Econ 219B
Psychology and Economics: Applications
(Lecture 13)

Stefano DellaVigna

April 23, 2008

Outline

1. Market Reaction to Biases: Introduction

2. Market Reaction to Biases: Pricing

3. Intro to Problem Set

4. Market Reaction to Biases: Behavioral Finance

1 Market Reaction to Biases: Introduction

- So far, we focused on consumer deviations from standard model

- Who exhibits these deviations?

  1. **Self-control and naivete'.** Consumers (health clubs, food, credit cards, smoking), Employees (retirement saving, benefit take-up), Students (homework)

  2. **Reference dependence.** Workers (labor supply, increasing wages), (inexperienced) traders (sport cards), Investors, Consumers (insurance), House owners

  3. **Social preferences.** Consumers (giving to charities), Employees (effort, strikes)

4. **Biased Beliefs.** Individual investors, CEOs, Consumers (purchases, betting)

5. **Inattention.** Individual investors, Consumers (eBay bidding, taxation)

6. **Menu Effects.** Individual investors, Consumers (loans, 410(k) plans)

7. **Social Pressure and Persuasion.** Voters, Employees (productivity), Individual investors (and analysts)

8. **Emotions.** Individual investors, Consumers

- What is missing from picture?

- Experienced agents
- Firms
- Broadly speaking, market interactions with ‘rational’ agents

- Market interactions
  - Everyone ‘born’ with biases
  - But: Effect of biases lower if:
    - learning with plenty of feedback
    - advice, access to consulting
    - specialization

* Competition 'drives out of market' (BUT: See last lecture)

- For which agents are these conditions more likely to be satisfied?

- Firms

- In particular, firms more likely to be aware of biases

- Implications? Study biases in the market

- Six major instances:
  - Interaction between firms and consumers (contract design, price choice — today)
  - Interaction between experienced and inexperienced investors (noise traders and behavioral finance — today or next week)
  - Interaction between managers and investors (corporate finance — next week)
  - Interaction between employers and employees (labor economics — briefly next week)
  - Interaction between politicians and voters (political economy — next week)
  - Institutional design (next week)

2 Market Reaction to Biases: Pricing

- Consider now the case in which consumers purchasing products have biases
- Firm maximize profits
- Do consumer biases affect profit-maximizing contract design?
- How is consumer welfare affected by firm response?
- Analyze first the case of consumers with $\left(\beta, \hat{\beta}, \delta\right)$ preferences

2.1 Self-Control

MARKET (I). INVESTMENT GOODS

- Monopoly
- Two-part tariff: L (lump-sum fee), p (per-unit price)
- Cost: set-up cost K, per-unit cost a

Consumption of investment good

Payoffs relative to best alternative activity:

- Cost c at $t = 1$, stochastic
  - non-monetary cost
  - experience good, distribution $F(c)$
- Benefit $b > 0$ at $t = 2$, deterministic

CONSUMER BEHAVIOR.

- Long-run plans at $t = 0$:

Consume $\iff \beta\delta(-p - c + \delta b) > 0$

$\iff c < \delta b - p$

- Actual consumption decision at $t = 1$:

Consume $\iff c < \beta\delta b - p$ (Time Inconsistency)

- Forecast at $t = 0$ of consumption at $t = 1$:

Consume $\iff c < \hat{\beta}\delta b - p$ (Naiveté)

FIRM BEHAVIOR. Profit-maximization

$$\max_{L,p} \delta \left\{ L - K + F\left(\beta\delta b - p\right)\left(p - a\right)\right\}$$

$$\text{s.t. } \beta\delta \left\{ -L + \int_{-\infty}^{\hat{\beta}\delta b - p} \left(\delta b - p - c\right) dF\left(c\right) \right\} \geq \beta\delta\overline{u}$$

- Notice the difference between β and $\hat{\beta}$

Solution for the per-unit price p^*:

$$
\begin{aligned}
p^* &= a && \text{[exponentials]} \\
&\quad - \left(1-\hat{\beta}\right)\delta b \frac{f\left(\hat{\beta}\delta b - p^*\right)}{f\left(\beta\delta b - p^*\right)} && \text{[sophisticates]} \\
&\quad - \frac{F\left(\hat{\beta}\delta b - p^*\right) - F\left(\beta\delta b - p^*\right)}{f\left(\beta\delta b - p^*\right)} && \text{[naives]}
\end{aligned}
$$

Features of the equilibrium

1. *Exponential agents* ($\beta = \hat{\beta} = 1$).
   Align incentives of consumers with cost of firm
   $\Longrightarrow$ marginal cost pricing: $p^* = a$.

$$
\begin{aligned}
p^* \;=\; & a & \text{[exponentials]} \\
& -\left(1-\hat{\beta}\right)\delta b \frac{f\left(\hat{\beta}\delta b - p^*\right)}{f\left(\beta\delta b - p^*\right)} & \text{[sophisticates]} \\
& -\frac{F\left(\hat{\beta}\delta b - p^*\right) - F\left(\beta\delta b - p^*\right)}{f\left(\beta\delta b - p^*\right)} & \text{[naives]}
\end{aligned}
$$

2. *Hyperbolic agents.* Time inconsistency
   $\Longrightarrow$ below-marginal cost pricing: $p^* < a$.

   (a) *Sophisticates* $(\beta = \hat{\beta} < 1)$: commitment.

   (b) *Naives* $(\beta < \hat{\beta} = 1)$: overestimation of consumption.

MARKET (II). LEISURE GOODS

Payoffs of consumption at $t = 1$:

- Benefit at $t = 1$, stochastic
- Cost at $t = 2$, deterministic

$\Longrightarrow$ Use the previous setting: $-c$ is "current benefit", $b < 0$ is "future cost."

Results:

1. *Exponential agents.*

   Marginal cost pricing: $p^* = a$, $L^* = K$ (PC).

2. *Hyperbolic agents* tend to overconsume. $\Longrightarrow$

   Above-marginal cost pricing: $p^* > a$. Initial bonus $L^* < K$ (PC).

EMPIRICAL PREDICTIONS

Two predictions for time-inconsistent consumers:

1. Investment goods (Proposition 1):
   (a) Below-marginal cost pricing
   (b) Initial fee (Perfect Competition)
2. Leisure goods (Corollary 1)
   (a) Above-marginal cost pricing
   (b) Initial bonus or low initial fee (Perfect Competition)

FIELD EVIDENCE ON CONTRACTS

- US Health club industry ($11.6bn revenue in 2000)
  - monthly and annual contracts
  - Estimated marginal cost: \$3-\$6 + congestion cost
  - Below-marginal cost pricing despite small transaction costs and price discrimination
- Vacation time-sharing industry ($7.5bn sales in 2000)
  - high initial fee: $11,000 (RCI)
  - minimal fee per week of holiday: $140 (RCI)

- Credit card industry ($500bn outstanding debt in 1998)
  - Resale value of credit card debt: 20% premium (Ausubel, 1991)
  - No initial fee, bonus (car / luggage insurance)
  - Above-marginal-cost pricing of borrowing

- Gambling industry: Las Vegas hotels and restaurants:
  - Price rooms and meals below cost, at bonus
  - High price on gambling

WELFARE EFFECTS

Result 1. Self-control problems + Sophistication $\Rightarrow$ First best

- Consumption if $c \leq \beta\delta b - p^*$
- Exponential agent:
  - $p^* = a$
  - consume if $c \leq \delta b - p^* = \delta b - a$
- Sophisticated time-inconsistent agent:
  - $p^* = a - (1 - \beta)\delta b$
  - consume if $c \leq \beta\delta b - p^* = \delta b - a$
- Perfect commitment device
- Market interaction maximizes joint surplus of consumer and firm

Result 2. Self-control + Partial naiveté $\Rightarrow$ Real effect of time inconsistency

- $p^* = a - [F(\delta b - p^*) - F(\beta\delta b - p^*)]/f(\beta\delta b - p^*)$

- Firm sets p^* so as to accentuate overconfidence

- Two welfare effects:
  - Inefficiency: $\text{Surplus}_{\text{naive}} \leq \text{Surplus}_{\text{soph.}}$
  - Transfer (under monopoly) from consumer to firm

- Profits are increasing in naivete' $\hat{\beta}$(monopoly)

- $\text{Welfare}_{\text{naive}} \leq \text{Welfare}_{\text{soph.}}$

- Large welfare effects of non-rational expectations

2.2 Self-Control 2

- Kfir and Spiegler (2004), Contracting with Diversely Naive Agents.

- Extend DellaVigna and Malmendier (2004):
  - incorporate heterogeneity in naiveté
  - allow more flexible functional form in time inconsistency
  - different formulation of naiveté

- Setup:
  1. Actions:
     - Action $a \in [0, 1]$ taken at time 2
     - At time 1 utility function is $u(a)$
     - At time 2 utility function is $v(a)$
  2. Beliefs: At time 1 believe:
     - Utility is $u(a)$ with probability θ
     - Utility is $v(a)$ with probability $1 - \theta$
     - Heterogeneity: Distribution of types θ
  3. Transfers:
     - Consumer pays firm $t(a)$
     - Restrictive assumption: no cost to firm of providing a

- Therefore:
  - Time inconsistency ($\beta < 1$) –> Difference between u and v
  - Naiveté ($\hat{\beta} > \beta$) –> $\theta > 0$
  - Partial naiveté here modelled as stochastic rather than deterministic
  - Flexibility in capturing time inconsistency (self-control, reference dependence, emotions)

- Main result:
- **Proposition 1.** There are two types of contracts:
  1. Perfect commitment device for sufficiently sophisticated agents ($\theta < \underline{\theta}$)
  2. Exploitative contracts for sufficiently naive agents ($\theta > \underline{\theta}$)

- Commitment device contract:
  - Implement $a_\theta = \max_a u(a)$
  - Transfer:
    * $t(a_\theta) = \max_a u(a)$
    * $t(a) = \infty$ for other actions

  - Result here is like in DM: Implement first best

- Exploitative contract:
  - Agent has negative utility:

$$u\left(a_{\theta}^{v}\right) - t\left(a_{\theta}^{v}\right) < 0$$

  - Maximize overestimation of agents:

$$a_{\theta}^{u} = \arg\max\left(u\left(a\right) - v\left(a\right)\right)$$

2.3 Bounded Rationality

- Gabaix and Laibson (2003), *Competition and Consumer Confusion*
- Non-standard feature of consumers:
  - Limited ability to deal with complex products
  - imperfect knowledge of utility from consuming complex goods
- Firms are aware of bounded rationality of consumers
  $\longrightarrow$ design products & prices to take advantage of bounded rationality of consumers

Three steps:

1. Given product complexity, given number of firms: What is the mark-up? Comparative statics.

2. Given product complexity: endogenous market entry. What is the mark-up? What is the number of firms?

3. Endogenous product complexity, endogenous market entry: What are mark-up, number of firms, and degree of product complexity?

We will go through 1, skip 2, and talk about the intuition of 3.

Example: Checking account. Value depends on

- interest rates
- fees for dozens of financial services (overdrafts, more than x checks per months, low average balance, etc.)
- bank locations
- bank hours
- ATM locations
- web-based banking services
- linked products (e.g. investment services)

Given such complexity, consumers do not know the exact value of products they buy.

Model

- Consumers receive noisy, *unbiased* signals about product value.
  - Agent a chooses from n goods.
  - True utility from good i:

$$Q_i - p_i$$

  - Utility signal

$$U_{ia} = Q_i - p_i + \sigma_i \varepsilon_{ia}$$

    σ_i is complexity of product i.
    ε_{ia} is zero mean, iid across consumers and goods, with density f and cumulative distribution F.
    (Suppress consumer-specific subscript a;
    $U_i \equiv U_{ia}$ and $\varepsilon_i \equiv \varepsilon_{ia}$.)

- Consumer decision rule: Picks the one good with highest signal U_i from $(U_i)_{i=1}^n$.
  (**Assumption!** What justifies this assumption?) Demand for good i

$$
\begin{aligned}
D_i &= P\left(U_i > \max_{j \neq i} U_j\right) \\
&= E\left[P\left[\text{for all } j \neq i, U_i > U_j | \varepsilon_i\right]\right] \\
&= E\left[\prod_{j \neq i} P\left[U_i > U_j | \varepsilon_i\right]\right] \\
&= E\left[\prod_{j \neq i} P\left[\frac{Q_i - p_i - \left(Q_j - p_j\right) + \sigma_i \varepsilon_i}{\sigma_j} > \varepsilon_j | \varepsilon_i\right]\right]
\end{aligned}
$$

$$
D_i = \int f\left(\varepsilon_i\right) \prod_{j \neq i} F\left(\frac{Q_i - p_i - \left(Q_j - p_j\right) + \sigma_i \varepsilon_i}{\sigma_j}\right) d\varepsilon_i
$$

Market equilibrium with exogenous complexity

Bertrand competition with

- Q_i : quality of a good,
  σ_i : complexity of a good,
  c_i : production cost
  p_i : price
- Simplification: Q_i, σ_i, c_i identical across firms. (*Problem: How should consumers choose if all goods are known to be identical?*)
- Firms maximize profit $\pi_i = (p_i - c_i)\, D_i$
- Symmetry reduces demand to

$$D_i = \int f(\varepsilon_i)\, F\left(\frac{p_j - p_i + \sigma\varepsilon_i}{\sigma}\right)^{n-1} d\varepsilon_i$$

Example of demand curves

Gaussian noise $\varepsilon \sim N\,(0{,}1)$, 2 firms

Demand curve faced by firm 1:

$$\begin{aligned}
D_1 &= P\,(Q - p_1 + \sigma\varepsilon_1 > Q - p_2 + \sigma\varepsilon_2) \\
&= P\left(p_2 - p_1 > \sigma\sqrt{2}\eta\right) \text{ with } \eta = (\varepsilon_2 - \varepsilon_1)\,/\sqrt{2} \text{ N(0,1)} \\
&= \Phi\left(\frac{p_2 - p_1}{\sigma\sqrt{2}}\right)
\end{aligned}$$

Usual Bertrand case $(\sigma = 0)$: infinitely elastic demand at $p_1 = p_2$

$$D_1 \in \left\{ \begin{array}{ccc} 1 & \text{if} & p_1 < p_2 \\ [0,1] & \text{if} & p_1 = p_2 \\ 0 & \text{if} & p_1 > p_2 \end{array} \right\}$$

Complexity case ($\sigma > 0$) : Smooth demand curve, no infinite drop at $p_1 = p_2$. At $p_1 = p_2 = p$ demand is $1/2$.

$$\max_{p_1} \Phi\left(\frac{p_2 - p_1}{\sigma\sqrt{2}}\right)[p_1 - c_1]$$

$$f.o.c. : -\frac{1}{\sigma\sqrt{2}}\phi\left(\frac{p_2 - p_1}{\sigma\sqrt{2}}\right)[p_1 - c_1] + \Phi\left(\frac{p_2 - p_1}{\sigma\sqrt{2}}\right) = 0$$

Intuition for non-zero mark-ups: Lower elasticity increases firm mark-ups and profits. Mark-up proportional to complexity σ.

Endogenous complexity

- Consider Normal case –> For $\sigma \to \infty$

$$\max_{p_1} \Phi\left(\frac{p_2 - p_1}{\sigma\sqrt{2}}\right)[p_1 - c_1] \to \max_{p_1} \frac{1}{2}[p_1 - c_1]$$

Set $\sigma \to \infty$ and obtain infinite profits by letting $p_1 \to \infty$

(Choices are random, Charge as much as possible)

- Gabaix and Laibson: Concave returns of complexity $Q_i(\sigma_i)$

  Firms increase complexity, unless "clearly superior" products in model with heterogenous products.

In a nutshell: market does not help to overcome bounded rationality. Competition may not help either

- More work on Behavioral IO:

- **Heidhus-Koszegi (2006, 2007)**

  - Incorporate reference dependence into firm pricing

  - Assume reference point rational exp. equilibrium (**Koszegi-Rabin**)

  - Results on

    * Price compression (consumers hate to pay price higher than reference point)

    * But also: Stochastic sales

- **Gabaix-Laibson (1996)**

  - Consumers pay attention to certain attributes, but not others (Shrouded attributes)

  - Form of limited attention

  - Firms charge higher prices on shrouded attributes (add-ons)

  - Similar to result in **DellaVigna-Malmendier (2004)**: Charge more on items consumers do not expect to purchase

- **Ellison (2006):** Early, very concise literature overview

- Future work: *Empirical Behavioral IO*

  - Document non-standard behavior

  - Estimate structurally

  - Document firm response to non-standard feature

3 Intro to Problem Set

- Problem set focused on financial markets
- Biases of investors and accountants
- Accounting — Information on company performance
  - accounting books
  - quarterly earnings announcement
- Two main focuses:
  - Optimal accounting rules
  - Stock price response to profitability information in accounting books

- What is right valuation of company?
  - Crucial to guarantee right allocation of capital
  - Denote $e_{t,k}$ earnings (profits) of company k in year t
  - Stock price = Discounted sum of future cash flows:
$$p_{t,k} = e_{t,k} + \frac{e_{t+1,k}}{1+r} + \frac{e_{t+2,k}}{(1+r)^2} + \ldots$$
  - Need forecasts of future profitability $e_{t,k}$

- Two main components:
  - Short-run earnings performance
  - Long-run performance
  - Analysts provide forecasts on both

- **Analysts.** Process information on companies and make it available (for a fee)
  - Sell-side. Work for brokerage firm (investment bank)
  - Buy-side. Work for mutual funds

  - Sell-side analysts:
    - more likely to have conflict of interest (Inv. Bank selling shares of target company)
    - data widely available (IBES, FirstCall)

- Analysts generate two main outputs:

  1. Earning forecasts $\hat{e}_{t,k}$
     - Dollar earning per share of company
     - Quarterly or annual
     - Forecast h years into the future: $h \simeq 3, 4$ years
  2. Long-term "growth rate" of earnings g_e

- Common forecasting model:

$$\begin{aligned} \hat{p}_{t,k} &= e_{t,k} + \frac{\hat{e}_{t+1,k}}{1+r} + \frac{\hat{e}_{t+2,k}}{(1+r)^2} + \ldots \\ &+ \sum_{t=0}^{\infty} \frac{1}{(1+r)^{h+t}} \hat{e}_{t+h,k} * g_e \end{aligned}$$

Company releases of information

- Each quarter: Announcement of accounting performance
  - Scheduled announcement, conference call
  - Release of accounting indicators
  - Special focus on earnings per share $e_{t,k}$
- Comparison of forecasted and realized earnings
- Measure of new information: earning surprise $e_{t,k} - \hat{e}_{t,k}$.
- Renormalize by price of share: $s_{t,k} = \left(e_{t,k} - \hat{e}_{t,k}\right)/p_{t,k}$
- Investors react to new information by updating stock price $p_{t,k}$

- **Problem set**
- Focus on response of stock prices to earning surprise
- Economic significance:
  - Processing of new information
    - Clean measure of information
    - Clean measure of response
  - Timing of release of information by company

- Identify in the data three anomalies:

- **Anomaly 1. Post-Earnings Announcement Drift. (Chan, Jegadeesh, and Lakonishok, 1996; Bernard and Thomas, 1989).**

  – Announcements of good news in earnings $e_{t,k}$ are followed by higher returns over next 2-3 quarters

  – Arbitrage should eliminate this

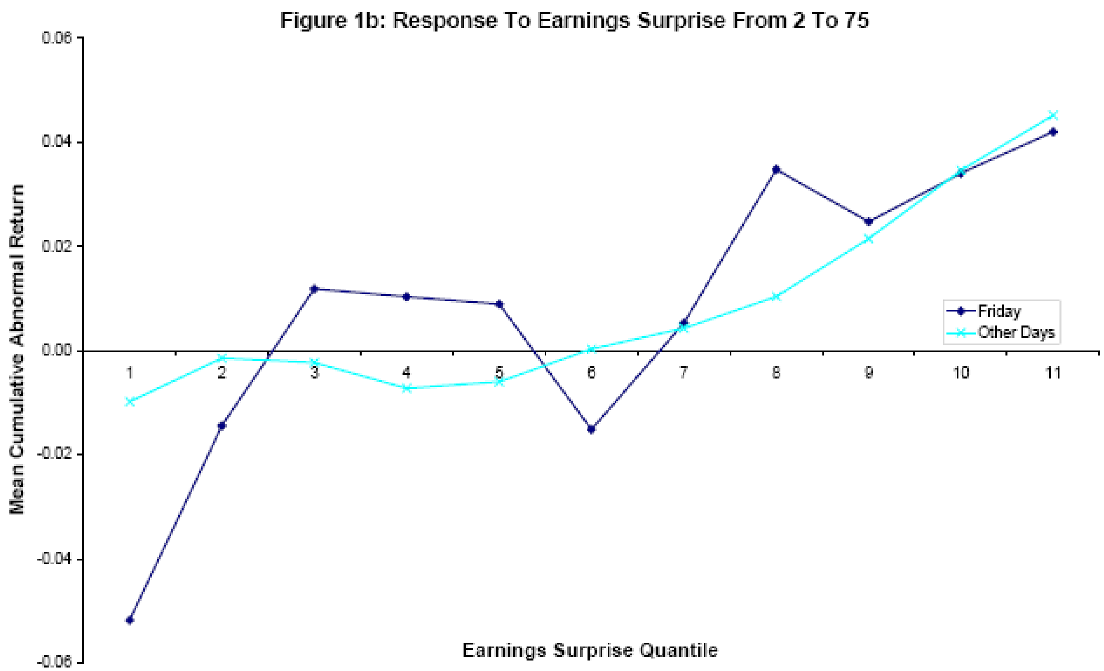
  – Interpretation: Investors inattentive initially, news incorporated slowly over time

- Measure new information using earnings surprise $s_{t,k}$

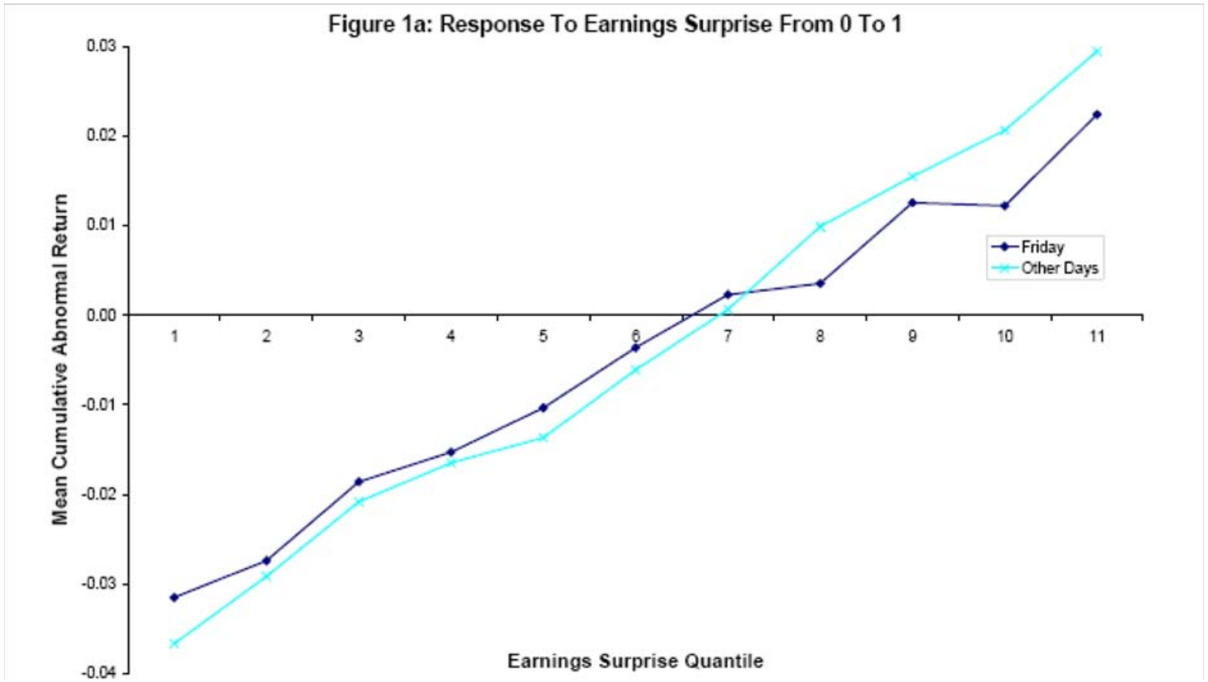
- Follow standard 'quantile' procedure: Divide into quantiles based on $s_{t,k}$

- Plot returns for each quantile

- Focus on light blue line for now (Figure from DellaVigna and Pollet, 2006)

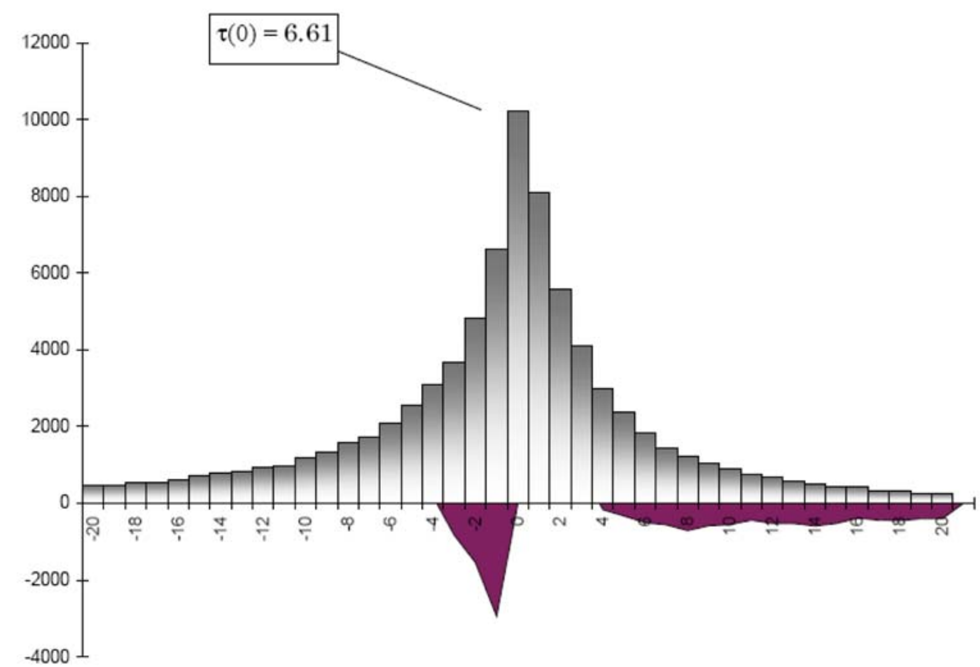


- **Anomaly 2. Less Immediate Response and more Drift when More distractions (DellaVigna-Pollet, forthc.; Hirshleifer-Lim-Teoh, 2007)**
  - Announcements on Friday (DVP) or with more competing news (HLT):
    - ∗ Drift stronger and Immediate response lower
    - ∗ Inattention: More distracted investors



- **Anomaly 3.** (Degeorge, Patel, and Zeckhauser, 1999)
  - CEOs shift the earnings so as to meet analyst expectations

Figure 6. Histogram of Forecast Error for Earnings Per Share: Exploring the threshold of "meet analysts' expectations"



- Similar result if earnings compared to earnings 4 quarters ago or compared to zero profits
- Interpretation:
  - Investors have 'bias': They penalize significantly companies that fail to meet thresholds
  - Managers cater to this bias by manipulating earnings

4 Methodology: Clustering Standard Errors

- Common econometric issue: Errors correlated across groups of observations

- Example 1. State-Year panel: Effect of abortion on crime

  1. Persistent shock to State over time (Autocorrelation)

  2. Correlation in shocks across State within year (Cross-Sectional correlation)

- Example 2. Earnings announcement panel

  1. Persistent shock to Company over time (Autocorrelation)

  2. Correlation in shocks across companies within date (Cross-Sectional correlation)

- OLS standard errors assume i.i.d. cross-sectionally and over time
- Clustered standard errors can take care of Issue 1 or 2 — not both:
  1. Cluster by State (Company):
     – Assume independence across States (companies)
     – Allow for any correlation over time within State (company)
  2. Cluster by year (date)
     – Assume independence across years (dates)
     – Allow for any correlation within a year (date) across States (companies)
- How does this work?

- Assume simple univariate regression:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it}$$

- OLS estimator:

$$\hat{\beta} = \beta + \left(x'x\right)^{-1} x'\varepsilon = \beta + \frac{Cov\left(x, \varepsilon\right)}{Var\left(x\right)}$$

- $Var\left(\hat{\beta}\right)$ under i.i.d. assumptions:

$$Var\left(\hat{\beta}\right)_{OLS} = \left(x'x\right)^{-1} \sum_{i,t} \left(x_{it}\hat{\varepsilon}_{it}\right)\left(\hat{\varepsilon}_{it}x_{it}\right)\left(x'x\right)^{-1} \frac{1}{NT} = \frac{\hat{\sigma}^2}{\sum x_{it}^2 NT}$$

- White-heteroskedastic:

$$Var\left(\hat{\beta}\right)_{Het} = \frac{1}{\sum_{it} x_{it}^2} \sum_{it} \frac{x_{it}^2 \hat{\varepsilon}_{it}^2}{\sum x_{it}^2}$$

- White-heteroskedastic:

$$Var\left(\hat{\beta}\right)_{Het} = \frac{1}{\sum_{it} x_{it}^2} \sum\nolimits_{it} \frac{\left(x_{it}\hat{\varepsilon}_{it}\right)^2}{\sum x_{it}^2}$$

  - Notice: Second sum is weighted average of $\hat{\varepsilon}_{it}^2$, with more weight given to observations with higher x_{it}^2
  - If high x_{it}^2 is associated with high $\hat{\varepsilon}_{it}^2$, $Var\left(\hat{\beta}\right)_{Het} > Var\left(\hat{\beta}\right)_{OLS}$

- Standard Errors Clustered by I (allow for autocorrelation):

$$Var\left(\hat{\beta}\right)_{Clust} = \frac{1}{\sum_{it} x_{it}^2} \sum\nolimits_{i} \frac{\left(\sum_t x_{it}\hat{\varepsilon}_{it}\right)^2}{\sum x_{it}^2}$$

  - First sum all the covariances $x_{it}\hat{\varepsilon}_{it}$ within a cluster
  - Then square up and add across the clusters
  - Notice: This is as if one cluster (one i) was one observation

- That is, this form of clustering allows

$$E(u_{it}u_{it'}|X_{it}X_{it'}) \neq 0$$

  - Correlation within cluster i

- Requires

$$E(u_{it}u_{i't'}|X_{it}X_{i't'}) = 0$$

  for $i \neq i'$

  - No correlation across clusters

- When is $Var\left(\hat{\beta}\right)_{Clust} > Var\left(\hat{\beta}\right)_{Het}$?
- Example: Assume $I = 2$, $T = 2$

$$Var\left(\hat{\beta}\right)_{Het} = \frac{1}{\sum_{it} x_{it}^2} \frac{(x_{11}\hat{\varepsilon}_{11})^2 + (x_{12}\hat{\varepsilon}_{12})^2 + (x_{21}\hat{\varepsilon}_{21})^2 + (x_{22}\hat{\varepsilon}_{22})^2}{\sum x_{it}^2}$$

- Compare to

$$\begin{aligned} Var\left(\hat{\beta}\right)_{Clust} &= \frac{1}{\sum_{it} x_{it}^2} \frac{(x_{11}\hat{\varepsilon}_{11} + x_{12}\hat{\varepsilon}_{12})^2 + (x_{21}\hat{\varepsilon}_{21} + x_{22}\hat{\varepsilon}_{22})^2}{\sum x_{it}^2} = \\ &= Var\left(\hat{\beta}\right)_{Het} + \frac{1}{\sum_{it} x_{it}^2} \frac{2x_{11}\hat{\varepsilon}_{11}\hat{\varepsilon}_{12}x_{12} + x_{21}\hat{\varepsilon}_{21}\hat{\varepsilon}_{22}x_{22}}{\sum x_{it}^2} \end{aligned}$$

  - Hence, $Var\left(\hat{\beta}\right)_{Clust} > Var\left(\hat{\beta}\right)_{Het}$ if $Ex_{i1}x_{i2} > 0$ and $E\hat{\varepsilon}_{i1}\hat{\varepsilon}_{i2} > 0$ –$>$Positive correlation within cluster among x variables and ε
  - Positive correlation –$>$ Standard errors understated if no clustering

- Calculation of Adjustment of Standard Errors due to Clustering
  - T observations within cluster
  - Within-cluster correlation of x_s: ρ_x
  - Within-cluster correlation of ε: ρ_ε
- Compare $Var\left(\hat{\beta}\right)_{Clust}$ and $Var\left(\hat{\beta}\right)_{OLS}$:

$$Var\left(\hat{\beta}\right)_{Clust} = Var\left(\hat{\beta}\right)_{OLS} * \left(1 + (T-1)\,\rho_x \varrho_\varepsilon\right)$$

  - Standard errors downward biased with OLS if $\rho_x \varrho_\varepsilon > 0$, or positive correlations (as above)
  - No bias if no correlation in *either* x or ε
  - Bias larger the larger is T
  - Illustrative case: Suppose all observations within cluster identical $(\rho_x = \rho_\varepsilon = 1)$ –> Bias $= T$

- Issues with clustering:

- Issue 1. **Number of clusters**

  - Convergence with speed I –> Need a large number of clusters I to apply LLN

  - Beware of papers that apply clustering with $<$20 clusters

  - **Cameron-Gelbach-Miller (2008):** Test with good finite sample properties even for $I \approx 10$

- Issue 2. **Cluster in only one dimension**

  - Clustering by I controls for autocorrelation

  - Clustering by T controls for cross-sectional correlation

  - How can control for both? Cannot really –> However: **Cameron-Gelbach-Miller (2006):** Two-way clustering, can do so

- Readings on clustered standard errors:
  - **Stata Manual** –> basic, intuitive
  - **Bertrand-Duflo-Mullainathan (QJE, 2004)** –> Excellent discussion of practical issues with autocorrelation in diff-in-diff papers, good intuition
  - **Peterson (2007)** –> Fairly intuitive, applied to finance
  - **Cameron-Trivedi (2006) and Wooldridge (2003)** –> More serious treatment
  - **Colin Cameron (Davis)'s website** –> Updates

5 Market Reaction to Biases: Behavioral Finance

- Who do 'smart' investors respond to investors with biases?
- First, brief overview of anomalies in Asset Pricing (from Barberis and Thaler, 2004)
  1. **Underdiversification.**
     (a) Too few companies.
        - Investors hold an average of 4-6 stocks in portfolio.
        - Improvement with mutual funds

     (b) Too few countries.
        - Investors heavily invested in own country.
        - Own country equity: 94% (US), 98% (Japan), 82% (UK)
        - Own area: own local Bells (Huberman, 2001)

(c) Own company

– In companies offering own stock in 401(k) plan, substantial investment in employer stock

2. **Naive diversification.**

– Investors tend to distribute wealth 'equally' among alternatives in 401(k) plan (Benartzi and Thaler, 2001; Huberman and Jiang, 2005)

3. **Excessive Trading.**

– Trade too much given transaction costs (Odean, 2001)

4. **Disposition Effect in selling**
   - Investors more likely to sell winners than losers

5. **Attention Effects in buying**
   - Stocks with extreme price or volume movements attract attention (Odean, 2003)

- Should market forces and arbitrage eliminate these phenomena?

- **Arbitrage**:
  - Individuals attempt to maximize individual wealth
  - They take advantage of opportunities for free lunches
- Implications of arbitrage: ‘Strange’ preferences do not affect pricing
- Implication: For prices of assets, no need to worry about behavioral stories
- Is it true?

- Fictitious example:
  - Asset A returns \$1 tomorrow with $p = .5$
  - Asset B returns \$1 tomorrow with $p = .5$

  - Arbitrage –> Price of A has to equal price of B
  - If $p_A > p_B$,
    * sell A and buy B
    * keep selling and buying until $p_A = p_B$
  - Viceversa if $p_A < p_B$

- Problem: Arbitrage is limited (de Long et al., 1991; Shleifer, 2001)
- In Example: can buy/sell A or B and tomorrow get fundamental value
- In Real world: prices can diverge from fundamental value

- Real world example. Royal Dutch and Shell
  - Companies merged financially in 1907
  - Royal Dutch shares: claim to 60% of total cash flow
  - Shell shares: claim to 40% of total cash flow
  - Shares are nothing but claims to cash flow
  - Price of Royal Dutch should be 60/40=3/2 price of Shell

- p_{RD}/p_S differs substantially from 1.5 (Fig. 1)

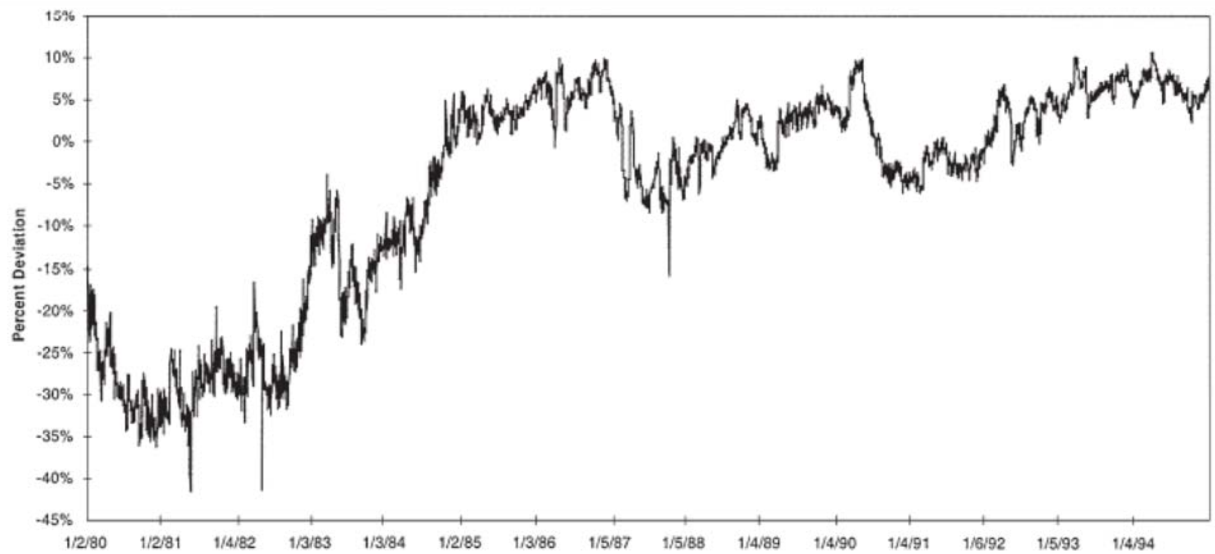


Fig. 1. Log deviations from Royal Dutch/Shell parity. Source: Froot and Dabora (1999).

- Plenty of other example (Palm/3Com)

- What is the problem?
  - Noise trader risk, investors with correlated valuations that diverge from fundamental value
  - (Example: Naive Investors keep persistently bidding down price of Shell)
  - In the long run, convergence to cash-flow value
  - In the short-run, divergence can even increase
  - (Example: Price of Shell may be bid down even more)

- **Noise Traders**
- DeLong, Shleifer, Summers, Waldman (*JPE* 1990)
- Shleifer, *Inefficient Markets*, 2000
- Fundamental question: What happens to prices if:
  - (Limited) arbitrage
  - Some irrational investors with correlated (wrong) beliefs
- First paper on Market Reaction to Biases
- *The* key paper in Behavioral Finance

The model assumptions

A1: arbitrageurs risk averse and short horizon

$\longrightarrow$ Justification?

* Short-selling constraints

  (per-period fee if borrowing cash/securities)

* Evaluation of Fund managers.

* Principal-Agent problem for fund managers.

A2: noise traders (Kyle 1985; Black 1986)

misperceive future expected price at t by

$\rho_t \overset{i.i.d.}{\sim} \mathcal{N}(\rho^*, \sigma_\rho^2)$

misperception *correlated* across noise traders ($\rho^* \neq 0$)

$\longrightarrow$ Justification?

* fads and bubbles (Internet stocks, biotechs)

* pseudo-signals (advice broker, financial guru)

* behavioral biases / misperception riskiness

What else?

- μ noise traders, $(1 - \mu)$ arbitrageurs
- OLG model
  - Period 1: initial endowment, trade
  - Period 2: consumption
- Two assets with identical dividend r
  - safe asset: perfectly elastic supply
    $\Longrightarrow$ price=1 (numeraire)
  - unsafe asset: inelastic supply (1 unit)
    $\Longrightarrow$ price?
- Demand for unsafe asset: λ^a and λ^n, with $\lambda^a + \lambda^n = 1$.

- CARA:

$$U(w) = -e^{-2(\gamma w)} \text{ (w wealth when old)}$$

$$E\left[U(w)\right] = \int_{\infty}^{\infty} -e^{-2\gamma w} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}(w-\overline{w})}$$

$$= -e^{-2\gamma(\overline{w}-\gamma\sigma_w^2)}$$

$$\Downarrow$$

$$\max E\left[U(w)\right] \quad \underset{\text{pos. mon. transf.}}{\curvearrowright} \quad \max \overline{w} - \gamma\sigma_w^2$$

Arbitrageurs:

$$\text{max}(w_t - \lambda_t^a p_t)(1 + r)$$

$$+\lambda_t^a (E_t[p_{t+1}] + r)$$

$$-\gamma\, (\lambda_t^a)^2\, Var_t(p_{t+1})$$

Noise traders:

$$\text{max}(w_t - \lambda_t^n p_t)(1 + r)$$

$$+\lambda_t^n (E_t[p_{t+1}] + \rho_t + r)$$

$$-\gamma\, (\lambda_t^n)^2\, Var_t(p_{t+1})$$

(Note: Noise traders know how to factor the effect of future price volatility into their calculations of values.)

f.o.c.

Arbitrageurs: $\frac{\partial E[U]}{\partial \lambda_t^a} \overset{!}{=} 0$

$$\lambda_t^a = \frac{r + E_t[p_{t+1}] - (1+r)p_t}{2\gamma \cdot Var_t(p_{t+1})}$$

Noise traders: $\frac{\partial E[U]}{\partial \lambda_t^n} \overset{!}{=} 0$

$$\lambda_t^n = \frac{r + E_t[p_{t+1}] - (1+r)p_t}{2\gamma \cdot Var_t(p_{t+1})}$$

$$+ \quad \frac{\rho_t}{2\gamma \cdot Var_t(p_{t+1})}$$

Interpretation

- Demand for unsafe asset function of:
  - $(+)$ expected return $(r + E_t[p_{t+1}] - (1 + r)p_t)$
  - (-) risk aversion (γ)
  - (-) variance of return $(Var_t(p_{t+1}))$

  - $(+)$ overestimation of return ρ_t (noise traders)

- Notice: noise traders hold more risky asset than arb. if $\rho > 0$ (and viceversa)

- Notice: Variance of prices come from noise trader risk. "Price when old" depends on uncertain belief of next periods' noise traders.

Impose general equilibrium: $\lambda^a + \lambda^n = 1$

Price

$$p_t = 1 + \frac{\mu(\rho_t - \rho^*)}{1+r} + \frac{\mu\rho^*}{r} - \frac{2\gamma\mu^2\sigma_\rho^2}{r(1+r)^2}$$

- Noise traders affect prices!

Interpretation

- Term 1: Variation in noise trader (mis-)perception
- Term 2: Average misperception of noise traders
- Term 3: Compensation for noise trader risk
- Special case: $\mu = 0$ (no noise traders)

Relative returns of noise traders

- Compare returns to noise traders R^n to returns for arbitrageurs R_a:

$$\Delta R = R^n - R^a = (\lambda_t^n - \lambda_t^a)\left[r + p_{t+1} - p_t(1+r)\right]$$

$$E(\Delta R) = \rho^* - \frac{(1+r)^2(\rho^*)^2 + (1+r)^2\sigma_\rho^2}{2\gamma\mu\sigma_\rho^2}$$

- Noise traders hold more risky asset if $\rho^* > 0$
- Return of noise traders can be higher if $\rho^* > 0$ (and not too positive)
- Noise traders therefore may outperform arbitrageurs if optimistic!
- (Reason is that they are taking more risk)

Welfare

- Sophisticated investors have higher utility
- Noise traders have lower utility than they expect
- Noise traders may have higher returns (if $\rho^* > 0$)
- Noise traders do not necessarily disappear over time

- Three fundamental assumptions

   1. OLG: no last period; short horizon

   2. Fixed supply unsafe asset (a cannot convert safe into unsafe)

   3. Noise trader risk systematic

- Noise trader models imply that biases affect asset prices:

   - Reference Dependence

   - Attention

   - Persuasion

- Here:
  - Biased investors
  - Non-biased investors
- Behavioral corporate finance:
  - Investors (biased)
  - CEOs (smart)
- Behavioral Industrial Organization:
  - Consumers (biased)
  - Firms (smart)

6 Next Lecture

- More Market Response to Biases
  - Managers: Corporate Decisions
  - Employers: Contracting
  - Politicians: Political Economy
  - Welfare Response to Biases
- Methodology of Field Psychology and Economics
- Concluding Remarks