Econ 219B
Psychology and Economics: Applications
(Lecture 4)

Stefano DellaVigna

February 11, 2015

Outline

1. Laboratory Experiments on Present Bias II
2. Methodology: Errors in Applying Present-Biased Preferences
3. Reference Dependence: Introduction
4. Reference Dependence: Housing I
5. Methodology: Bunching-Based Evidence of Reference Dependence
6. Reference Dependence: Housing II
7. Reference Dependence: Tax Elusion
8. Reference Dependence: Goals
9. Reference Dependence: Mergers

1 Laboratory Experiments on Present Bias II

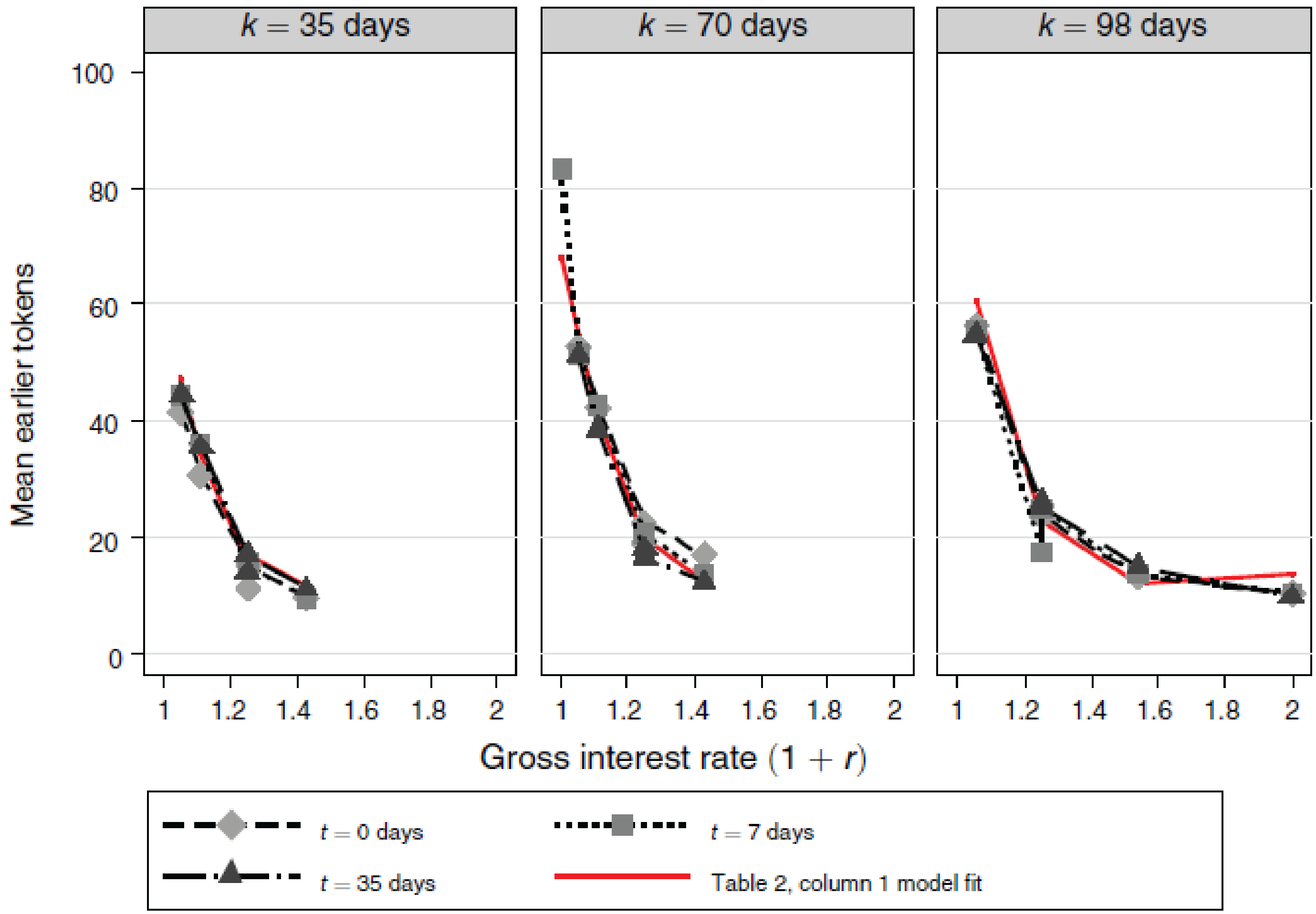
- Recent improved experimental design: Andreoni and Sprenger (AS, *AER* 2012)
- To deal with *Problem 1 (Credibility)*, emphasize credibility
  - All sooner and later payments, including those for $t = 0$, were placed in subjects' campus mailboxes.
  - Subjects were asked to address the envelopes to themselves at their campus mailbox, thus minimizing clerical errors
  - Subjects were given the business card of Professor James Andreoni and told to call or e-mail him if a payment did not arrive
- Potential drawback: Payment today take places at end of day
  - Other experiments: post-dated checks

- To deal with *Problem 3 (Concave Utility)*, design to estimate concavity:
  - Subject allocate share of money to earlier versus later choice
  - -> That is, interior solutions, not just corner solutions
  - Vary interest rate between earlier and later choice to back out concavity
- Example of choice screenshot

January 21, February 25 | January 21, April 1 | January 21, April 29 | January 28, March 4 | January 28, April 8

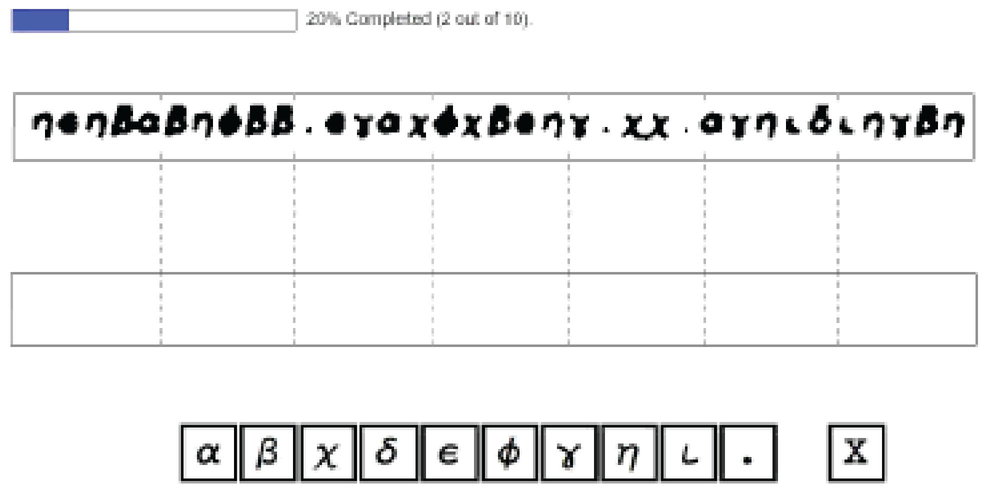
	Divide Tokens between January 28 (1 week(s) from today), and April 8 (10 week(s) later)	January 28	April 8
1	Allocate 100 tokens: 83 tokens at \$0.20 on January 28, and 17 tokens at \$0.20 on April 8	\$16.60	\$3.40
2	Allocate 100 tokens: 51 tokens at \$0.19 on January 28, and 49 tokens at \$0.20 on April 8	\$9.69	\$9.80
3	Allocate 100 tokens: 43 tokens at \$0.18 on January 28, and 57 tokens at \$0.20 on April 8	\$7.74	\$11.40
4	Allocate 100 tokens: 21 tokens at \$0.16 on January 28, and 79 tokens at \$0.20 on April 8	\$3.36	\$15.80
5	Allocate 100 tokens: 14 tokens at \$0.14 on January 28, and 86 tokens at \$0.20 on April 8	\$1.96	\$17.20

- Main result: No evidence of present bias



- What about *Problem 2 (Money vs. Consumption)*?
  - One solution: Do experiments with goods to be consumed right away:
    * Low- and High-brow movies (Read and Loewenstein, 1995)
    * Squirts of juice for thirsty subjects (McClure et al., 2005)
  - Problem: Harder to invoke linearity of utility when using goods as opposed to money
- Augenblick, Niederle, and Sprenger (*QJE* Forthcoming): Address problem by having subjects intertemporally allocate effort
  - 102 subjects have to complete boring task

Panel A: Job 1- Greek Transcription

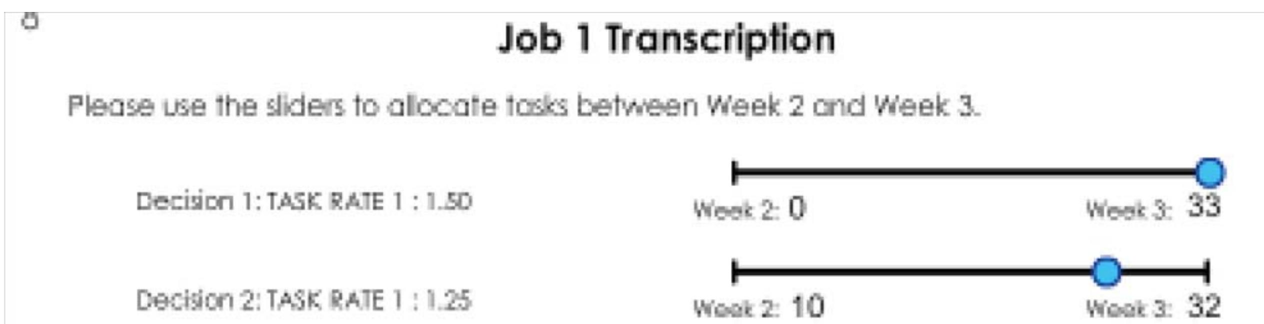


- - Experiment over multiple weeks, complete online
  - Pay largely at the end to reduce attrition
  - Week 1: Choice allocation of job between weeks 2 and 3
  - Week 2: Choose again allocation of job between weeks 2 and 3
  - –> Do subjects revise the choice?
  - As in AS, choice of interior solution, and varied ‘interest rate’ between periods

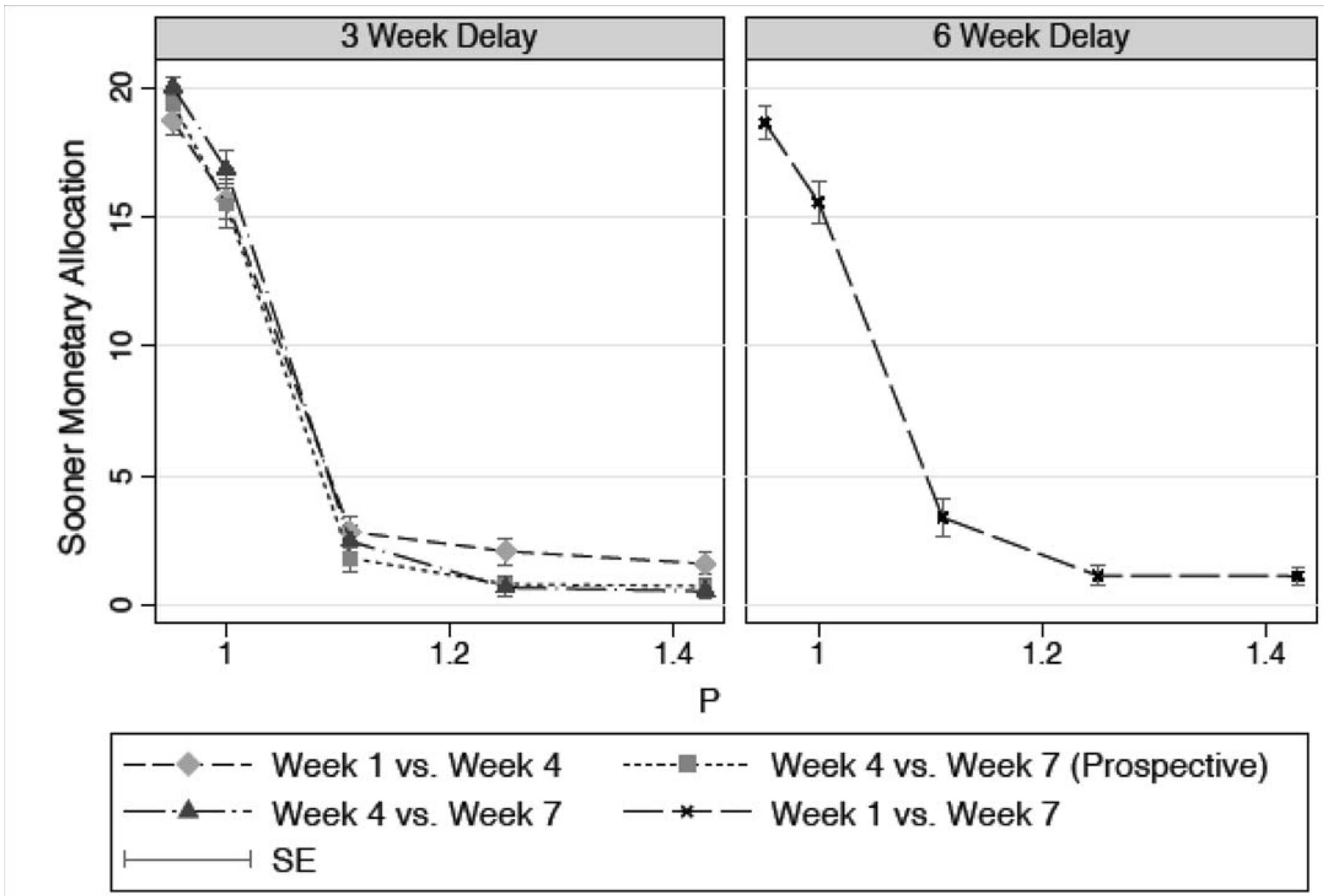
– Also do monetary discounting

Table 1: Summary of Longitudinal Experiment

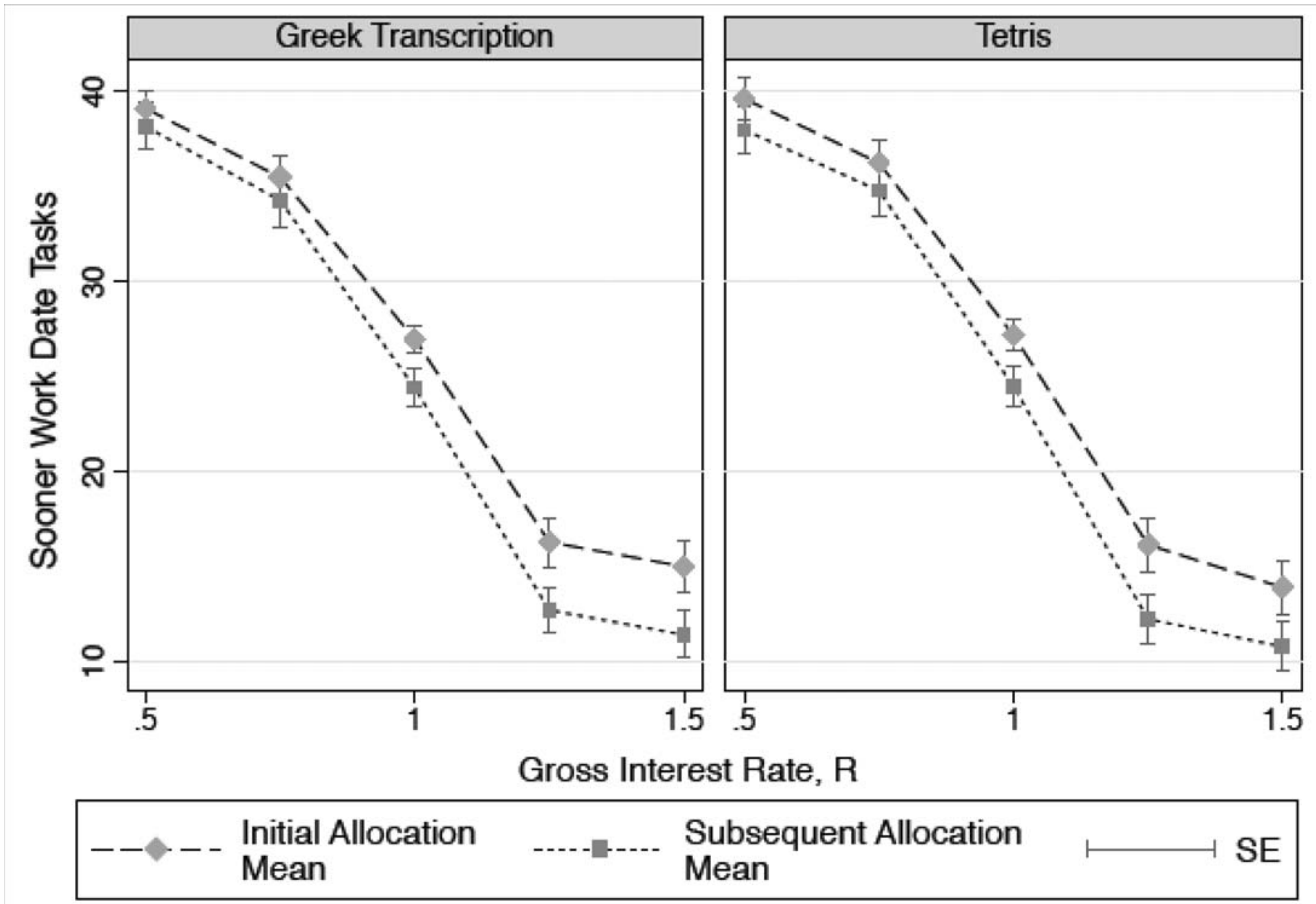
	10 Effort Allocations	Minimum Work	Allocation-That-Counts Chosen	Complete Work	Commitment Choice	Receive Payment
Week 1 (In Lab):	x	x				
Week 2 (Online):	x	x	x	x		
Week 3 (Online):		x		x		
Week 4 (In Lab):	x	x			x	
Week 5 (Online):	x	x	x	x		
Week 6 (Online):		x		x		
Week 7 (In Lab):						x



- Result 1: On monetary discounting no evidence of present-bias



- Result 2: Clear evidence on effort allocation



- Result 3: Estimate of present-bias given that can back out shape of cost of effort function $c(e)$

	Monetary Discounting		Effort Discounting		
	(1) All Delay Lengths	(2) Three Week Delay Lengths	(3) Job 1 Greek	(4) Job 2 Tetris	(5) Combined
Present Bias Parameter: β	0.974	0.988	0.900	0.877	0.888
	(0.009)	(0.009)	(0.037)	(0.036)	(0.033)
Daily Discount Factor: δ	0.998	0.997	0.999	1.001	1.000
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)
Monetary Curvature Parameter: α	0.975	0.976			
	(0.006)	(0.005)			
Cost of Effort Parameter: γ			1.624	1.557	1.589
			(0.114)	(0.099)	(0.104)
# Observations	1500	1125	800	800	1600
# Clusters	75	75	80	80	80
Job Effects					Yes

- **Dean and Sautmann (2014):** Provide direct evidence on *Problem 2 (Money vs. Consumption)*
  - Elicit time preferences with standard money now versus money in the future questions

Table 1: A Price List Experiment

Set A		Set B	
today	in 1 week	in 1 week	in 2 weeks
a_0	a_1	b_1	b_2
CFA 50	CFA 300	CFA 50	CFA 300
CFA 100	CFA 300	CFA 100	CFA 300
CFA 150	CFA 300	CFA 150	CFA 300
CFA 200	CFA 300	CFA 200	CFA 300
CFA 250	CFA 300	CFA 250	CFA 300
CFA 300	CFA 300	CFA 300	CFA 300
CFA 350	CFA 300	CFA 350	CFA 300
CFA 400	CFA 300	CFA 400	CFA 300

- – Observe shocks to ability to borrow and marginal utility of income
  – Do those affect the choices in price list?
  – If so, clearly we are not capturing δ, but rather r or u'
  – Estimate MRS from questions above, relate to adverse income shock

Table 5: Consumption shocks and MRS_t.

	MRS (A) OLS	MRS (A) OLS	MRS (A) OLS	MRS (A) OLS	MRS (A) IV	MRS (A) IV
Adv. event (0/1)	0.284 *	0.263 *				
	(0.124)	(0.124)				
Adv. event expense			0.256 +	0.237 +	1.707 *	1.579 *
			(0.147)	(0.141)	(0.695)	(0.797)
Constant	4.588 **	4.678 **	4.665 **	4.755 **	4.579 **	4.663 **
	(0.041)	(0.074)	(0.009)	(0.059)	(0.101)	(0.130)
Ind FE	yes	yes	yes	yes	yes	yes
Time FE		yes		yes		yes
Observations	2547	2547	2543	2543	2543	2543

Standard errors clustered at the individual level (OLS) or bootstrapped (IV, ML) (in p

*Significance levels + p<0.10, * p<0.05, ** p<0.01*

- – Related to savings shock

Table 7: Income, spending, and MRS_t.

	MRS (A) OLS		MRS (A) OLS		MRS (A) OLS		MRS (A) OLS		MRS (A) IV		MRS (A) IV		MRS (A) ML	
Labor income					-0.185		-0.189		-0.153		-0.159		-0.324	*
					(0.142)		(0.143)		(0.163)		(0.142)		(0.135)	
Nonlabor income					-0.330		-0.321		-0.268		-0.265		-0.281	
"endogenous"					(0.251)		(0.258)		(0.261)		(0.270)		(0.351)	
Nonlabor income	**-0.409**	**	**-0.409**	**	**-0.382**	**	**-0.384**	**	**-0.378**	*	**-0.380**	*	**-0.407**	**
"exogenous"	**(0.142)**		**(0.149)**		**(0.125)**		**(0.133)**		**(0.171)**		**(0.149)**		**(0.199)**	
Other spending					0.268	*	0.245	+	0.192		0.177		0.236	
					(0.128)		(0.131)		(0.141)		(0.132)		(0.135)	
Adv. event expense	**0.252**	+	**0.233**	+	**0.251**		**0.222**		**1.683**	+	**1.562**	*	**0.357**	+
	(0.145)		**(0.139)**		**(0.182)**		**(0.183)**		**(0.761)**		**(0.769)**		**(0.250)**	
Constant	4.69	**	4.782	**	4.56	**	4.67	**	4.527	**	4.622	**	2.737	**
	(0.011)		(0.059)		(0.093)		(0.125)		(0.144)		(0.145)		(0.145)	
Ind FE	yes		yes		yes		yes		yes		yes		yes	
Time FE			yes				yes				yes			
Observations	2540		2540		2390		2390		2390		2390		1437	

Standard errors clustered at the individual level (OLS) or bootstrapped (IV, ML) (in parentheses).
*Significance levels + p<0.10, * p<0.05, ** p<0.01*

- **Carvalho, Meier, Wang (2014):** Replicates both of the previous findings
  - Measures time preferences with money and real effort
  - 1,191 participants rendomized into
    * Surveyed before payday (financially constrained)
    * Surveyed after payday (not constrained)
  - Real effort task (clever):
    * Complete shorter survey within 5 days
    * Complete longer survey within 35 days
    * Multiple choices withvarying length of sooner survey

- Replicates Dean and Sautmann result on financial choices

Table 3: Intertemporal Choices about Monetary Rewards

	\$ Amount of Sooner Reward	
	Coefficient	Standard Error
{Before Payday} * {Immediate Rewards}	10.6	3.83***
{Before Payday} * Interest Rate	2.7	3.24
{Before Payday} * Delay Time	-1.4	1.06
{Before Payday}	-6.3	9.80
{Immediate Rewards}	-5.3	2.75*
Experimental Interest Rate	-47.3	2.33***
Delay Time	-0.7	0.72
Constant	304.3	6.83***

Notes: This table reports results from an OLS regression where the dependent variable is the dollar amount of the sooner payment. “Immediate Rewards” is an indicator variable that is 1 if the mailing date of the sooner payment is today. “Delat Time” is the time interval between the sooner and later payments. The sample is restricted to the 1,060 subjects who made all 12 choices in the task with monetary rewards. $N = 12{,}720$.

- Replicates Augenblick et al. on real effort

Table 4: Intertemporal Choices about Real Effort

	Monthly Discount Rate
{Before Payday} * {Immediate Task}	-0.03
	[0.025]
{Before Payday}	0.02
	[0.027]
{Immediate Task}	0.09
(5-day deadline for short-sooner survey)	[0.018]***
Constant	0.31
	[0.019]***

Notes: This table reports estimates from an interval regression where the dependent variable is the interval measure of the individual discount rate (IDR). Two IDRs are estimated for each subject; one for each time frame. "Immediate Task" is an indicator variable for the "5 days (sooner) x 35 days (later)" time frame. Standard errors clustered at the individual level. The sample is restricted to the 1,025 subjects who made all 10 choices in the non-monetary intertemporal task. $N = 2{,}050$.

2 Methodology: Errors in Applying Present-Biased Preferences

- Present-Bias model very successful
- Quick adoption at cost of incorrect applications
- Four common errors

- **Error 1. Procrastination with Sophistication**
  - 'Self-Control leads to Procrastination'
  - This is not accurate in two ways
  - *Issue 1.*
    * (β, δ) Sophisticates do not delay for long (see our calibration)
    * Need Self-control $+$ Naiveté (overconfidence) to get long delay
  - *Issue 2.* (Definitional issue) We distinguished between:
    * Delay. Task is not undertaken immediately
    * Procrastination. Delay systematically beyond initial expectations
    * Sophisticates and exponentials do not procrastinate, they *delay*

- **Error 2. Naives with Yearly Decisions**
  - 'We obtain similar results for naives and sophisticates in our calibrations'
  - Example 1. Fang, Silverman (*IER*, 2009)
  - Single mothers applying for welfare. Three states:
    1. Work
    2. Welfare
    3. Home (without welfare)
  - Welfare dominates Home – So why so many mothers stay Home?

Choice at $t-1$	Choice at t		
	Welfare	Work	Home
Welfare			
Row %	84.3	3.5	12.3
Column %	76.7	6.3	17.9
Work			
Row %	5.3	79.3	15.3
Column %	2.6	76.4	12.1
Home			
Row %	28.3	12.0	59.7
Column %	20.7	17.3	70.0

- – Model:
    - ∗ Immediate cost ϕ (stigma, transaction cost) to go into welfare
    - ∗ For ϕ high enough, can explain transition
    - ∗ Simulate Exponentials, Sophisticates, Naives

- However: Simulate decision at **yearly** horizon.
- BUT: At yearly horizon naives do not procrastinate:
  - Compare:
    - Switch now
    - Forego *one year* of benefits and switch next year
- Result:
  - Very low estimates of β
  - Very high estimates of switching cost ϕ
  - Naives are same as sophisticates

Parameters		(1)		(2)		(3)	
		Time Consistent		Present-Biased (sophisticated)		Present-Biased (Naive)	
		Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Preference Parameters							
Discount Factors	β	1	n.a.	0.33802	0.06943	0.355	0.0983
	δ	0.41488	0.07693	0.87507	0.01603	0.868	0.02471
Net Stigma	$\phi^{(1)}$	7537.04	774.81	8126.19	834.011	8277.46	950.77
(by type)	$\phi^{(2)}$	10100.9	1064.83	10242.01	955.878	10350.20	1185.27
	$\phi^{(3)}$	13333.2	1640.18	12697.25	1426.40	12533.69	1685.92

- – Conjecture: If allowed daily or weekly decision, would get:
  - ∗ Naives fit much better than sophisticates
  - ∗ β much closer to 1
  - ∗ ϕ much smaller

– Example 2. Shui and Ausubel (2005) –> Estimate Ausubel (1999)

* Cost k of switching from credit card to credit card
* Again: Assumption that can switch only every quarter
* Results of estimates (again):
  · Quite low β
  · Naives do not do better than sophisticates
  · Very high switching costs

Table 4: Estimated Parameters [a]

	Sophisticated Hyperbolic	Naive Hyperbolic	Exponential
β	0.7863 (0.00192)	0.8172 (0.003)	
δ	0.9999 (0.00201)	0.9999 (0.0017)	0.9999 (0.00272)
k	0.02927 \$293 (0.00127)	0.0326 \$326 (0.00139)	0.1722 \$1,722 (0.0155)

- **Error 3. Present-Bias over Money**
  - We discussed problem applied to experiments
  - Same problem applies to models
    * Notice: Transaction costs of switching k in above models are real effort, apply immediately
    * Effort cost c of attending gym also 'real' (not monetary)
    * Consumption-Savings models: Utility function of consumption c, not income I

- **Error 4. Getting the Intertemporal Payoff Wrong**
  - 'Costs are in the present, benefits are in the future'
  - (β, δ) models very sensitive to timing of payoffs
  - Sometimes, can easily turn investment good into leisure good
  - Need to have strong intuition on timing
  - Example: Paper on nuclear plants as leisure goods
    * Immediate benefits of energy
    * Delayed cost to environment
  - BUT: 'Immediate' benefits come after 10 years of construction costs!

3 Reference Dependence: Introduction

- Kahneman and Tversky (*EMA* 1979) — Anomalous behavior in experiments:

  1. *Concavity over gains.* Given \$1000, A=(500,1) $\succ$ B=(1000,0.5;0,0.5)
  2. *Convexity over losses.* Given \$2000, C=(-1000,0.5;0,0.5) $\succ$ D=(-500,1)
  3. *Framing Over Gains and Losses.* Notice that A=D and B=C
  4. *Loss Aversion.* (0,1) $\succ$ (-8,.5;10,.5)
  5. *Probability Weighting.* (5000,.001) $\succ$ (5,1) and (-5,1) $\succ$ (-5000,.001)

- Can one descriptive model theory fit these observations?

- **Prospect Theory** (Kahneman and Tversky, 1979)

- Subjects evaluate a lottery $(y, p; z, 1-p)$ as follows: $\pi(p)\, v(y-r) + \pi(1-p)\, v(z-r)$

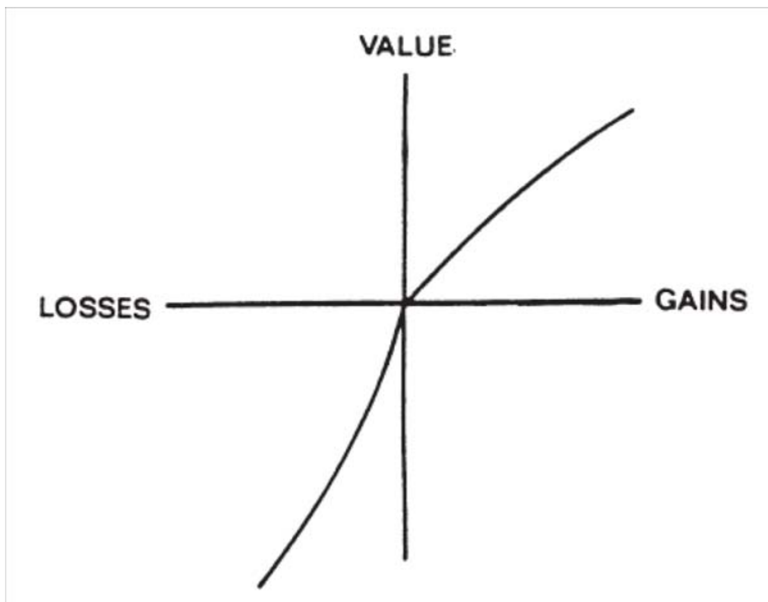
- Five key components:

  1. Reference Dependence

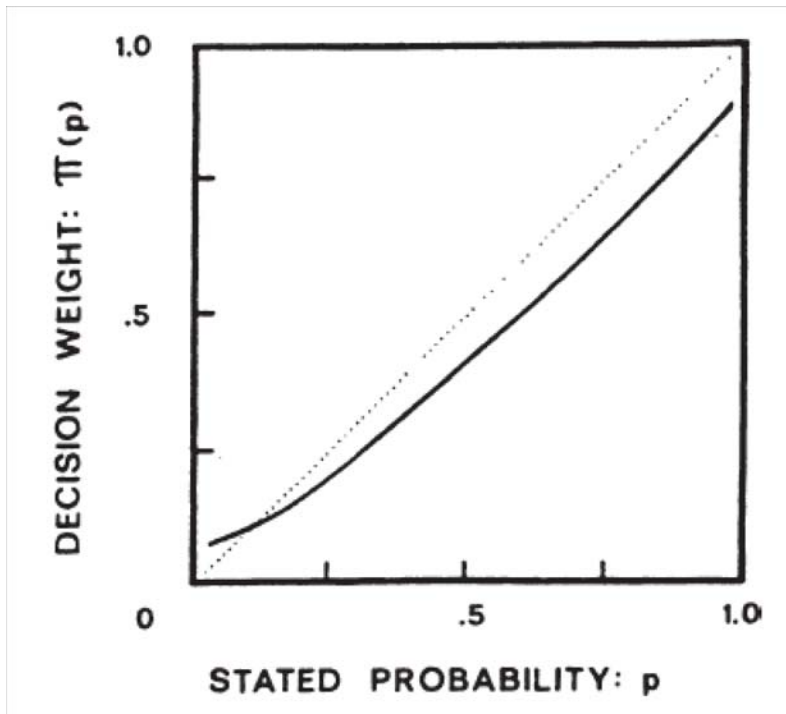
     – Basic psychological intuition that changes, not levels, matter (applies also elsewhere)

     – Utility is defined over differences from reference point r −> Explains Exp. 3

2. Diminishing sensitivity.
   - Concavity over gains of v –> Explains (500,1)$\succ$(1000,0.5;0,0.5)
   - Convexity over losses of v –> Explains (-1000,0.5;0,0.5)$\succ$(-500,1)
3. Loss Aversion –> Explains (0,1) $\succ$ (-8,.5;10,.5)



4. Probability weighting function π non-linear –> Explains (5000,.001) $\succ$ (5,1) and (-5,1) $\succ$ (-5000,.001)



- Overweight small probabilities + Premium for certainty

5. Narrow framing (Barberis, Huang, and Thaler, 2006; Rabin and Weizsäcker, 2011)

   – Consider only risk in isolation (labor supply, stock picking, house sale)

   – Neglect other relevant decisions

- Tversky and Kahneman (1992) propose calibrated version

$$v(x) = \begin{cases} (x-r)^{.88} & \text{if } x \geq r; \\ -2.25\left(-(x-r)\right)^{.88} & \text{if } x < r, \end{cases}$$

and

$$w(p) = \frac{p^{.65}}{\left(p^{.65} + (1-p)^{.65}\right)^{1/.65}}$$

- Reference point r?

- Open question – depends on context

- Koszegi-Rabin (2006 on): personal equilibrium with rational expectation outcome as reference point

- Most field applications use only (1)+(3), or (1)+(2)+(3)

$$v(x) = \begin{cases} x - r & \text{if } x \geq r; \\ \lambda(x - r) & \text{if } x < r, \end{cases}$$

- Assume backward looking reference point depending on context

4 Reference Dependence: Housing I

- Start from old-school reference-dependence paper

- Two typical ingredients:

  1. Backward-looking reference points (status quo, focal point, or past outcome)

  2. 'Informal' test – No model

- **Genesove-Mayer (QJE, 2001)**

  1. For houses sales, natural reference point is previous purchase price

     – Validation: 75% of home owners remember exactly the purchase price of their home (survey evidence from our door-to-door surveys)

  2. Loss Aversion –> Unwilling to sell house at a loss

     – Will ask for higher price if at a loss relative to pruchase price

- Evidence: Data on Boston Condominiums, 1990-1997
- Substantial market fluctuations of price

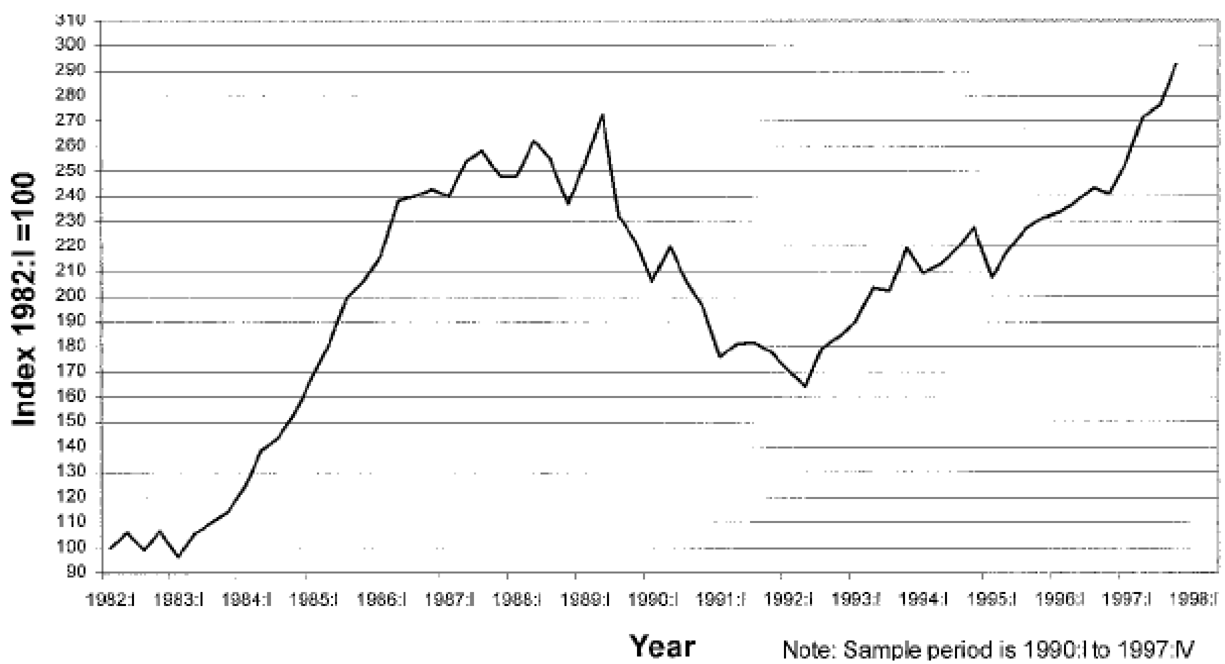


FIGURE I
Boston Condominium Price Index

- Observe:
  - Listing price $L_{i,t}$ and last purchase price P_0
  - Observed Characteristics of property X_i
  - Time Trend of prices δ_t
- Define:
  - $\hat{P}_{i,t}$ is market value of property i at time t
- Ideal Specification:

$$\begin{aligned} L_{i,t} &= \hat{P}_{i,t} + m\mathbf{1}_{\hat{P}_{i,t}<P_0}\left(P_0 - \hat{P}_{i,t}\right) + \varepsilon_{i,t} \\ &= \beta X_i + \delta_t + v_i + mLoss^* + \varepsilon_{i,t} \end{aligned}$$

- However:
  - Do not observe $\hat{P}_{i,t}$, given v_i (unobserved quality)
  - Hence do not observe $Loss^*$

- Two estimation strategies to bound estimates. *Model 1:*

$$L_{i,t} = \beta X_i + \delta_t + m\mathbf{1}_{\hat{P}_{i,t}<P_0} (P_0 - \beta X_i - \delta_t) + \varepsilon_{i,t}$$

  - This model overstate the loss for high unobservable homes (high v_i)
  - Bias upwards in $\hat{m}$, since high unobservable homes should have high $L_{i,i}$

- *Model 2:*

$$L_{i,t} = \beta X_i + \delta_t + \alpha (P_0 - \beta X_i - \delta_t) + m\mathbf{1}_{\hat{P}_{i,t}<P_0} (P_0 - \beta X_i - \delta_t) + \varepsilon_{i,t}$$

- Estimates of impact on sale price

TABLE II

LOSS AVERSION AND LIST PRICES

DEPENDENT VARIABLE: LOG (ORIGINAL ASKING PRICE),

OLS equations, standard errors are in parentheses.

Variable	(1) All listings	(2) All listings	(3) All listings	(4) All listings	(5) All listings	(6) All listings
LOSS	0.35 (0.06)	0.25 (0.06)	0.63 (0.04)	0.53 (0.04)	0.35 (0.06)	0.24 (0.06)
LOSS-squared			−0.26 (0.04)	−0.26 (0.04)		
LTV	0.06 (0.01)	0.05 (0.01)	0.03 (0.01)	0.03 (0.01)	0.06 (0.01)	0.05 (0.01)
Estimated value in 1990	1.09 (0.01)	1.09 (0.01)	1.09 (0.01)	1.09 (0.01)	1.09 (0.01)	1.09 (0.01)
Estimated price index at quarter of entry	0.86 (0.04)	0.80 (0.04)	0.91 (0.03)	0.85 (0.03)		
Residual from last sale price		0.11 (0.02)		0.11 (0.02)		0.11 (0.02)
Months since last sale	−0.0002 (0.0001)	−0.0003 (0.0001)	−0.0002 (0.0001)	−0.0003 (0.0001)	−0.0002 (0.0001)	−0.0003 (0.0001)
Dummy variables for quarter of entry	No	No	No	No	Yes	Yes
Constant	−0.77 (0.14)	−0.70 (0.14)	−0.84 (0.13)	−0.77 (0.14)	−0.88 (0.10)	−0.86 (0.10)
R^2	0.85	0.86	0.86	0.86	0.86	0.86
Number of observations	5792	5792	5792	5792	5792	5792

- Effect of experience: Larger effect for owner-occupied

TABLE IV
LOSS AVERSION AND LIST PRICES: OWNER-OCCUPANTS VERSUS INVESTORS
DEPENDENT VARIABLE: LOG (ORIGINAL ASKING PRICE)
OLS equations, standard errors are in parentheses.

Variable	(1) All listings	(2) All listings	(3) All listings	(4) All listings
LOSS × owner-occupant	0.50	0.42	0.66	0.58
	(0.09)	(0.09)	(0.08)	(0.09)
LOSS × investor	0.24	0.16	0.58	0.49
	(0.12)	(0.12)	(0.06)	(0.06)
LOSS-squared × owner-occupant			−0.16	−0.17
			(0.14)	(0.15)
LOSS-squared × investor			−0.30	−0.29
			(0.02)	(0.02)
LTV × owner-occupant	0.03	0.03	0.01	0.01
	(0.02)	(0.02)	(0.01)	(0.01)
LTV × investor	0.053	0.053	0.02	0.02
	(0.027)	(0.027)	(0.02)	(0.02)
Dummy for investor	−0.02	−0.02	−0.03	−0.03
	(0.014)	(0.01)	(0.01)	(0.01)
Estimated value in 1990	1.09	1.09	1.09	1.09
	(0.01)	(0.01)	(0.01)	(0.01)
Estimated price index at quarter of entry	0.84	0.80	0.86	0.82
	(0.05)	(0.04)	(0.04)	(0.04)
Residual from last sale price		0.08		0.08
		(0.02)		(0.02)

- Some effect also on final transaction price

TABLE VI

LOSS AVERSION AND TRANSACTION PRICES

DEPENDENT VARIABLE: LOG (TRANSACTION PRICE)

NLLS equations, standard errors are in parentheses.

Variable	(1) All listings	(2) All listings
LOSS	0.18	0.03
	(0.03)	(0.08)
LTV	0.07	0.06
	(0.02)	(0.01)
Residual from last sale price		0.16
		(0.02)
Months since last sale	−0.0001	−0.0004
	(0.0001)	(0.0001)
Dummy variables for quarter of entry	Yes	Yes
Number of observations	3413	3413

- Lowers the exit rate (lengthens time on the market)

TABLE VII

HAZARD RATE OF SALE

Duration variable is the number of weeks the property is listed on the market. Cox proportional hazard equations, standard errors are in parentheses.

Variable	(1) All listings	(2) All listings	(3) All listings	(4) All listings
LOSS	−0.33	−0.63	−0.59	−0.90
	(0.13)	(0.15)	(0.16)	(0.18)
LOSS-squared			0.27	0.28
			(0.07)	(0.07)
LTV	−0.08	−0.09	−0.06	−0.06
	(0.04)	(0.04)	(0.04)	(0.04)
Estimated value	0.27	0.27	0.27	0.27
in 1990	(0.04)	(0.04)	(0.04)	(0.04)
Residual from		0.29		0.29
last sale		(0.07)		(0.07)

- – Overall, plausible set of results that show impact of reference point

5 Methodology: Bunching-Based Evidence of Reference Dependence

- How does one identify reference-dependence?
- Some Cases: Key role for *diminishing sensitivity* and *probability weighting*
  - Disposition effect: Diminishing sensitivity –> more prone to sell winners (part of effect)
  - Insurance: Prob. weighting –> propensity to get low deductible
- Most Cases: Key role for *loss aversion*
- Common element for several papers:
  - Well-defined, backward-looking reference point r
  - Optimal effort choice e^*

- Cost of effort $c(e)$
- Return of effort e, reference point r

- Individual maximizes

$$\begin{aligned} \max_e e + \eta [e - r] - c(e) \text{ for } e &\geq r \\ \max_e e + \eta\lambda [e - r] - c(e) \text{ for } e &< r \end{aligned}$$

- Derivative of utility function:

$$\begin{aligned} 1 + \eta - c'(e^*) \text{ for } e &\geq r \\ 1 + \lambda\eta - c'(e^*) \text{ for } e &< r \end{aligned}$$

  - Discontinuity in marginal utility of effort
  - Implication 1 –> Bunching at $e^* = r$
  - Implication 2 –> Missing mass of distribution for $e < r$ compared to $e > r$

- Older literature does not purse this, new literature does
  - Bunching is much harder to explain with alternative models
  - Shift in mass can generally be well identified too under assumptions of contiuity of distribution
- Examine four related applications:
  1. Housing (where test is not formalized)
     - Effort: How hard to 'push' the house
     - Reference point: Purchase price
  2. Tax filing
     - Effort: Tax elusion
     - Reference point: Withholding amount
  3. Marathon running

  - Effort: Running
  - Reference point: Round goal

4. Merger
  - Effort: Pushing for higher price
  - Reference point: 52-week high

- Two more related cases next lecture:

5. Labor supply
  - Effort: Work more hours
  - Reference point: Expected daily earnings?

6. Job search
  - Effort: Search for a job
  - Reference point: Recent average earnings

6 Reference Dependence: Housing II

- Return to Housing case, formalize intuition.
  - Seller chooses price P at sale
  - Higher Price P
    - lowers probability of sale $p(P)$ (hence $p'(P) < 0$)
    - increases utility of sale $U(P)$
  - If no sale, utility is $\bar{U} < U(P)$ (for all relevant P)

- Maximization problem:

$$\max_P p(P)U(P) + (1 - p(P))\bar{U}$$

- F.o.c. implies

$$MG = p(P^*)U'(P^*) = -p'(P^*)(U(P^*) - \bar{U}) = MC$$

- Interpretation: Marginal Gain of increasing price equals Marginal Cost

- S.o.c are

$$2p'(P^*)U'(P^*) + p(P^*)U''(P^*) + p''(P^*)(U(P^*) - \bar{U}) < 0$$

- Need $p''(P^*)(U(P^*) - \bar{U}) < 0$ or not too positive

- Reference-dependent preferences with reference price P_0 (with pure gain-loss utility):

$$v(P|P_0) = \begin{cases} P - P_0 & \text{if } P \geq P_0; \\ \lambda(P - P_0) & \text{if } P < P_0, \end{cases}$$

  – (in this case, think of $\bar{U} < 0$)

  – Can write as

$$\begin{aligned} p(P) &= -p'(P)(P - P_0 - \bar{U}) \text{ if } P \geq P_0 \\ p(P)\lambda &= -p'(P)(\lambda(P - P_0) - \bar{U}) \text{ if } P < P_0 \end{aligned}$$

  – Plot Effect on MG and MC of loss aversion

- Compare $P^*_{\lambda=1}$ (equilibrium with no loss aversion) and $P^*_{\lambda>1}$ (equilibrium with loss aversion)

- Case 1. Loss Aversion λ increase price $(P^*_{\lambda=1} < P_0)$

- Case 2. Loss Aversion λ induces bunching at $P = P_0$ $(P^*_{\lambda=1} < P_0)$

- Case 3. Loss Aversion has no effect ($P^*_{\lambda=1} > P_0$)

- General predictions. When aggregate prices are low:
  - High prices P relative to fundamentals
  - Bunching at purchase price P_0
  - Lower probability of sale $p(P)$
  - Longer waiting on market
- Important to tie housing evidence to model

- **Gagnon-Bartsch, Rosato, and Xia (2010)**: Re-analyze data
  - Some evidence on bunching
  - Did not do shifting test
  - Would be great to redo with data from recent recession

7 Reference Dependence: Tax Elusion

- **Alex Rees-Jones (2014)**
- Important setting which can also differentiate from alternative model of reference points:
  - Utility has fixed jump (but no kink)
  - Prediction of bunching
  - BUT no prediction of shift in distribution
- Slides courtesy of Alex
- Other relevant paper: **Engstrom, P., Nordblom, K., Ohlsson, H., & Persson, A. (AEJ: Policy, forthcoming)**
  - Similar evidence, but focus on claiming deductions

Decision environment

Consider the decisions made in the process of filing tax returns.

Some tax-relevant behaviors are predetermined.

- E.g., withholding, labor supply.

But, conditional on predetermined behavior, the taxpayer can:

1. Work to claim tax shelters for past behavior.
2. Pursue additional tax shelters.

Sheltering reduces current tax payment, at a cost:

- Evasion: e.g., income underreporting.
  - Costs: expected future penalties, accounting effort, stigma, etc.
- Avoidance: e.g., legal pursuit of credits, deductions.
  - Costs: effort and attention.

Model of sheltering decisions

$$\max_{s \in \mathbb{R}^+} \quad \underbrace{m(-b^{PM} + s)}_{\text{utility over money}} \quad - \quad \underbrace{c(s)}_{\text{cost of sheltering}}$$

- b^{PM} —"pre-manipulation" balance due, with PDF f_b^{PM}.
  - Determined by past labor supply decisions, tax payments, and many other factors.
  - **Primary assumption:** f_b^{PM} is continuous.
- s — tax dollars sheltered.
  - Assumes that sheltering can be precisely targeted.
- $c(\cdot)$ — increasing, convex, and twice continuously differentiable cost of sheltering.

Simple example with smooth utility

Consider a model abstracting from income effects:

$$\max_{s \in \mathbb{R}^+} \quad \underbrace{(w - b^{PM} + s)}_{\text{linear utility over money}} \quad - \quad \underbrace{c(s)}_{\text{cost of sheltering}}$$

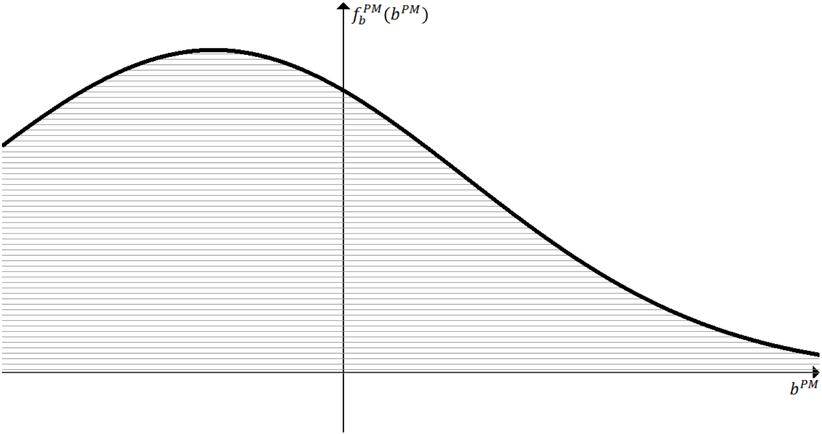
Optimal sheltering is determined by the first-order condition:

$$1 - c'(s^*) = 0$$

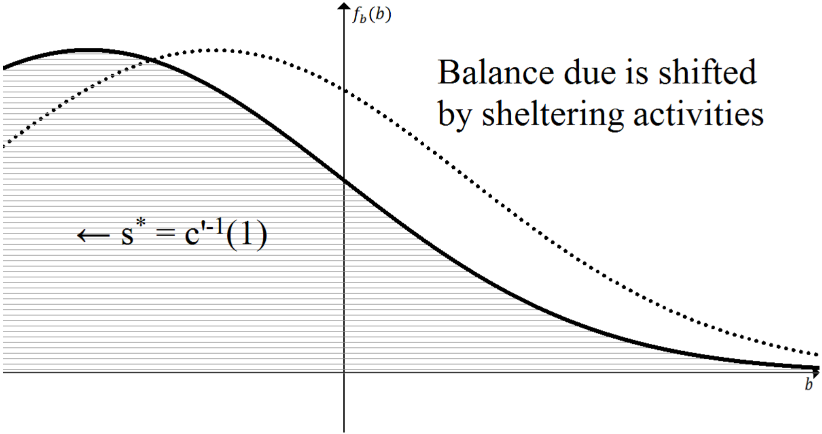
Optimal sheltering solution: $s^* = c'^{-1}(1)$.

$\rightarrow$ Distribution of balance due, $b \equiv b^{PM} - s^*$, is a horizontal shift of the distribution of b^{PM}.

PDF of pre-manipulation balance due



PDF of final balance due after sheltering



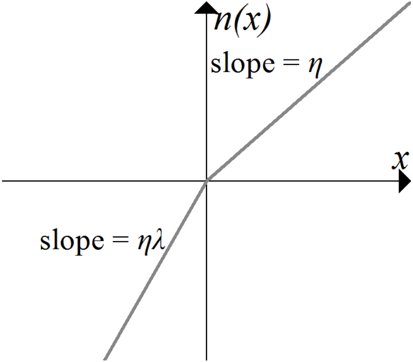
Loss-averse case

$$\max_{s \in \mathbb{R}^+} \underbrace{m(-b^{PM} + s)}_{\text{utility over money}} - \underbrace{c(s)}_{\text{cost of sheltering}}$$

Loss-averse utility specification:

$$\underbrace{(w - b^{PM} + s)}_{\text{consumption utility}} + \underbrace{n(-b^{PM} + s - r)}_{\text{gain-loss utility}}$$

$$n(x) = \begin{cases} \eta x & \text{if } x \geq 0 \\ \eta\lambda x & \text{if } x < 0 \end{cases}$$



Optimal loss-averse sheltering

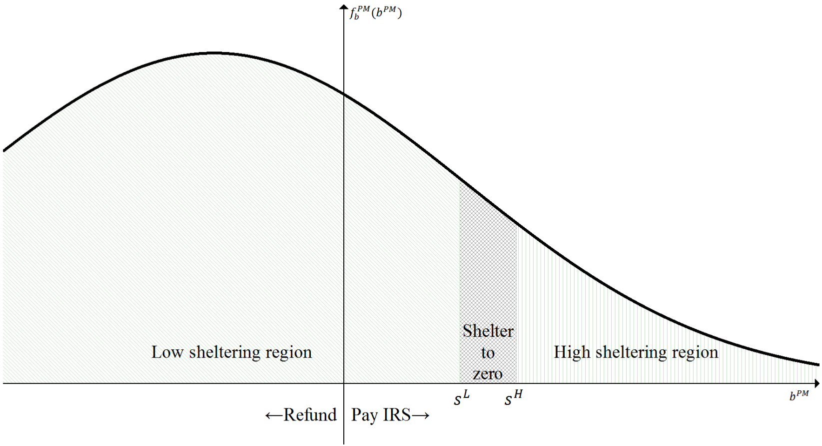
This model generates an optimal sheltering solution with different behavior across three regions:

$$s^*(b^{PM}) = \begin{cases} s^H & \text{if } b^{PM} > s^H - r \\ b^{PM} + r & \text{if } b^{PM} \in \left[s^L - r, s^H - r\right] \\ s^L & \text{if } b^{PM} < s^L - r \end{cases}$$

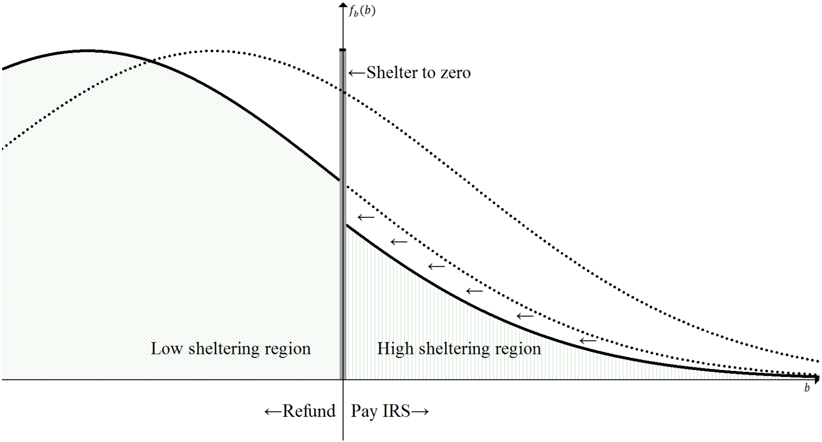
where $s^H \equiv c'^{-1}(1 + \eta\lambda)$ and $s^L \equiv c'^{-1}(1 + \eta)$.

- Sufficiently large $b^{PM} \rightarrow$ high amount of sheltering.
- Sufficiently small $b^{PM} \rightarrow$ low amount of sheltering.
- For an intermediate range, sheltering chosen to offset b^{PM}.

PDF of pre-manipulation balance due



PDF of final balance due after loss-averse sheltering



Revenue effect of loss framing: $s^H - s^L$.

Goals of empirical analysis

We will now test these two predictions in IRS tax records, and quantify the revenue effect each implies.

Bunching prediction: Excess mass at gain/loss threshold.

Shifting prediction: Dist. of losses shifted relative to gains.

Need to address potential confounds:

- Nonrefundable credits
- Extremely accurate tax forecasting
- Fixed costs in the loss domain
- Interactions with tax preparers
- Avoidance of underwitholding penalties
- Liquidity constraints

Data description

Dataset: 1979-1990 SOI Panel of Individual Returns.

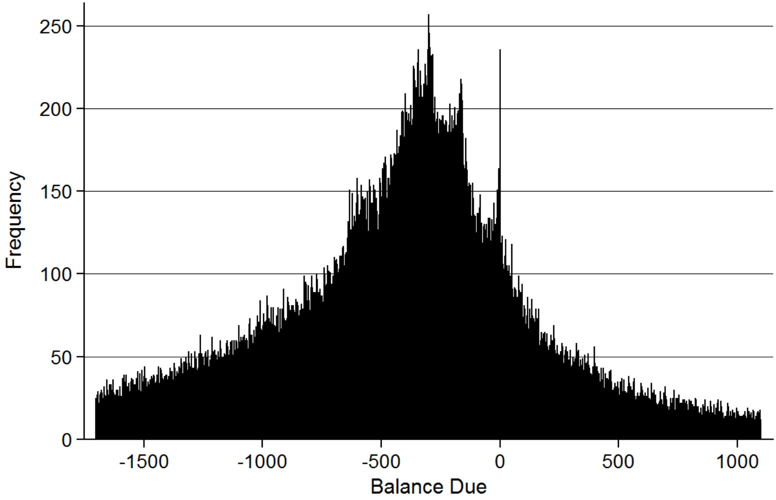
- Contains most information from Form 1040 and some related schedules.
- Randomized by SSNs.

Exclude observations filed from outside of the 50 states + DC, drawn from outside the sampling frame, observations before 1979.

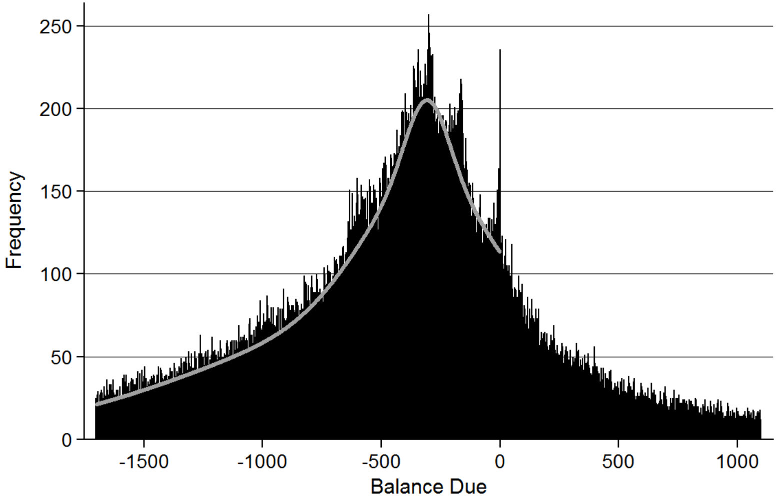
Exclude individuals with zero pre-credit tax due, individuals with zero tax prepayments.

Primary sample: $\approx 229k$ tax returns, $\approx 53k$ tax filers.

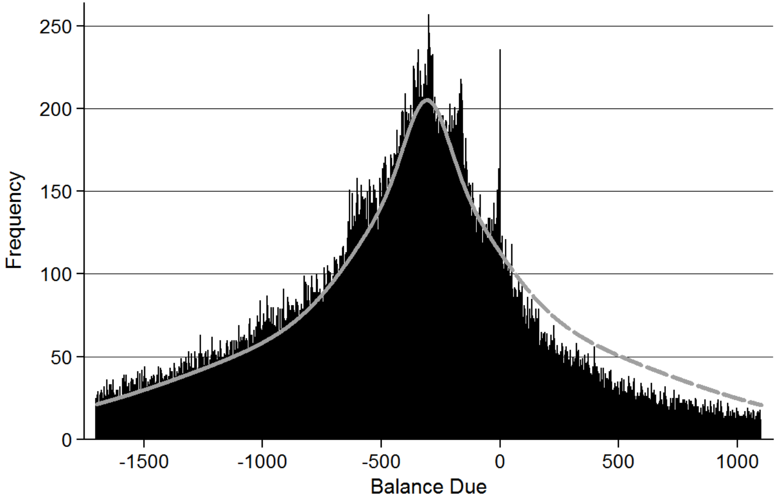
First look: distribution of nominal balance due



First look: distribution of nominal balance due



First look: distribution of nominal balance due



Quantifying excess mass

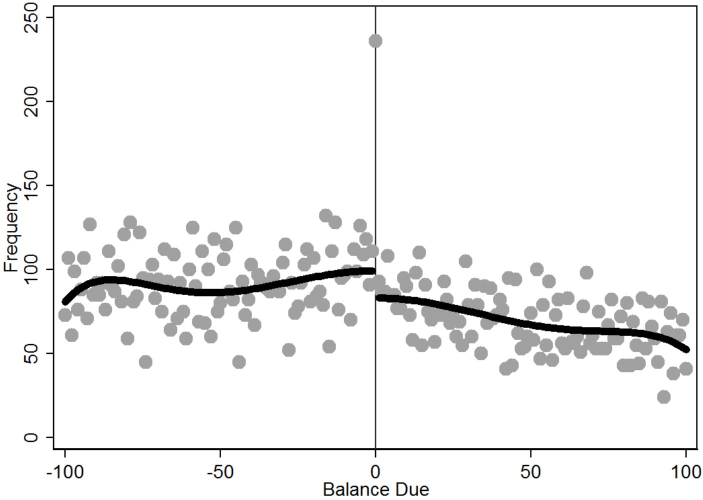
Approach motivated by Chetty, Friedman, Olsen, and Pistaferri (2011), who studied bunching behavior in an alternate setting.

$$C_j = \alpha + \left[\sum_{i=1}^{7} \beta_i \cdot b_j^i \right] + \gamma \cdot I(b_j = 0) + \delta \cdot I(b_j > 0) + \epsilon_j$$

Fits the histogram local to the referent with a 7th-order polynomial.

- All values expressed in 1990 dollars.

Distribution of balance due near gain/loss threshold



	(1) All AGI groups	(2) 1st AGI quartile	(3) 2nd AGI quartile	(4) 3rd AGI quartile	(5) 4th AGI quartile
γ: I(balance due $=0$)	136.43***	46.57***	26.79***	21.06***	42.01***
	(18.46)	(8.25)	(6.95)	(5.66)	(4.15)
δ : I(balance due >0)	-16.26*	-3.50	-4.20	-3.42	-5.14**
	(9.41)	(4.21)	(3.54)	(2.89)	(2.12)
α : Constant	99.57***	33.43***	27.21***	21.94***	16.99***
	(5.45)	(2.44)	(2.05)	(1.67)	(1.23)
Balance-due polynomial	X	X	X	X	X
N: Bins in histogram	201	201	201	201	201
Observations	16348	5725	4553	3602	2468
R^2	0.490	0.479	0.259	0.209	0.489

Notes: Standard errors in parentheses. Similar estimates generated with bootstrapped standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table with bootstrapped SEs

Results robust to alternative orders of the polynomial.

- Similar or stronger significance patterns for polynomials of order one through ten.
- BIC selects 2nd-order polynomial, yields similar results.

These estimates can be used to bound $s^H - s^L$.

Estimates of shifting in loss domain

The estimates we've focused on thus far have been based on the bunching prediction.

Now we will assess the shifting prediction.

- Complementary approach: estimates $(s^H - s^L)$ from a different feature of the data.
- Different strengths and weaknesses.

Pros: uses more of the data, less danger that individuals near zero are non-representative.

Cons: will rely more on functional form restrictions, more susceptible to systematic differences in unobserved variables.

Excluding data at gain/loss threshold, loss-averse sheltering implies:

$$f_b(x) = \begin{cases} f_b^{PM}(x+\kappa) & \text{if } x < r \\ f_b^{PM}(x+\kappa+\tilde{s}) & \text{if } x > r \end{cases}$$

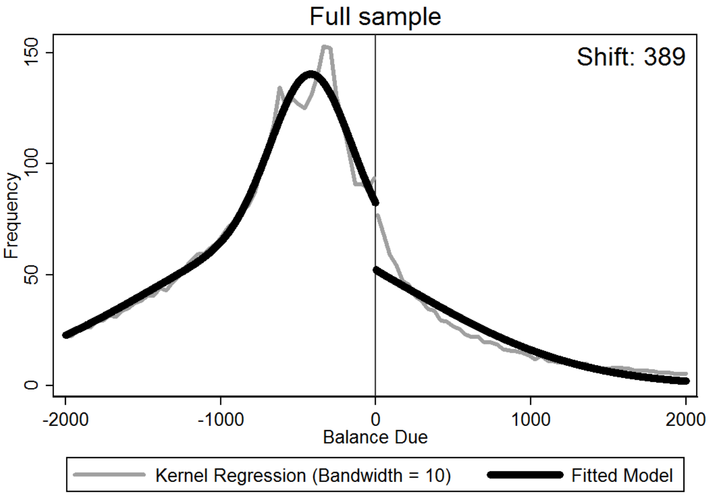
$$\kappa \equiv s^L, \; \tilde{s} \equiv s^H - s^L$$

Empirical approach: Use NLLS to fit a mixture of normal distributions to the histogram, directly modeling shift.

$$C_j = Obs \cdot \left[\sum_{i=1}^{2} \frac{p_i}{\sigma_i} \phi \left(\frac{b_j + \tilde{s} \cdot I(b > 0) - \mu}{\sigma_i} \right) \right] + \epsilon_j$$

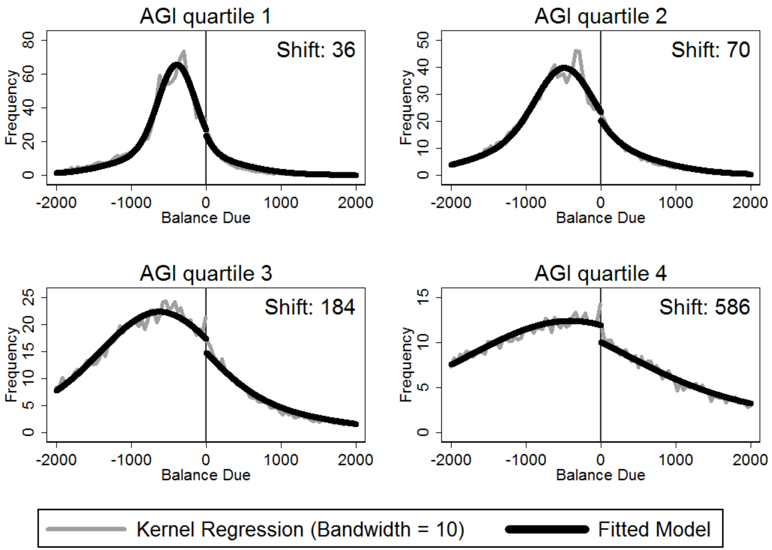
- Common mean assumed to preserve symmetry.
- Similar estimates generated by fitting skew-normal distribution, but fit is worse.

Fit of predicted distributions



Estimate table

Fit of predicted distributions



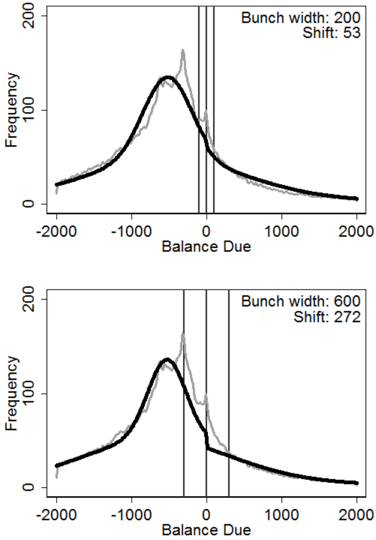
Rationalizing differences in magnitudes

What drives the differences in the bunching and shifting estimates?

Primary explanation: assumption that sheltering can be manipulated to-the-dollar.

- Possible for some types of sheltering: e.g. direct evasion, choosing amount to give to charity, targeted capital losses.
- Not possible for many types of sheltering.
- Excess mass at zero will “leave out” individuals without finely manipulable sheltering technologies.
- Potential solution: permit diffuse bunching “near” zero.

Fit of predicted distributions

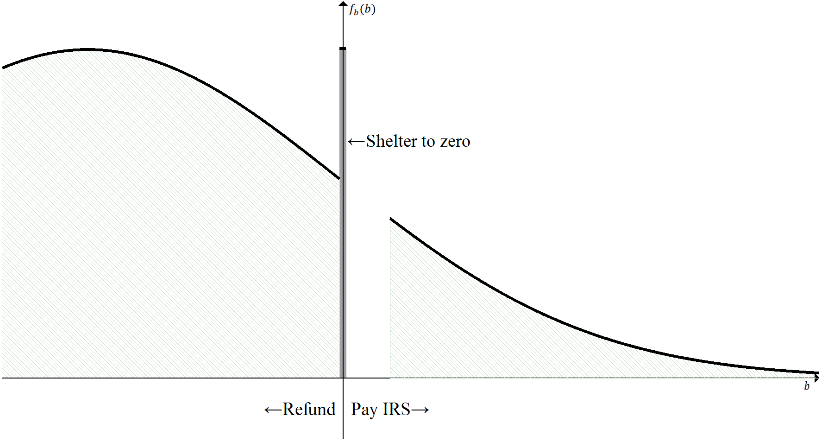


Sheltering-relevant behaviors at zero balance due

	(1)	(2)	(3)	(4)	(5)	(6)
	Adjustments		*Itemized Deduction*		*Credits*	
	> 0	Amount	> 0	Amount	> 0	Amount
Balance due $= 0$	0.09***	1138.38*	0.01	2015.49*	0.01	535.50
	(0.03)	(619.59)	(0.03)	(1112.42)	(0.03)	(493.06)
Balance due > 0	0.05***	259.35***	-0.00	429.42***	-0.01***	27.97
	(0.00)	(76.24)	(0.00)	(99.31)	(0.00)	(29.76)
Filing-year fixed effects	X	X	X	X	X	X
Balance-due polynomial	X	X	X	X	X	X
Lagged-AGI polynomial	X	X	X	X	X	X
N	148325	33935	148325	62441	148325	54223

Notes: OLS regressions with standard errors clustered at the individual level. Monetary quantities expressed in 1990 dollars. Xs indicate the presence of filing-year fixed effects, a third-order polynomial in lagged AGI, or a third-order polynomial in balance due interacted with I(balance due > 0) to allow for discontinuity at zero. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Distribution with fixed cost in loss domain

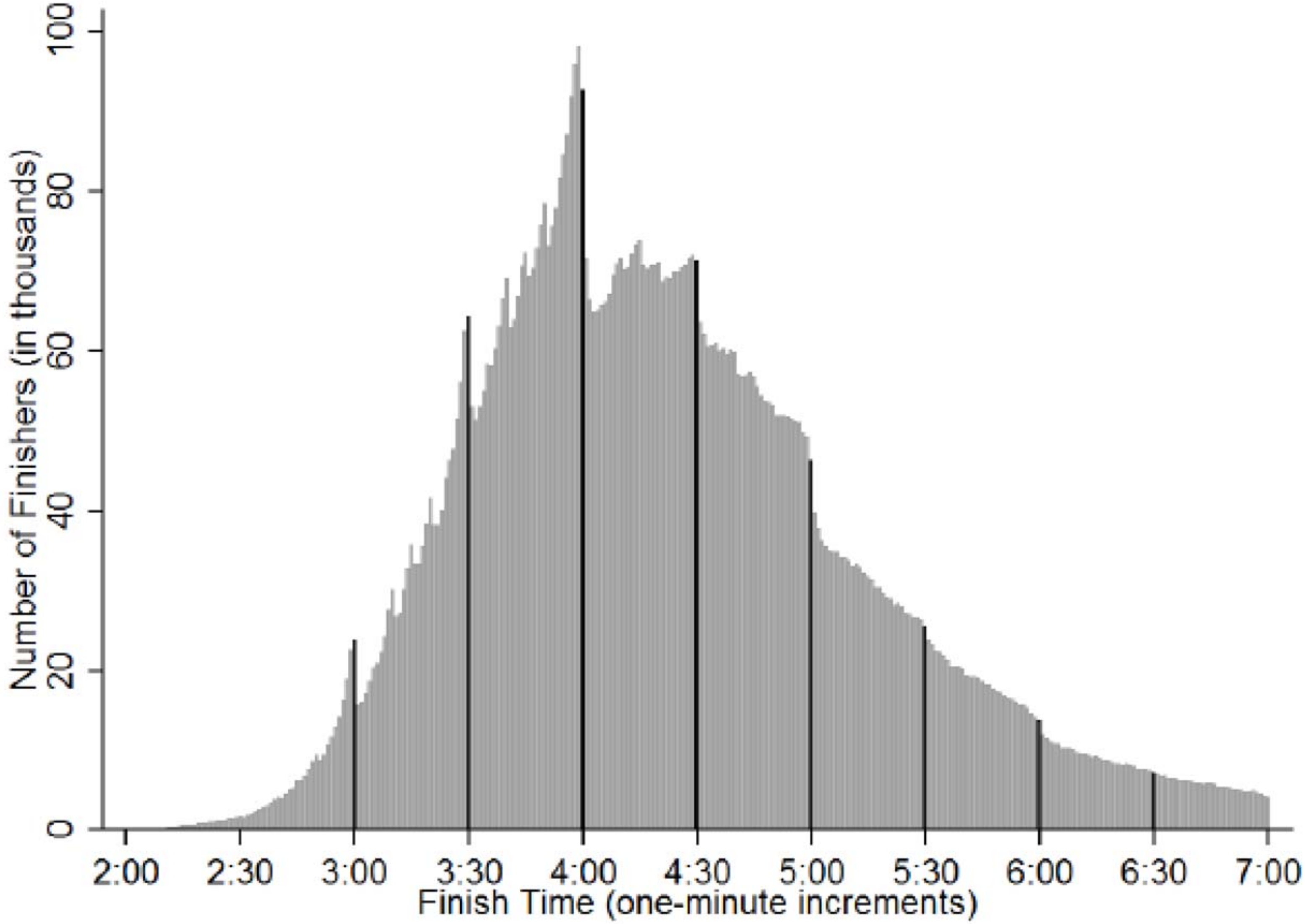


Back

8 Reference Dependence: Goal Setting

- **Allen, Dechow, Pope, Wu (2014)**
- Reference point can be a goal
- Marathon running: Round numbers as goals
- Similar identification considering discontinuities in finishing times around round numbers

Figure 2: Distribution of marathon finishing times ($n = 9,378,546$)

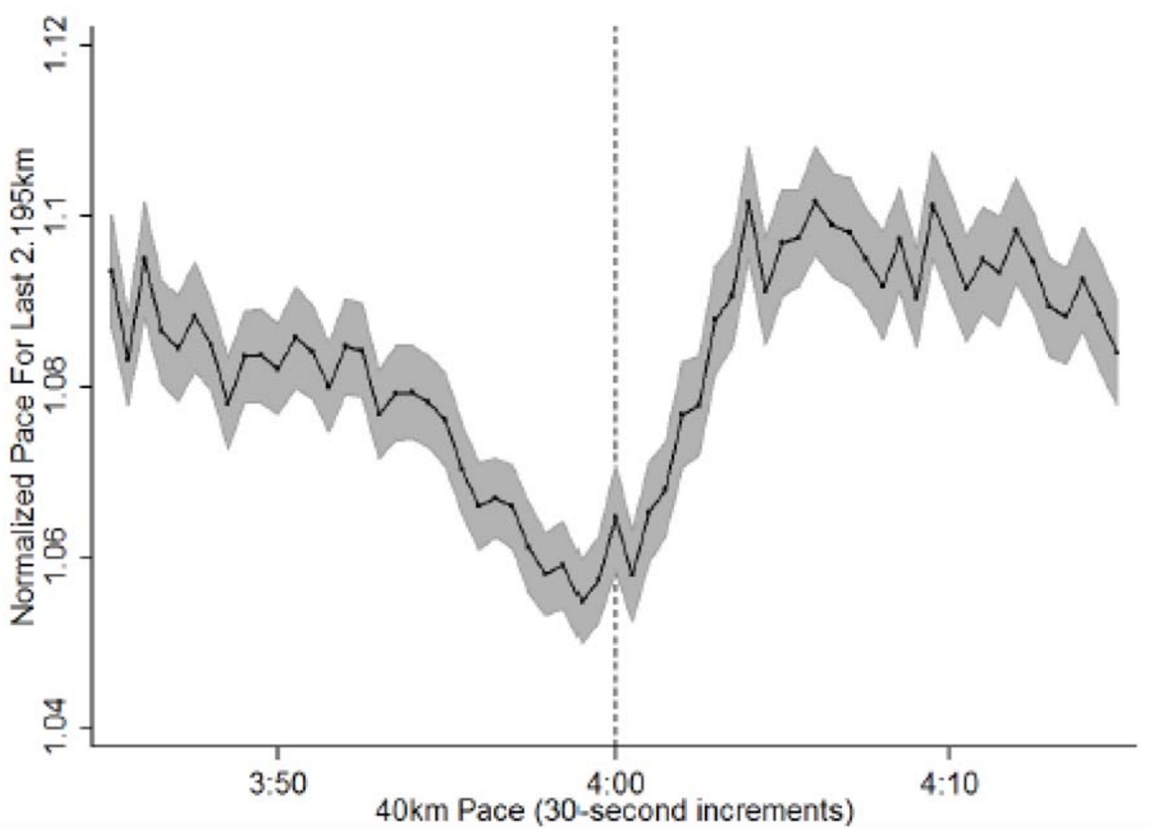


NOTE: The dark bars highlight the density in the minute bin just prior to each 30 minute threshold.

- Channel of effects: Speeding up if behind and can still make goal

Figure 8: Normalized pace for last 2.195 kilometers as a function of 40 kilometer pace

(a) Runners on 3:45 to 4:15 pace through 40 kilometers



- Evidence strongly consistent with model
  - Missing distribution to the right
  - Some bunching
- Hard to back out loss aversion given unobservable cost of effort

9 Reference Dependence: Mergers

- Baker, Pan, Wurgler (*JF* 2012)

- On the appearance, very different set-up:

  - Firm A (Acquirer)

  - Firm T (Target)

- After negotiation, Firm A announces a price P for merger with Firm T

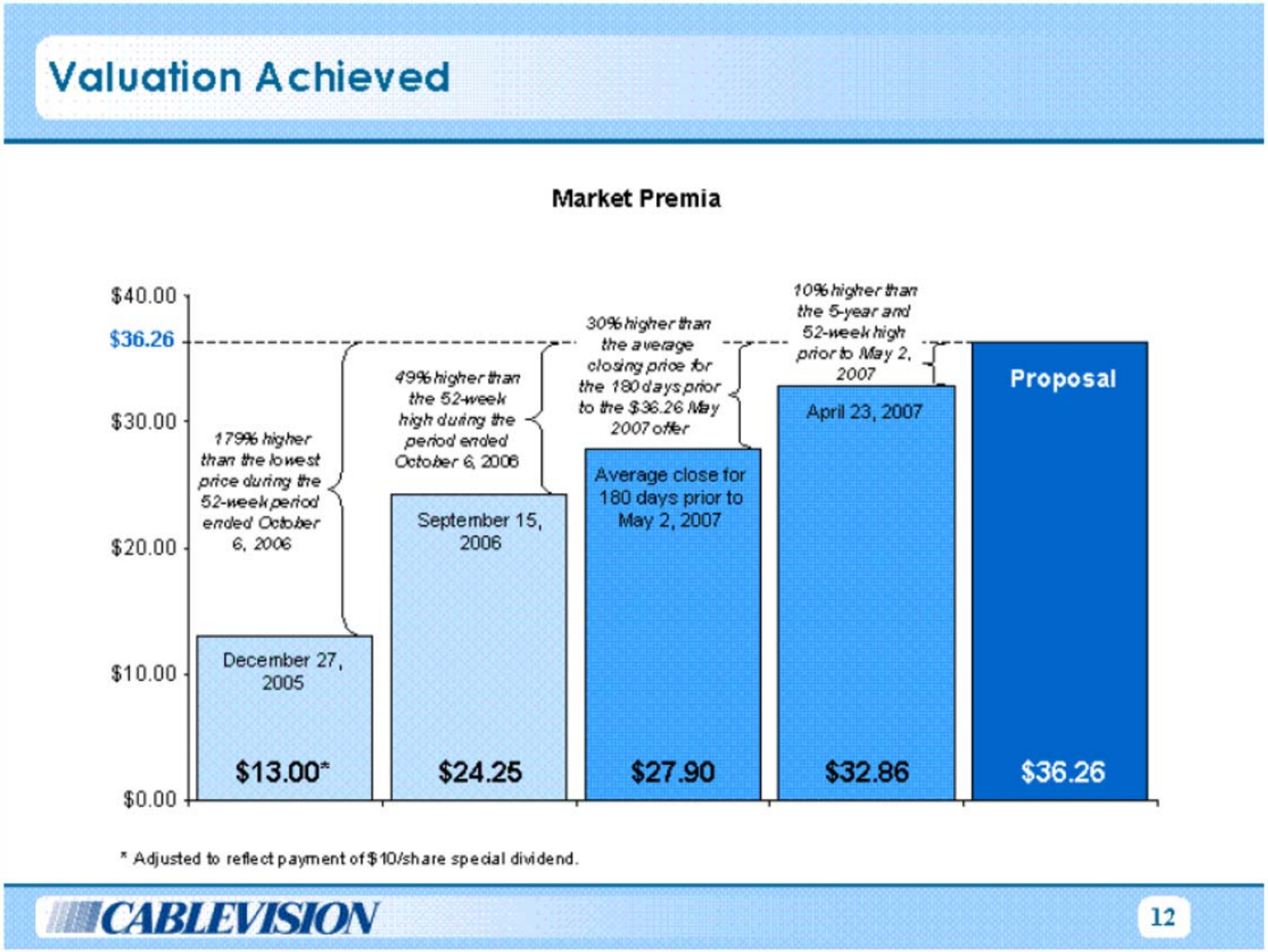
  - Price P typically at a 20-50 percent premium over current price

  - About 70 percent of mergers go through at price proposed

  - Comparison price for P often used is highest price in previous 52 weeks, P_{52}

  - Example of how Cablevision (Target) trumpets deal

Figure 1. Slide from Cablevision Presentation to Shareholders, October 24, 2007. The management of Cablevision recommended acceptance of a $36.26 per share cash bid from the Dolan family. The slide compares this bid price to various recent prices including 52-week highs.



- Assume that Firm T chooses price P, and A decides accept reject
- As a function of price P, probability $p(P)$ that deal is accepted (depends on perception of values of synergy of A)
- If deal rejected, go back to outside value $\bar{U}$
- Then maximization problem is same as for housing sale:

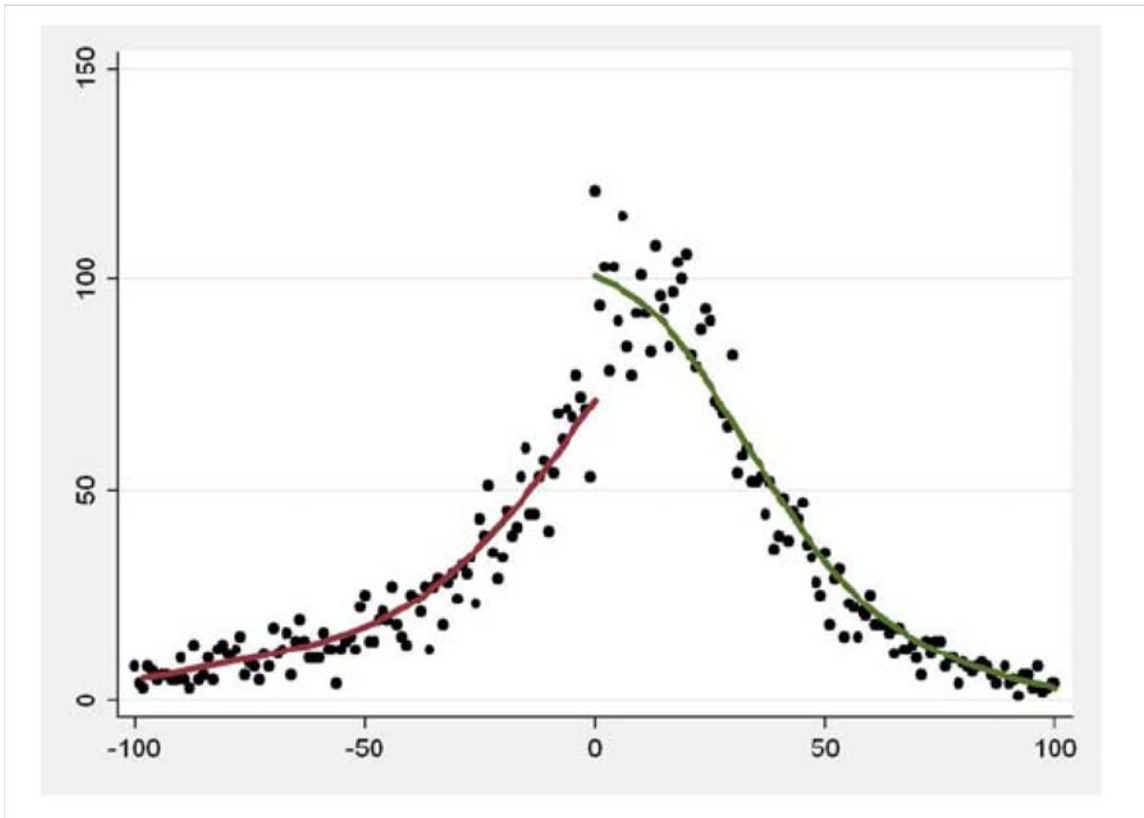
$$\max_P p(P)U(P) + (1 - p(P))\bar{U}$$

- Can assume T reference-dependent with respect to

$$v(P|P_0) = \begin{cases} P - P_{52} & \text{if } P \geq P_{52}; \\ \lambda(P - P_{52}) & \text{if } P < P_{52}, \end{cases}$$

- Obtain same predictions as in housing market
- (This neglects possible reference dependence of A)
- Baker, Pan, and Wurgler (2009): Test reference dependence in mergers
  - Test 1: Is there bunching around P_{52}? (GM did not do this)
  - Test 2: Is there effect of P_{52} on price offered?
  - Test 3: Is there effect on probability of acceptance?
  - Test 4: What do investors think? Use returns at announcement

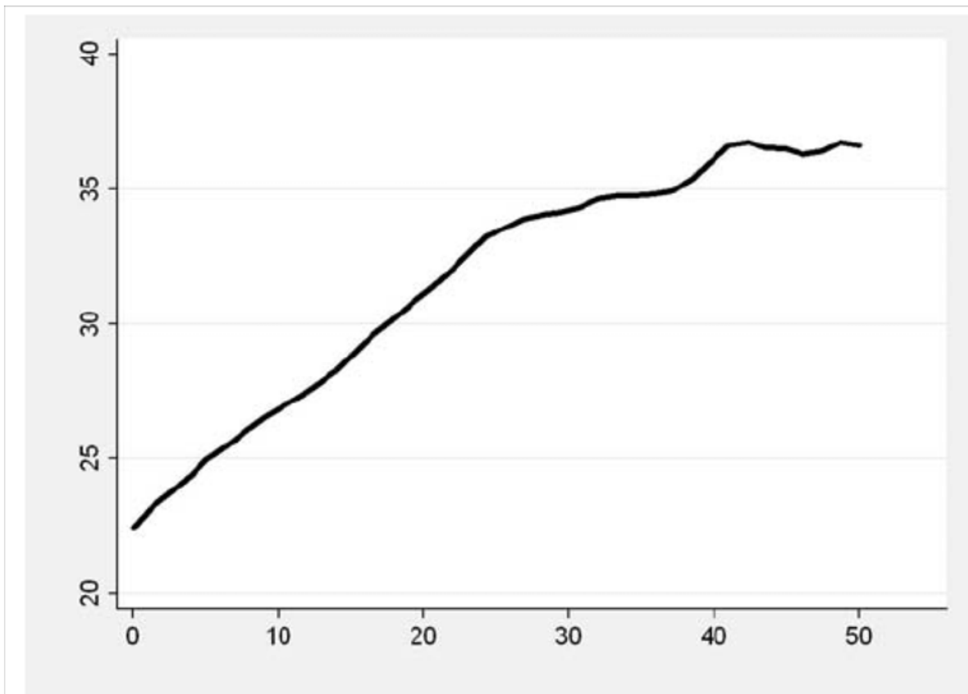
- Test 1: Offer price P around P_{52}
  - Some bunching, missing left tail of distribution



- Notice that this does not tell us how the missing left tail occurs:
  - Firms in left tail raise price to P_{52}?
  - Firms in left tail wait for merger until 12 months after past peak, so P_{52} is higher?
  - Preliminary negotiations break down for firms in left tail
- Would be useful to compare characteristics of firms to right and left of P_{52}

- Test 2: Kernel regression of price offered P (Renormalized by price 30 days before, P_{-30}, to avoid heterosked.) on P_{52} :

$$100 * \frac{P - P_{-30}}{P_{-30}} = \alpha + \beta \left[100 * \frac{P_{52} - P_{-30}}{P_{-30}} \right] + \varepsilon$$



- Test 3: Probability of final acquisition is higher when offer price is above P_{52} (Skip)
- Test 4: What do investors think of the effect of P_{52}?
  - Holding constant current price, investors should think that the higher P_{52}, the more expensive the Target is to acquire
  - Standard methodology to examine this:
    * 3-day stock returns around merger announcement: $CAR_{t-1,t+1}$
    * This assumes investor rationality
    * Notice that merger announcements are typically kept top secret until last minute –> On announcement day, often big impact

- Regression (Columns 3 and 5):

$$CAR_{t-1,t+1} = \alpha + \beta \frac{P}{P_{-30}} + \varepsilon$$

where P/P_{-30} is instrumented with P_{52}/P_{-30}

Table 8. Mergers and Acquisitions: Market Reaction. Ordinary and two-stage least squares regressions of the 3-day CAR of the bidder on the offer premium.

$$r_{t-1\to t+1} = a + b\frac{Offer_{it}}{P_{i,t-30}} + e_{it}$$

$$\left(\frac{Offer_{it}}{P_{i,t-30}} - 1\right)\cdot 100 = a + b_1 \min\left(\left(\frac{52WeekHigh_{i,t-30}}{P_{i,t-30}} - 1\right)\cdot 100, 25\right) + b_2 \max\left(0, \min\left(\left(\frac{52WeekHigh_{i,t-30}}{P_{i,t-30}} - 1.25\right)\cdot 100, 50\right)\right) + b_3 \max\left(\left(\frac{52WeekHigh_{i,t-30}}{P_{i,t-30}} - 1.75\right)\cdot 100, 0\right) + e_{it}$$

where r is the market-adjusted return of the bidder for the three-day period centered on the announcement date, *Offer* is the offer price from Thomson, P is the target stock price from CRSP, and *52WeekHigh* is the high stock price over the 365 calendar days ending 30 days prior to the announcement date. The first, second, and fourth columns use ordinary least squares. The third and the fifth columns instrument for the offer premium using *52WeekHigh*. Robust t-statistics with standard errors clustered by month are in parentheses.

	OLS	*OLS*	*IV*	*OLS*	*IV*
	1	2	3	4	5
Offer Premium:					
b	-0.0186***	-0.0204***	-0.215***	-0.0443***	-0.253***
	(-2.64)	(-2.74)	(-3.48)	(-4.21)	(-4.39)

- Results very supportive of reference dependence hypothesis – Also alternative anchoring story

10 Next Lecture

- Reference-Dependent Preferences
  - Labor Supply
  - Job Search
  - Finance
- Problem Set 2 due next week