Econ 219B
Psychology and Economics:
Applications
(Lecture 5)

Stefano DellaVigna

February 18, 2004

Outline

1. Welfare Programs (Present Bias)

2. Labor Supply: A Framework

3. Labor Supply: Estimation

4. Labor Supply: Bike Messengers

5. Labor Supply: Final Thoughts

6. Loss Aversion and Experience

1 Welfare Programs (Present Bias)

- Fang, Silverman (2002, 2003)

- Stylized Facts:

  - limited transition from welfare to work

  - (more importantly) large share of mothers staying home and not claiming benefits

- Examines decisions of single mothers with kids. Three states: Welfare (leisure + benefits), Work (wages), Home (leisure)

- Mothers stay home because of one-time social disapproval of claiming benefits

- Naiveté crucial here

Table 2: Transition Matrix, Never-married Women with at Least One Child

Choice (t-1)	Choice (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

of those who chose welfare in period t, 76.7% had chosen welfare in the previous period. Of those who chose work in period $t-1$, 79.3% went on to choose it again in period t. Decisions to remain at home are considerably less persistent. Of those who chose to stay home in period $t-1$, 59.7% chose it again in period t.

6 Results

6.1 Estimates of Θ'

The parameters of the government benefits and fertility functions (Θ'), estimated in the first stage, are presented in Tables 9 and 12 of the appendix, respectively. As has been often noted, there is considerable variation in benefits levels across states. In our sample, the estimated average annual benefit for a mother with two children ranges from \$4,856 (1987 dollars) to \$9,490. Patterns of welfare participation vary with the level of benefits in ways consistent with optimizing behavior. In our sample, residents of the 5 states with the highest benefits spend 56 percent of the period observed on welfare; in the 5 states with the lowest benefits the participation rate is 37 percent.

The estimate of the fertility function's parameters suggests that the probability of an additional birth is decreasing with age and with the number of children. The estimate also indicates that, relative to those who stay home, the probability of an additional birth is lower for workers and higher for those on welfare. We note, however, that our simple exogenous model of subsequent valid in this more realistic model, and that in practice the two discount parameters are separately identified with reasonable precision.

6.3 Parameter Estimates and Simulations

Table 4 presents estimates of the parameters of the model under the assumption that agents are naive. Estimation of the model with sophisticated agents remains in progress. The estimated present-bias factor $\beta = 0.61$ and the estimated standard discount factor $\delta = 0.92$ together imply a one-year ahead discount rate of 78%. Inferential studies such as Hausman (1979), and Warner and Pleeter (2001) estimate (one-year ahead) discount rates ranging from 0 to 89% depending on the characteristics of the individual and intertemporal trade-offs at stake. Experimental studies have estimated this figure to be approximately 40% in an average population.

Table 4: Parameter Estimates, Naïve Agents

		parameter	point estimate	std. error
utility parameters	time discounts	β	0.61	0.33
		δ	0.92	0.05
	net stigma	ϕ	4046.74	1123.81
	home production	e_0	3953.13	545.79
		e_1	370.55	150.52
		e_2	-148.1	56.09
		η	5101.51	522.17
wage & skill parameters	constant	ln(r) + ha0	8.22	0.15
	yrs. of school	α_1	0.037	0.012
	experience	α_2	0.115	0.016
	experience^2	α_3	-0.0064	0.001
	1st yr. exper.	α_4	0.086	0.041
	exper. decay	α_5	0.191	0.091
continuation values	no. children	ω_1	510.04	479.97
	no. children^2	ω_2	-6143.43	1294.87
	experience	ω_3	29.03	43.36
	experience^2	ω_4	107.39	38.16
	welfare lag	ω_5	-5325.95	4066.26
	work lag	ω_6	1147.05	1256.76
variance/ covariance	std. dev. ε_0	$\sigma_{\varepsilon 0}$	3174.12	901.47
	std. dev. ε_1	$\sigma_{\varepsilon 1}$	0.342	0.099
	std. dev. ε_2	$\sigma_{\varepsilon 2}$	5050.12	909.82
	cov(ε_0,ε_2)	$\sigma_{\varepsilon 0 \varepsilon 2}$	-2550.08	674.2
	std. dev. meas err.	σ_{me}	0.272	0.12

N=4487 log likelihood = -3821.45

2 Labor Supply: A Framework

- Camerer et al. (1997), Farber (2003), Fehr and Goette (2002)

- Daily labor supply by cabbies and bike messengers

- Framework:

  – effort e (no. of hours)

  – hourly wage w

  – Returns of effort: $w * e$

  – Cost of effort $c(e) = Ce^2/2$ convex within a day

- Standard model: Agents maximize

$$we - c(e) = we - \frac{Ce^2}{2}$$

- Model with reference dependence:

- Threshold R of earnings agent wantes to achieve

- Loss aversion for outcomes below threshold:

$$U(p) = \begin{cases} we - R & \text{if } we \geq R \\ \lambda(we - R) & \text{if } we < R \end{cases}$$

  with $\lambda > 1$ loss aversion coefficient

- Referent-dependent agent maximizes

$$\begin{array}{ll} we - R - c(e) & \text{if } e \geq R/w \\ \lambda(we - R) - c(e) & \text{if } e < R/w \end{array}$$

- Derivative with respect to e:

$$\begin{array}{ll} w - Ce & \text{if } e \geq R/w \\ \lambda w - Ce & \text{if } e < R/w \end{array}$$

- Three cases.

  1. Case 1 ($\lambda w - CR/w < 0$).

     – Optimum at $e^* = \lambda w/C < R/w$

2. Case 2 ($\lambda w - CR/w > 0 > w - CR/w$).

   – Optimum at $e^* = R/w$

3. Case 3 ($w - CR/w > 0$).

   - Optimum at $e^* = w/C > R/w$

- **Standard theory ($\lambda = 1$).**

- Interior maximum: $e^* = w/C$ (Cases 1 or 3)

- Labor supply

- Combine with labor demand: $e^* = a - bw$, with $a > 0, b > 0$.

- Optimum:

$$L^S = w^*/C = a - bw^* = L^D$$

or

$$w^* = \frac{a}{b + 1/C}$$

and

$$e^* = \frac{a}{bC + 1}$$

- Comparative statics with respect to a (labor demand shock): $a \uparrow$ –> $e^* \uparrow$ and $w^* \uparrow$

- On low-demand days (low w) work less hard

- Save effort for high-demand days

- **Model with reference dependence ($\lambda > 1$)**:
  - Case 1 or 3 still exist
  - BUT: Case 2. Kink at $e^* = R/w$ for $\lambda > 1$

- Labor supply

- Combine with labor demand: $e^* = a - bw$, with $a > 0, b > 0$.

- Consider Case 2

- Optimum:

$$L^S = R/w^* = a - bw^* = L^D$$

and

$$w^* = \frac{a + \sqrt{a^2 + 4Rb}}{2b}$$

- Comparative statics with respect to a (labor demand shock):

  - $a \uparrow -> e^* \uparrow$ and $w^* \uparrow$ (Cases 1 or 3)

  - $a \uparrow -> e^* \downarrow$ and $w^* \uparrow$ (Case 2)

- Case 2: On low-demand days (low w) need to work harder to achieve reference point $R ->$ Work harder

- Opposite prediction to standard theory

- (Neglected negligible wealth effects)

3 Labor Supply: Estimation

3.1 Camerer et al (1997)

- Issues with labor supply estimation in Camerer:
  1. Division bias in regressing hours on log wages
     - IV wage using other workers' wage (Camerer)
     - Hazard regression on hours and total earnings (Farber)

2. Are the authors really capturing demand shock or supply shock?

   – Consider standard model above

   – Increase in C (rain) –> $e^* \downarrow$ and $w^* \uparrow$

   – Negative correlation between e^* and w^*

   – Standard issue with estimating demand and supply function

   – Econometric issue: Shocks to both demand and supply

   – Illustrate: Graddy, Fulton fish market

3. What determines the reference point R?

   – Camerer et al.: *Daily* target of earning

   – Does it depend on form of payment?

   – More generally: Intended good performance over a short-enough time frame that allows for keeping track of progress

     * Cab drivers?

     * Stadium vendors?

     * Education?

     * Charitable contributions?

     * Unemployed people

3.2 Oettinger (1999)

- Stadium vendors participation decision

- No data on within-day effort measure

- Data on supply decision across days

- 127 vendors in 81 games

- Observation of:

  - earnings per match

  - vendor participation

- Standard theory:
  - On low-demand games fewer vendors show up
  - Show up on high-demand days

- Model with reference dependence:
  - Same!
  - If framing over homestands, more refined test

- Results. Table 5:
  - OLS estimates
  - 2SLS estimates
  - Supply and demand issue
  - Is instrument credible?

TABLE 4

ESTIMATES OF REDUCED-FORM LOG EARNINGS EQUATION

	DEFINITION OF ACTIVE STATUS			
	Narrow		Broad	
	(1)	(2)	(3)	(4)
	A. Coefficient Estimates and Standard Errors			
Monday–Thursday day game	−.0565	.1435	−.0492	.1550
	(.0689)	(.0477)	(.0672)	(.0428)
Monday–Thursday night game	−.3058	−.0607	−.3095	−.0645
	(.0517)	(.0455)	(.0548)	(.0465)
Friday (night) game	−.0312	.0480	−.0280	.0463
	(.0582)	(.0406)	(.0594)	(.0406)
Saturday (night) game	.1117	.1152	.1091	.1115
	(.0458)	(.0357)	(.0460)	(.0369)
Promotional date	.1550	.0266	.1702	.0393
	(.0533)	(.0342)	(.0565)	(.0375)
Opponent in first place	.0692	−.0556	.0582	−.0602
	(.0658)	(.0490)	(.0640)	(.0503)

Home team games out of first	−.0404	−.0347	−.0305	−.0260
	(.0248)	(.0150)	(.0220)	(.0132)
Daytime high temperature	.0069	.0047	.0106	.0071
	(.0027)	(.0018)	(.0036)	(.0029)
24-hour rainfall > .25 inch	.1242	.1084	.1247	.1086
	(.0643)	(.0470)	(.0685)	(.0469)
Log of attendance	. . .	.5680	. . .	.5600
		(.0606)		(.0635)
Inverse Mills ratio (selectivity correction)	.1736	.1523	.1051	.0818
	(.0715)	(.0712)	(.0669)	(.0656)
	B. χ^2 Statistic p-Values and Degrees of Freedom			
Individual vendor dummies	<.0001	<.0001	<.0001	<.0001
	[125]	[125]	[126]	[126]
Opponent dummies	<.0001	<.0001	<.0001	.0002
	[12]	[12]	[12]	[12]
Observations	3,579	3,579	3,580	3,580
R^2	.650	.670	.649	.669

NOTE.—The estimated covariance matrix allows for an arbitrary error covariance structure across vendors at any given game but assumes independent errors across games, after allowing for vendor fixed effects. One earnings observation is lost under the narrow definition of active status because there is one vendor who participated at only one game, which took place more than 30 days after the date of hire. All the specifications also include as explanatory variables the log of the number of (other) active vendors, the number of games the opposing team is out of first place, and indicators for the season (before Memorial Day or after Labor Day) and for whether the home team was in first place.

TABLE 5

Estimates of Structural Probit Model for Participation

	Definition of Active Status			
	Narrow		Broad	
	(1)	(2)	(3)	(4)
	A. Coefficient Estimates and Standard Errors			
Predicted log hourly earnings	.7644	.7282	.6125	.6045
	(.1990)	(.2173)	(.1819)	(.1934)
Monday–Thursday day game	−.6815	−.7347	−.6258	−.6897
	(.1716)	(.1404)	(.1612)	(.1494)
Monday–Thursday night game	.1624	.0638	.1869	.0966
	(.1735)	(.1882)	(.1665)	(.1942)
Friday (night) game	.4105	.3842	.3783	.3629
	(.2094)	(.2111)	(.1803)	(.1901)
Saturday (night) game	.2923	.2927	.2739	.2729
	(.1714)	(.1581)	(.1539)	(.1462)
Opponent in first place	. . .	.1203	. . .	.1504
		(.1022)		(.1015)
Home team games out of first	. . .	−.0173	. . .	−.0321
		(.0268)		(.0233)

Daytime high temperature	−.0031	.0078	−.0066	.0002
	(.0054)	(.0042)	(.0060)	(.0039)
24-hour rainfall > .25 inch	−.2690	−.0860	−.2500	−.0853
	(.1288)	(.1087)	(.1239)	(.0883)
	B. χ^2 Statistic p-Values and Degrees of Freedom			
Individual vendor dummies	<.0001	<.0001	<.0001	<.0001
	[123]	[123]	[124]	[124]
Opponent dummies	. . .	<.0001	. . .	<.0001
		[12]		[12]
Vendor demographic indicators × day/time/season dummies	<.0001	<.0001	<.0001	<.0001
	[36]	[36]	[36]	[36]
Sample average elasticity of participation with respect to hourly earnings	.568	.546	.757	.759
Elasticity of participation with respect to hourly earnings at covariate sample means	.511	.480	.677	.675
	(.132)	(.143)	(.199)	(.216)
Observations	6,029	6,029	8,601	8,601
Log likelihood	−2,953.1	−2,907.5	−3,319.9	−3,250.9

Note.—The estimated covariance matrix allows for an arbitrary error covariance structure across vendors at any given game but assumes independent errors across games, after allowing for vendor fixed effects. The sample sizes are slightly smaller than the total number of active observations in tables 1 and 2 because the inclusion of vendor fixed effects eliminates vendors who either always participated or never participated. All the specifications include as explanatory variables indicators for the season (before Memorial Day or after Labor Day). The specifications in cols. 2 and 4 also include as explanatory variables the number of games the opposing team is out of first place and an indicator for whether the home team is in first place.

TABLE 6

Estimates of the Aggregate Participation Model

Dependent Variable: Log of Aggregate Participation

	OLS		2SLS		
	(1)	(2)	(3)	(4)	(5)
	Coefficient Estimates and Standard Errors				
Log of average hourly earnings of participating vendors	.2378	.0858	.5346	.6209	.6457
	(.0986)	(.1107)	(.1508)	(.1525)	(.2064)
Monday–Thursday day game	−.3764	−.4024	−.3640	−.3604	−.3997
	(.0650)	(.0596)	(.0692)	(.0718)	(.0724)
Monday–Thursday night game	.0870	−.0086	.1838	.2120	.1587
	(.0594)	(.0580)	(.0723)	(.0742)	(.0847)
Friday (night) game	.1772	.1515	.2040	.2118	.2114
	(.0586)	(.0514)	(.0630)	(.0653)	(.0646)
Saturday (night) game	.0735	.0841	.0408	.0312	.0286
	(.0587)	(.0508)	(.0635)	(.0657)	(.0636)
Opponent in first place	. . .	.0410	. . .	. . .	.0272
		(.0613)			(.0745)
Home team games out of first	. . .	−.0586	. . .	. . .	−.0313
		(.0212)			(.0269)
Daytime high temperature	.0008	.0057	−.0002	−.0005	.0041
	(.0028)	(.0028)	(.0029)	(.0031)	(.0034)
24-hour rainfall > .25 inch	−.1080	.0027	−.1520	−.1648	−.0734
	(.0621)	(.0613)	(.0679)	(.0703)	(.0774)
	Included as Controls?				
Opponent indicators	no	yes	no	no	yes
Measures of team quality	no	yes	no	no	yes
	Exclusion Restrictions (Instruments for Log Earnings)				
Promotional date indicator	. . .	. . .	yes	yes	yes
Log attendance	. . .	. . .	no	yes	yes
Opponent indicators	. . .	. . .	yes	no	no
Measures of team quality	. . .	. . .	yes	no	no
	Overidentification Test				
p-value	. . .	. . .	.060	.021	.400
Degrees of freedom	. . .	. . .	16	1	1
	Test of Joint Significance of Instruments in First Stage of Regression				
p-value	. . .	. . .	.0013	<.0001	<.0001
Degrees of freedom	. . .	. . .	17	2	2
Observations	81	81	81	81	81
R^2	.727	.847	.692	.669	.774

Note.—All the specifications also include as explanatory variables the log of the total number of active vendors and indicators for the season (before Memorial Day or after Labor Day).

4 Labor Supply: Bike Messengers

4.1 Fehr and Goette (2002)

- Bike Messengers I

- Slides courtesy of Lorenz Goette

Motivation

- **Important intution from standard economic model**
  - Work more when wages are exceptionally high.
  - Relevant in many applications.
  - Astonishingly little evidence.

  ➢ Data may be bad, or model may be wrong.

- **Focus of this study**
  - Combine evidence from two field experiments with bicycle messengers.

  - Experiment 1: Temporary, but large, wage increase
    - Measure the labor supply responses on hours and effort margins.

  - Experiment 2: Distinguish between different models
    - Standard vs. “Non-Standard” Model

The Experimental Setup in this Study

Bicycle Messengers in Zurich, Switzerland

- Data: Delivery records of Veloblitz and Flash Delivery Services, 1999 - 2000.
    - Contains large number of details on every package delivered.

    ➢ Observe hours (shifts) and effort (revenues per shift).

- Work at the messenger service
    - Messengers are paid a commission rate *w* of their revenues r_{it}. (w = „wage“). Earnings wr_{it}
    - Messengers can freely choose the number of shifts and whether they want to do a delivery, when offered by the dispatcher.

    ➢ suitable setting to test for intertemporal substitution.

- Highly volatile earnings
    - Demand varies strongly between days

    ➢ Familiar with changes in intertemporal incentives.

Experiment 1

- **The Temporary Wage Increase**
  - Messengers were randomly assigned to one of two treatment groups, A or B.
    - N=22 messengers in each group

  - Commission rate w was increased by 25 percent during four weeks
    - Group A: September 2000
      (Control Group: B)
    - Group B: November 2000
      (Control Group: A)

- **Intertemporal Substitution**
  - Wage increase has no (or tiny) income effect.
  - Prediction with time-separable prefernces, t= a day:
    - ➢ Work more shifts
    - ➢ Work harder to obtain higher revenues

  - Comparison between TG and CG during the experiment.
    - Comparison of TG over time confuses two effects.

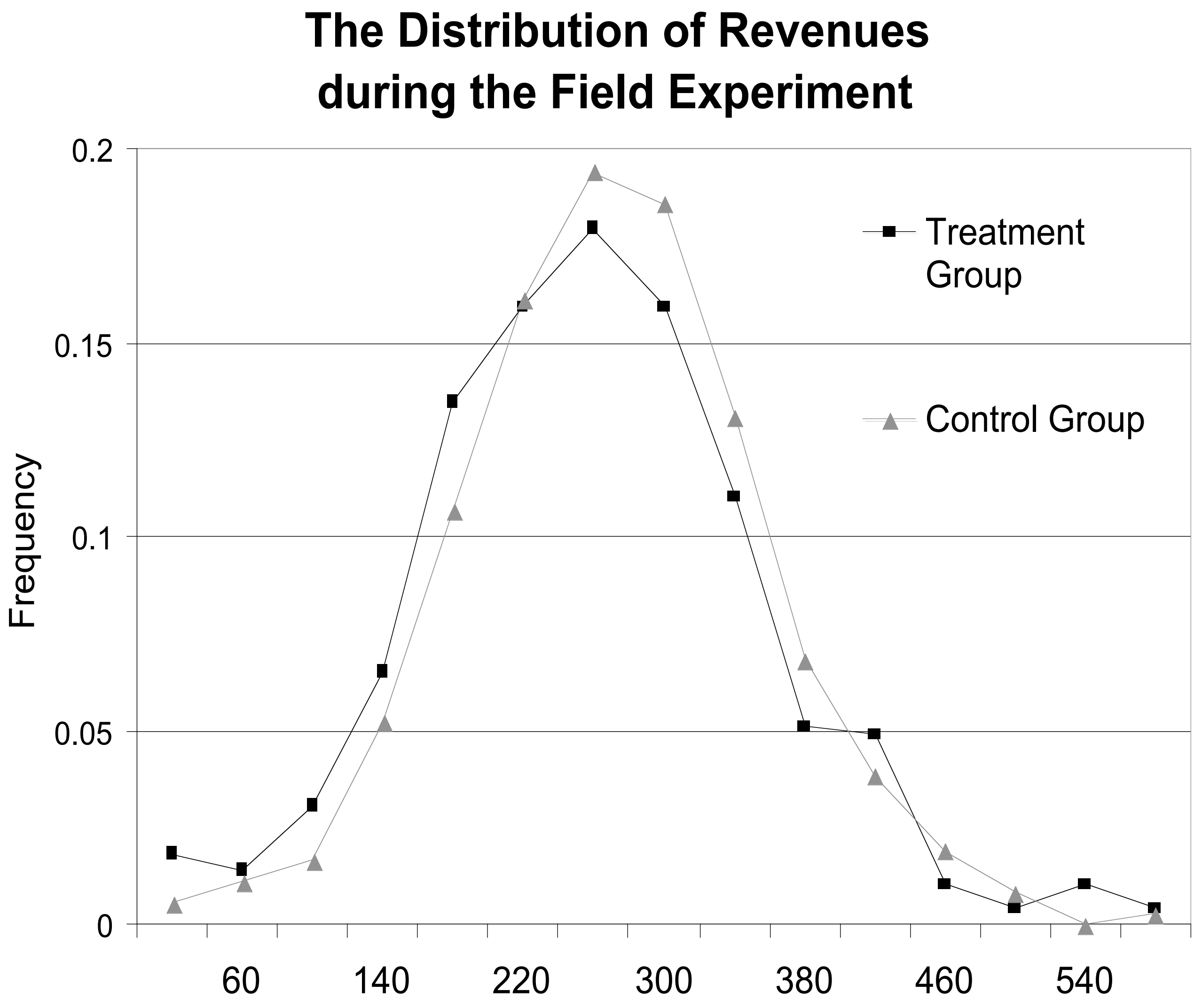
Results for Hours

- Treatment group works 12 shifts, Control Group works 9 shifts during the four weeks.
- Treatment Group works significantly more shifts ($X^2(1) = 4.57, p<0.05$)
- Implied Elasticity: 0.8

Figure 6: The Working Hazard during the Experiment

Results for Effort: Revenues per shift

- Treatment Group has lower revenues than Control Group: - 6 percent. (*$t = 2.338$, $p < 0.05$*)
- Implied *negative* Elasticity: -0.25



- Distributions are significantly different (KS test; *$p < 0.05$*);

Results for Effort, cont.

- **Important caveat**
  - Do lower revenues relative to control group reflect lower effort or something else?

- **Potential Problem: Selectivity**
  - Example: Experiment induces TG to work on bad days.
  - More generally: Experiment induces TG to work on days with unfavorable states
    - ➢ If unfavorable states raise marginal disutility of work, TG may have lower revenues during field experiment than CG.

- **Correction for Selectivity**
  - Observables that affect marginal disutility of work.
    - ➢ Conditioning on experience profile, messenger fixed effects, daily fixed effects, dummies for previous work **leave result unchanged.**
  - Unobservables that affect marginal disutility of work?
    - Implies that reduction in revenues only stems from sign-up shifts in addition to fixed shifts.
    - ➢ **Significantly lower revenues on fixed shifts, not even different from sign-up shifts.**

Corrections for Selectivity

- **Comparison TG vs. CG without controls**
  - Revenues 6 % lower (s.e.: 2.5%)

- **Controls for daily fixed effects, experience profile, workload during week, gender**
  - Revenues are 7.3 % lower (s.e.: 2 %)

- **+ messenger fixed effects**
  - Revenues are 5.8 % lower (s.e.: 2%)

- **Distinguishing between fixed and sign-up shifts**
  - Revenues are 6.8 percent lower on fixed shifts (s.e.: 2 %)
  - Revenues are 9.4 percent lower on sign-up shifts (s.e.: 5 %)

➢ **Conclusion: Messengers put in less effort**
  - Not due to selectivity.

Measuring Loss Aversion

- **A potential explanation for the results**
  - Messengers have a daily income target in mind
  - They are loss averse around it
  - Wage increase makes it easier to reach income target

  ➢ That's why they put in less effort per shift

- **Experiment 2: Measuring Loss Aversion**
  - Lottery A: Win CHF 8, lose CHF 5 with probability 0.5.
    - 46 % accept the lottery
  - Lottery C: Win CHF 5, lose zero with probability 0.5; or take CHF 2 for sure
    - 72 % accept the lottery
  - Large Literature: Rejection is related to loss aversion.

- **Exploit individual differences in Loss Aversion**
  - Behavior in lotteries used as proxy for loss aversion.

  ➢ Does the proxy predict reduction in effort during experimental wage increase?

Measuring Loss Aversion

- **Does measure of Loss Aversion predict reduction in effort?**
  - Strongly loss averse messengers reduce effort substantially: Revenues are 11 % lower (s.e.: 3 %)
  - Weakly loss averse messenger do not reduce effort noticeably: Revenues are 4 % lower (s.e. 8 %).
  - No difference in the number of shifts worked.

➢ **Strongly loss averse messengers put in less effort while on higher commission rate**

- Supports model with daily income target

➢ **Others kept working at normal pace, consistent with standard economic model**

- Shows that not everybody is prone to this judgment bias (but many are)

Concluding Remarks

- **Our evidence does not show that intertemporal substitution in unimportant.**
  - Messenger work more shifts during Experiment 1
  - But they also put in less effort during each shift.

- **Consistent with two competing explanantions**

  - Preferences to spread out workload
    - ➢ But fails to explain results in Experiment 2

  - Daily income target and Loss Aversion
    - ➢ Consistent with Experiment 1 and Experiment 2

    - ➢ Measure of Loss Aversion from Experiment 2 predicts reduction in effort in Experiment 1

    - ➢ Weakly loss averse subjects behave consistently with simplest standard economic model.

    - ➢ Consistent with results from many other studies.

4.2 Goette and Huffman (2004)

- Bike Messengers II

- Data on within-day revenue of bike messengers

- Firms in San Francisco (1998-2003) and Basel (Sw., 2001-2003)

- Pay by commission on revenue

- Increase of commission by 5% after 12-14 weeks

- Effect of increased wage on revenue (effort) within the day

- Hourly revenue increases early in the day, but *decreases* later, when (presumably) target is met

- Consistent with reference dependence

Figure 2a

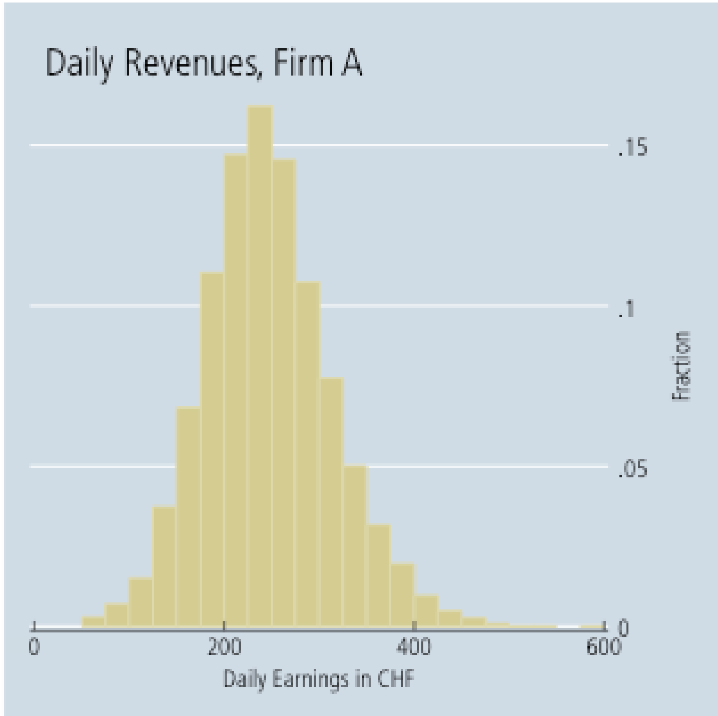


Figure 2b

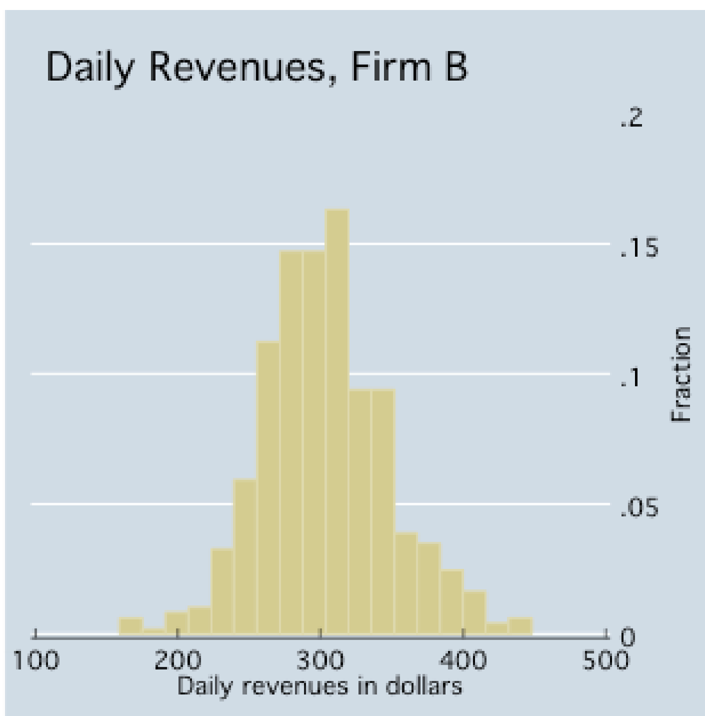


Figure 3a: The Impact of a 5 Percentage Point Increase in the Commission Rate on Messengers' Hourly Revenues, Firm A
(+/- 2s.e. of estimate, adjusted for clustering on days)

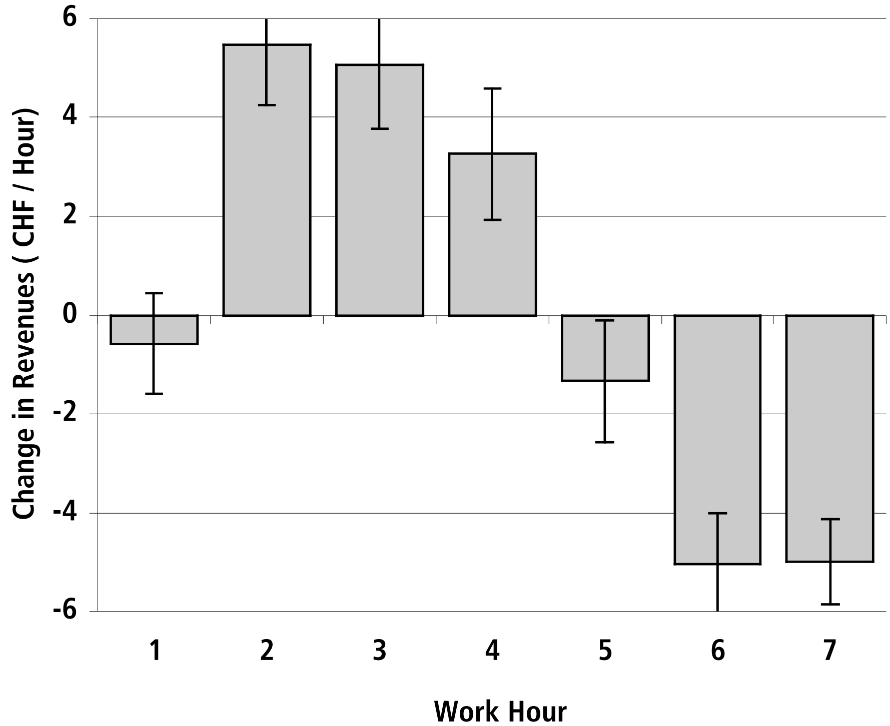


Figure 3b: The Impact of a 5 Percentage Point Increase in the Commission Rate on Messengers' Hourly Revenues, Firm B
(+/- 2s.e. of estimate, adjusted for clustering on days)

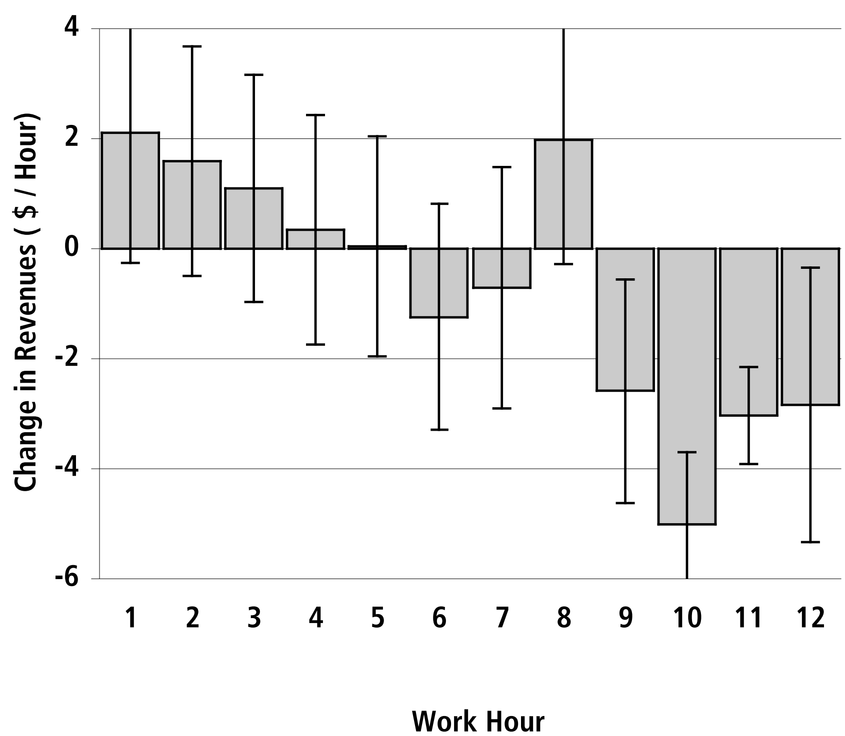


**Figure 5: Experience and Revenues per Hour:
Regression Estimates of impact on hourly revenues (Regression from Table 4, column 3, +/- s.e. of estimate, adjusted for clustering on day)**

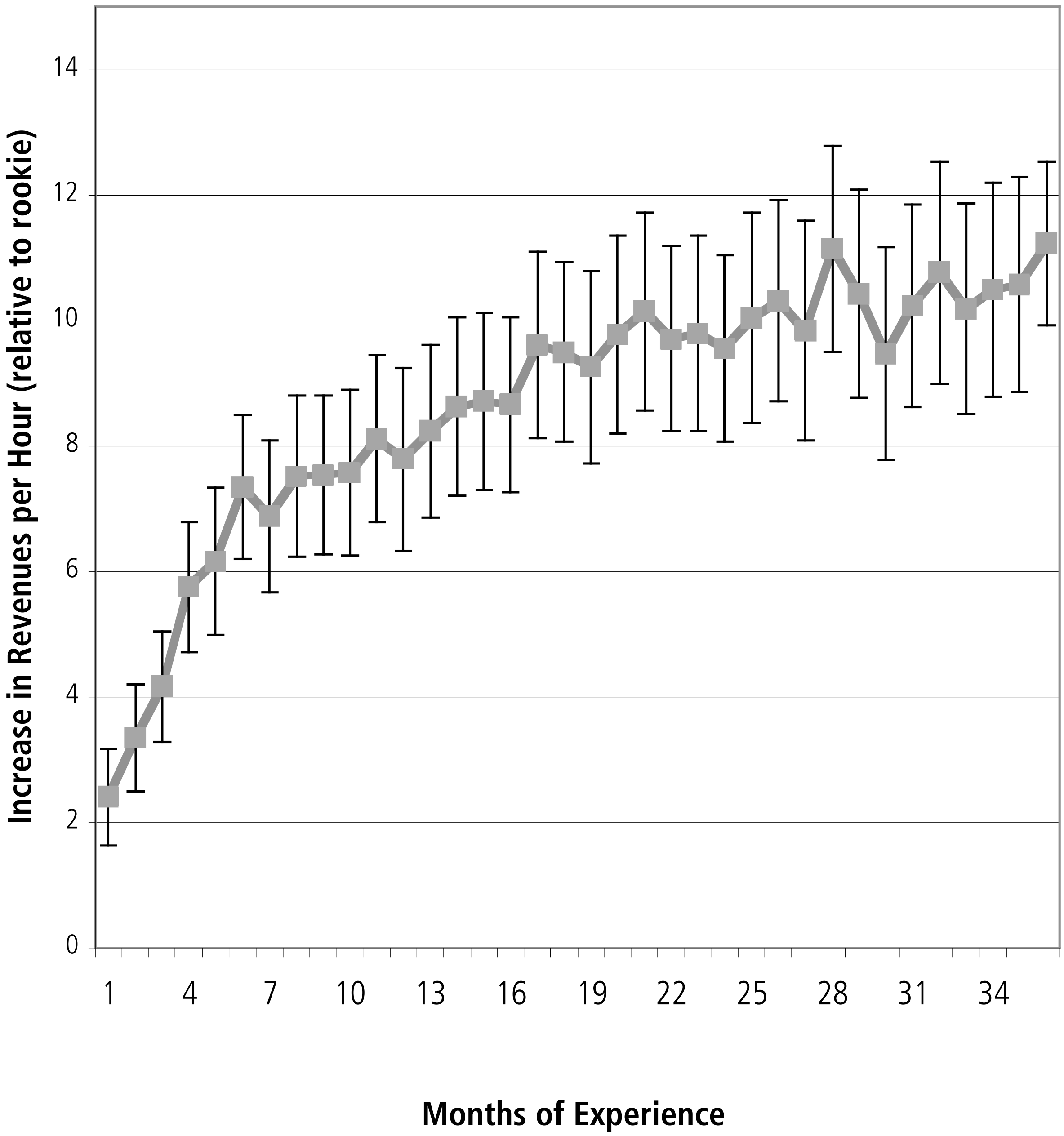
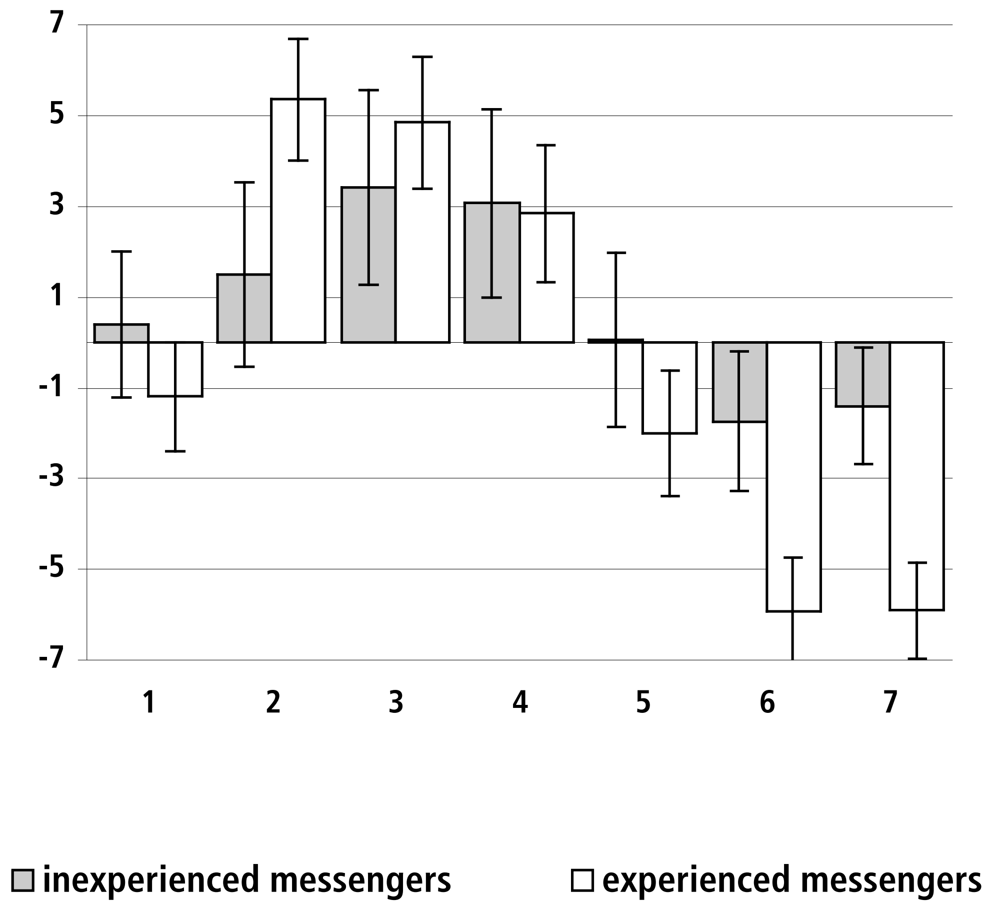


Figure 6: The Impact of a 5 Percentage Point Increase in the Commission Rate on Messengers' Hourly Revenues, Inexperienced vs. Experienced Messengers, Firm A

(+/- s.e. of estimate, adjusted for clustering on days)



5 Labor Supply: Final Thoughts

- What identifies the three papers?

- Cab Drivers: Shocks in daily earnings
  - S? D?
  - Have to hope it comes through Demand for cabs

- Stadium vendors: instruments for Demand
  - attendance, quality of opponent
  - May affect also Supply?
  - First cut: higher attendance –>higher supply–> downward bias

- But: could go other way – prefer to watch game at home

- Bike Messengers: Exogenous variation in prices
  - Randomized variation (paper I)
  - Natural experiment on piece rate (paper II)
  - Control for total supply of messengers

- Do test on homestand
  - Take one-week homestand with 2 teams
  - Assume both teams are bad
  - Higher participation of vendors because of loss aversion

- Marginal and inframarginal cab drivers
  - Drivers A work every day
  - Drivers B work only occasionally and part-time
  - Low-wage days: Only drivers A work
  - High-wage days: Drivers A and B work
  - Data on participation?

6 Loss Aversion and Experience

- Important open issue: effect of stakes and experience on biases

- Effect of experience in previous papers:

  - Camerer et al.: Experienced agents are less likely to exhibit loss aversion

  - Table: Unstable coefficients on expeirence in cab paper

  - Do experience tas drivers make more money overall? (level effect)

  - In 401(k) investment experience decreases effect of default

6.1 List (QJE, 2003)

- Experience and Market Anomialies

- Field experiment:

  - hybrid of lab experiment and field data studies

  - Used for loss aversion, auctions, charitable giving

  - Take advantage of your interests/passions!

- Experiments I, II

- Protocol:
  - Get people to fill survey
  - Hand them memorabilia card A (B) as thank-you gift
  - After survey, show them memorabilia card B (A)
  - "Do you want to switch?"
  - "Are you going to keep the object?"
  - Experiments I, II with different object
- Prediction of Endowment effect: too little trade

- Objection 1: Is it experience or is it just sorting?

- Experiment III with follow-up of experiment I

- People with serious endowment effect do not go to markets

- But: Last Table

- Objection 2. Are inexperienced people indifferent between different cards?

- People do not know own preferences

- People have never seen these cards

- Last Table

- Objection 3. What are people learning about?

- Getting rid of loss-aversion?

- Learning better value of cards?

- If do not know value, adopt salesman technique

- Is learning localized or do people generalize the learning to other goods?

TABLE I
SELECTED CHARACTERISTICS OF PARTICIPANTS

	Sportscard market I		Pin market	Sportscard market II
	Dealers mean (std. dev.)	Nondealers mean (std. dev.)	Consumers mean (std. dev.)	Nondealers mean (std. dev.)
Trading experience	14.82	5.66	6.98	6.84
	(11.0)	(6.42)	(13.63)	(7.98)
Years of market experience	10.36	6.95	5.05	7.13
	(6.75)	(9.37)	(5.64)	(9.05)
Income	4.26	4.04	4.06	4.36
	(1.92)	(2.06)	(2.25)	(1.82)
Age	34.68	34.70	31.48	34.83
	(11.98)	(14.06)	(13.68)	(12.51)
Gender (percent male)	0.93	0.86	0.48	0.89
	(0.25)	(0.34)	(0.50)	(0.32)
Education	3.42	3.84	3.10	3.85
	(1.42)	(1.49)	(1.53)	(1.50)
Good B	0.527	0.527		—
	(0.50)	(0.50)		
Good D	—	—	0.50	—
			(0.50)	
Good F	—	—	—	0.53
				(0.50)
N	74	74	80	53

a. *Trading experience* represents the number of trades made in a typical month.
b. *Years of market experience* denotes years that the subject has been active in the market.
c. *Income* denotes categorical variable (1–8): 1) Less than $10,000, 2) $10,000 to $19,999, 3) $20,000 to $29,999, 4) $30,000 to $39,999, 5) $40,000 to $49,999, 6) $50,000 to $74,999, 7) $75,000 to $99,999, 8) $100,000 or over.
d. *Age* denotes actual age in years.
e. *Gender* denotes categorical variable: 0 if female, 1 if male.
f. *Education* denotes categorical variable 1) Eighth grade or less, 2) High School, 3) 2-Year College, 4) Other Post-High School, 5) 4-Year College, 6) Graduate School Education.
g. *Good B (D) (F)* denotes the subject's initial endowment, and =1 if the subject was endowed with *Good B (D) (F),* 0 otherwise.

suggest undertrading occurred. Given that subjects were randomly allocated either good A or good B, equivalence of WTA and WTP would imply that approximately half of the goods were improperly allocated and should be traded. The actual percentages of subjects who chose to trade are 32.8 percent (23 of 70) and 34.6 percent (27 of 78), suggesting that $WTA > WTP$. These figures suggest that once endowed with one of the goods the subjects were close to two times more likely to select that good (computed as ½ $((P_{A|A}/P_{A|B}) + (P_{B|B}/P_{B|A}))$.

TABLE II
SUMMARY TRADING STATISTICS FOR EXPERIMENT I: SPORTSCARD SHOW

Variable	Percent traded	p-value for Fisher's exact test
Pooled sample (n = 148)		
Good A for Good B	32.8	<0.001
Good B for Good A	34.6	
Dealers (n = 74)		
Good A for Good B	45.7	0.194
Good B for Good A	43.6	
Nondealers (n = 74)		
Good A for Good B	20.0	<0.001
Good B for Good A	25.6	

a. Good A is a Cal Ripken, Jr. game ticket stub, circa 1996. Good B is a Nolan Ryan certificate, circa 1990.
b. Fisher's exact test has a null hypothesis of no endowment effect.

Although these results are suggestive, they may be an artifact of the sampling procedure—by chance subjects who preferred good A (good B) may have been endowed with good A (good B), leading to false inference. To amend this situation, I test the null hypothesis of no endowment effect by using a Fisher's exact test, which has a hypergeometric distribution under the null. The result of the exact test presented in row 1, column 2 of Table II, strongly suggests that the null hypothesis should be rejected ($p <$ .001) for the pooled sample, implying that an endowment effect exists. This evidence, which is consistent with past experimental studies, is at odds with conventional economic theory, which assumes that indifference curves are completely reversible when transactions costs are zero [Knetsch 1989].

Panels two and three in Table II present split subsamples and tell an intuitive story consistent with the research hypothesis—dealers tend to trade more than nondealers, regardless of which good they were initially endowed. For example, whereas 43.6 percent and 45.7 percent of dealers chose to execute a trade, only 20–25 percent of nondealers chose to trade. These proportions suggest that nondealers were nearly 3.5 times more likely to select the good which they were endowed, whereas dealers were only 1.25 times more likely to choose their endowed good. A Fisher's exact test shows that for nondealers the null hypothesis of no endowment effect should be rejected at the $p < .001$ level.

TABLE III
NONDEALER SUMMARY STATISTICS FOR EXPERIMENT I: SPORTSCARD SHOW

Variable	Percent traded	p-value for Fisher's exact test
Experienced nondealers (n = 30)	46.7	0.32
Inexperienced nondealers (n = 44)	6.80	<0.001

a. Experienced nondealers are those consumers who trade 6 or more times per month (5.66 is the mean level of monthly trades for nondealers). Inexperienced nondealers trade less than 6 times per month.
b. Fisher's exact test has a null hypothesis of no endowment effect.

Alternatively, the null hypothesis cannot be rejected at conventional significance levels in the dealer treatments ($p = .19$). This result provides initial evidence that experienced consumers' utility functions may not reflect an endowment effect.

To investigate this finding further, I present Table III, which provides a breakdown of the nondealer data based on the level of trading experience of each subject. I split the sample of experienced and inexperienced nondealers according to the central tendency of the data. Experienced nondealers are those who trade 6 or more times in a typical month, where 6 is a shade above the mean level of monthly trades (5.66). Inexperienced nondealers are those subjects who trade fewer than six times per month. The results are compelling. For experienced nondealers, 14 of 30 (46.7 percent) opted to trade. This figure is very close to the dealers' trading strategy observed above, and using a Fisher's exact test the null hypothesis cannot be rejected at conventional significance levels ($p = 0.32$). For inexperienced nondealers the endowment effect is large: only 6.8 percent (3 of 44) of inexperienced subjects opted to trade, and the hypothesis of no endowment effect is rejected at the $p < 0.001$ level. This latter finding suggests that once inexperienced consumers are endowed with a good, they are thirteen times more likely to keep that good. This average increase in the likelihood that the subject chooses a good once endowed with it is slightly higher than that observed in Knetsch [1989].

Although analysis of the raw data provides evidence that supports the main conjecture of the study, there has been no attempt to control for other factors that may affect the propensity to trade. These other subject-specific variables can be adequately accounted for in a basic econometric model:

$$trade = g(\alpha + \beta' X), \tag{1}$$

TABLE V
SUMMARY TRADING STATISTICS FOR EXPERIMENT II: PIN TRADING STATION

Variable	Percent traded	p-value for Fisher's exact test
Pooled sample (n = 80)		
Good C for Good D	25.0	<0.001
Good D for Good C	32.5	
Inexperienced consumers (<7 trades monthly; n = 60)	25.0	<0.001
Experienced consumers (≥7 trades monthly; n = 20)	40.0	0.26
Inexperienced consumers (<5 trades monthly; n = 50)	18.0	<0.001
Experienced consumers (≥5 trades monthly; n = 30)	46.7	0.30

a. Good C is a cloisonné Valentine's Day pin portraying Mickey and Minnie Mouse, circa 2000. Good D is a cloisonné St Patrick's Day 2000 portraying Mickey Mouse, circa 2000.
b. Experienced consumers are those consumers who trade 7 (or 5) or more times per month (6.55 is the mean level of monthly trades). Inexperienced consumers trade less than 7 (or 5) times per month.
c. Fisher's exact test has a null hypothesis of no endowment effect.

sportscard market: i) as Table V illustrates, an overall endowment effect exists at the $p < .001$ level; but individual behavior converges to the neoclassical prediction as trading experience intensifies (see the bottom two panels in Table V); and ii) the regression results presented in Table VI, which include expansions to the cubic, support these conclusions. Regression estimates also suggest that women tend to trade less than men, but the difference is only marginally significant. This finding may have been absent in the sportscard market because the sample was largely comprised of men. Although gender and the endowment effect appear linked, future research is necessary before any firm conclusions can be reached concerning this relationship.[6]

IV. EXPERIMENTAL DESIGN III

Although both sets of field results are consonant with the notion that neoclassical expectations are met when trading experience intensifies, it remains an open question as to whether

6. To examine whether information asymmetry is driving the results, I ran identical trading exercises using coffee mugs and candy bars on the floor of a sportscard show in Tucson, AZ. I find results consistent with the above findings. These results will be reported elsewhere [List 2002].

TABLE IX
SUMMARY STATISTICS FOR EXPERIMENT III: FOLLOW-UP SPORTSCARD SHOW

	Increased number of trades	Stable number of trades	Decreased number of trades
No trade in Experiment I; trade in Experiment III	13	1	2
No trade in Experiment I; no trade in Experiment III	8	7	11
Trade in Experiment I; Trade in Experiment III	4	0	0
Trade in Experiment I; No trade in Experiment III	2	0	5
N	27	8	18

a. Columns denote changes in subjects' trading experience over the year; rows denote subjects' behavior in the two field trading experiments.
b. Fifty-three subjects participated in both Experiment I and the follow-up experiment.

treatment, what is necessary is a within-person analysis, which by definition controls for individual-specific heterogeneity that is left uncontrolled in a cross-sectional analysis.

A first straightforward test of whether experience and trading activity are positively associated within subjects is to examine individual trading rates over time. Table IX summarizes the four possible outcomes across three trading dimensions. The raw data show that over the course of the year, many subjects experienced a growth in their personal number of trades: 27 subjects (51 percent) increased their monthly trading rate, whereas 18 (34 percent) and 8 (15 percent) subjects decreased or had flat trading rates compared with the previous year. This result suggests that a slight majority of subjects gained trading experience over the year. At a superficial level, this is weak evidence in favor of the research hypothesis.

A closer examination of the data in Table IX suggests that 42 of 53 (79.2 percent) subjects did not execute a trade in the initial experiment (summation of rows 1 and 2). Of those 42 subjects, data in rows 1 and 2 of Table IX indicate that 21, 13, and 8 reported an increase, decrease, and no change in their monthly trading rate compared with the previous year. Of the 21 subjects who increased their trading rate over the year, 13 (62 percent) chose to trade in the follow-up experiment. This percentage compares favorably to the two of thirteen subjects (15.4 percent) or

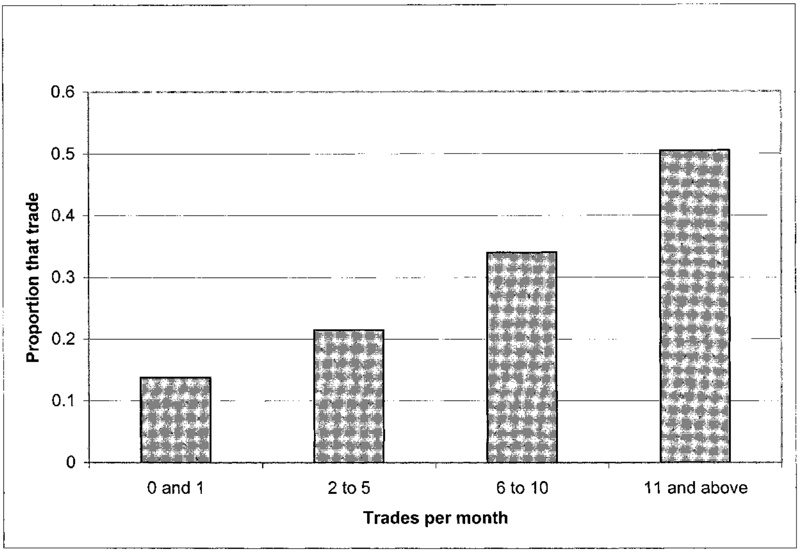


FIGURE I
Summary of Trading Results

panel data logit models a clear result is that a significant relationship exists between trading experience and the probability of executing a trade, but diminishing returns are again evident.

V. Further Experimental Evidence

As previously mentioned, in the follow-up sportscard field experiment I also obtained data from nineteen dealers. The endowment effect can again be rejected in these data, as ten of the nineteen dealers (52.6 percent) chose to trade their endowed good. Overall, therefore, I find a substantial amount of evidence that suggests individual behavior converges to the neoclassical prediction as trading experience intensifies. This major insight is perhaps best illustrated in Figure I, which pools the data across the three field trading treatments—a total of 300 subjects. Figure I, which makes the trade probability a function of previous trading experience, clearly illustrates that individual behavior converges to the neoclassical prediction as consumers gain experience.

V. A. Statements of Value in Auctions

A well-known experimental result is that institutions influence behavior; thus, a test of whether experience influences the

TABLE XI
SELECTED CHARACTERISTICS OF TUCSON SPORTSCARD PARTICIPANTS

	Dealers		Nondealers	
	WTA mean (std. dev.)	WTP mean (std. dev.)	WTA mean (std. dev.)	WTP mean (std. dev.)
Bid or offer	8.15	6.27	18.53	3.32
	(9.66)	(6.90)	(19.96)	(3.02)
Trading experience	16.67	15.78	4.00	3.73
	(19.88)	(13.71)	(5.72)	(3.46)
Years of market experience	10.23	10.57	5.97	5.60
	(5.61)	(8.13)	(5.87)	(6.70)
Income	3.46	3.40	3.37	3.40
	(2.17)	(2.03)	(2.14)	(2.24)
Age	29.20	31.00	28.40	29.00
	(12.20)	(14.70)	(14.90)	(15.30)
Gender (percent male)	0.87	0.90	0.90	0.90
	(0.35)	(0.31)	(0.31)	(0.31)
Education	3.36	3.40	3.03	3.23
	(1.77)	(2.03)	(1.73)	(1.81)
N	30	30	30	30

a. *Trading experience* represents the number of trades made in a typical month.
b. *Years of market experience* denotes years that the subject has been active in the market.
c. *Income* denotes categorical variable (1–8): 1) Less than $10,000, 2) $10,000 to $19,999, 3) $20,000 to $29,999, 4) $30,000 to $39,999, 5) $40,000 to $49,999, 6) $50,000 to $74,999, 7) $75,000 to $99,999, 8) $100,000 or over.
d. *Age* denotes actual age in years.
e. *Gender* denotes categorical variable: 0 if female, 1 if male.
f. *Education* denotes categorical variable 1) Eighth grade or less, 2) High School, 3) 2-Year College, 4) Other Post-High School, 5) 4-Year College, 6) Graduate School Education.

VI. EVIDENCE FROM NONMEMORABILIA COLLECTORS

Even though the data in each field experiment reveal similar insights, the scope of the study may be interpreted narrowly due to the nature of the sample used—memorabilia collectors. In this section I rectify this potential shortcoming by i) presenting new evidence from a laboratory experiment that indicates the findings

WTA/WTP-tuple that can be summarized accordingly: $\partial WTP/\partial y = 1 - WTP/WTA$, where y is income (see, e.g., Bateman et al. [1997]). As such, taken literally, the disparity observed suggests that, roughly, if a dealer's income increased by $100, she would spend an additional $23.07 on sheets of University of Wyoming basketball trading cards. Likewise, if a nondealer's income increased by $100, she would spend an additional $82.10 on sheets of University of Wyoming basketball trading cards. Running the risk of making too much of a few point estimates rather than relying on inference gained from the statistical tests, I view these estimates as implausibly large.

6.2 List (EMA, 2004)

- Field experiment on sport cards

- Similar to experiment I in List (2003), except that objects are mugs and chocolate

- Trading in four groups:

  1. Mug: "Switch to Chocolate?"

  2. Chocolate: "Switch to Mug?"

  3. Neither: "Choose Mug or Chocolate?"

  4. Both: "Switch to Mug or Chocolate?"

- Large endowment effect for inexperienced card dealers

- No endowment effect for experienced card dealers!

- Learning generalizes beyond original domain

TABLE I

SELECTED CHARACTERISTICS OF PARTICIPANTS

	Dealers Mean (Std. dev.)	Nondealers Mean (Std. dev.)	Nondealers Mean (Std. dev.)
Trading intensity	11.81 (10.9)	4.94 (6.58)	6.88 (6.39)
Yrs. of market experience	9.88 (9.79)	7.15 (9.83)	7.21 (8.03)
Income	4.15 (1.75)	4.10 (1.69)	4.18 (1.81)
Age	36.55 (13.1)	34.54 (14.41)	37.04 (14.1)
Gender (% male)	0.94 (0.24)	0.85 (0.35)	0.82 (0.39)
Education	3.54 (1.40)	3.44 (1.33)	3.54 (1.54)
Sample Sizes			
Private			
Treatment $E_{candybar}$	30	31	---
Treatment E_{both}	32	30	---
Treatment $E_{neither}$	35	33	---
Treatment E_{mug}	32	30	---
Public			
Treatment $E_{candybar}$	---	---	33
Treatment E_{both}	---	---	28
Treatment $E_{neither}$	---	---	29
Treatment E_{mug}	---	---	35

Notes:

1. *Trading intensity* represents the number of trades made in a typical month.
2. *Yrs. of market experience* denotes years that the subject has been active in the market.
3. *Income* denotes categorical variable (1-8): 1) Less than $10,000, 2) $10,000 to $19,999, 3) $20,000 to $29,999, 4) $30,000 to $39,999, 5) $40,000 to $49,999, 6) $50,000 to $74,999, 7) $75,000 to $99,999 8) $100,000 or over.
4. *Age* denotes actual age in years.
5. *Gender* denotes categorical variable: 0 if female, 1 if male.
6. *Education* denotes categorical variable (1-6): 1) Eighth grade or less, 2) High School, 3) 2-Year College, 4) Other Post-High School, 5) 4-Year College, 6) Graduate School Education.
7. "Private" and "Public" sample sizes denote the number of subjects in Experiments 1A and 1B, respectively.

TABLE II

SUMMARY OF EXPERIMENTAL RAW DATA

	Number of Subjects Choosing Candy Bar	Number of Subjects Choosing Mug	Pearson χ^2
Panel A. Nondealers (Private)			
Treatment $E_{candybar}$	25 (81%)	6 (19%)	19.21 (3 df)
Treatment E_{both}	18 (60%)	12 (40%)	
Treatment $E_{neither}$	15 (45%)	18 (55%)	
Treatment E_{mug}	7 (23%)	23 (77%)	
Panel B. Nondealers (Public)			
Treatment $E_{candybar}$	29 (88%)	4 (12%)	34.79 (3 df)
Treatment E_{both}	16 (57%)	12 (43%)	
Treatment $E_{neither}$	17 (59%)	12 (41%)	
Treatment E_{mug}	6 (17%)	29 (83%)	
Panel C. Dealers (Private)			
Treatment $E_{candybar}$	14 (47%)	16 (53%)	0.54 (3 df)
Treatment E_{both}	14 (44%)	18 (56%)	
Treatment $E_{neither}$	18 (51%)	17 (49%)	
Treatment E_{mug}	14 (44%)	18 (56%)	

Panel D. Trading Rates	Preferred Exchange	p-value for Fisher's exact test
Pooled nondealers (n = 129)	0.18 (0.38)	<0.01
Inexperienced consumers (< 6 trades monthly; n = 74)	0.08 (0.27)	<0.01
Experienced consumers (≥ 6 trades monthly; n = 55)	0.31 (0.47)	<0.01
Intense consumers (≥ 12 trades monthly; n = 16)	0.56 (0.51)	0.64
Pooled dealers (n = 62)	0.48 (0.50)	0.80

Notes:

1. The Pearson Chi-Square tests in Panel A are distributed with 3 degrees of freedom and each have a null hypothesis of Hicksian preferences.
2. Data in Panel B are pooled from Treatments $E_{candybar}$ and E_{mug}. For non-dealers, data from "public" and "private" are pooled. Standard deviations are in parentheses.
3. Experienced consumers are those consumers who trade 6 or more times per month (6 is roughly the mean level of monthly trades). Intense consumers trade 12 or more times per month (12 is roughly the mean plus one standard deviation).
4. Fisher's exact test in Panel B has a null hypothesis of no endowment effect.

TABLE III

SUMMARY EMPIRICAL ESTIMATION RESULTS

Variable	Nondealers	Dealers
Constant	-2.49(0.87)*	-3.16(1.4)*
Treatment E_{mug}	3.02(0.48)*	0.61(0.58)
Treatment E_{mug} **trading intensity*	-0.16(0.05)*	-0.03(0.04)
Trading intensity	0.02(0.012)	0.05(0.03)
Years of market experience	-0.03(0.02)	0.07(0.04)
Income	0.14(0.09)	0.39(0.13)*
Age	0.01(0.01)	0.04(0.20)
Gender	0.87(0.47)	-0.24(0.80)
Education	-0.09(0.12)	-0.21(0.17)
N	121	58

Notes:

1. Dependent variable equals 1 if subject departed the experiment with a mug, 0 otherwise. *Treatment* E_{mug} =1 if agent was initially endowed with a mug, 0 otherwise; Gender = 1 if male, 0 otherwise.
2. Standard errors are in parentheses beside coefficient estimates.
3. Sample sizes may not match sample sizes in Table 1 due to some respondents not responding to the income question on the survey. Reported results omit these observations. If means are used to fill in the missing observations, results are not quantitatively or qualitatively different from the results reported.
4. "*" denotes coefficient estimate is significant at the $p < .05$ level.

6.3 Haigh and List (JF, 2004)

- Experienced traders vs. students

- Compare attitude to risk

- Invest up to 100 points in a gamble:

  - $p = 1/3$: get 350 points

  - $p = 2/3$: get 0 points

  - Not explained too well (Sarah)

- Conversion rates 1:1 (students) and 4:1 (investors)

- Nine rounds

- Frequent feedback: invest every period, with feedback every period

- Infrequent feedback: invest every 3 periods, with feedback every 3 periods

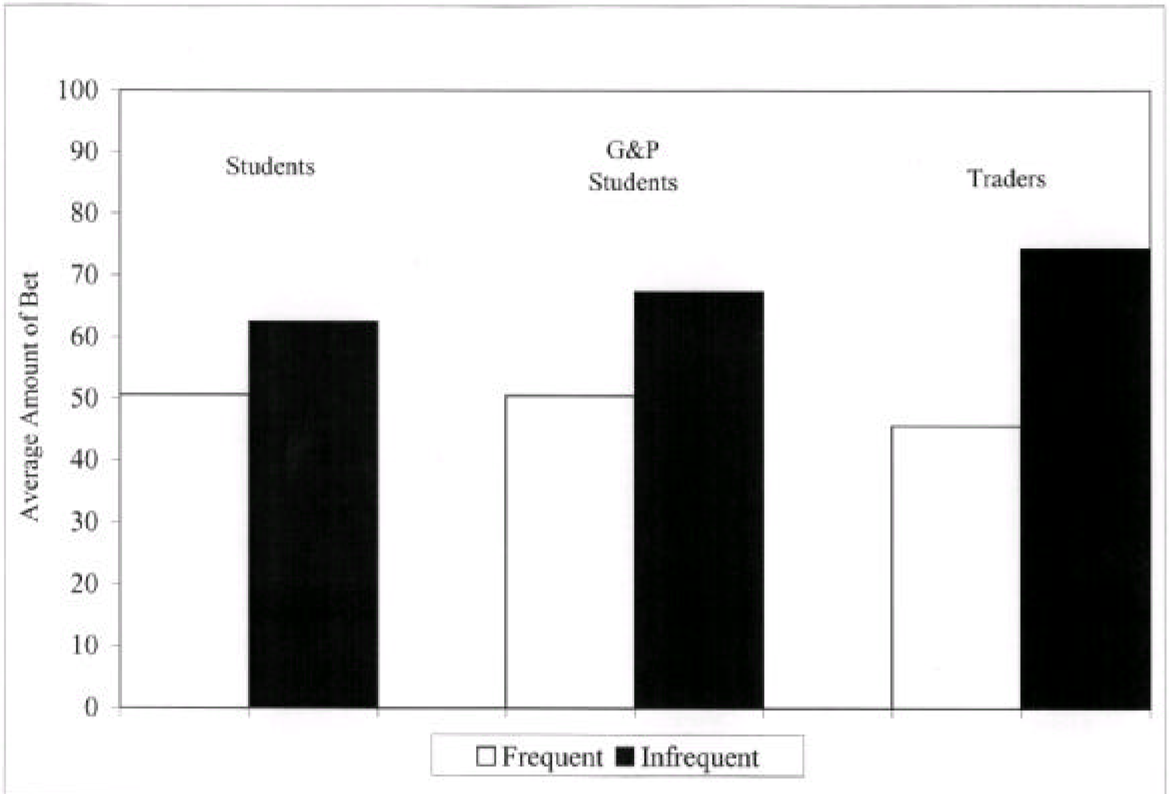
- Myopic Loss Aversion: Invest more with infrequent feedback (losses less frequent)

- Standard Theory: Does not matter

- Myopic Loss Aversion effect for both students and traders

- Effect strong for traders!

- Why?
  - Traders ‘trained’ that losses are bad (Vikram, Rob L.)
  - Do people learn better in ‘important’ tasks?



Note: G&P denotes Gneezy and Potters (1997).

Figure 1. Comparing betting patterns.