Field Experiments

For a long time, when economists spoke of "field experiments" they meant "social experiments" that were organized like clinical trials, ideally with randomized assignment of subjects into some social program intended e.g. to help unemployed workers find new employment.

These have some of the same advantages and disadvantages as clinical trials; how long to conduct the experiment is an issue when the control group involves "denial of treatment," subjects in different conditions may withdraw from the program differentially, some treatments may have to be organized by region, not by individual, so full randomization isn't possible, etc.

An interesting "how to" discussion is found in

Gueron, J. M. (1999) *The Politics of Random Assignment: Implementing Studies and Impacting Policy.* New York: **Manpower Demonstration Research Corporation.** http://www.mdrc.org/publications/45/workpaper.html

and for a bibliography see e.g.

Greenberg, D. and Shroder, M. (1997) *Digest of Social Experiments*, 2d ed. Washington, D.C.: Urban Institute Press.

Lately, influenced by laboratory experiments, the phrase "field experiment" has started to mean any experiment conducted outside of the laboratory.

The idea is to look for robustness to factors such as

–Subject pool
  • E.g. some naturally occurring populations (such as kibbutz residents, eBay traders, etc., i.e. not university students, or typical experimental subjects)
–Information that the subjects have
  • E.g. experts versus novices
–Commodity
–Task
–Stakes
–Environment

Often the aim is to re-examine some phenomenon observed in the laboratory and see how robust it is to some aspect of being 'taken into the field'.

An active early-adopter of this approach is John List at Maryland:

List, J.A. "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," *Econometrica* (2004)

List, J.A. "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field," *Quarterly Journal of Economics* (2004)

Haigh, M. and J.A. List "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis," *Journal of Finance* (2004)

List, J.A. "Does Market Experience Eliminate Market Anomalies?," *Quarterly Journal of Economics* (2003), 118, 41-71.

List, J.A. "Preference Reversals of a Different Kind: The More is Less Phenomenon," *American Economic Review* (2002), 92 (5): pp. 1636-1643.

List, J.A. and Lucking-Reiley, D. "The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign," *Journal of Political Economy* (2002), 110(1), pp. 215-233.

List, J.A. "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards," *American Economic Review* (2001), 91(5): pp. 1498-1507.

List, J.A. and Lucking-Reilly, D. "Demand Reduction in a Multi-Unit Auction: Evidence from a Sportscard Field Experiment," *American Economic Review* (2000), 90(4), pp. 961-972.

List suggests the following taxonomy:

- conventional lab experiment
  - one that employs a standard subject pool of students, an abstract framing, and an imposed set of rules
- synthetic field experiment
  - same as a conventional lab experiment but with a non-standard subject pool
- framed field experiment
  - same as a synthetic field experiment but with field context in either the commodity, task, or information set that the subjects can use
- natural field experiment
  - same as a framed field experiment but where the environment is the one that the subjects naturally undertake these tasks, such that the subjects do not know that they are in an experiment
- social experiment
- natural experiment

List, J.A. "Does Market Experience Eliminate Market Anomalies?," *Quarterly Journal of Economics* (2003), 118(1), pp. 41-71.

(This paper was originally circulated as "The Effect of Market Experience on the WTA/WTP Disparity: Evidence from the Field.")

i)   trading patterns of sports memorabilia at a sportscard show in Orlando, FL, and

ii)   trading patterns of collector pins in a market constructed by Walt Disney World at the Epcot Center in Orlando, FL.

"…these markets are natural settings for an experiment on the relationship between market experience and the endowment effect, as they provide natural variation across individual levels of expertise. In the sportscard show field experiments, I conduct some of the treatments with professional dealers and others with ordinary consumers. The design was used to capture the distinction between consumers who have intense trading experience (dealers) and those who have less trading experience (nondealers)."

"A major advantage of this particular field experimental design is that my laboratory is the marketplace: subjects would be engaging in similar activities whether I attended the event or went to the opera. In this sense, I am gathering data in the least obtrusive way possible while still maintaining the necessary control to execute a clean comparison between treatments. This highlights the naturalness of this particular setting, and the added realism associated with my field experiments."

The experiment is intended to see if the endowment effect will occur for sports memorabilia, in a marketplace for such items.

Aside from the challenges that endowment effects raise for models of individual preferences, there's also been a big debate, particularly in the environmental literature, about how to evaluate stated preferences for environmental goods if reported WTA>>WTP.

Horowitz and McConnell (2002; *Journal of Environmental Economics and Management*) document 45 studies that examine the WTA/WTP disparity. A majority of these studies find a large endowment effect:

Trading Patterns (e.g., Knetsch, 1989, *AER*)
89% of those originally endowed with a mug chose to keep the mug
90% of those endowed with a chocolate bar decided to keep the chocolate bar

Is the effect evident in market settings, or does market experience eliminate the endowment effect?

Others have attempted to examine this conjecture in n-shot lab studies (e.g., Knez et al., 1985, *AER*; Coursey et al., 1987, *QJE*; Shogren et al., 1994 *AER*)

Results are generally mixed

II. Design I—Orlando Sportscard Show (Dec. 1999)

1. **Subjects are randomly endowed with one good**
2. **Fewer than 50% of the subjects should trade their good if an endowment effect exists**

- **Good A is a Kansas City Royals game ticket stub dated June 14, 1996, was issued for admission to the baseball game in which Cal Ripken Jr. broke the world record for consecutive games played.**

- **Good B is a dated certificate commemorating the game that Nolan Ryan achieved what only 20 previous baseball players had done, winning 300 games (dated July 31, 1990), was distributed by the Milwaukee Brewers to fans in attendance of the ballgame.**

Each participant's experience typically followed three steps:

(1) **completing a survey**

(2) **considering the potential trade**

(3) **exit interview**

Experimental Results I

An Endowment Effect is Evident

Variable	% 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**	**(H_o: No Endowment Effect)**

- **Subjects are nearly two times more likely to keep the endowed good**

A Link Appears to Exist Between the Endowment Effect and Trading Experience

Dealers (n = 74)		
Good A for Good B	45.7	p = 0.194
Good B for Good A	43.6	

Nondealers (n = 74)		
Good A for Good B	20.0	p < 0.001
Good B for Good A	25.6	

- **Nondealers nearly 3.5 times more likely to select the good which they were endowed**

- **Dealers are 1.25 times more likely to choose their endowed good.**

Variable	% Traded	p-value for Fisher's exact test
Experienced nondealers (n = 30) (6 or more trades per month)	46.7	0.32
Inexperienced nondealers (n = 44)	6.80	< 0.01

- Inexperienced consumers are 13 times more likely to keep their endowed good.

Experiment II—Pin Trading (May 2000)

Collector pin market in *Walt Disney World's Epcot Center* in Orlando, FL.:

- **Higher percentage of female traders**
- **Quite popular**
  1. Pins have changed hands since the first modern Olympic Games in Athens in 1896.

  2. In the Atlanta Olympic Games, more than 1.2 million people visited the two Olympic Pin Trading Centers, where an estimated 3 million pins changed hands during the Games

     - **Good C is a cloisonné (enamel) pin of Mickey and Minnie Mouse which was issued on Valentines Day, 2000. The pin retailed for approximately $20 and sold-out within days of its issuance.**

     - **Good D is a cloisonne pin of Mickey Mouse which was issued on St Patrick's Day, 2000. Also retailed for approximately $20 and sold-out within days of its issuance.**

Experimental Results II

An Endowment Effect is Evident

Variable	% 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	

- **Subjects are nearly 2.5 times more likely to keep endowed good.**

A Link Appears to Exist Between the Endowment Effect and Trading Experience

Variable	% Traded	p-value for Fisher's exact test
Experienced nondealers (n = 20) (7 or more trades per month)	**40.0**	**0.26**
Inexperienced nondealers (n = 60)	**25.0**	**< 0.001**
Experienced nondealers (n = 30) (5 or more trades per month)	**46.7**	**0.30**
Inexperienced nondealers (n = 50)	**18.0**	**<0.001**

Are results driven by treatment or selection?

- **Some subjects may have a static "taste" for trading, whereas others do not have a preference for trading**

- **Results are then capturing static preferences**

Experiment III—Return to Original Sportscard Show

In Sept. 2000:

- **List emailed and/or telephoned the 148 original subjects of the Nov. 1999 Orlando Sportscard show**

- **108 promised to attend**
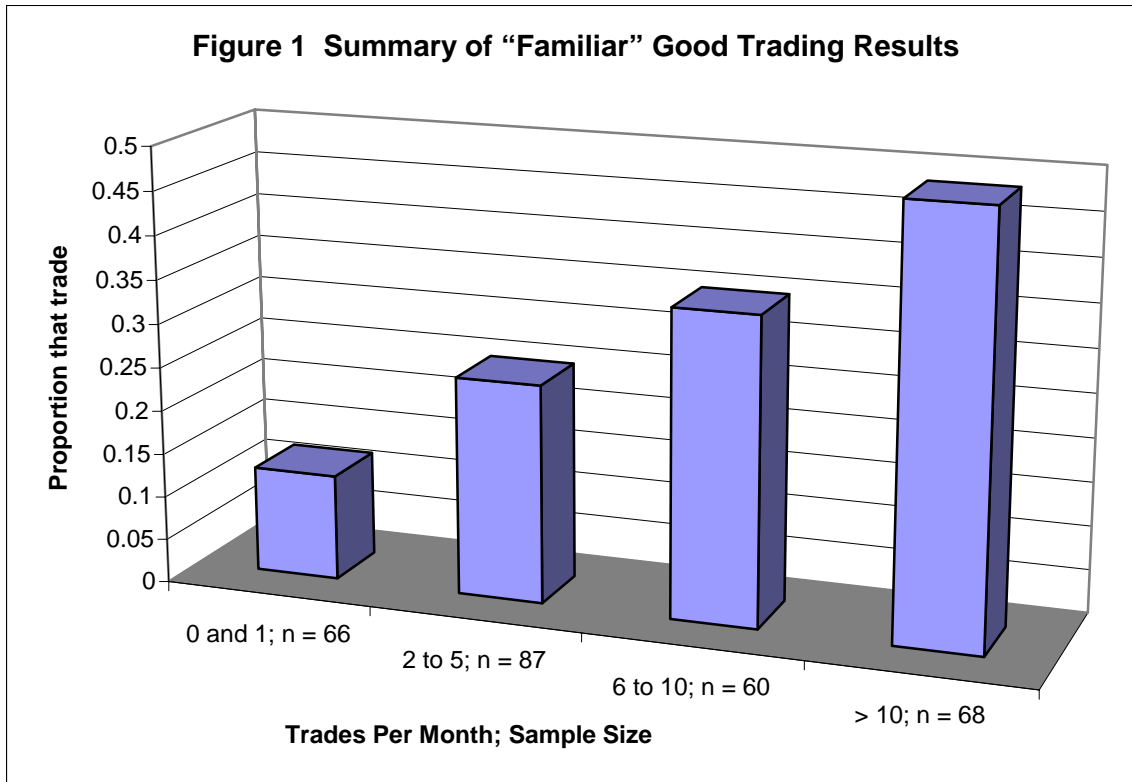
Nov. 4 and 5, 2000:

- **Only 72 showed-up**
  1. **53 nondealers**
  2. **19 dealers**

Two unique goods:

- **Good E: an attractive autographed 5X8 photo of Byron "Mex" Johnson**

- **Good F: an official National League baseball autographed by Byron "Mex" Johnson**

Johnson was a Negro League baseball player for the Kansas City Monarchs from 1937-1940

Figure 1 Summary of "Familiar" Good Trading Results

- Y-axis: **Proportion that trade** (0 to 0.5)
- X-axis: **Trades Per Month; Sample Size**
  - 0 and 1; n = 66
  - 2 to 5; n = 87
  - 6 to 10; n = 60
  - > 10; n = 68

5. Experimental Results III

Results consistent with the data in both the first sportscard field experiment and the pin field study:

- **Inexperienced subjects (less than 7 trades per month) executed a trade at a rate of 28% (9 of 32)**

- **11 of 21 (52.3%) experienced consumers chose to trade**

Are the results robust to a change in the trading mechanism?

Experiment IV—Tucson Sportscard Show

- **Examine explicit statements of WTA and WTP in actual auctions on the floor of a sportscard show**

- **Unique goods are "mint" and "near-mint" graded sheets of University of Wyoming basketball trading cards distributed to fans in attendance of "Midnight Madness" on the campus of the University of Wyoming in October, 1994**

- **50% of subjects compete in an auction for the sheet (provide WTP values)**

- **50% of subjects are given sheet and compete in an auction to sell the sheet (provide WTA values)**

Random *n*th-price auction:

1. Each bidder submits a bid (offer)

2. Each bid (offer) is rank-ordered from lowest to highest

3. Monitor selects a random number (n) uniformly-distributed between 2 and Z (Z bidders)

4. In the WTP case, the monitor sells one unit of the good to each of the (n-1) highest bidders at the *n*th-price; in the WTA case, the monitor buys one unit each from the (n-1) lowest bidders and pays the *n*th-lowest bid

Experimental Results IV

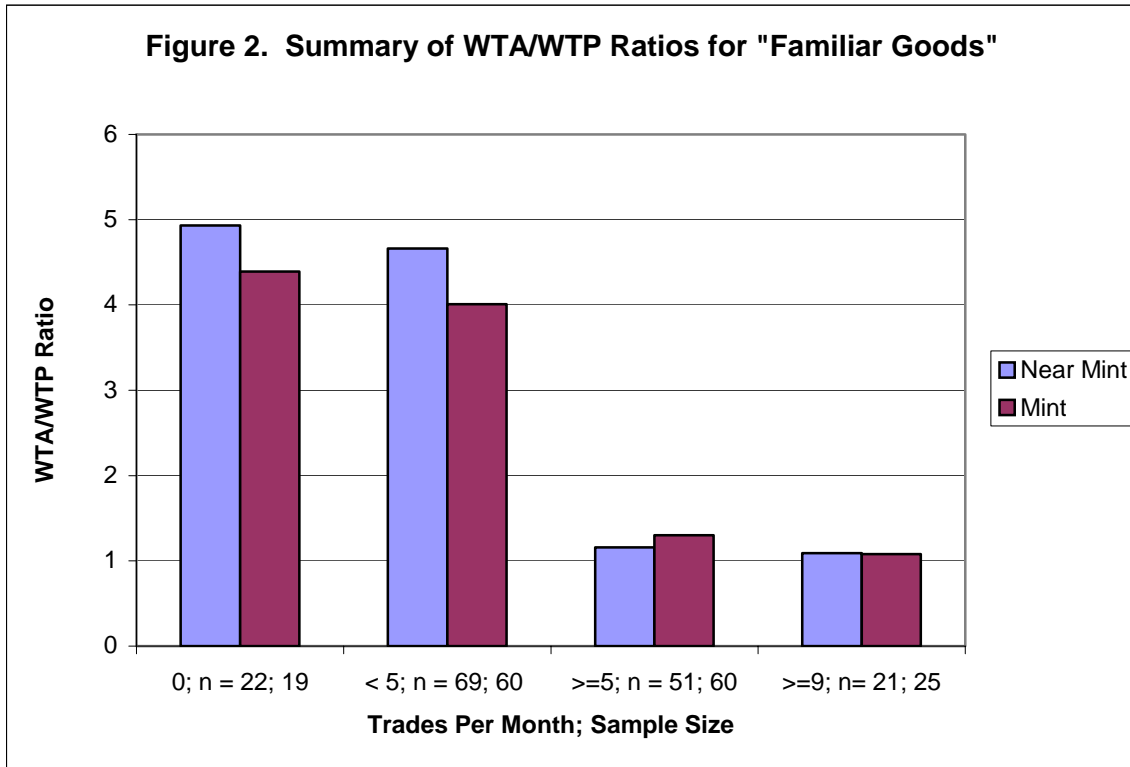Figure 2. Summary of WTA/WTP Ratios for "Familiar Goods"
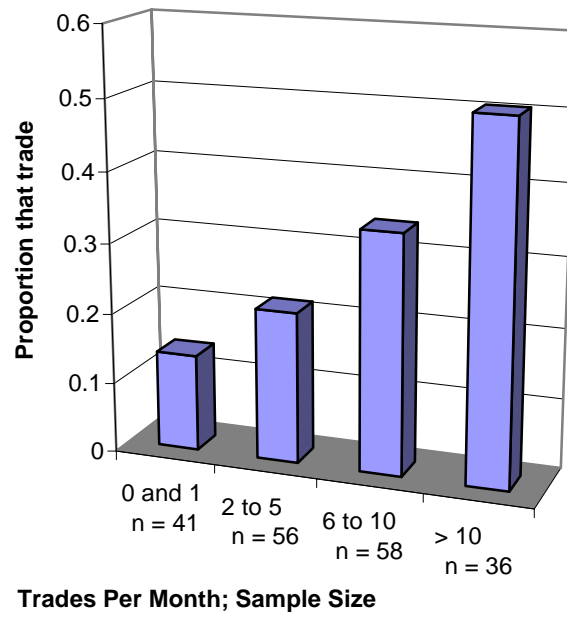
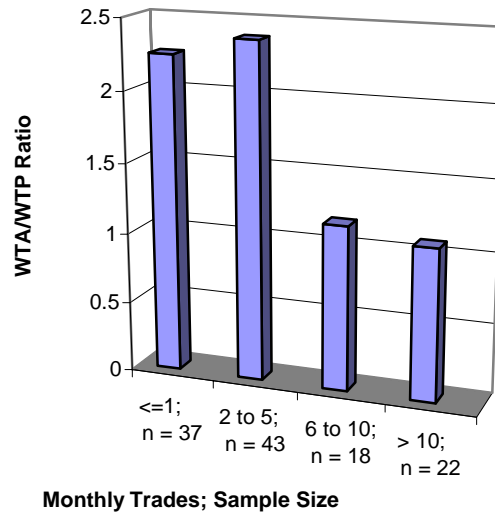Figure 3. Summary of "Unfamiliar" Trading Results

Figure 4. Summary of WTA/WTP Ratios for "Unfamiliar" Goods

What are some hypotheses that might account for the different results between experienced and inexperienced traders, or between the sportscard market and the lab?

Could sports memorabilia at a memorabilia market be "goods for trade" rather than consumption?

(List attempted to control for this with the survey: almost none of the subjects reported that they intended to sell or trade the items immediately—but maybe a sports card market at which John is giving these things away isn't the best place to sell them…)

List, J.A. "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," *Econometrica* (2004)
JLISTEMArevision.pdf

Experiment V—Tucson Sportscard Show, June 2001.

- **Examine trading and bidding behavior on the floor of a sportscard show using coffee mugs and chocolate**

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 ($\geq$ 6 trades monthly; n = 55)	0.31 (0.47)	<0.01
Intense consumers ($\geq$ 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.

List's WTA/WTP Conclusions

- **For the average consumer, the field data suggest there is an overall endowment effect**

- **Within both institutions—observed trading rates and explicit value revelation—evidence that individual trading intensity matters**

What could be driving the results?
1. **Avoid dealing with better informed agents and avoid making decisions when ill-informed**

2. **Some subjects derive utility from trading**

3. **Preference uncertainty**

4. **There is evidence that the "pain of given it up" is the main effect of endowment.**

5. **Marketlike experience may be leading to a lessening of this "sell-side" pain**

List has used sports card market participants to explore a number of other individual choice phenomena (including what he calls 'more is less' preference reversals…), and collective phenomena (including racial discrimination)

Other kinds of field experiments…

To assess direct mail appeals

To assess personal qualities of field subjects, as either a pre or post treatment effect…

The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign

John A. List

University of Maryland

David Lucking-Reiley

University of Arizona

We design a field experiment to test two theories of fund-raising for threshold public goods: Andreoni predicts that publicly announced "seed money" will increase charitable donations, whereas Bagnoli and Lipman predict a similar increase for a refund policy. Experimentally manipulating a solicitation of 3,000 households for a university capital campaign produced data confirming both predictions. Increasing seed money from 10 percent to 67 percent of the campaign goal produced a nearly sixfold increase in contributions, with significant effects on both participation rates and average gift size. Imposing a refund increased contributions by a more modest 20 percent, with significant effects on average gift size.
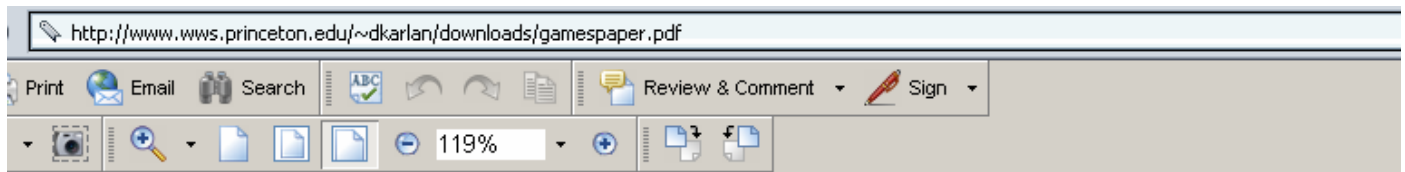
Using Experimental Economics to Measure Social Capital and Predict Financial Decisions[*]

Dean S. Karlan
Department of Economics
Research Program in Development Studies
Princeton University
dkarlan@princeton.edu

August 3rd, 2003

Abstract: Questions remain as to whether results from experimental economics games are generalizable to real decisions in non-laboratory settings. Furthermore, important questions persist about whether social capital can help solve seemingly missing credit markets. I conduct two experiments, a Trust game and a Public Goods game, and a survey to measure social capital. I then examine whether behavior in the games predicts repayment of loans to a Peruvian group lending microfinance program. Since the structure of these loans relies heavily on social capital to enforce repayment, this is a relevant and important test of the games, as well as of other measures of social capital. I find that individuals identified as "trustworthy" by the Trust game are in fact less likely to default on their loans.

The Effect of Educational Vouchers on Academic and Non-Academic Outcomes: Using Experimental Economic methods to study a Randomized Natural Experiment" Eric Bettinger and Robert Slonim

Abstract

This paper uses experimental economic methods and program evaluation techniques to assess the effects of a school voucher program on academic and non-academic outcomes.
Evidence on educational vouchers has largely focused on test scores, yet the rhetoric of voucher advocates suggests that any effects on test scores may be only a small part of their potential benefits …. Thus, test scores may not fully capture the effects of these differences. To test voucher advocates' claims, we present evidence from a privately-run, educational voucher program in Toledo, Ohio. The program focused on low-income families with children in grades K-8 and used a lottery to determine voucher recipients. We exploit the lottery to identify the effect of the voucher by comparing the outcomes of lottery winners and losers. Like other studies, we measure the effects of the voucher on traditional measures of student learning (e.g. test scores and grade repetition); however, we also attempt to measure the effects of the voucher on other non-academic outcomes, including altruism and patience. We measure these effects using protocols developed from the field of experimental economics. Our results provide mixed evidence on the voucher. We find little evidence that voucher winners have higher test scores than unsuccessful applicants. We do find, however, that voucher winners were more altruistic towards charitable organizations but were not more patient than unsuccessful applicants.

Sosis, Richard and Bradley J. Ruffle (2003) "Religious Ritual and Cooperation: Testing for a Relationship on Israeli Religious and Secular Kibbutzim," *Current Anthropology*, 44:5, 713-722, available at:
http://www.journals.uchicago.edu/CA/journal/issues/v44n5/035701/035701.html