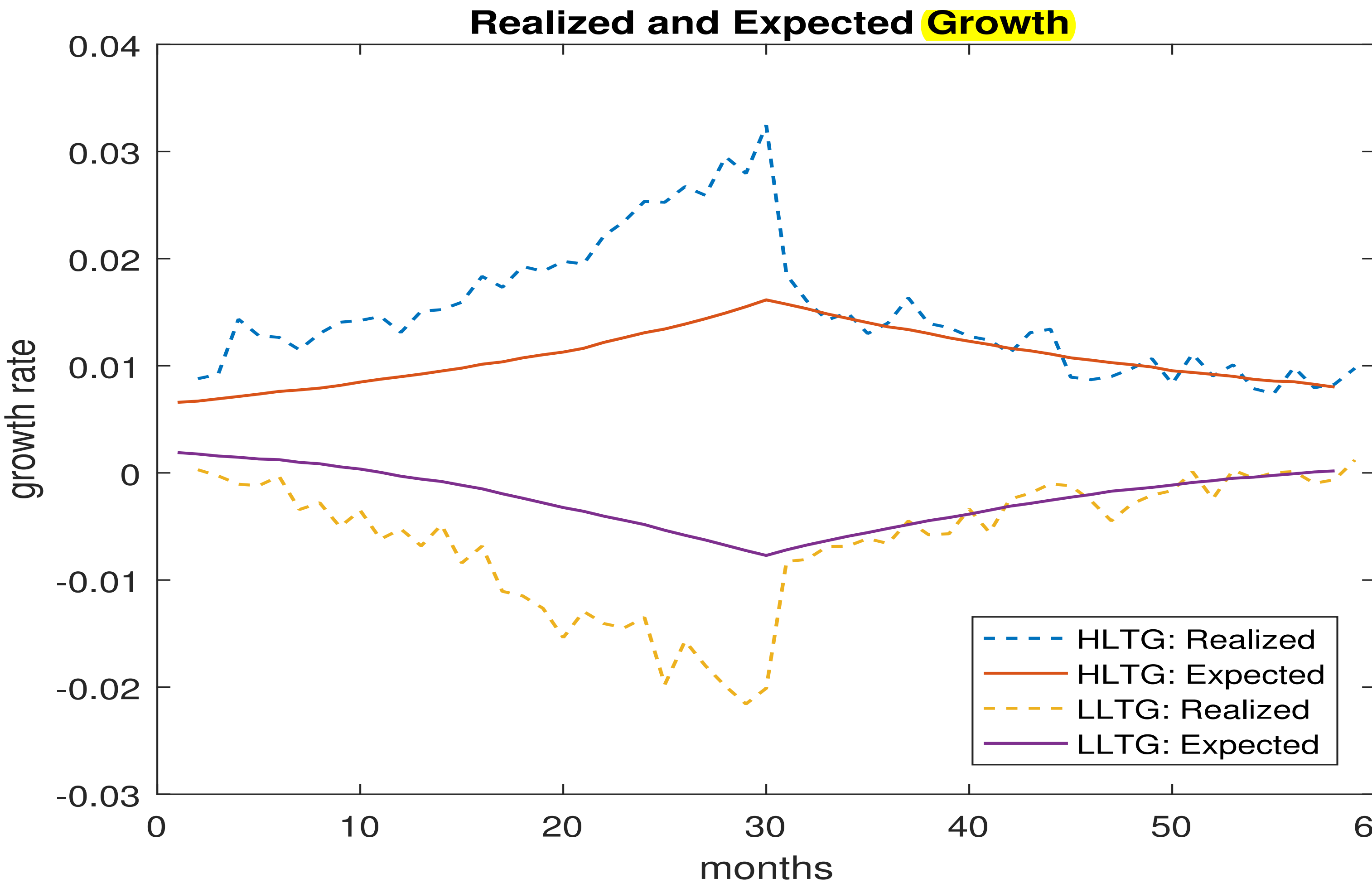
$$\begin{pmatrix} x_{i,t} \\ pd_{i,t} \end{pmatrix} = \begin{pmatrix} \delta_0 \\ A \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 \\ B_g & -B_r \end{pmatrix}}_{\mathbf{Z}} \begin{pmatrix} g_{i,t} \\ r_{i,t} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma^g_\varepsilon & 0 \\ 0 & \sigma^{pd}_\varepsilon \end{pmatrix}}_{\mathbf{\Sigma}} \begin{pmatrix} \varepsilon^g_{i,t} \\ \varepsilon^{pd}_{i,t} \end{pmatrix}$$

- Given the Kalman gain $\mathbf{K}_t = \mathbf{Z}\mathbf{P}_t\mathbf{T}'(\mathbf{Z}\mathbf{P}_t\mathbf{Z}' + \mathbf{\Sigma}\mathbf{\Sigma}')^{-1}$, the filter for g_{it} is:

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

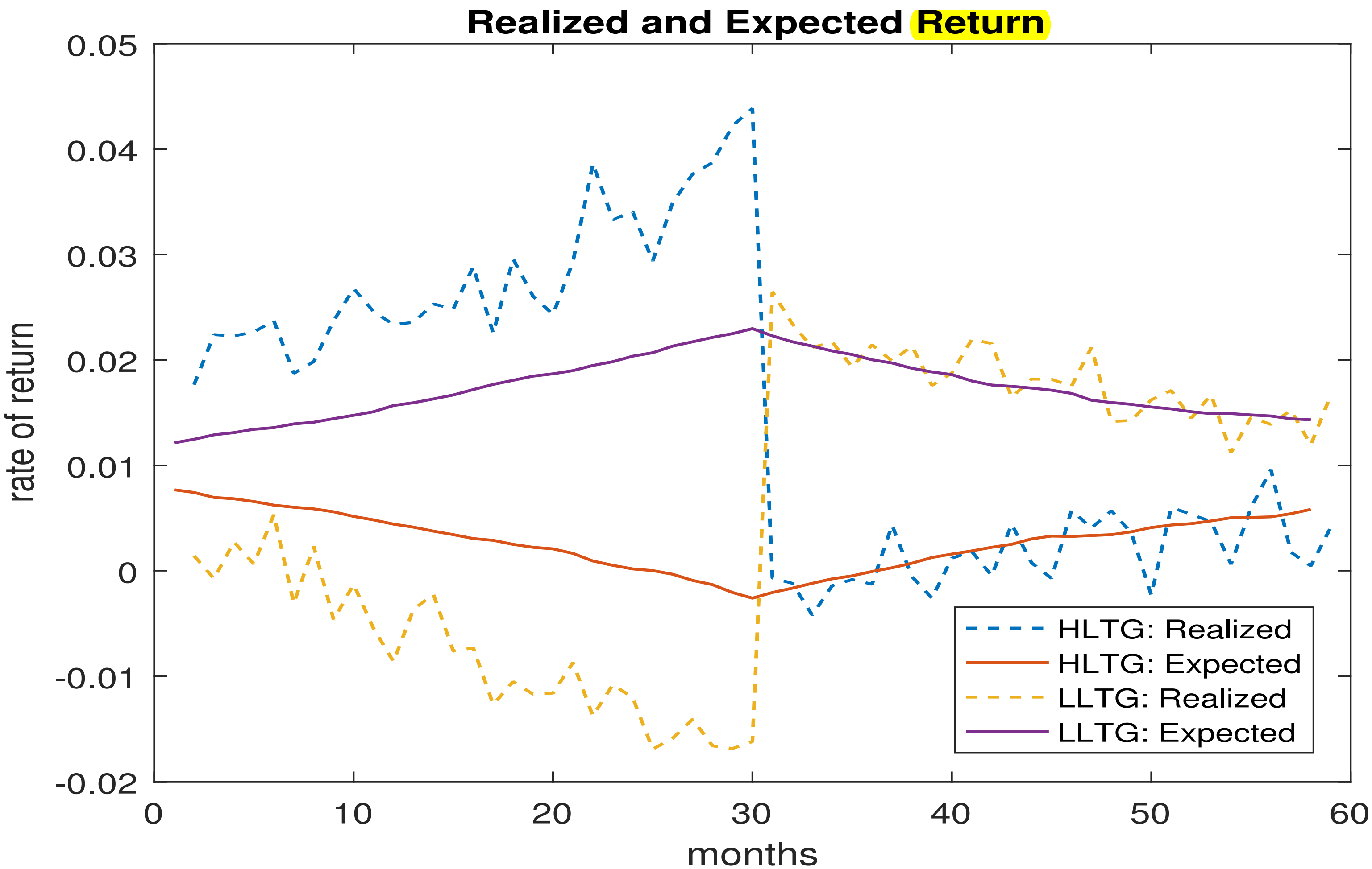
- Higher $\widehat{g}_{i,t}$ occurs if:
  1. Realized growth is higher than expected: $x_{i,t} > E_{t-1}(x_{i,t})$
  2. Realized return is higher than expected: $pd_{i,t} > E_{t-1}(pd_{i,t})$
- The second effect comes from a <u>decrease</u> in *expected* return r_{it}.
  - High growth stocks $\Longrightarrow$ forecast lower future return

Reverse Causality: Learning from Prices – Growth

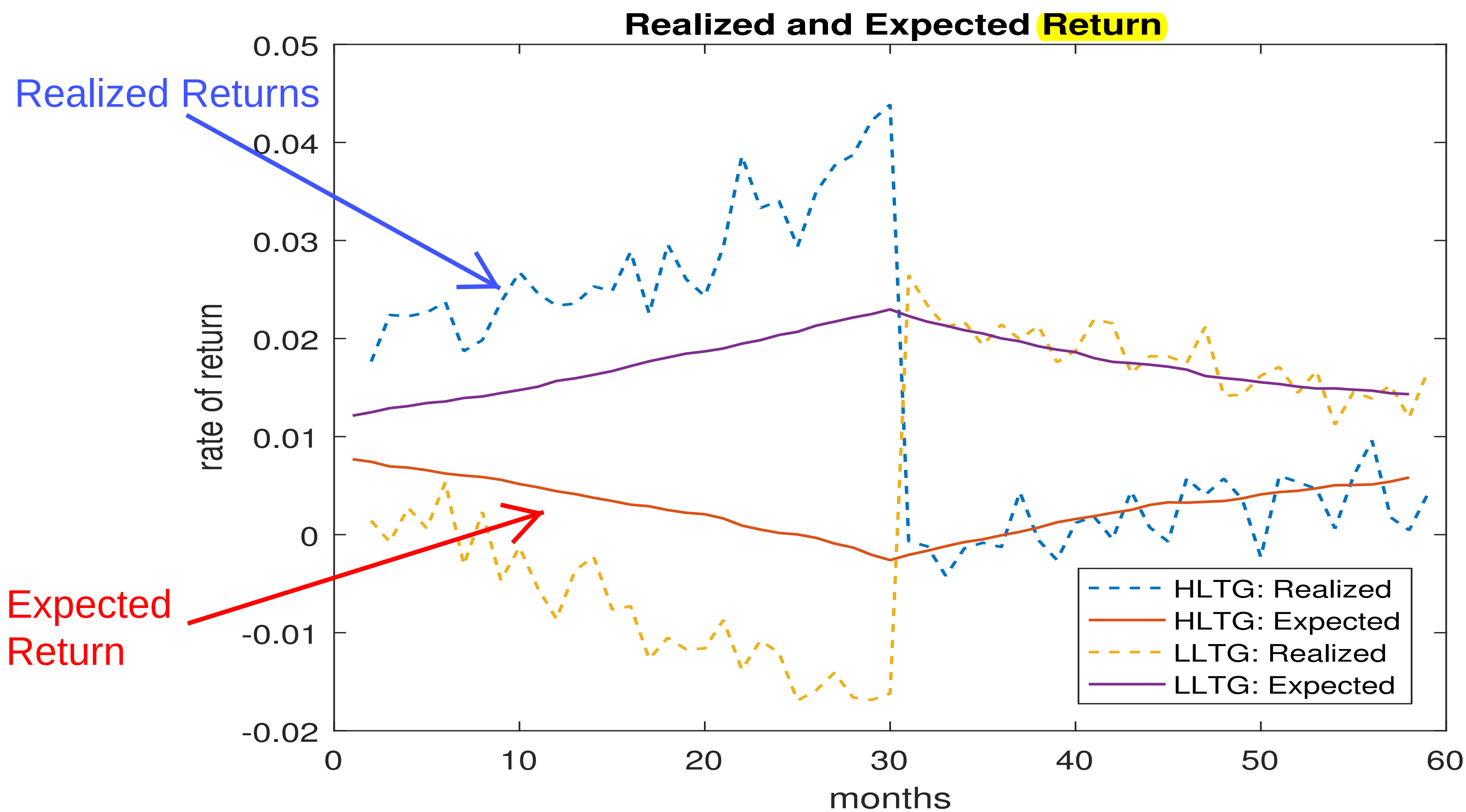


- Same implications for realized and expected growth as in previous cases.

Reverse Causality: Learning from Prices – Returns



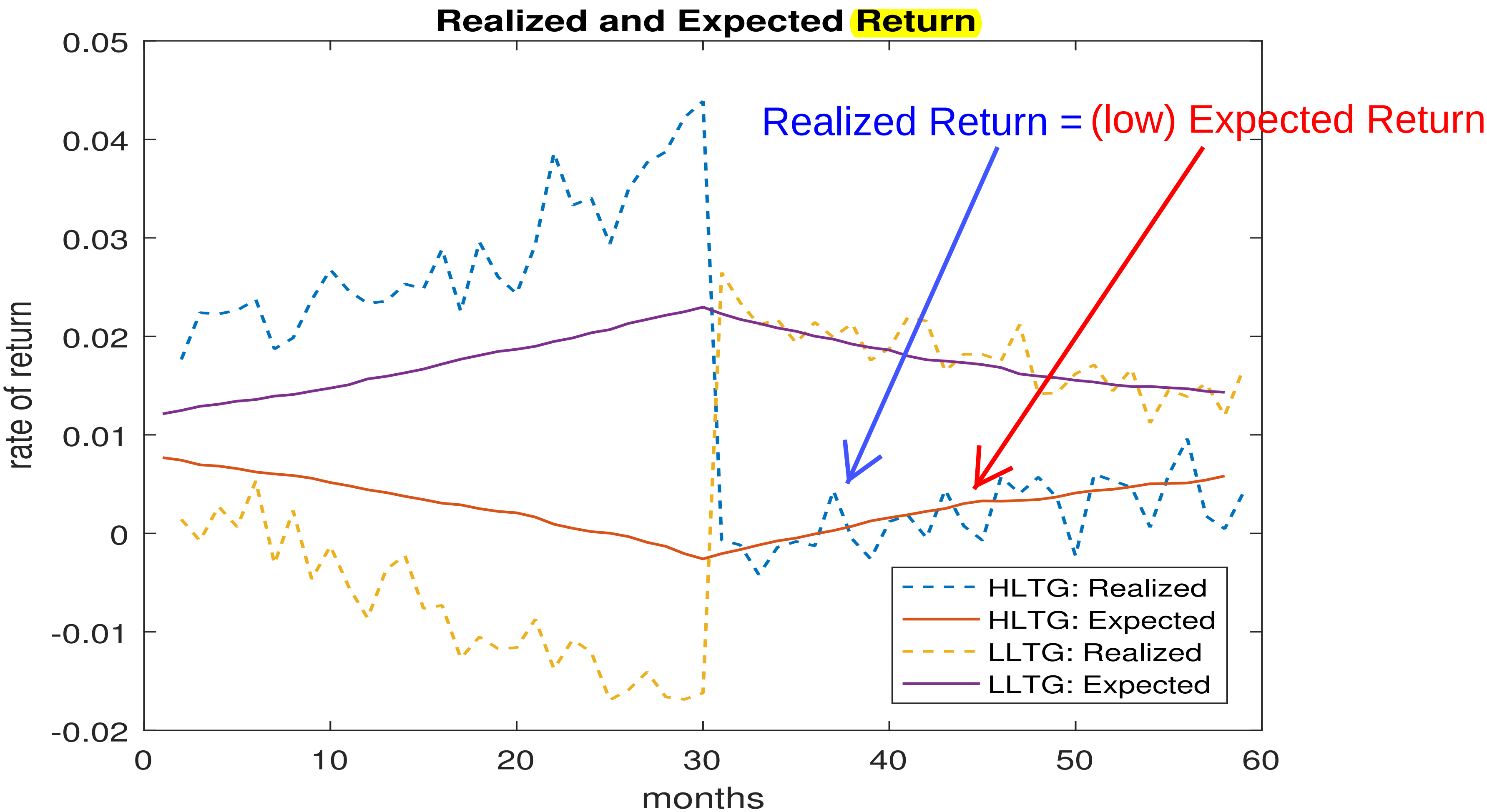
Reverse Causality: Learning from Prices – Returns



- HLTG portfolio:

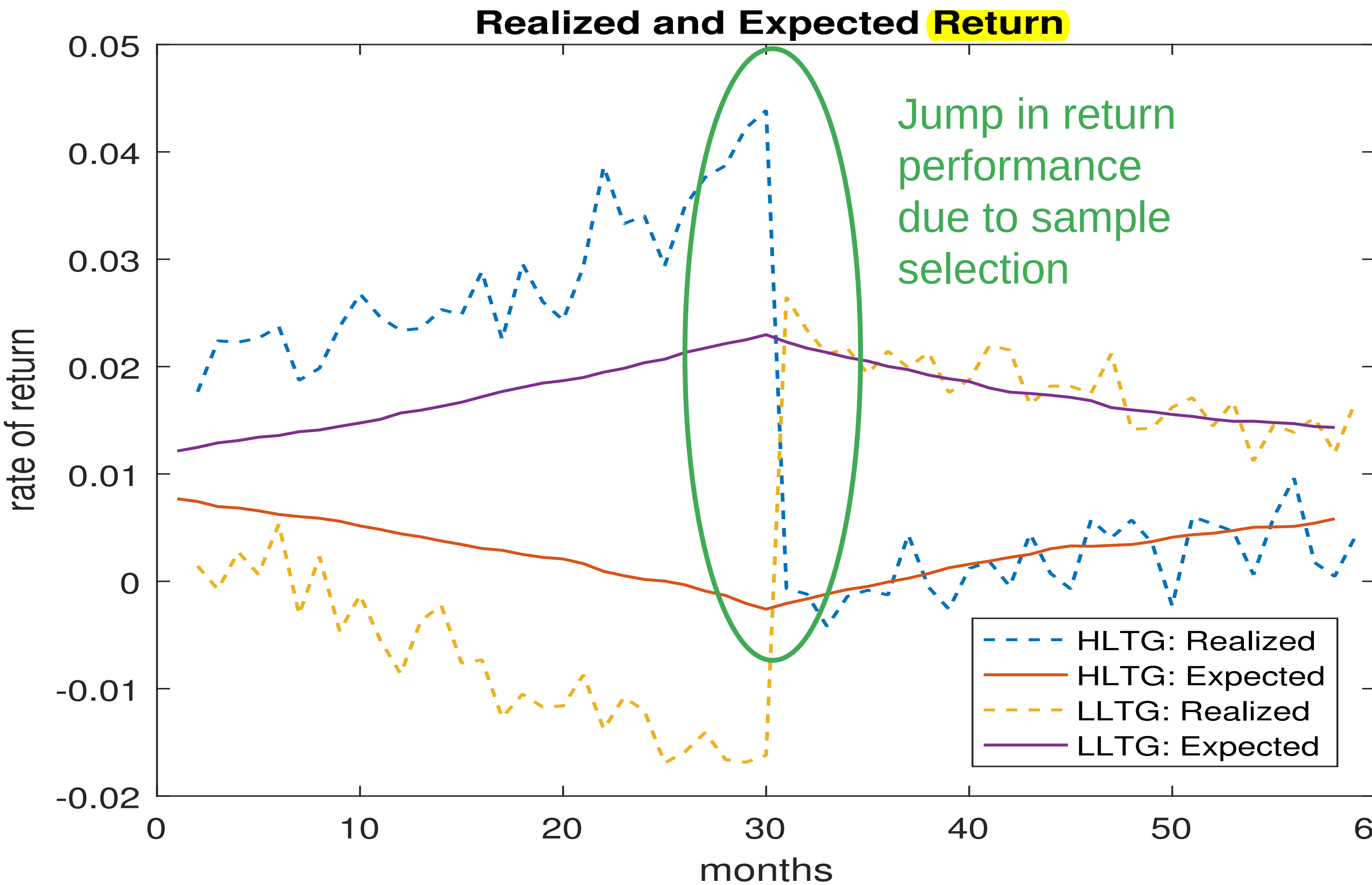
  1. Pre-ranking: Higher realized returns and *declining* expected return

Reverse Causality: Learning from Prices – Returns



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

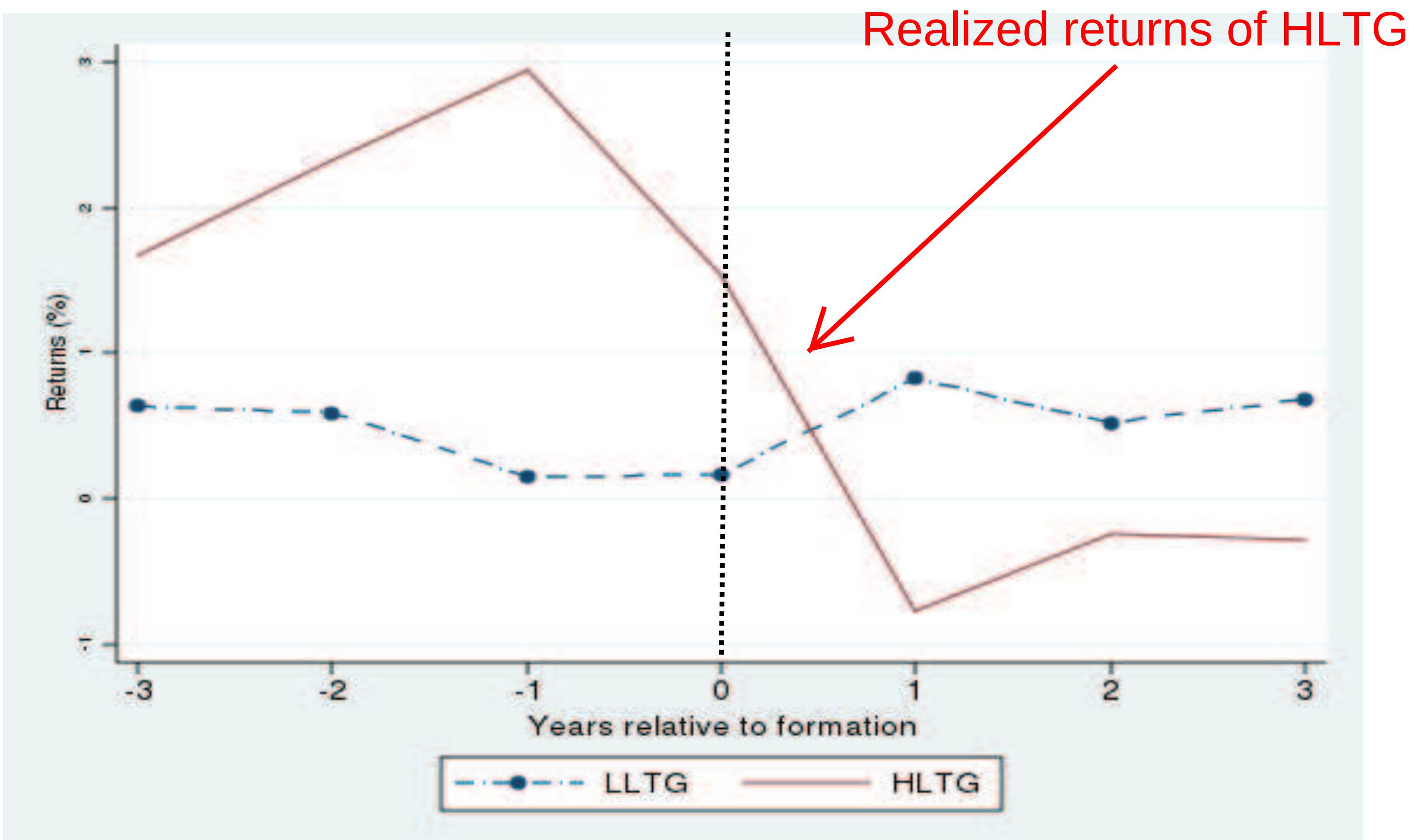
Reverse Causality: Learning from Prices – Returns



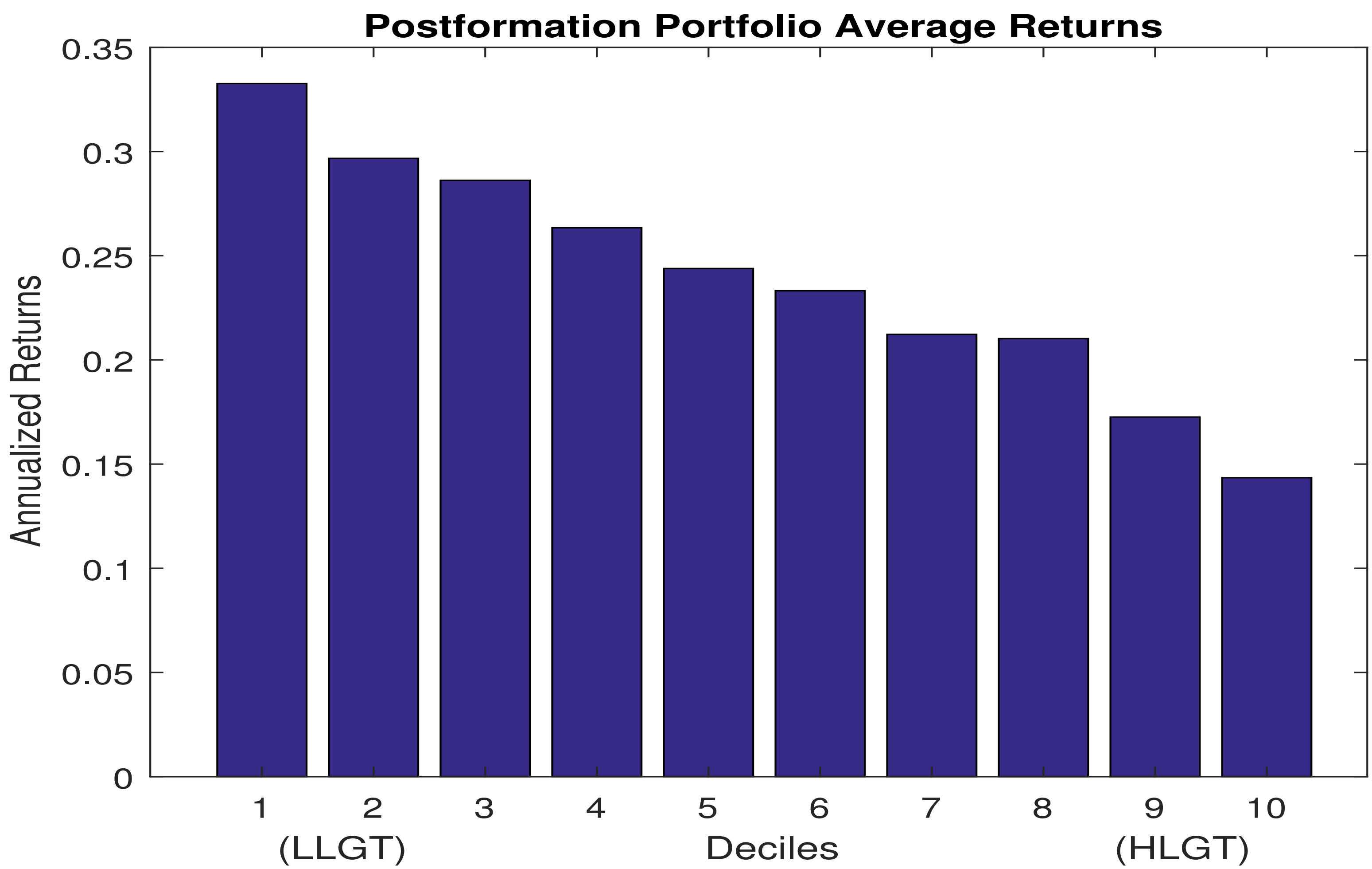
- HLTG portfolio:
  1. Pre-ranking: Higher realized returns and *declining* expected return
  2. Post-ranking: Very low realized returns, consistent with low expected return

Reverse Causality: Learning from Prices – Returns

Figure 5 BGLS: Twelve-day Returns on Earnings Announcements for LTG Portfolios. (DATA)

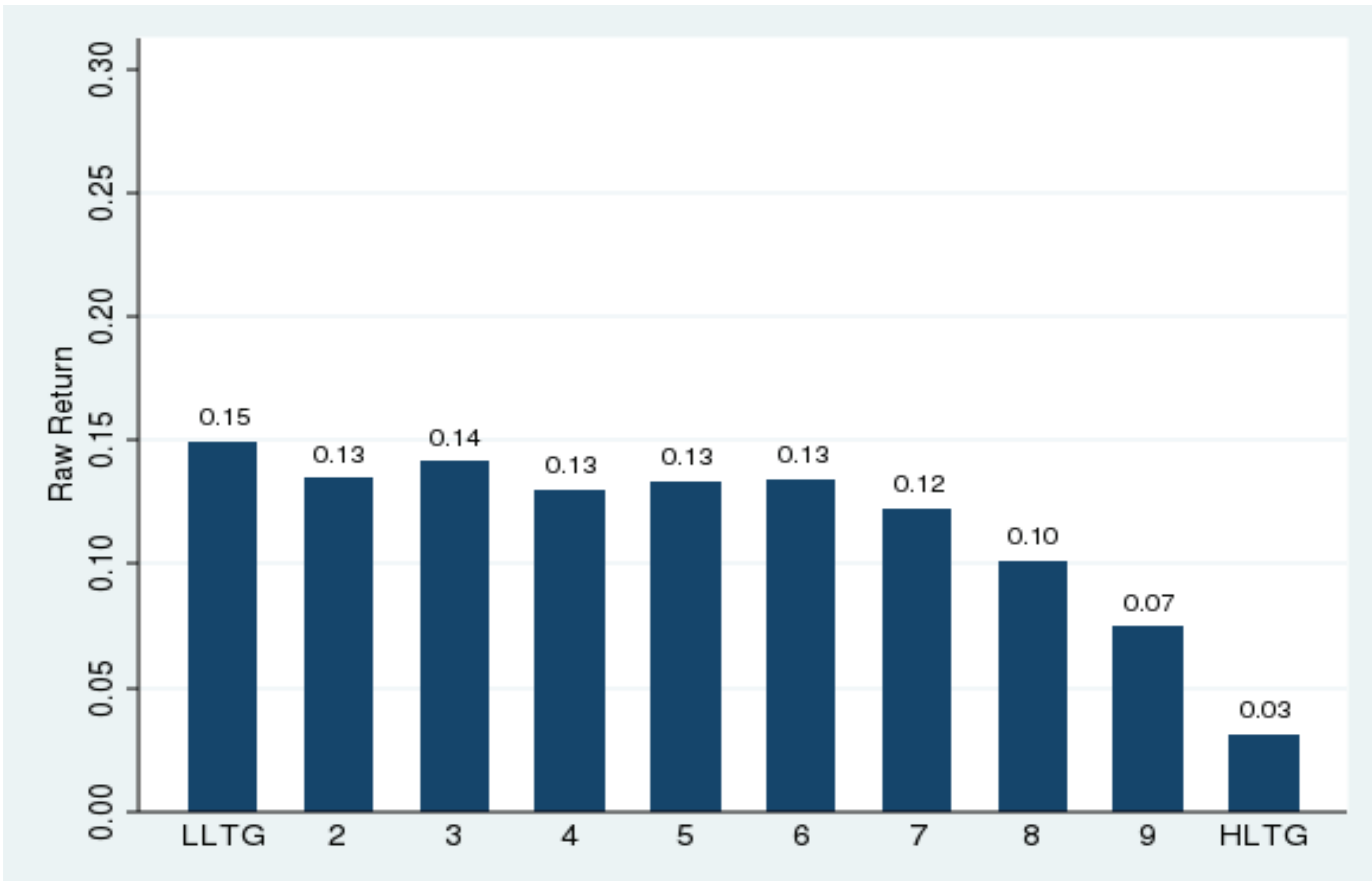


Reverse Causality: Learning from Prices – Average Returns



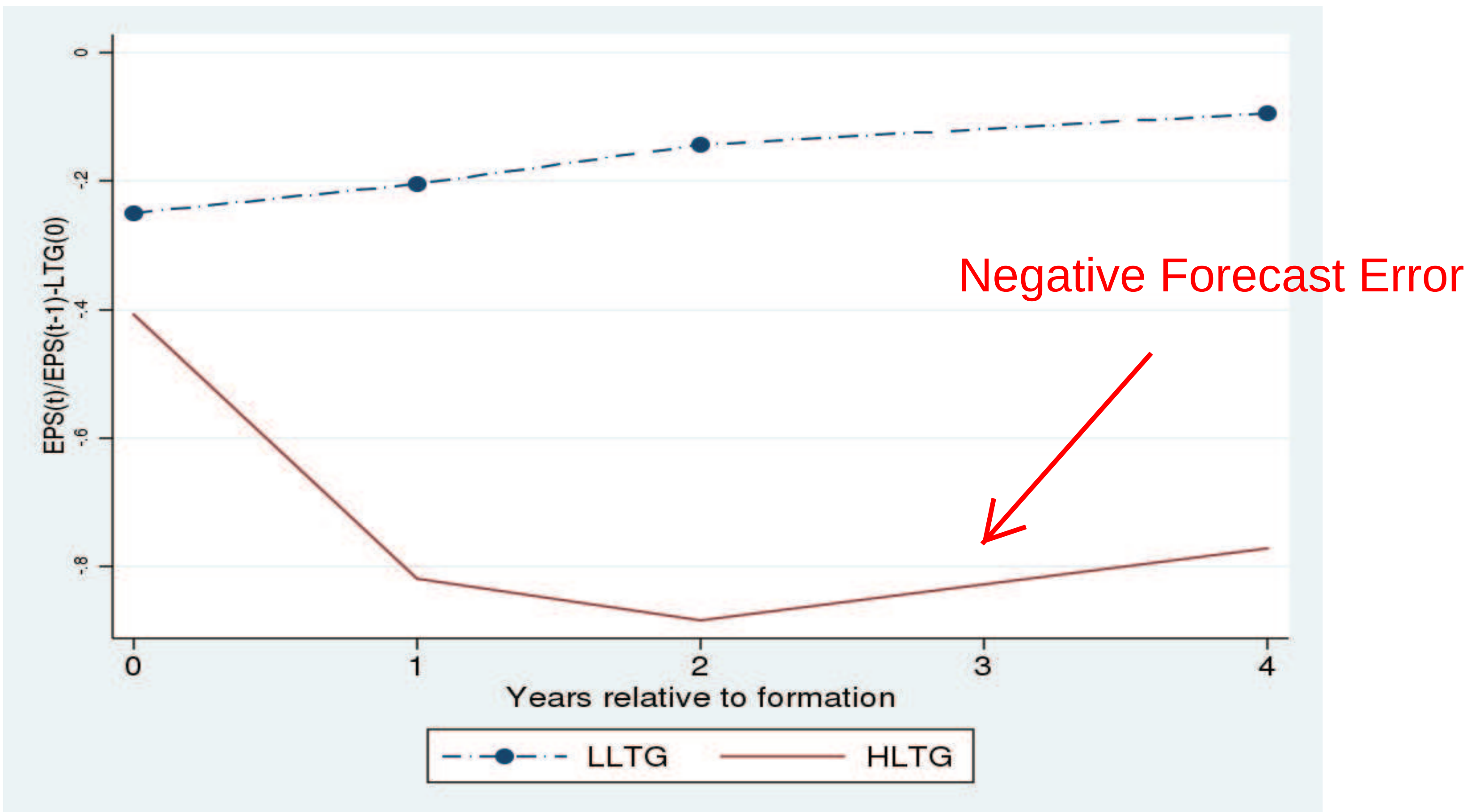
Reverse Causality: Learning from Prices – Average Returns

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)



Final Comments

- Main evidence for the model is on beliefs: forecast errors and over-reaction.
  - What does this evidence imply for the magnitude of θ?
  - Is θ calibrated from psychology able to *quantitatively* rationalize beliefs dynamics and asset prices?
    - The model is "portable": Can we use estimates of θ from other studies?
    - In JF paper, θ 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 θ?
  - Is θ calibrated from psychology able to *quantitatively* rationalize beliefs dynamics and asset prices?
    - The model is "portable": Can we use estimates of θ from other studies?
    - In JF paper, θ 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)