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

Stefano DellaVigna

April 22, 2020

Outline

1. Methodology: Structural Behavioral Economics
2. Behavioral Finance
3. Behavioral Corporate Finance
4. Behavioral Labor
5. Behavioral Development

Section 1

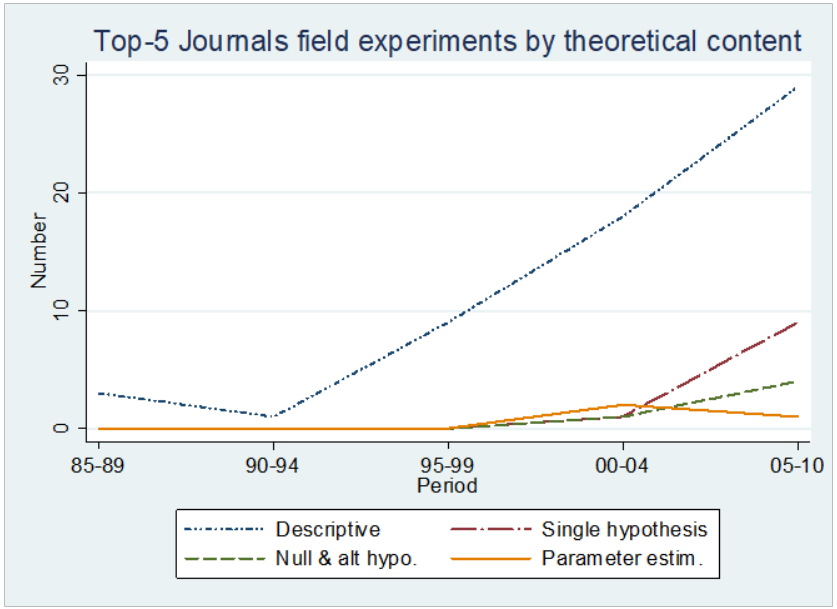
Methodology: Structural Behavioral Economics

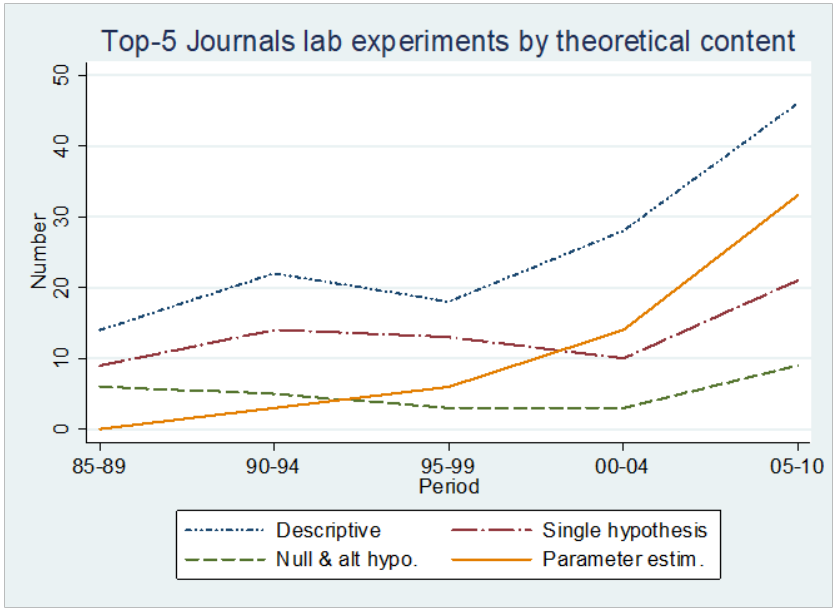
Overview

- Structural estimation in behavioral economics
  - Use model for estimation
  - Estimate key model parameters
- What do we mean by structural?
  "Estimation of a model on data that recovers parameter estimates (and c.i.s) for some key model parameters"
- Chapter in *Handbook of Behavioral Economics*

Overview

- Chapter Focused on applications to field evidence. Structural lab evidence is way ahead (**Card, DellaVigna, Malmendier, JEP 2012**)





Overview Advantages

- Six advantages to Structural Behavioral Econ:
  1. (Calibration) It builds on, and expands, great behavioral tradition of calibrating models: Are magnitudes right?
  2. (Model and Assumptions) Helps to better understand the model and clarifies implicit assumptions
  3. (Stability) Helps to understand whether key behavioral parameters are stable, including out of sample
  4. (Out of Sample) Allows for out of sample predictions which can be tested
  5. (Design) Can lead to better experimental design
  6. (Welfare and Policy) It allows for welfare evaluation and policy counterfactuals

Overview Limitations

- Three limitations to Structural Behavioral Econ:
  1. (Not the Right Tool) Not all questions lend themselves obviously to parameter estimation
  2. (Complexity and Time Costs) It will, generally, take long, and there is higher possibility of errors
  3. (Robustness) Need extra work to make sure estimates are robust, and which assumptions are driving them

Calibration

- Importance of calibrating models is lesson ONE from behavioral economics
- Example 1: Inertia in retirement savings
  - Standard model can explain *qualitative* pattern given switching costs k
  - But magnitudes? Costs would need to be ridiculous (**O'Donoghue and Rabin, 1998**)
  - Instead, procrastination plausible for naïve β-δ model even with β very close to 1 (O'Donoghue and Rabin, 1999; 2001)
  - Problem set 1: Extend O'Donoghue and Rabin calibration with more realistic assumptions (stochastic k)
- Example 2: **Rabin (EMA 2000)** calibration theorem on risk

Calibration

- The lesson of calibration is that it is important to check for reasonable values of the parameters
- Example. **DellaVigna et al., 2017** on job search
  - Exp discounting and $\beta\delta$ model have about same fit
  - BUT δ model has 15-day δ=.9 (implausible impatience)
  - β model instead has β =.6 (in range of other estimates)
  - estimated values for exp discounting outside the range of plausible, not so for $\beta\delta$ model

Model and Assumptions

- Point 1. Strengthens evidence-theory debate
  - Sketch of model will not suffice
  - Full specification in order to do estimation, forced to work out details
  - Run countless simulations at different parameter values
  - $\rightarrow$Leads to understanding model better
  - Does model really do what you thought it would do?
  - $\rightarrow$Can even lead to theoretical break-through
  - **Barseghyan, Molinari, O'Donoghue, Teitelbaum** in working out AER 2013 estimation of insurance choice got result on non-identification of loss aversion with KR preferences

Model and Assumptions

- Point 2. Better empirical test
- Example 1: **Genesove and Mayer (QJE)** is pioneering application of reference dependence to housing
  - Individuals hate to sell house at loss relative to purchase price
    - Yet, GM did not work out a model of reference dependence
    - If one writes one, one does not get the GM specification
    - One does get, though, the prediction of bunching at the last house price sale (which they did not test)
  - Example 2: Similar issues (as well known now) for **Camerer et al. (1997)** cab drivers paper

Model and Assumptions

- Point 3. Clarify needed assumptions
- Real-effort experiment (e.g., **Gneezy et al., 2003; Gill et al., 2016**)

$$Effort_{n,s,r} = \beta T + \gamma \boldsymbol{X_{n,s}} + \varepsilon_{n,s,r}$$

- What assumption are behind such specification?
  - Effort is outcome of maximization decision, $max_{e_i} s(T_i) e_i - \frac{\exp(\gamma e_i)}{\gamma} \eta_i$
  - Assume η log-normal, $ln(\eta_i) \sim N(\gamma k(X_i), \gamma^2 \sigma^2)$.
  - Taking first order conditions,

$$e_i = \frac{1}{\gamma} log[s(T_i)] - k(X_i) + \epsilon_i.$$

  - Can estimate with Non-linear Least Squares, almost like OLS (nl in Stata) – Easy!

Model and Assumptions

- Do you buy the needed assumptions?
- Can assume alternative functional form assumptions
- **Power Cost Function:**

$$c(e) = \frac{e^{1+\gamma}}{1+\gamma}$$

- Implied expression for effort is **(DellaVigna, List, Malmendier, and Rao, 2016)**

$$log(e_i) = \frac{1}{\gamma} log\left[s\left(T_i\right)\right] - k\left(X_i\right) + \epsilon_i$$

Stability

- Behavioral economics has convergence on some parsimonious models:
  - Beta-delta model of time preference (Laibson, 1997; O'Donoghue and Rabin, 1999)
  - Reference-dependence model of risk preferences (Kahneman and Tversky, 1979; Koszegi and Rabin, 2006
- For these models, is there reasonable agreement in parameters across settings?
- Case 1. Beta-delta model
  - Evidence from the field provided evidence of present bias
  - BUT Laboratory evidence (**Andreoni and Sprenger AER**) estimated β close to 1
  - Design with effort choice a la **Augenblick, Niederle, and Sprenger, QJE** appears to solve the puzzle: $\beta = 0.9$ over effort, not on money since timing of money is fungible

Stability

- Case 2. Reference-dependent model
- Most of the focus is on loss aversion parameter λ and on reference point r.
- But probability weighting $\pi(p)$ plays an important role
  - overweighting of small probabilities
  - **Sydnor (2012) and Barseghyan et al:** helps explain home insurance purchases
  - **Barberis (2018):** can explain preferences for IPOs which are skewed
  - Evidence from laboratory lottery choices is strong: $\pi(0.01) = 0.06$

Stability

- Inspired by this, simple design: compare effect of two incentives
  - A. Piece rate of p
  - B. Piece rate of $100p$, paid with probability 0.01
  - With probability weighting, B should be more effective

Table 2b. Evidence for Overwighting of Small Probabilities, Studies with Probabilistic Incentives

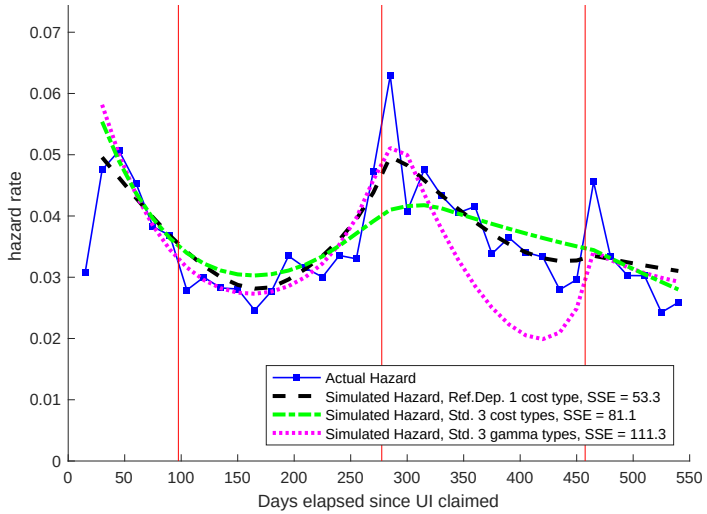
Paper	Subjects	Effort Task	Sample Size	Treatments (Certain Reward vs. Probabilistic Reward with low p)	Effort with Certain Reward, Mean (S.D.)	Effort with Probabilistic Reward, Mean (S.D.)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B. Field Studies Comparing Certain Reward to Probabilistic Reward						
DellaVigna and Pope (forthcoming)	Mturk	Button Presses	555 (P), 558 (C)	1% chance of winning US$1 (P) vs. fixed payment of US$0.01 (F) per 100 presses	2029 (27.47)	1896 (28.44)
Halpern et al. (2011)	Resident Physicians in a US Database	Survey Response	358 (P), 400 (C)	0.4% chance of winning US$2500 (P) vs. fixed payment of US$10 (F) for response	0.558 (0.497)	0.511 (0.500)
Thirumurthy et al. (2016)	Men aged 21 to 39 years old in Kenya	Uptake of Circumcision	302 (P), 308 (C)	Mixed lottery with expected retail value of US$12.50 (P) vs. food voucher worth US$12.50 (F)	0.084 (0.278)	0.033 (0.179)
Diamond and Loewy (1991)	Undergraduates in State University	Recycling	78 (P), 113 (C)	5% chance of winning $5 and 1% chance of winning $25 (P) vs. $0.50 voucher for campus store (F)	0.212 (0.409)	0.308 (0.462)
Dolan and Rudisill (2014)	16 to 24 year olds in England	Return Test Kit via Mail	247 (P), 549 (C)	10% chance of a 50 GBP Tesco voucher (P) vs. 5 GBP Tesco voucher (F)	0.732 (0.443)	0.706 (0.455)

- Do not find evidence of probability weighting!

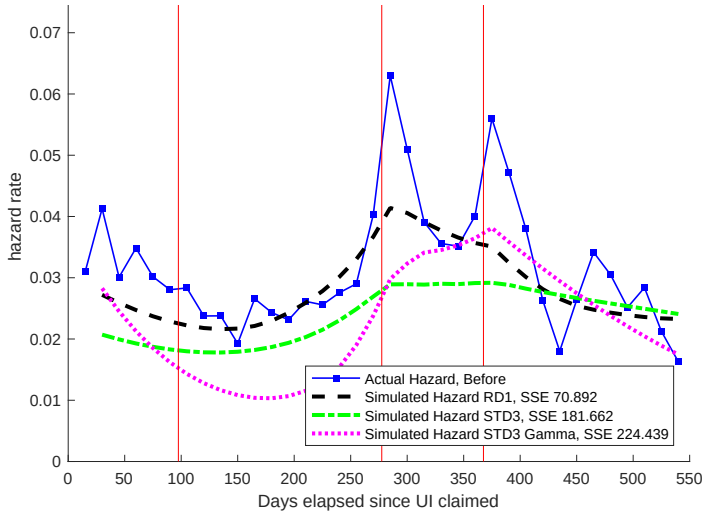
Out of Sample

- Structual estimates allow for out of sample predictions (McFadden et al. 1977; Todd and Wolpin, 2006)
- Some examples in behavioral economics
- Example 1. Social Image. **DellaVigna et al. (2017 RES)**.
  - Estimate impact of get-out-the-vote intervention based on flyer experiment
- Example 2. Job Search. **DellaVigna et al. (2017 QJE)**
  - Estimate reference dependent model has good fit, even with just 1 type
  - A specific "standard" model (3-type heterogeneity in elasticity) can also fit well
  - Examine out of sample, consider smaller reform 2 years prior
  - Also consider group with lower earnings which had different reform

Stability



(a) Out-of-sample predictions of models for unemployment system 2 years prior to reform and empirical hazard



Experimental Design

- Structural estimation is most often used on observational data (e.g. consumption/savings)
  - In fact, sometimes used as substitute for clear identification (not recommended)
- BUT estimation is perfect for experiments
- Advantage 1. Get best of both worlds
  1. Get reduced-form results / treatment effects
  2. PLUS estimate parameters based on experimental results

Experimental Design

- Advantage 2: estimation informs the exp. design
- Idea:
  1. set up model + estimation before running experiment
  2. Create simulated data set (possibly using pilot data)
  3. Attempt to estimate on this data
- Often you will realize that
  - you need an extra treatment in design...
  - or more sample in one treatment...
  - or you are badly underpowered
  - $\rightarrow$Change design! (Cannot do this in obs. studies)
- Different from reduced-form power studies, as this is about estimating the parameters

Experimental Design

- Example 1: Time preference experiments a la **Andreoni and Sprenger** or **Augenblick et al.**
  - Designed so as to estimate time preferences
  - Also, design to estimate confounding parameters (curvature of utility and cost of effort function)
- Example 2: Charity and Social Pressure Experiment **(DLM QJE)**
  - Add entire set of experiments to estimate cost of sorting in/out of the home
- Example 3. Limited attention and taxation
  - **Allcott and Taubinsky (AER 2016)** on energy inattention
  - **Taubinsky and Rees-Jones (2016)** on limited attention to taxes

Experimental Design

- What design tricks facilitate estimation?
- Trick 1. "Price out" behavioral parameters comparing to an intervention in $
  - **DLM (2012)**– survey treatment where advertise $ payment
  - **Andreoni and Sprenger (AER 2011); Augenblick, Niederle, and Sprenger (QJE 2014)** –variation in interest rate
- Trick 2. Within-subject experiment
  - Often allows to extract more information
  - BUT trade-off with simplicity of design
  - **Allcott and Taubinsky (AER 2016)** on energy inattention
  - **Taubinsky and Rees-Jones (2016)** on limited attention to taxes
  - Real Effort experiments and identification of cost of effort **(DellaVigna, List, Malmendier, and Rao, 2016)**

Welfare and Policy

- Advantage of estimating model is... you can use it!
  - Compute welfare of setting versus counterfactuals
  - Estimate effect of potential policies
- Some examples:
  - **DellaVigna, List, and Malmendier (QJE 2012)** on charity
  - **Handel (AER 2013)** on health insurance
  - **DellaVigna, List, Malmendier, Rao (REStud 2017)** on voting
  - **Bernheim, Fradkin, Popov (AER 2015)** on retirement saving
  - **Allcott and Taubinsky (AER 2015)** on energy

Overview Limitations

- Three limitations to Structural Behavioral Econ:
  1. (Not the Right Tool) Not all questions lend themselves obviously to parameter estimation
  2. (Complexity and Time Costs) It will, generally, take long, and there is higher possibility of errors
  3. (Robustness) Need extra work to make sure estimates are robust, and which assumptions are driving them

Limitation 1: Not Right Tool

- Not all questions lend themselves obviously to parameter estimation
  - Exploratory question on which we do not have a good model
    - Example: **Dana, Weber, and Kuang (2007)** on moral wriggleroom
    - Example: Framing effects (eg **Benartzi and Thaler, 2002**)
  - Many reduced form/policy question
    - **Bhargava and Manoli AER**
  - Models and Axioms
    - Models can provide comparative statics test of different models, or axioms
    - Can do so without structural estimation, do not need to specify all assumptions

Limitation 2: Complexity and Time Costs

- Estimation can be relatively easy, but it typically is not
- You will learn a lot about optimization methods!
- And... there will be bugs in your code
  - Test very extensively, have other people peer review the code
  - Use simulations: simulate and estimate
  - Example of error: **Apesteguia and Ballester (JPE 2018)** point flaw in earlier Harrison paper
- Keep in mind the objective: Complexity of the model and estimation is not the aim, it is a necessary evil
- BUT structural model can be simple
  - If rich data provides necessary variation (*sufficient statistic* approach)
  - OR if data collection / experiment is set up to make *estimation simple*

Limitation 3: Robustness to Assumptions

- What about pitfalls?
- Assume you fully specify model and estimate it
  - You spend years of life doing it
  - Model implies welfare and policy implications
  - Now you sell the implications for policy and welfare
- **Warning 1.** Estimation and implications are only as good as assumptions going into it
  - Test much robustness
- **Warning 2.** Standard errors do not acknowledge model mis-specification $\rightarrow$ Point estimates are likely too precise
- **Warning 3.** Do you do as well out of sample as in sample?

Limitation 3: Robustness to Assumptions

- How to avoid pitfall: **Allcott and Taubinsky (AER 2016)**
- Consumers make choices between incandescent and CFL

	Choice A **Philips 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack**	**Choice B** **Philips 60-Watt Incandescent Light Bulbs, 4-Pack**
		

Decision Number	**Choice A** 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack	**Choice B** 60-Watt Incandescent Light Bulbs, 4-Pack
	Purchase **Choice A for free**	Purchase **Choice B for $10**
1)	○	○
	Purchase **Choice A for free**	Purchase **Choice B for $8**
2)	○	○
	Purchase **Choice A for free**	Purchase **Choice B for $6**
3)	○	○

- One group receives information on energy savings, other not

Limitation 3: Robustness to Assumptions

- Moderate Shift in demand curve due to information
- What if not sure about some key assumptions?

		(1)	(2)	(3)
Row	Scenario	Optimal Subsidy ($/pckg)	Welfare Effect of Ban ($/pckg)	Effect of Ban (Percent of "Perceived Surplus")
1	Base	3	-0.44	-41
	If censored, assume ...			
2	WTP={$12,-$12}	3	-0.34	-36
3	WTP={$20,-$20}	3	-0.60	-47
4	self-reported hypothetical WTP	3	-0.61	-43
	Scale average marginal bias to match ...			
5	consumers who pass review "quiz"	3	-0.41	-38
6	consumers w/ "correct" endline beliefs	3	-0.13	-12
7	Balanced Treatment group	3	-0.48	-45
8	10 percent confidence bound	1	-0.92	-86
9	90 percent confidence bound	(Ban)	0.05	4
	Additional Distortion Computed from Excess Mass Test			
10	Excess mass consumers have $v = 7.66$	8	1.22	114

Common Methods of Estimation

1. Minimum Distance Estimator / Method of Moments
2. Maximum Likelihood
3. Others (OLS / NLLS / "Bunching")

Minimum Distance

- Minimum Distance Steps
  1. Pick the observed empirical moments to match, $\hat{m}$
  2. Solve/ Simulate the model at a given set of parameters θ and generate the same moments, $m(\theta)$
  3. Find the set of parameters that minimize the distance between the empirical and model-generated moments
     $\hat{\theta} = argmin(m(\theta) - \hat{m})' W (m(\theta) - \hat{m})$
     - Example: **Laibson, Repetto, Maxted, and Tobacman (2015)**
       - Study consumption-savings of households
       - Step 1: Many possible moments, pick important ones
       - Step 2 (hardest for them): Write full consumption-savings problem as function of model parameters
       - Step 3 (easy): Grid search of parameters such that would have good fit

Minimum Distance

- Advantages of Minimum Distance:
  - Transparent in what identifies: moments you pick Perfect for experiments:
    - Moment 1 is X in control, Moment 2 is X in treatment
  - Also, allows you to post the moments even if data is confidential
- Disadvantages of Minimum Distance:
  - Does not use all the information in the data, only what is contained in moments
  - Sensitive to choice of moments

Maximum Likelihood

- Maximum Likelihood Steps
  1. Specify fully the statistical model, deriving likelihood $L(x|\theta)$, where x is data
  2. Maximize likelihood given the data, picking parameters $\hat{\theta}$
- **Augenblick and Rabin (RES, 2019)**
  - Model implies that optimal effort is given by

$$e^* = \arg\max_e \quad \delta^{T-k} \cdot (e \cdot w) - \frac{1}{\beta^{\mathbf{1}(k=t)}} \cdot \frac{1}{\beta_h^{\mathbf{1}(p=1)}} \cdot \delta^{t-k} \cdot \frac{1}{\varphi \cdot \gamma}(e+10)^{\gamma}.$$

  - Leads to first order condition

$$e^* = \left(\frac{\delta^{T-k} \cdot \varphi \cdot w}{\frac{1}{\beta^{\mathbf{1}(k=t)}} \cdot \frac{1}{\beta_h^{\mathbf{1}(p=1)}} \cdot \delta^{t-k}}\right)^{\frac{1}{\gamma-1}} - 10.$$

  - Assuming additive noise in observed effort yields likelihood

$$L(e_j) = \phi\left(\frac{e_j^* - e_j}{\sigma}\right)$$

Maximum Likelihood

- Advantages of Maximum Likelihood:
  - Uses all the information in the data
- Disadvantages of Minimum Distance:
  - Identification is less transparent
  - Observations with a very low likelihood could be driving the results

II. Randomness

- Consider a typical model solution, eg., Nash eq., or Market Equilibrium
- Solution often is one equilibrium action, or one price
- Yet, in reality we always a distribution of outcomes
- Where does the randomness come from?
- That will be key step to take model into econometrics.
- Three broad categories
  1. Random utility (McFadden logit)
  2. Random coefficients
  3. Implementation error

Section 2

Behavioral Finance

Behavioral Finance Anomalies

- How do 'smart' investors respond to investors with biases?
- First, brief overview of anomalies in Asset Pricing (from Barberis and Thaler, 2004)

1. **Underdiversification.**
   1. Too few companies.
      - Investors hold an average of 4-6 stocks in portfolio.
      - Improvement with mutual funds
   2. 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)
   3. 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)

6. **Inattention to Fees.**

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

6. **Inattention to Fees.**

Should market forces and arbitrage eliminate these phenomena?

Arbitrage

- By assumption:
  - 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 $\rightarrow$ 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 $p_A < p_B$

But...

- 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

Royal Dutch and Shell

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

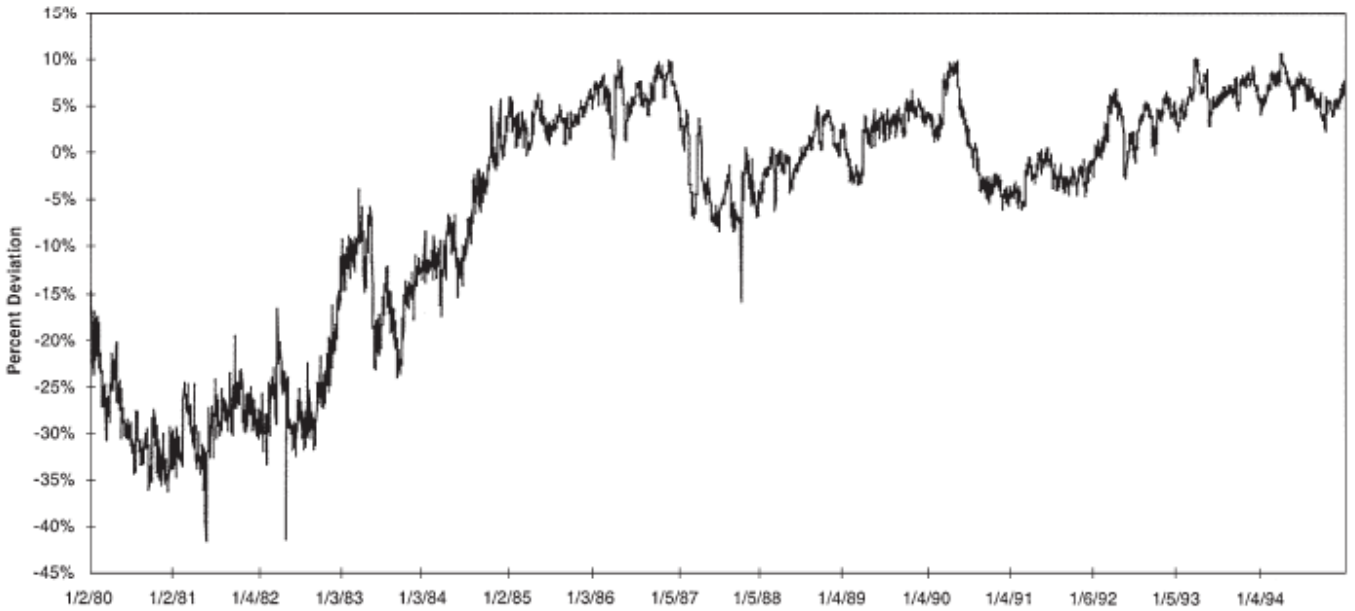


Fig. 1. Log deviations from Royal Dutch/Shell parity. Source: Froot and Dabora (1999).

- Plenty of other examples (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

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.

Model Assumptions

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 $\lambda^n \mu + \lambda^a (1-\mu) = 1$.
- CARA: $U(w) = -e^{-2\gamma w}$ (w wealth when old)

$$
\begin{aligned}
E\left[U(w)\right] &= \int_{-\infty}^{\infty} -e^{-2\gamma w} \cdot \frac{1}{\sqrt{2\pi\sigma_w^2}} \cdot e^{-\frac{1}{2\sigma^2}(w-\bar{w})^2} dw \\
&= -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_w^2}} \cdot e^{-\frac{4\gamma w\sigma_w^2 + w^2 + \bar{w}^2 - 2w\bar{w}}{2\sigma_w^2}} dw \\
&= -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_w^2}} \cdot e^{-\frac{(w-[\bar{w}-2\gamma\sigma_w^2])^2 - 4\gamma^2\sigma_w^4 + 4\gamma\sigma_w^2\bar{w}}{2\sigma_w^2}} dw \\
&= -e^{-\frac{4\gamma\sigma_w^2\bar{w} - 4\gamma^2\sigma_w^4}{2\sigma_w^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_w^2}} \cdot e^{-\frac{(w-[\bar{w}-2\gamma\sigma_w^2])^2}{2\sigma_w^2}} dw \\
&= -e^{-\frac{4\gamma\sigma_w^2\bar{w} - 4\gamma^2\sigma_w^4}{2\sigma_w^2}} = -e^{-2\gamma\left(\bar{w} - \gamma\sigma_w^2\right)}
\end{aligned}
$$

$$\Downarrow$$

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

Arbitrageurs:

$$\max(w_t - \lambda_t^a p_t)(1 + r)$$
$$+\lambda_t^a(E_t[p_{t+1}] + r)$$
$$-\gamma\left(\lambda_t^a\right)^2 Var_t(p_{t+1})$$

Noise traders:

$$\max(w_t - \lambda_t^n p_t)(1 + r)$$
$$+\lambda_t^n(E_t[p_{t+1}] + \rho_t + r)$$
$$-\gamma\left(\lambda_t^n\right)^2 Var_t(p_{t+1})$$

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

f.o.c.

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

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

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

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

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: $\lambda^n \mu + \lambda^a (1-\mu) = 1$ to obtain

$$\begin{aligned} 1 &= \frac{r + E_t[p_{t+1}] - (1+r)p_t}{2\gamma \cdot Var_t(p_{t+1})} + \mu \frac{\rho_t}{2\gamma \cdot Var_t(p_{t+1})} \text{ or} \\ p_t &= \frac{1}{1+r}\left[r + E_t[p_{t+1}] - 2\gamma \cdot Var_t(p_{t+1}) + \mu\rho_t\right] \end{aligned}$$

- To solve for p_t, we need to solve for $E_t[p_{t+1}] = E[p]$ and $Var_t(p_{t+1})$

$$\begin{aligned} E[p] &= \frac{1}{1+r}\left[r + E_t[p] - 2\gamma \cdot Var_t(p_{t+1}) + \mu E[\rho_t]\right] \\ E[p] &= 1 + \frac{-2\gamma \cdot Var_t(p_{t+1}) + \mu\rho^*}{r} \end{aligned}$$

- Rewrite p_t plugging in

$$\begin{aligned} p_t &= 1 - \frac{2\gamma \cdot Var_t(p_{t+1})}{r} + \frac{\mu\rho^*}{r(1+r)} + \frac{\mu\rho_t}{1+r} \\ Var[p_t] &= Var\left[\frac{\mu\rho_t}{1+r}\right] = \frac{\mu^2}{(1+r)^2} Var(\rho_t) = \frac{\mu^2}{(1+r)^2}\sigma_\rho^2 \end{aligned}$$

- Rewrite p_t

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

- Noise traders affect prices!
- Term 1: Variation in noise trader (mis-)perception
- Term 2: Average misperception of noise traders
- Term 3: Compensation for noise trader risk

Relative returns of noise traders

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

$$\begin{aligned} \Delta R &= R^n - R^a = \left(\lambda_t^n - \lambda_t^a\right)\left[r + p_{t+1} - p_t\left(1+r\right)\right] \\ E\left(\Delta R|\rho_t\right) &= \rho_t - \frac{\left(1+r\right)^2 \rho_t^2}{2\gamma\mu\sigma_\rho^2} \\ E\left(\Delta R\right) &= \rho^* - \frac{\left(1+r\right)^2 \left(\rho^*\right)^2 + \left(1+r\right)^2 \sigma_\rho^2}{2\gamma\mu\sigma_\rho^2} \end{aligned}$$

- 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

Section 3

Behavioral Corporate Finance

Behavioral Corporate Finance

- **Version 1. Behavioral Managers, Rational Investor**
  - **Malmendier and Tate** on mergers, cash-flow sensitivity
- **Version 2. Rational Managers, Behavioral Investors**
- **Baker, Ruback, and Wurgler (2005)**. Firm has to decide how to finance investment project:
  1. internal funds (cash flow/retained earnings)
  2. bonds
  3. stocks

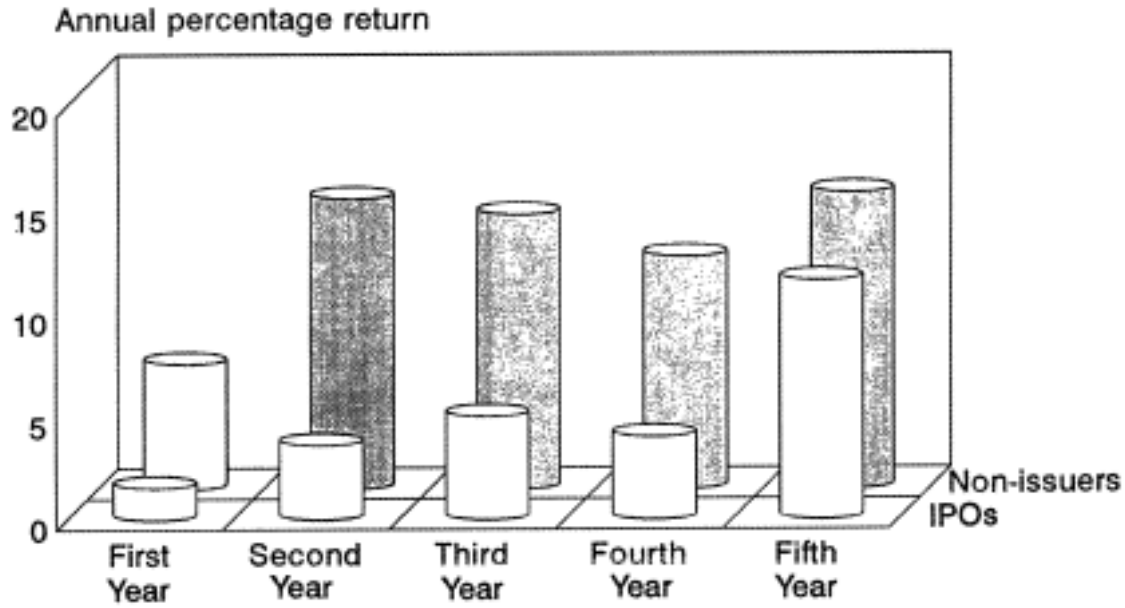
Findings

- Fluctuation of equity prices due to noise traders
- Managers believe that the market is inefficient
  - Issue equity when stock price exceeds perceived fundamental value
  - Delay equity issue when stock price below perceived fundamental value
- Consistent with
  - Survey Evidence of 392 CFO's (Graham and Harvey 2001): 67% say under/overvaluation is a factor in issuance decision
  - Insider trading

- Go over quickly one important example

Long-run performance of equity issuers

- Market Timing prediction: Companies issuing equity underperform later
- **Loughran-Ritter (1995)**: Compare matching samples of
  - companies doing IPOs
  - companies not doing IPOs but have similar market cap.



Similar finding with SEOs

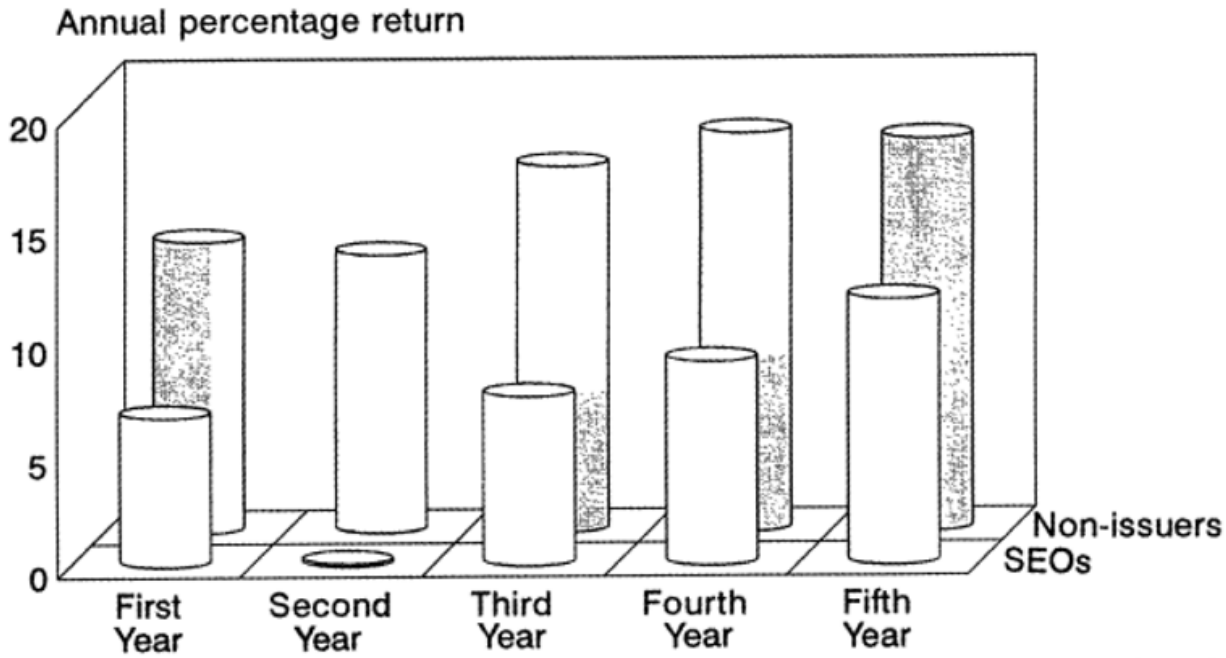


Figure 2. The average annual raw returns for 4,753 initial public offerings (IPOs), and their matching nonissuing firms (top), and the average annual raw returns for 3,702 seasoned equity offerings (SEOs), and their matching nonissuing firms (bottom), during the five years after the issue. The equity issues are from 1970 to 1990. Using the first closing postissue market price, the equally weighted average buy-and-hold return for the year after the issue is calculated for the issuing firms and for their matching firms (firms with the same market capitalization that have not issued equity during the prior five years). On each anniversary of the issue date, the equally weighted average buy-and-hold return during the next year for all of the surviving issuers and their matching firms is calculated. For matching firms that get delisted (or issue equity) while the issuer is still trading, the proceeds from the sale on the delisting date are reinvested in a new matching firm for the remainder of that year (or until the issuer is delisted). The numbers graphed above are reported in Table III.

Section 4

Behavioral Labor

History by Field

- Fields of applications of Behavioral Economics
- 1980s and 1990s
  - Behavioral Finance (Asset Pricing)
  - Behavioral Household Finance
- 2000 on
  - Behavioral Corporate Finance
  - Behavioral Development Economics
  - Behavioral Industrial Organization
- Last 5-10 years
  - Behavioral Public Finance
  - Behavioral Health Economics
  - Behavioral Political Economy
  - Behavioral Education Economics

Early Behavioral Labor

- What about Behavioral Labor Economics?
- Work since early days of behavioral econ:
  - Inter-industry wage differentials (Thaler, 1989)
  - Fairness constraints (Kahneman-Knetsch-Thaler, 1986)
  - Gift Exchange (Akerlof, 1982, 1984)
- Yet, did not quite pick up pace
- Surprising because of deep psych roots in labor econ
  - Institutionalists influenced by Veblen
  - Richard Ely: "exploring the human problem by the psychological method"
  - Albert Rees / John Dunlop: key role of fairness in labor negotiation

Five Areas

- **Behavioral labor economics**:
  1. Worker Effort (present bias/pay equity/gift exchange)
  2. Wage setting (wage compression/nominal wage rigidity)
  3. Job Search (present bias/reference dependence/overconfidence)
  4. Labor supply (reference dependence/gender norms)
  5. Educational choices (present bias/inattention/social norms)

Present Bias

- What determines effort at the workplace?
  - Explicit and implicit incentives (e.g., promotions)
  - But also motivation and social preferences
- Factor 1: **Present bias**
  - Work entails immediate effort costs and delayed benefit
  - Present bias –> Not work hard enough
  - Kaur, Kremer, Mullainathan (JPE) in the field
  - Augenblick, Niederle, Sprenger (QJE) with real effort task

Social Preferences

- Factor 2: **Social Preferences**
  - If motivation important for effort, may justify paying higher wages to trigger reciprocity (Akerlof 1981)
- Two dimensions of motivation:
  - Vertical Social Preferences (toward employer)
  - Horizontal Social Preferences (toward co-employees) – Card et al (2012)
- Two dimensions of effort:
  - Extra effort that helps the employer (e.g., fix production defect) – eg, Hjort (2014)
  - Boycott that hurts employer (e.g., stealing stationery) – eg, Krueger-Mas (2004)

Other Factors

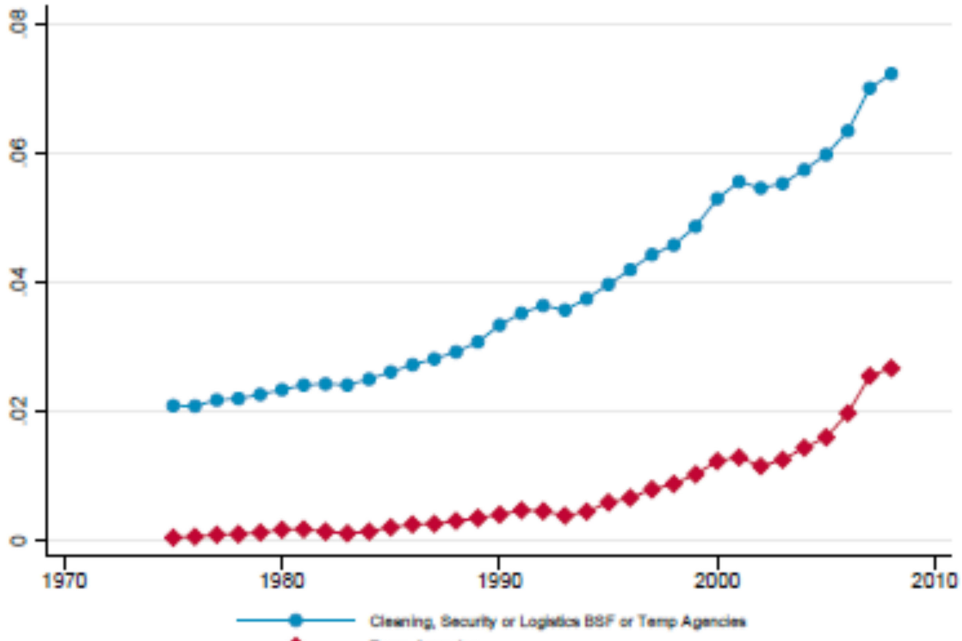
- Other factors
- **Non-financial rewards** as motivators (eg Akerlof and Kranton; Besley-Ghatak; Ashraf, Bandiera and Jack)
  - Evidence mixed on importance
- **Crowd-out** of motivation from financial rewards (Gneezy and Rustichini QJE)
  - Not much evidence
- **Overconfidence** of workers, matters for eg, retention of truckers (Hoffman and Burks)

Wage Setting

- Implications of effort findings for wage setting
- **Present bias** –> Commitment devices
  - Why so uncommon? (Laibson AEA)
- **Vertical social preferences** –>
  - Gift exchange / efficiency wages maybe
  - Constraints on Manager pay (Mas, 2015)
- **Horizontal social preferences** –>
  - Pay compression / wage differentials across firms (Card, Heining, Kline QJE)
  - Drives outsourcing (Goldschmidt and Schmieder QJE)
- **Overconfidence** of workers –> Stock options (Bergman and Jenter, 2007; Cowgill and Zitzewitz, 2015)

Wage Setting, Outsourcing

- Pay compression:
  - Some firms pay more than others
  - This applies also to low skill workers
  - As competition strengthens save $ by outsourcing low-skill workers (Goldschmidt and Schmieder QJE)
- Germany : Share of outsourced workers much increased



Wage Setting, Outsourcing

- More outsourcing in firms with higher pay, controlling for all else
- Preferences for equity have indirect effects

Table III: The Effect of Proxies for Wage Premia on the Probability of Outsourcing

	All Establishments				Establishment Panel Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Log Estab Size	0.0084***					
	(0.00016)					
Log Avg Estab Wage		0.00044				
		(0.00032)				
AKM Effect			0.0046***			
			(0.00057)			
Wage Premium to FSCL workers over BSF firms				0.0015***		
				(0.00026)		
Collective Agreement					0.0091***	
					(0.0013)	
Pay Wages Above Standard						0.0029**
						(0.0014)
Observations	2086507	2086505	1892408	1769077	68577	68595
Mean of Dep Var	.012	.012	.011	.014	.02	.02
Mean of Indep Var	4.788	4.285	.003	1.162	.81	.34

Wage Setting, Nominal Wage Rigidity

- **Kahneman, Knetsch and Thaler (1986)**
- Telephone surveys in Canada in 1984 and 1985 $\rightarrow$ Ask questions on fairness

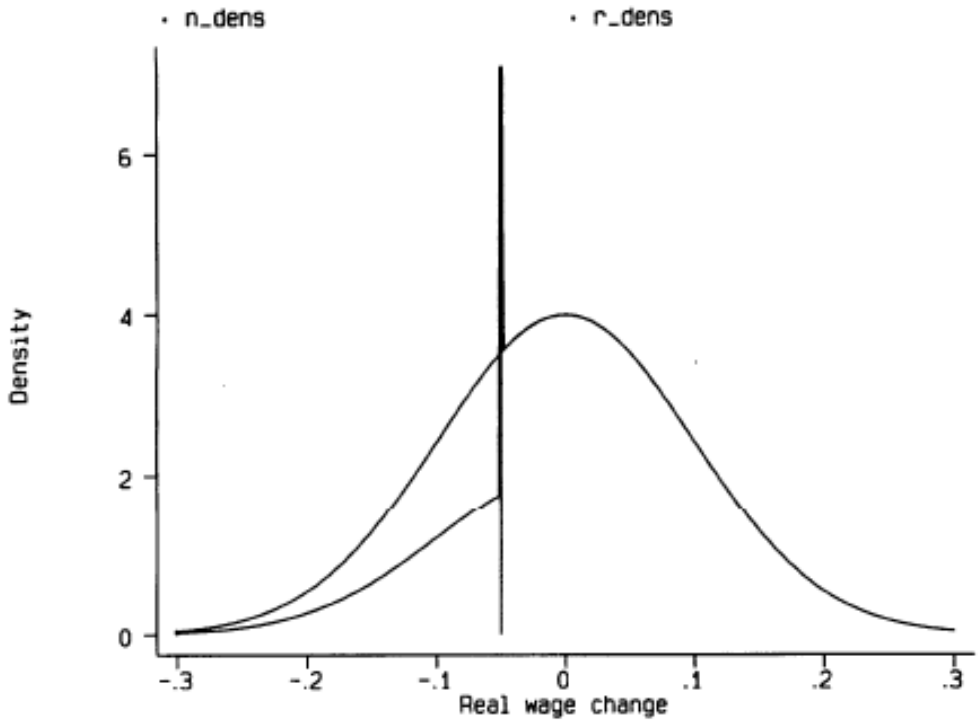
Question 4A. A company is making a small profit. It is located in a community experiencing a recession with substantial unemployment but no inflation. There are many workers anxious to work at the company. The company decides to decrease wages and salaries 7% this year.
($N = 125$) Acceptable 38% Unfair 62%

Question 4B. ...with substantial unemployment and inflation of 12%...The company decides to increase salaries only 5% this year.
($N = 129$) Acceptable 78% Unfair 22%

  - A real and nominal wage cut is not fair (Question 4A)
  - A real (but not nominal) wage cut is fair (Question 4B)
- **Bewley (1999)** interviews of managers and workers
  - They hate nominal wage cuts –> Expect employers to avoid nominal wage decreases ($w_t - w_{t-1} < 0$)

Wage Setting, Nominal Wage Rigidity

- Examine discontinuity around 0 of nominal wage changes
- Prediction of theory:



Wage Setting, Nominal Wage Rigidity

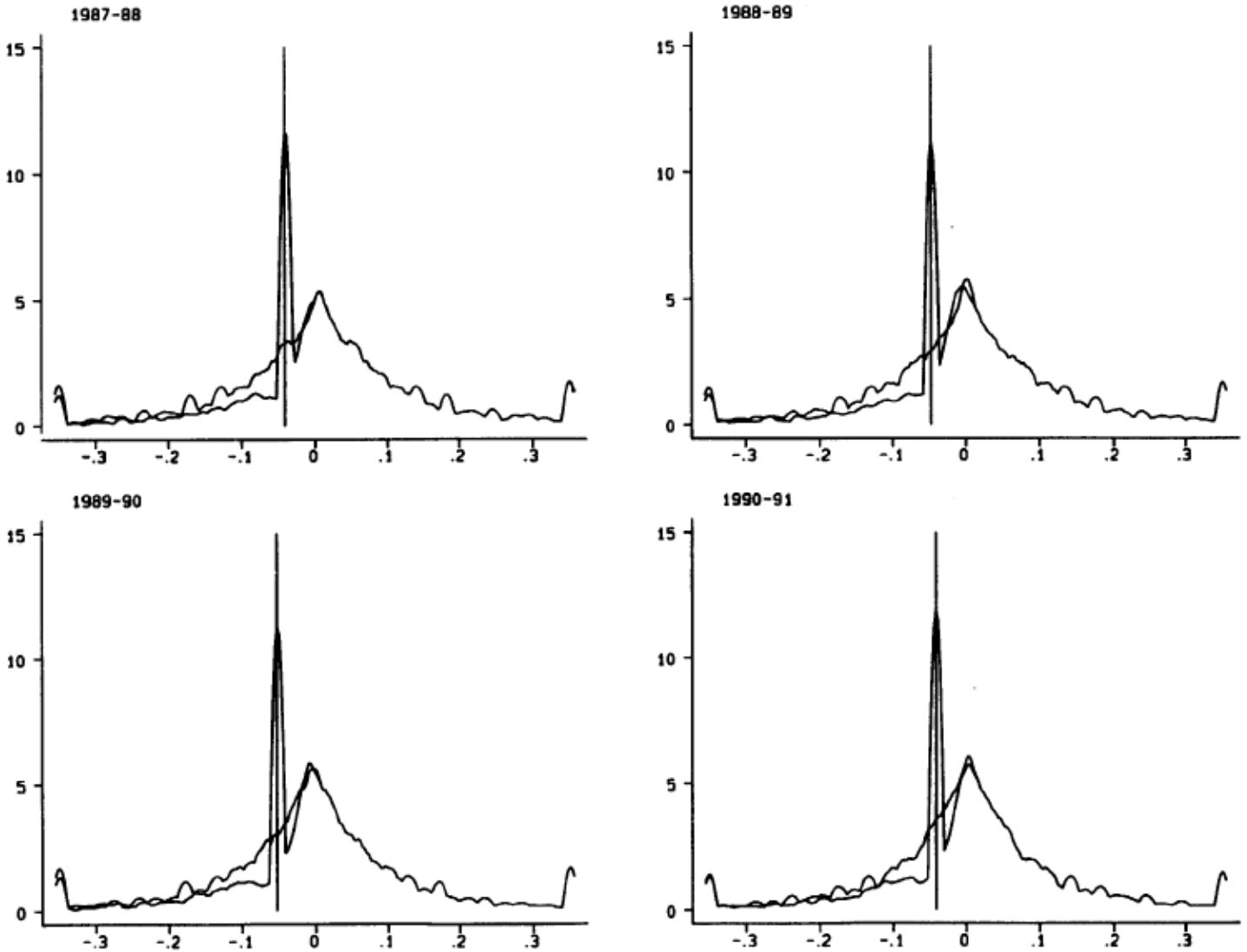
- Data sources:
  - 1979-1993 CPS.
    - Rolling 2-year panel
    - Restrict to paid by the hour and to same 2-digit industry in the two years
    - Restrict to non-minimum wage workers
  - PSID 4-year panels 1976-79 and 1985-88
- Use Log Wage changes: $\log w_t - \log w_{t-1}$
- Issue with measurement error and heaping at $\log w_t - \log w_{t-1} = 0$
- Construct counterfactual density of LogWage changes
  - Assume symmetry
  - Positive log wage changes would not be affected

Wage Setting, Nominal Wage Rigidity

- Plots using kernel estimates of density (local smoother)
- Compare the actual distribution and the predicted one
- Evidence from the CPS year-by-year
- Problem more severe in years with lower inflation

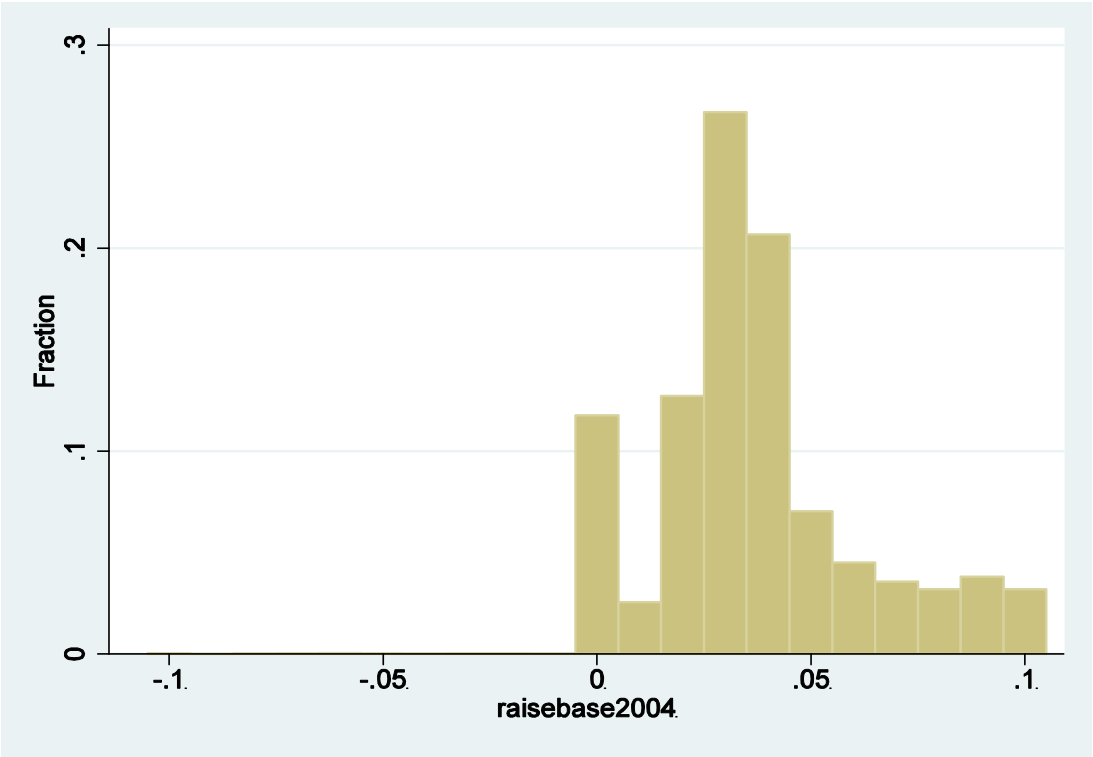
- Large effect of nominal rigidities
- Effect on firings?

Real Wage Changes, 1987-88 to 1998-91



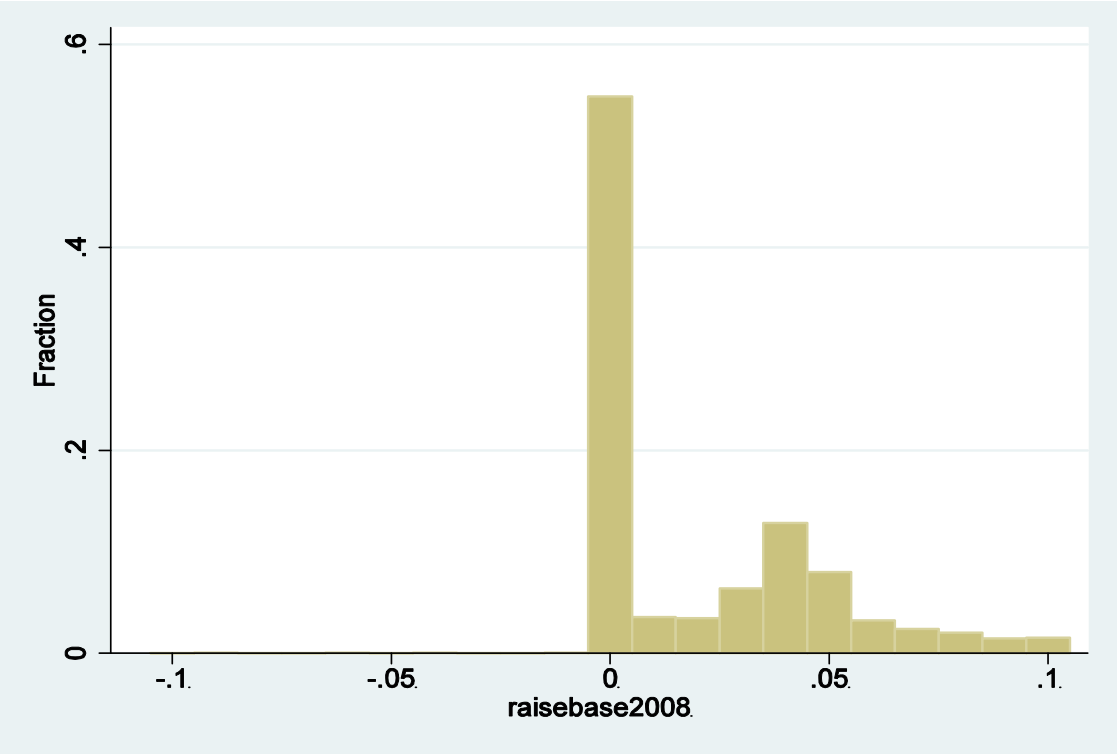
Hipsman (2011): 2003 & 2004

- Administrative data from several firms
  - Base pay % increase among those employed in 2003 and 2004
  - 58 (0.34%) cuts, 1,964 (10.18%) freezes, 15,091 (88.18%) raises



2007 & 2008

- Base pay % increase among those employed in 2007 and 2008
- 46 (0.36%) pay cuts, 6,913 (54.58%) pay freezes, 5,707 (45.06%) pay raises



Conclusions

- Card and Hyslop had *underestimated* the degree of nominal rigidity
- Important implications for labor markets when low inflation
  - If no pay cut, what margin of adjustment?
  - Firing?
  - Less hiring?
- Key under-researched topic in behavioral macro

Job Search, Simple Model

- Job search model with optimal search effort, no reservation wage
- An **unemployed** worker's value function is

$$V_t^U = \max_{s_t \in [0,1]} u(y_t) - c(s_t) + \delta \left[s_t V_{t+1}^E + (1 - s_t) V_{t+1}^U \right]$$

- Solution for **optimal search**:

$$c'(s_t^*) = \delta \left[V_{t+1}^E - V_{t+1}^U \right]$$

- Where could standard model be too restrictive?

Job Search, Present Bias

- **Present Bias**: DellaVigna and Paserman (JOLE 2005): An **unemployed** worker's value function is

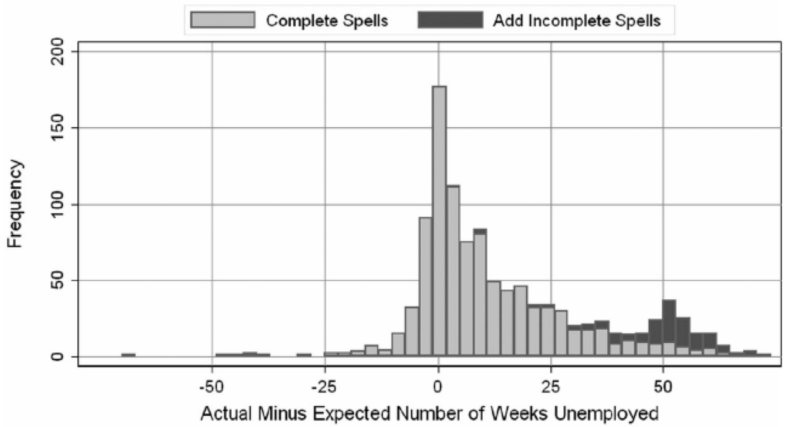
$$V_t^U = \max_{s_t \in [0,1]} u(y_t) - c(s_t) + \beta\delta \left[s_t V_{t+1}^E + (1 - s_t) V_{t+1}^U \right]$$

  - **Search effort**: search too little if $\beta < 1$
  - **Reservation wage** unaffected (under naivete')
- Can lead to different policy implications (Paserman EJ 2008), but observationally hard to distinguish
- **Reference Dependence** (DellaVigna et al 2017): Flow utility $u(y_t, r_t)$ also depends on reference point r_t:

$$u(y_t, r_t) = \begin{cases} v(y_t) + \eta(v(y_t) - v(r_t)) & \text{if} \quad y_t \geq r_t \\ v(y_t) + \eta\lambda(v(y_t) - v(r_t)) & \text{if} \quad y_t < r_t \end{cases}$$

Job Search, Overconfidence

- Are beliefs of unemployed workers right?
- Allow for individuals to have wrong beliefs on average: $\widetilde{p}(s)$ replaces s (Spinnewijn 2014)



Source: Spinnewijn (2015). Unemployed job seekers in Michigan & Maryland surveyed between 1996-1998.

Job Search, Overconfidence

- An **unemployed** worker's value function is

$$V_t^U = \max_{s_t \in [0,1]} u(y_t) - c(s_t) + \delta \left[\widetilde{p}(s_t) V_{t+1}^E + (1 - \widetilde{p}(s_t)) V_{t+1}^U \right]$$

- Important distinction:
  - Overconfidence in level ($\widetilde{p}(s) = s + o$) does not matter for search effort
  - Over/under-confidence about marginal return of effort ($\widetilde{p}'(s) \neq 1$) affects effort cost
- Clear evidence for the first form of overconfidence
- Very little known about the second (control overconfidence, **Spinnewijn, JEEA 2015**)

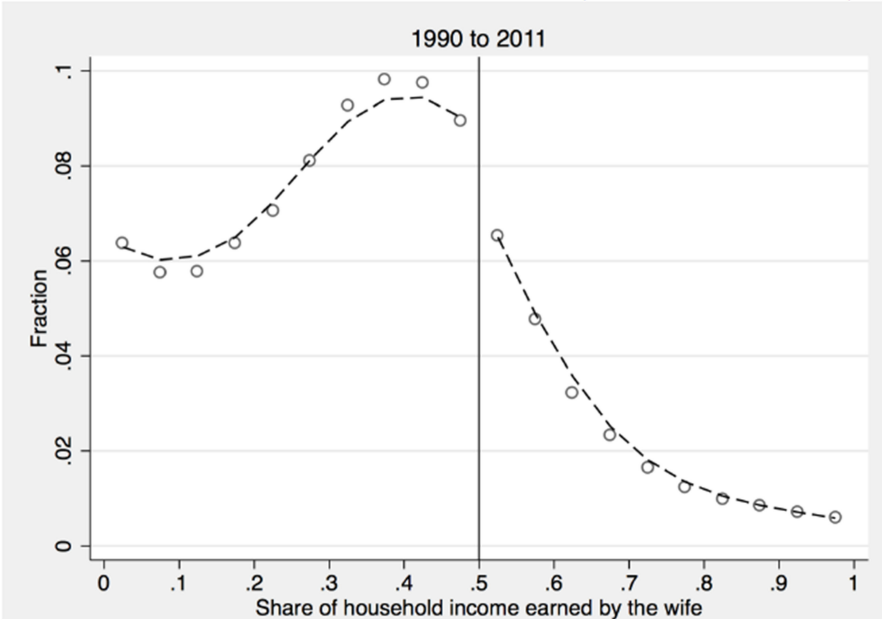
Labor Supply, High Frequency

- **(High-frequency) Labor Supply**
- Target earnings and Reference dependence: Cab drivers (Camerer et al., 1997; Farber, 2005, 2015; Crawford and Meng, 2011; Farber, 2015), Bike Messengers (Fehr and Goette, 2008)
- **Thakral and To (2019)** clarifies facts:
  - Yes, drivers more likely to stop when they earned more income, Effect size moderate overall
  - Importantly: only recent earnings (last 2-3 hours) matter, not earnings very early on in day –> Evidence on reference point formation

Labor Supply, Gender Norms

- Gender norms in Labor Supply (**Bertrand, Kamenica, Pan QJE**)
  - Husband dislikes if spouse earns more
  - Marriages less likely to occur if wife earns more

Figure II: Distribution of relative income (Census Bureau data)



- **Hermle (2019)**: What preferences exactly underlie this?

Educational Choices

- Facts on benefits of attending school:
  - Returns to school are high (10%-12%/yr of school)
  - If anything, college premium has increased over time
- Then why doesn't (nearly) everyone go?
- Active debate. Relevant behavioral factors:
  1. Excessive discounting (**Oreopoulos, 2007**)
  2. Poor information, misperceived returns to college (**Jensen, 2010; Wiswall and Zafar, 2015**)
  3. Complexity. FAFSA forms too hard (**Bettinger, Long, Oreopoulos, Sanbonmatsu, 2012**)
  4. Social norms. Acting white (**Fryer and Austen-Smith; Bursztyn, Jensen**)

Section 5

Behavioral Development

Government Intervention

- Slides from **Behavioral Development Economics,** in *Handbook of Behavioral Economics,* Vol 2), by Michael Kremer, Gautam Rao, Frank Schilbach

Topics covered (organized by development economics)

(1) Introduction

(2) **High rates of return without rapid growth (Euler equation puzzle)**

- (A) Euler Puzzle
- (B) Present bias
- (C) Reference-dependent preferences
- (D) Other behavioral factors (e.g. biased beliefs)

(3) Health

(4) Savings

(5) Risk and insurance

(6) Technology adoption

(7) Labor

(8) Firms

(9) Social preferences, culture, and development

(10) The psychology of poverty

High returns to capital in many contexts Banerjee and Duflo (2005)

- Borrowing at very high rates (70 to 100% annual rates and more)
  - Small-time fruit vendors in Chennai who borrow at daily rates of 5% (Karlan et al., 2018)
- High returns to small-business grants (de Mel et al., 2008)
- High returns to inventories (Kremer et al., 2013)
- Predictable large increases in prices between seasons (Burke et al., 2018)

Euler equation

- Suppose production function $F(K)$ with $F'(K) \geq 0$ and $F''(K) \leq 0$.
- Standard Euler equation links consumption growth to marginal return to capital:

$$u'(c_t) = \delta F'(K_t) u'(c_{t+1}) \tag{1}$$

- Implies (unrealistically) high consumption growth rates.
  - If log utility, $F'(K) = 50\%$ annually, and $\delta = 0.96$, then $\frac{\dot{c}}{c} = 44\%$.
  - If constant intertemporal elasticity of substitution utility with $\sigma = 2$, then $\frac{\dot{c}}{c} = 20\%$.
  - Still implies 38-fold consumption growth in 20 years.
- Need high "tax" or discount rate to resolve puzzle.
  - Implicit taxes due to corruption or redistributive pressures by from extended family members
  - Allowing for realistic values of such taxes does not resolve puzzle (Jakiela and Ozier (2015)).

Stochastic income and risk aversion?

- Maybe people don't invest because investments (e.g. fertilizer) are risky? Suppose income in period t is:

$$Y_t = Y_0 + \epsilon_t + \sum_{i=1}^{n} \mu_{i,t} F_i(K_{i,t}), \tag{2}$$

  where n assets/capital goods, arbitrary pattern of correlation.

- Stochastic Euler equations:

$$u'(c_t) = \delta \mathbb{E}_t[\mu_{i,t} F_i'(K_{i,t}) u'(c_{t+1})], \quad i = 1, 2, ..., n \tag{3}$$

- Given initial capital stock, risk aversion will:
  (i) reduce investment in assets which co-vary positively with consumption
  (ii) increase investment in assets which co-vary negatively with consumption

But: Optimal to build buffer stock savings (Deaton, 1991; Carroll, 1997).

- If patient, risk averse, and subject to large shocks, agents want to accumulate large buffer stock savings.
  - At any one time, only a few people should have low buffer stock.
- For majority with large buffer stock, consumption should not move much with:
  - high-frequency income shocks
  - predictable income changes (e.g. seasons)
- Implies that even if returns to fertilizer highly correlated with income in season, only modestly correlated with lifetime income and thus consumption
  - Beta of fertilizer investment (correlation of return with overall consumption) will be modest, and risk aversion will only modestly reduce fertilizer investment

Model with patient consumers seems to make incorrect predictions.

- In reality:
  - Liquid buffer stocks are often modest (Deaton, 1991).
  - Consumption co-varies with income, including predictable income (Townsend, 1995).
  - Karlan et al. (2014) find that rainfall insurance increases fertilizer use.
- A model will impatient agents can create these predictions.
- Thus with either deterministic or stochastic Euler equation, matching the data requires a high effective discount rate.

High discount rates?

- Maybe $\delta = 50\%$?
- Standard exponential discounting model has only one parameter for all time horizons.
  - Euler equation typically considers short horizons (≤ 1 year).
  - In exponential discounting model, high short-run discount rate implies that distant future is discounted at extremely high rates.
- Absurd implications
  - $\delta = 0.5$ implies would not give up \$1 today for \$1 billion in 30 years.
  - No one would own land, get an education, etc.

Implications of present-biased preferences

- Predictions behavior of present-biased agents (Angeletos et al., 2001):
  - Rapidly spend down liquid assets, becoming effectively liquidity constrained
  - Build up (or hold) a stock of illiquid assets that pay off in distant future
  - Leave high rate of return investments on the table, if effectively liquidity constrained
  - Not be able to smooth consumption; consumption will co-move with income shocks, even with predictable income variation
- The sophistication of the present biased actor will determine the degree of procrastination and demand for commitment devices (O'Donoghue and Rabin, 1999, 2001).
- Implies modified Euler equation (Harris and Laibson, 2001)

Topics covered

(1) Introduction

(2) High rates of return without rapid growth (Euler equation puzzle)

(3) **Health**

  (A) Under-investment in preventive health
  (B) Present bias
  (C) Biased beliefs
  (D) Incorrect mental models

(4) Savings

(5) Risk and insurance

(6) Technology adoption

(7) Labor

(8) Firms

(9) Social preferences, culture, and development

(10) The psychology of poverty

Under-investment in preventive health

- Widely studied case of under-investment in high-return opportunities: low investment in preventive health (e.g. vaccinations, deworming, bed nets, water treatment, hypertension)
- Recent literature established several stylized facts regarding health behavior in developing countries (Dupas, 2011; Kremer and Glennerster, 2011; Dupas and Miguel, 2017).
  (1) Low willingness to pay (WTP) for preventive health
  (2) High expenditures for treatments of acute conditions
  (3) High sensitivity of health investments to price and convenience

Demand for preventative health: low WTP and high price sensitivity

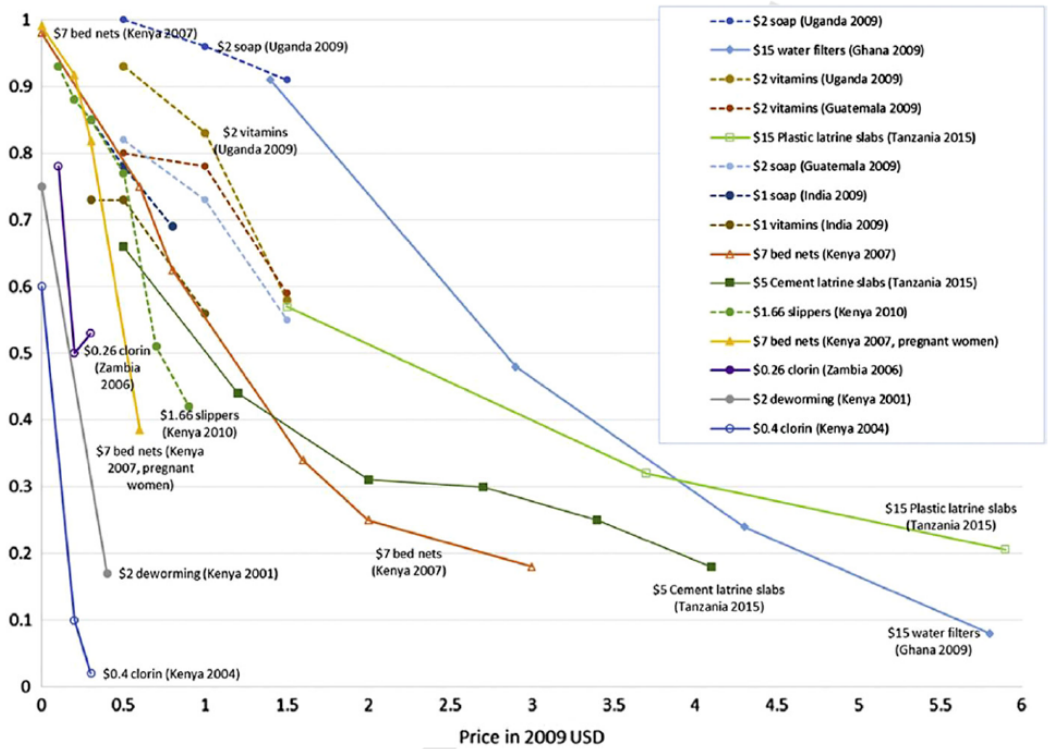


Figure: Share of individuals taking up the product as function of price (from Dupas and Miguel (2017))

High price sensitivity of demand for preventative health investments

- High price-sensitivity even in cases of substantial long-run benefits:
  - Deworming medication (Miguel and Kremer, 2004); mosquito nets (Cohen and Dupas, 2010); water treatment (Ashraf et al., 2010).
  - Example: estimated private financial benefit of deworming is $142 (Baird et al., 2016), yet $0.30 per child cost-sharing fee decreased take up 80 percent (Miguel and Kremer, 2004).
- High sensitivity also for monetary and non-monetary incentives:
  - Large impacts of small (and time-limited) incentives (lentils) for vaccination (Banerjee et al., 2010) or collecting HIV tests (Thornton, 2008)
  - Prima facie evidence against liquidity constraints (though not conclusive)
- If individuals are given more time to purchase, then lower price sensitivity, but demand still fairly sensitive to price (Dupas, 2011a).

Present bias and procrastination

- Driven by the immediate *utility costs* of the investment:
  - Examples: hassle and psychic costs of going to doctor, walking to farther-away water source, using dilute chlorine solution, changing diet, learning painful news about health status, taking medication.
  - Not financial costs unless severely liquidity constrained
- Procrastination requires both present bias and some degree of naivete.
  - Prefer to do painful task tomorrow, mis-predict that they will do it tomorrow.
- Consistent with:
  - (i) effect of time-limited incentives: e.g. Banerjee et al. (2010)
  - (ii) effect of reducing hassle costs: e.g. water dispensers (Ahuja et al., 2010)
- *Note:* Would not procrastinate on acute condition, since benefits immediate

Present bias and liquidity constraints

- Present bias can lead to liquidity constraints (Angeletos et al., 2001)
- Once liquidity-constrained:
  - High-return preventive investments may be left unexploited.
  - Monetary expenditures might now translate into (almost) immediate utility costs, since need to cut back on other consumption in order to, e.g. pay for doctor visit.
- Consistent with:
  - Evidence on effects of increased liquidity (Dupas and Robinson, 2013)
  - High impact of small discounts to fertilizer around time of harvest (Duflo et al., 2011)

Biased beliefs

- Making good decisions regarding health requires forming accurate beliefs about numerous variables. Difficult due to uncertainty and heterogeneity across individuals (Arrow, 1963).
- Inaccurate beliefs (e.g. misperceived returns to health investments) could help explain under-investment in health. Some evidence of inaccurate beliefs regarding health in developing societies (e.g. Delavande and Kohler (2009); Godlonton et al. (2016)).
- Information interventions appear to have large impacts on health outcomes in some contexts and small to null in others Dupas (2011); Dupas and Miguel (2017).
  - Other behavioral biases might be at play in situations of low impacts of info.
  - Motivated beliefs (e.g. deriving utility from belief that one is healthy) could matter as well.
  - More work is required to understand the determinants of success in various contexts.

Incorrect causal theories or mental models

- Individuals may interpret what they observe through the wrong causal model or theory (Schwartzstein, 2014; Gagnon-Bartsch et al., 2018).
- Incorrect mental models that may be important for health outcomes in developing societies include superstitious beliefs or beliefs in magical theories of sickness and health which include witchcraft.
- Ashraf et al. (2017) illustrate this issue in the case of maternal risk in Zambia and a wide-spread belief about martial infidelity and complications during childbirth
- Parents across the world confidently hold wrong beliefs about need to re-hydrate children in response to diarrhea. Datta and Mullainathan (2014): 30 to 50 percent of women in their sample (in India) recommended *decreasing* fluid intake of infants to treat diarrhea

Evidence against importance of some ideas from psychology in the field

- Little evidence for real-world development importance of some psychological effects frequently invoked by practitioners to justify policy:
  - **Sunk-cost fallacy:** No evidence that higher prices cause greater product use Ashraf et al. (2010); Cohen and Dupas (2010).
  - **Crowd-out of intrinsic motivation:** Little evidence that extrinsic incentives crowd out intrinsic motivations in real world development contexts or that paying more leads to substantially less-motivated workers (Dal Bó et al., 2013; Ashraf et al., 2014, 2018).

Topics covered

(1) Introduction
(2) High rates of return without rapid growth (Euler equation puzzle)
(3) Health
(4) **Savings**
(5) Risk and insurance
(6) Technology adoption
(7) Labor
(8) Firms
(9) Social preferences, culture, and development
(10) The psychology of poverty

"Standard" barriers to saving

- Savings are necessary to self-insure against risks and to finance lumpy investments
- "Standard" barriers to savings include:
  - Lack of access to formal savings products
  - Prohibitive costs of opening a banking account etc.
- Dupas et al. (2018) find small effects of providing bank accounts to poor individuals, suggesting other (potentially behavioral) constraints may play a role in reducing savings

Commitment savings devices

- A key prediction of present bias: households accumulate few liquid savings over time, while building up substantial illiquid wealth. Consistent with savings patters across the world (Angeletos et al., 2001; Banerjee and Duflo, 2007; Morduch et al., 2009)
- Ashraf et al. (2006): evidence for demand for commitment devices in the domain of savings which evidences present-bias (as discussed in Section 3.2).
- A key open question surrounding the usefulness of commitment devices is the optimal trade-off between commitment and flexibility. Too stringent commitment reduces take-up and too flexible commitment does not overcome self-control problems.
- Dupas and Robinson (2013) find that a softer savings device increases spending on preventative care relative to a control group and a more stringent alternative.

Designing financial products for behavioral agents: Default effects

- Setting default choices is a cheap but often highly powerful tool in changing behavior.
- For instance, setting the default to automatic enrollment as opposed to non-enrollment has substantial impacts on individuals' retirement choices, particularly for lower-income individuals (Chetty, 2015; Chetty et al., 2014; Madrian and Shea, 2001)
- Blumenstock et al. (2018): setting opt-in defaults increase the savings of Afghanistan workers. Additionally, they argue the underlying mechanism involves present bias as well as the hassle costs of thinking through different options.

Designing financial products for behavioral agents: Attention

- Inattention can distort individuals' decision making in spheres ranging from savings to medical adherence and as such can have large costs.
- Karlan et al. (2016) study the impact of reminders on savings and consumption choices and find that reminders increase the salience of savings goals.
- Many reminder interventions in health (e.g. Pop-Eleches and et al. (2011))
- Potential negative externalities if attention is a limited resource. Need more evidence on whether reminders remain effective in the long term

Topics covered

(1) Introduction
(2) High rates of return without rapid growth (Euler equation puzzle)
(3) Health
(4) Savings
(5) **Risk and insurance**
(6) Technology adoption
(7) Labor
(8) Firms
(9) Social preferences, culture, and development
(10) The psychology of poverty

Low take-up of insurance

- Many people in developing countries exposed to very risky income streams (e.g. farming)
- Yet low take up of actuarially fair weather insurance (Cole et al., 2013).
  - Basis risk? (Clarke, 2016; Mobarak and Rosenzweig, 2012; Giné et al., 2008)
- Low take-up of health insurance (Thornton et al., 2010)
  - Administrative issues?

Potential explanations for low demand: Non-standard preferences

- Casaburi and Willis (2018): insurance meant to shift resources across states, yet most actual insurance contracts involve transferring resources over time
  - Eliminating the intertemporal component increases insurance take-up dramatically.
  - Important role for liquidity constraints, present bias

- Could loss aversion/prospect theory play a role?
  - Reference-dependent preferences increase risk aversion over moderate stakes and may lead thus cause *over*-insurance (Sydnor, 2010).
  - But premia might be seen as losses, thus curbing insurance demand (Eckles and Volkman-Wise, 2011).
  - Diminishing sensitivity away from reference point could lead to risk-seeking behavior in loss domain.

Potential explanations for low insurance demand: non-standard beliefs

- **Projection bias:** In good states of the world, agents may underestimate their marginal utility in bad states of the world (Loewenstein et al., 2003).
- **Recency effects:** Agents might place disproportionate weight on events from the recent past (Hogarth and Einhorn, 1992; Fuster et al., 2010; Chang et al., 2018; Karlan et al., 2014).
- **Motivated reasoning:** If individuals directly derive utility from beliefs about their future well-being, they may seek to maintain biased beliefs about their current health or the likely future state of the world.
- **Beliefs in higher powers:** Individuals' beliefs might deviate in more dramatic ways from standard probability assessments. Beliefs in higher powers might suppress insurance demand (Auriol et al., 2018).

Topics covered

(1) Introduction
(2) High rates of return without rapid growth (Euler equation puzzle)
(3) Health
(4) Savings
(5) Risk and insurance
(6) **Technology adoption**
(7) Labor
(8) Firms
(9) Social preferences, culture, and development
(10) The psychology of poverty

Technology adoption

- Various examples with apparently non-optimal technology choice:
  - Pineapple farming in Ghana, HYV seeds, seaweed pod size, fertilizer, contraceptives, soccer ball manufacturing techniques, layout of equipment in textile factories
- Do external analysts correctly understand payoffs?
- Do decision makers have adequate information?

Technology adoption: attention and complexity

- Inattention and wrong mental models (Hanna et al., 2014)
  - Production function is complex and attention is costly.
  - Individuals will pay attention to the dimensions they think are important.
  - If start off thinking something is not important (wrong mental model), will not pay attention and will never learn, even with data that would otherwise lead to revision of beliefs.

- Complexity of information
  - Provision of simplified information about seaweed pod size (Hanna et al., 2014), water safety (Bennear et al., 2013) or business practices (Drexler et al., 2014) may be more effective than providing full information.
  - Downsides of presenting simplified information: heterogeneity in population; external analysts might misunderstand decision problem

Technology adoption: present bias and loss aversion

- Present bias (Duflo et al., 2011)
  - If adoption requires costly experimentation, individuals might procrastinate since benefits are often much delayed.
  - Could benefit from simplification (if learning is costly).
  - Is there demand for commitment for technology adoption (training)?
  - Time-limited discounts around harvest highly effective at increasing take-up of fertilizer
- Loss aversion
  - Conjecture: relevant reference point when trying something new is the status quo. Possibility of losses with respect to the status quo will trigger loss aversion
  - Possibility of insurance or informal risk-sharing to improve outcomes?

Behavioral social learning

- Rational social learning will often lead society to right long-run choice if some can get past initial experimentation costs
- Banerjee (1992) herd behavior: model converges on optimal technology if:
  - observe output
  - observe size of investment
  - smooth loss function makes choices reveal signals
- Why might individuals not converge on optimal technology? We distinguish:
  (1) Barriers to sharing or seeking information
  (2) Barriers to correctly interpreting information

Barriers to sharing and seeking information: Social-image concerns

- The degree of communication between people is endogenous. Providing and soliciting information is a decision.
- People may be hesitant to ask for or provide information when doing so signals effort or ability (Chandrasekhar et al. (2018); Banerjee et al. (2018);
  - Implies seeding info more broadly can reduce learning
- People may not be willing to provide information to others for free if they paid for it or put in effort to get it.

Topics covered

(1) Introduction
(2) High rates of return without rapid growth (Euler equation puzzle)
(3) Health
(4) Savings
(5) Risk and insurance
(6) Technology adoption
(7) **Labor**
(8) Firms
(9) Social preferences, culture, and development
(10) The psychology of poverty

Distinct features of labor markets in developing economies

- Labor markets in developing economies are different to labor markets in rich countries in three key ways that make behavioral biases potentially more important:
  - High levels of informality
  - High levels of casual labor
  - High degree of self employment

Factory discipline as commitment device

- Clark (1994) argues workers want factory discipline as a commitment device.
  - Much rosier view
- Kaur et al. (2015)
  - About a third of data-entry workers choose dominated commitment contract over piece rate contract
  - Offering dominated contract increases output.
  - Substantial heterogeneity; some evidence of learning
  - With asymmetric information, firms may screen out undesirable workers with factory discipline or steep incentives, reducing overall welfare
  - Justification for legislation limiting hours, etc.?

Wage rigidities

- The share of the population employed in agriculture is much higher in poor countries than in rich countries. And most farms employ outside workers for short spells using informal contracts Kaur (2019).
- Agricultural labour markets have many features that ostensibly should make them efficient: many small buyers and sellers of labor, no formal unions and little to no enforcement of minimum wages
- Despite this, even in these decentralized informal markets, nominal wage rigidities and limited dispersion of wages across workers persist. (Kaur, 2019; Breza et al., 2018b,a).

Why do wage rigidities persist?

- Wage rigidities seem persistent even in the absence of enforced minimum wages or formal institutions like unions.

- These rigidities appear to be enforced via **social sanctions**:
  - Breza et al. (2018a): nominal wage rigidities persist in part due to workers turning down public offers of jobs with wages below the prevailing market wage which workers accept when those offers are made in private.
  - Breza et al. (2018b): when coworker productivity is difficult to observe, then introducing pay inequality reduces worker output.

Wages and incentives to do good

- **Incentives in public and non-profit sectors:**
  - Some evidence of positive effects of financial incentives on public/non-profit sector worker productivity (Duflo et al., 2012; Muralidharan and Sundararaman, 2011)
  - But providing incentives to multi-tasking agents is difficult (Holmstrom and Milgrom, 1991).
  - Additionally, financial incentive programs tend to be politically unpopular and therefore are rarely scaled by governments (Finan et al., 2017)

- **Crowd-out intrinsic motivation:**
  - Lab evidence suggests extrinsic rewards can reduce intrinsic motivation (Deci, 1971; Bénabou and Tirole, 2003)
  - But very limited field evidence of substantial crowding out (Lacetera et al., 2013)

Selection of workers

- Does offering higher wages, which might attract more talent, negatively select on the pro-social motivation of workers?
- Majority of evidence suggests no negative selection. Dal Bó et al. (2013) in Mexico and Ashraf et al. (2018), in Zambia.
- Evidence consistent with underlying correction of cognitive ability and pro-sociality (Falk et al., 2018).
- However, Deserranno (2019) finds that posting job notices with a higher implied pay attracts candidates who donate less money in dictator games, and who perceive lower social benefits to the job at the time of applying.

Female labor force participation (FLFP)

- 52% of women in poor countries participate in the labour force compared to 78% of men (Duflo et al., 2012)
- Standard explanations emphasize biological reasons which, it is typically argued, engender differences in the specialization of the sexes between wage work and domestic work.
- Leaves much of the variation in FLFP unexplained, even conditional on income per capita.
- Behavioral explanations include:
  - Low self-efficacy (McKelway, 2018)
  - Social norms suppressing FLFP (Bursztyn et al., 2018a)

Topics covered

(1) Introduction
(2) High rates of return without rapid growth (Euler equation puzzle)
(3) Health
(4) Savings
(5) Risk and insurance
(6) Technology adoption
(7) Labor
(8) **Firms**
(9) Social preferences, culture, and development
(10) The psychology of poverty

Behavioral firms

- Is it reasonable to assume firms (as opposed to individuals) make choices that maximize profits? Are there reasons to believe firms in developing economies are more behavioral?
  - Here: broad definition of "behavioral": deviations from profit maximizing behavior
- Lucas (1978) span of control model and Chicago critique of behavioral economics:
  - Behavioral firms will be weeded out of the market.
  - Even if only 5% of people don't have behavioral biases, they will become managers of firms.
- Distortions in developing countries prevent efficient firms from growing and displacing less efficient ones.
- Self-employed individuals in developing countries are not just behavioral consumers, they are behavioral firms – or at least behavioral managers.

Reasons developing economy firms could be more behavioral

(1) Lower competitive pressures due to:

  (i) Import restrictions

  (ii) Restriction of new entrants into markets based on regulation, financial constraints, and agency problems

Reasons developing economy firms could be more behavioral (cont'd)

(2) Smaller firm sizes which limit the scope for within-firm competition that causes non-behavioral agents to rise to management:

(i) Smaller firm sizes as discussed in the previous chapter potentially due to:

- Taxation and regulation (e.g. labor regulation), predation
- Credit market issues (But profitable firms should grow over time?)
- Correlation between firm size and family structure (Ilias (2006); Bertrand et al. (2008))
- Difficulty of cooperation?

(ii) Implications

- Firms may only replace self employment when productivity advantage becomes large enough to outweigh these costs.
- Reduces ability of innovations to spread, incentives to innovate
- Reduces replacement of inefficient producers

Behavioral firms: low levels of trust

- Once we start considering behavioral biases in firm decision-making, many unexplored and potentially important areas of research arise.

- Example: Low levels of trust and missing firm growth:
  - Firms in developing countries are small and standard explanations do not completely account for just how small these firms tend to be.
  - Low levels of trust associated with smaller firm sizes Cingano and Pinotti (2012); Algan and Cahuc (2014)
  - Non-Western countries are more likely to emphasize loyalty to one's group (Haidt, 2013), which might in turn limit cooperation with out-group members.

Behavioral firms: management practices

- Improved management practices have been shown to increase firm profitability in developing country contexts (Bloom et al., 2013; Bruhn et al., 2018).
- Why are such services not demanded and offered more?
- Firms that fail to adopt these profitable practices are not necessarily weeded out of the market.

New research horizons associated with behavioral firms

- Lots of unexplored areas waiting to be explored:
  - The nature of the objective function of small (family) businesses
  - Demand forecasting/estimating by firms
  - Optimality of pricing or product choices amongst firms
  - Inventory management
  - Firm labor and capital-investment decisions
  - Technology adoption

Topics covered

(1) Introduction
(2) High rates of return without rapid growth (Euler equation puzzle)
(3) Health
(4) Savings
(5) Risk and insurance
(6) Technology adoption
(7) Labor
(8) Firms
(9) **Social preferences, culture, and development**
(10) The psychology of poverty

Trust, cooperation, and development

- Trust and cooperation important for economic and political outcomes
  - e.g. Algan and Cahuc (2014) review
- Developing countries have lower levels of trust and positive reciprocity
  - Falk et al. (2018) using global survey
- Is this a cause or consequence of development?

Trust, cooperation, and development (cont'd)

- Good reasons to think that variation in trust and reciprocity have deep historical roots
  - Enke (2018): historical tightness of kinship predicts modern-day in-group favoritism, willingness to cheat on and distrust outsiders, local rather than broader institutions.
  - Nunn and Wantchekon (2011): long-term consequences of slave trade
  - Henrich et al. (2010): evolution of fairness and punishment facilitated trust and cooperation, allowing for large-scale societies
    - E.g., moralizing gods and cooperation with strangers?
    - Market integration and fairness; community size and punishment
- But likely also in part a consequence of development, e.g. market exposure and well-functioning legal institutions might themselves increase trust.

Social image and norms

- Frontier of behavioral research on (pro)social behavior is on social image
  - Desire to conform to social norms
  - And also to impress (in socially sanctioned ways)
  - Visibility of actions can matter a great deal
- Some recent applications
  - Bursztyn et al. (2018b) on conspicuous consumption in Indonesia
  - Chandrasekhar et al. (2018, 2015); Banerjee et al. (2018) on social learning
- Much more work to be done in developing-country settings
  - Including on how norms change, e.g. gender norms

Shaping social preferences and norms

- Important to understand policies which can improve inter-group behaviors
  - Rao (2019) on integration in schools
  - Blouin and Mukand (2017) on post-conflict Rwanda
  - Lowe (2018) on different types of contact
  - Okunogbe (2018) on consequences of national service in Nigeria
  - Miguel (2004) on national identity in Tanzania
  - Role of policy and culture (Miguel and Gugerty, 2005)

- And policies which can influence certain social norms
  - La Ferrara et al. (2012); Jensen and Oster (2009): TV effects on fertility, gender attitudes
  - Bursztyn et al. (2018a) on female labor force participation in Saudi Arabia

Topics covered

(1) Introduction
(2) High rates of return without rapid growth (Euler equation puzzle)
(3) Health
(4) Savings
(5) Risk and insurance
(6) Technology adoption
(7) Labor
(8) Firms
(9) Social preferences, culture, and development
(10) **The psychology of poverty**

Scarcity and cognitive function

- Mullainathan and Shafir (2013) argue that poverty impedes cognitive function through scarcity. They argue scarcity engenders an increased focus on money and as such the "bandwidth" available for other tasks is reduced.

- Mani et al. (2013): empirical evidence in support of this hypothesis
  - Lab study: inducing thoughts about money lowered the cognitive function of the poor and not the wealthy.
  - Complementary field study exploited within person variation; sugar cane farmers in India had significantly worse cognitive performance before harvest as in contrast to right after harvest.

- Potentially very important results but methodological limitations (e.g. potential learning effects in second study) and (so far) lack of successful replications
  - Carvalho et al. (2016): no differences in cognitive function and decision-making around payday among US workers

Conclusion

- Ideas from behavioral economics help explain important puzzles in development, with important limitations
- Taking behavioral development economics seriously will involve testing specific mechanisms and providing calibrations and estimations where possible (DellaVigna, 2018).
- Many unanswered questions remain and we hoped to have pointed at some of those in the preceding slides. So much more exciting work to be done!
- We did not cover some important topics in development to which behavioral economics may be fruitfully applied (e.g. education, political economy, economics of the family).

Section 6

Next Lecture

Next Lecture

- Behavioral Public
- Teaching Evaluations