The B.E. Journal of Economic Analysis & Policy

Frontiers

Volume 8, *Issue* 1 2008 *Article* 47

The Market: Catalyst for Rationality and Filter of Irrationality

John A. List* Daniel L. Millimet†

*University of Chicago and NBER, jlist@uchicago.edu
†Southern Methodist University, millimet@mail.smu.edu

Recommended Citation
John A. List and Daniel L. Millimet (2008) "The Market: Catalyst for Rationality and Filter of Irrationality," *The B.E. Journal of Economic Analysis & Policy*: Vol. 8: Iss. 1 (Frontiers), Article 47.
Available at: http://www.bepress.com/bejeap/vol8/iss1/art47

The Market: Catalyst for Rationality and Filter of Irrationality*

John A. List and Daniel L. Millimet

Abstract

Assumptions of individual rationality and preference stability provide the foundation for a convenient and tractable modeling approach. While both of these assumptions have come under scrutiny in distinct literatures, the two lines of research remain disjointed. This study begins by explicitly linking the two literatures while providing insights into whether market experience mitigates one specific form of individual rationality—consistent preferences. Using field experimental data gathered from more than 800 experimental subjects, we find evidence that the market is a catalyst for this type of rationality. The study then focuses on aggregate market outcomes by examining empirically whether individual rationality of this sort is a prerequisite for market efficiency. Using a complementary field experiment, we gathered data from more than 380 subjects of age 6-18 in multi-lateral bargaining markets at a shopping mall. We find that our chosen market institution is a filter of irrationality: even when markets are populated solely by irrational buyers, aggregate market outcomes converge to the intersection of the supply and demand functions.

KEYWORDS: rationality, market learning, field experiment, competitive market theory

---

*Corresponding author: John List, Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 6063. Telephone: 773.702.9811; fax: 773.702.8490. The authors are grateful to George Akerlof, William Harbaugh, Liesl Koch, Marc Nerlove, and Vernon Smith for providing helpful insights during this research project. Mike Price provided valuable research assistance.

I. Introduction

There are no assumptions more prevalent in economics than rationality and preference stability. Invoking both produces the foundation for a convenient and tractable modeling approach. A testament to the importance that rationality and preference stability hold in economics is the rich literature that tests these assumptions. To explore individual rationality, scholars have combined controlled laboratory experiments with Varian's [1982] theory of revealed preference (e.g., Sippel, 1997; Harbaugh et al., 2001; Andreoni and Miller, 2002).[1] Generally these studies find that between 10 and 75 percent of experimental subjects violate the Generalized Axiom of Revealed Preference (GARP).[2] Furthermore, in spite of the deep-rooted Hobbesian notion of exogenous and stable individual preferences, a recent "revolution" has taken place – building on earlier work by Adam Smith, Alexis de Tocqueville, and John Stuart Mill – that acknowledges the evolution of individual preferences over time (Georgescu-Roegen, 1950; Stone, 1954; Houttaker and Taylor, 1966; Wales, 1971; Philips, 1972; Brown and Heien, 1972; Manser, 1976; Pollak, 1978).

Interestingly, while both of these assumptions have come under scrutiny in distinct literatures, the literatures remain disjointed. In particular, tests for preference stability inevitably involve data collection at multiple points in time and presume that agents are rational. Such studies, dating back to at least Koopmans [1964], hypothesize a strong role for market experience in the evolution of individual preferences.[3] Similarly, early experimental tests on rationality involved data collection on consumption decisions at multiple points in time and presumed that agents had stable preferences (see, e.g., Battalio et al. [1973]). Since these early experimental studies on rationality were conducted, researchers have become aware of the potentially confounding influence of endogenous preferences and have since designed experiments to circumvent this possibility (e.g., Sippel [1997]). Studies pertaining to endogenous preferences, however, have yet to consider the potentially confounding effects of irrational behavior, or, more precisely, evolving or *endogenous rationality*.

---

[1] The use of experiments to test individual behavior actually has a much longer history (see, e.g., May [1954], MacCrimmon and Toda [1969], Battalio et al. [1973]).

[2] Sippel [1997] uses students of law or economics and ten budget sets for eight goods. Mattei [2000] uses college students and other adults, along with 20 budget sets for eight goods. Harbaugh et al. [2001] use 7- and 11-year-old participants and eleven budget sets for two goods. Andreoni and Miller [2002] use college economics students and eleven budget sets defined in a modified dictator game where subjects had to decide how much money to keep and how much to give to charity.

[3] Bowles [1998] provides an excellent survey of recent arguments and presents empirical evidence that suggests preferences are determined, in part, by markets and other economic institutions.

If factors such as market experience influence the development of rational behavior, previous tests of preference stability would be unable to distinguish between the market's impact on preferences and the market's effect on the ability of agents to *implement* those preferences. One of our main goals in this study is to make this distinction clear by using one specific type of rationality test that is fundamental in economics: preference consistency. Specifically, we seek to answer two questions in the first part of the paper: (i) Do individuals in the marketplace exhibit rationality in their choice behavior? and (ii) Is rational choice behavior influenced by experience in the marketplace?

To provide initial insights into these questions, we design and implement two experiments in distinct locations—one in Arizona from November 2000-June 2001, and one in Illinois from November 2006-June 2007. Each of the two experiments includes *two rounds* (seven months apart) of a controlled field experiment on a panel of children 6-18 years old at a shopping mall. Since each round of the experiment involves subjects making consumption choices under different budget constraints, the level of rationality exhibited by subjects during each round is identifiable, even if respondents' preferences change between rounds of the experiment. Moreover, since a number of our subjects were at the shopping mall simply to enjoy the mall with their parents, while others were there to participate in the market for sports collectibles (sportscard show), we are able to analyze the impact of sportscard market experience on rational choice behavior. Importantly, we circumvent the potential non-randomness of sportscard market experience by exogenously inducing a random sample of subjects simply shopping at the mall to enter the sportscard market. This treatment represents a particularly demanding test of the impact of market experience on learning since it represents a test of how experience in one well-defined market affects rational choice behavior in a *separate*, quite distinct market, rather than a test of how experience in a particular market induces certain heuristics that can be applied in future transactions in the *same* market over similar tasks.[4]

Empirical results generated from observing more than 800 experimental subjects indicate that (i) only about 31% of agents exhibit behavior consistent with rational choice theory, and (ii) market experience facilitates the development of such behavior. Specifically, we find that subjects self-selecting into the sportscard market prior to the initial experimental round had, on average, two fewer GARP violations (representing about two-thirds of a standard deviation); and while close to 73% of sportscard market participants exhibited behavior

[4] Experiences in one arena that enable individuals to apply more rational choice behavior in a separate arena has been referred to as "learning more rational rules" (Slonim, 1999).

consistent with utility maximization, only 28% of non-participants exhibited such behavior.

More importantly, the randomization of subjects into the marketplace combined with the panel nature of our data permits tests of the impact of *exogenous* market participation on rational choice behavior. We find the number of GARP violations to be 0.4 lower, and the probability of zero violations to be 6% higher, in round two among subjects induced into market participation. In addition, subjects induced into market participation experienced a *reduction* in GARP violations across survey rounds of about 0.9 (representing a decline of about one-fourth of a standard deviation), while non-participants experienced virtually no change. Finally, we document gains to the intensity of market experience: exogenously induced new, full-time entrants into the market experienced a reduction in GARP violations between four and six times larger on average than that experienced by comparable new, one-time entrants. These findings are robust to controls for various observable attributes of subjects, the possible endogeneity of market participation, possible sample selection bias due to nonrandom attrition, issues of sample size, location of the experiment, and several econometric modeling assumptions.

While these findings might have some import both positively and normatively, Becker's [1962] insight that several fundamental features of economics, such as correctly sloped supply and demand schedules, could result even when agents are irrational is an important consideration. In this sense, it is quite possible that those interested in market outcomes might find little to be aroused by in our experimental results, as irrational individual choice behavior may have little disruptive impact on market level outcomes. In other words, while *individual* irrationalities are observed in experimental markets, *market* irrationalities may not be observed. To address this issue, we link actual behavior from the GARP experiment to behavior in a market field experiment to examine a third question: Is individual rationality of the sort considered herein a prerequisite for market efficiency in multi-lateral decentralized markets? Such an analysis is in the spirit of Forsythe et al. [1992] and Gode and Sunder [1993].

Empirical results indicate that even in decentralized bargaining markets populated entirely by irrational children buyers, market outcomes are efficient and prices follow expectations after a few rounds of play. One major difference between markets populated by inexperienced and experienced children is rent allocation: whereas the allocation of rents favors sellers in markets populated entirely by inexperienced and irrational children, it is distributed more symmetrically in markets populated by experienced and rational children. Thus, we conclude that individual irrationality as measured in our experiment does not unduly influence aggregate efficiency in bilateral bargaining markets.

The remainder of our study is organized as follow. Section II describes the experimental protocols in the two locations. Section III analyzes the results. Section IV examines efficiency properties of markets populated entirely by irrational buyers. Section V concludes.

II. Experimental Design

A. Experimental Design I

Our initial experimental treatments were run at various sportscard shows at a shopping mall in a large urban setting in Arizona during the period November 2000 to June 2001; we refer to these data as the "Arizona sample" or "Experiment I". In the first set of treatments carried out in November of 2000, the monitor approached *young* individuals in and near the marketplace and inquired about their interest in participating in an experiment that would take about ten minutes. The decision to solicit youths as experimental subjects stemmed from, among other reasons, our desire to examine the role of the market in the development of rational choice behavior. A pilot survey revealed that many youths had limited individual exposure to organized markets.[5]

If the individual agreed to be an experimental participant, the monitor began the four steps of the experiment: first, the subject began by filling out a brief survey in which information on age, gender, years of sportscard market experience, buying, selling, and trading intensity in the sportscard market, and the number of monthly visits to the mall were obtained. After completing the survey, Step 2 began: the subject was physically given the experimental sheets and instructions for the GARP experiment. Our GARP treatments closely follow Harbaugh et al. [2001], as we present our subjects with 11 different choice sets (over boxes of juice and bags of chips) on 11 separate sheets of paper and inform each subject to choose their most preferred bundle on each sheet, after which the monitor will choose one sheet to execute for real payment.[6]

Figure 1 graphically depicts the 11 choice sets. In our design, a GARP violation occurs when a bundle x is chosen when a bundle y is available, where bundle y has at least as much of all goods and strictly more of at least one good than a third bundle z, and z has been directly or indirectly revealed preferred to x.

---

[5] For example, more than 95% of children age 10 and under stated that besides the sportscard market they had very limited individual opportunity in other organized markets.

[6] As stressed in the introduction, while we believe GARP is the most fundamental definition of rationality—in simple static choice settings GARP is necessary and sufficient for integrability—one could use several other measures of rationality. For example, how well one satisfies dominance or Bayes rule might be important. In this spirit, it is possible that subjects who have few (many) violations in our setting have many (few) violations of rationality for other tests.

A bundle *z* is directly revealed preferred to another bundle *x* if (i) *z* is chosen when *x* is available, or (ii) *z* is chosen when another bundle containing at least as much of all goods as in *x* and strictly more of at least one good is available. Bundle *z* is indirectly revealed preferred to bundle *x* if a string of directly preferred relations suggests that *z* is preferred to *x*. If an individual's choices do not violate GARP, they are consonant with the individual maximizing a continuous, concave, strongly monotonic utility function.

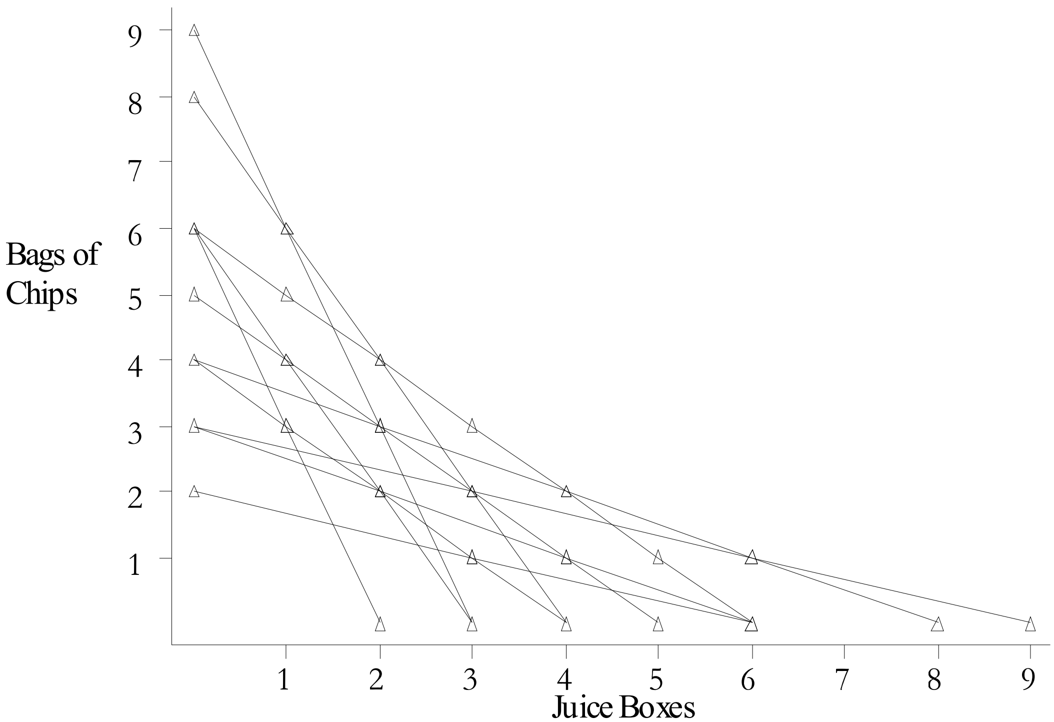

Figure 1. Choice Sets

Several examples were carried out to ensure that the subject understood the details of the experiment. No time limit was imposed. In Step 3, the monitor informed the subject which sheet was to be actually executed and the subject received his or her chosen bundle from that sheet. Step 4 closed the experiment and included "parting gifts" for subjects in certain treatments (explained more fully below).

To provide the necessary variation in our subject pool, we conducted some of our GARP treatments with subjects who were simply at the mall for enjoyment (denoted *MALL*), while others were conducted with participants who were at the mall to participate in the sportscard market (denoted *SPORTS*). This aspect of our design was used to capture the distinction between consumers who had previous market experience and those who (may) have had no rigorous market experience. While performance differentials across these groups lend insights into the rational choice behavior of individuals in and out of the sportscard market, self-selection of individuals into the sportscard market prohibits one from drawing conclusions concerning the causal effect of market experience on rational behavior.

To circumvent this difficult issue, we further delineated the *MALL* subjects into two groups: *MALLgift* and *MALLnogift*. The delineation was changed at the top of each hour, so subjects' treatment type was determined exogenously based on the time they visited the mall. For the *MALLgift* subjects, we provided a "parting gift" of approximately $25 worth of sportscards and memorabilia in Step 4 of the experiment. In each subject's gift bundle, we included several items designed to engage the subject in the marketplace. The monitor informed subjects in this treatment group that the gifts were theirs to keep, and they could sell or trade the gifts in the marketplace or take them home. The monitor stressed that dealers at the show were interested in the goods and that the goods had a book value of approximately $25. Furthermore, in an effort to ensure that the local dealers would have an interest in buying and trading with these subjects, prior to the show the monitor discussed various items with the dealers attending the show to determine the appropriate composition of the gift packages. In practice, these agents engaged in buying, selling, and trading in the live marketplace.

Treatment *MALLnogift* was identical to the *MALLgift* treatment except that in Step 4 the monitor closed the experiment by thanking the subject for his or her participation, and no "parting gift" was provided. Likewise, in the *SPORTS* treatment, since subjects were already at the mall to attend the sportscard show, in Step 4 the monitor closed the experiment by thanking the subject for his or her participation, and no "parting gift" was given.

To gather the experimental evidence necessary to analyze the learning issue, we returned to the same mall the following June and ran similar GARP treatments using the same subject pool, identical experimental procedures, but different goods (instead of chips and juice boxes we used packs of gum and candy bars). To recruit the same subjects, in May one of the authors personally telephoned and/or e-mailed the 277 subjects that participated in the first November sportscard show experiment. He was able to contact and obtain agreement to meet him at the June sportscard show from more than 170 subjects. As a friendly reminder, within two weeks of the experiment he called/e-mailed the 173 subjects that agreed to participate; despite this reminder only 104 subjects attended the sportscard show.

Table 1A provides a summary of the experimental design and the number of subjects in each treatment. For example, the *MALLgift* treatment in November included 110 subjects and of those 110 subjects 42 returned for the second GARP experiment, which we denote treatment *MALLgiftII*. Given the relatively low return rates – to be expected given the age of the subjects and the length of time between treatments – non-random attrition is a well-founded concern; we will return to this point often.

Table 1A. Individual Choice (GARP) Experimental Design (Arizona Sample)

	November 2000	**June 2001**
***MALL* subject with free gift bag**	*MALLgift* $n = 110$	*MALLgiftII* $n = 42$
***MALL* subject without free gift bag**	*MALLnogift* $n = 109$	*MALLnogiftII* $n = 31$
Sportscard subject without free gift bag	*SPORTS* $n = 58$	*SPORTSII* $n = 31$

Notes: Each cell represents one unique treatment. For example, "*MALLgift*" in row 1, column 1 denotes that 110 subjects were at the mall for reasons other than the sportscard show and in Step 4 they were given a $25 bundle of sportscards and memorabilia. No subject participated in more than one treatment.

Of course, over the six-month period between the first and second treatments subjects in the *MALLgift* treatment could have learned to avoid preference inconsistency via several distinct routes: (i) natural aging process; (ii) market experience garnered through buying, selling, and trading in the marketplace to earn the highest profits from their gifts;[7] or, (iii) market experience garnered through buying, selling, and trading in the marketplace *and* by subsequently joining the sportscard market as a "regular" participant. By comparing learning rates across these cells with their counterparts in treatments *MALLnogift* (*MALLnogiftII*) and *SPORTS* (*SPORTSII*), we are able to explore the power of experience in the sportscard market.

[7] In some cases, agents had to buy or trade for certain goods to make their bundle more attractive to buyers.

B. Experimental Design II

While experiments are extremely beneficial in situations where it is difficult, if not impossible, to identify naturally-occurring sources of exogenous variation in the treatment of interest, some important limitations arose with Experiment I. First, the sample size is small, and the return rate of subjects is low. Second, little information is available about the experimental subjects aside from age, gender, and frequency of mall visits. Finally, in the case of field experiments, the generalizability of the findings to other locations or populations remains an open question.

Table 1B. Individual Choice (GARP) Experimental Design (Illinois Sample)

	November 2006	**July 2007**
***MALL* subject with free gift bag**	*MALLgift* *n* = 255	*MALLgiftII* *n* = 157
***MALL* subject without free gift bag**	*MALLnogift* *n* = 267	*MALLnogiftII* *n* = 146
Sportscard subject without free gift bag	*SPORTS* *n* = 20	*SPORTSII* *n* = 13

Notes: Each cell represents one unique treatment. For example, "*MALLgift*" in row 1, column 1 denotes that 110 subjects were at the mall for reasons other than the sportscard show and in Step 4 they were given a $25 bundle of sportscards and memorabilia. No subject participated in more than one treatment.

To tackle some of the outstanding issues, we replicated the above experiment at various sportscard shows at a shopping mall in a large urban setting in Illinois during the period November 2006 to June 2007. We refer to this as the "Illinois sample" or "Experiment II". The design of Experiment II was, in every way, identical to the Arizona experiment except in three respects. First, the Illinois experiment was a larger scale experiment, involving over-sampling of the experimental treatment and control groups. Specifically, whereas the Arizona sample contained 277 subjects during round one (79% belonging to the experimental treatment and control groups), the Illinois sample contained 542 subjects (96% belonging to the experimental treatment and control groups).

Second, to induce a greater rate of return of experimental subjects in the second round, we offered all returning participants a $20 award. As a result, the return rate of the experimental subjects exceeded 58% (versus only 33% in the Arizona sample). See Table 1B.

Finally, we collected additional information on attributes of the subjects; these include race, grade in school, type of school attended (public versus private), number of siblings, subjective expectations related to the weather in the subsequent week and whether their teachers like them, number of television hours watched per week, favorite subject in school, relationship to the adult accompanying them to the mall, and various characteristics of the accompanying adult (see Appendix B for a copy of the survey).

III. Experimental Results

A. Pooled Sample

Our initial analysis pools the data from the two experiments. In this section, we document five salient findings. For clarity, we begin by stating each result and proceed to discuss the evidence supporting the result. Before turning our focus to these findings, however, Table A1 in Appendix A provides descriptive statistics of the variables in common across the two experiments. Across our treatments, the percentage of females ranged from 0% to 52% and the number of mall visits per month ranged from roughly four to eight. The average age was approximately 10.5 years; there were at least ten, and no more than 49, subjects in the first round of each age between six and sixteen years old (and six 17-year-old subjects as well).

Our data yield a first finding:

Result 1: 69% of agents exhibit at least one GARP violation.

Table 2 provides support for this result, detailing summary characteristics of the GARP results for the 1239 subject-round observations. On average, subjects had over 3.6 GARP violations, with almost 69% of subjects having at least one violation.[8] Focusing on the first round only (when attrition is a non-issue) makes no difference; 69% of subjects had at least one violation in the first round as well. This result is consistent with Sippel [1997] who reports that 57% participants violated GARP, as well as Harbaugh et al. [2001] who find that roughly 40% of their 42 eleven-year-old subjects had at least one GARP violation.

[8] Alternatives to violation counts such as Afriat's continuous measure of the distance from GARP were also considered; the data pattern is similar to violation counts so we suppress further discussion.

Table 2. Summary of GARP Violations: Pooled Sample

Group	GARP Violations	H_O: Equality of Means Across Participants & Non-Participants	Observations
Full Sample	3.63 (3.31) [68.60%]		1239 [Round 1: *n* = 819] [Round 2: *n* = 420]
<u>Group 1 (*MALLgift*)</u>:			
Round 1			
Non-Participants	4.18 (3.50) [71.78%]		365
Round 2			
Non-Participants	3.40 (4.10) [60.00%]	$t = 0.15$ [$p = 0.88$]	5
Participants	3.21 (2.59) [74.74%]		194
<u>Group 2 (*SPORT*)</u>:			
Round 1			
Participants	1.92 (2.98) [37.18%]		78
Round 2			
Participants	1.64 (2.24) [40.91%]		44
<u>Group 3 (*MALLnogift*)</u>:			
Round 1			
Non-Participants	3.95 (3.41) [72.34%]		376
Round 2			
Non-Participants	3.62 (3.35) [69.18%]	$t = 0.69$ [$p = 0.49$]	159
Participants	3.06 (3.00) [61.11%]		18

Notes: Figures represent mean GARP violations. Standard deviations are in parentheses. Percentage of observations with non-zero GARP violations or *p*-values is in brackets. Group 1 includes non-participants in the market at the time of round one who were randomly given $25 worth of free cards after round one was completed. Group 2 includes subjects who were already in the market at the time of round one. Group 3 includes non-participants in the market at the time of round one who were not given any free cards after round one was completed. Participants include both regular market participants and one-time participants (partial compliers).

Figure 2 shows the distribution of GARP violations separately by location, as well as using the pooled sample. Restricting attention to the sub-sample with at least one GARP violation, the distribution has a mode at three violations and is fairly uniform from six to ten GARP violations.

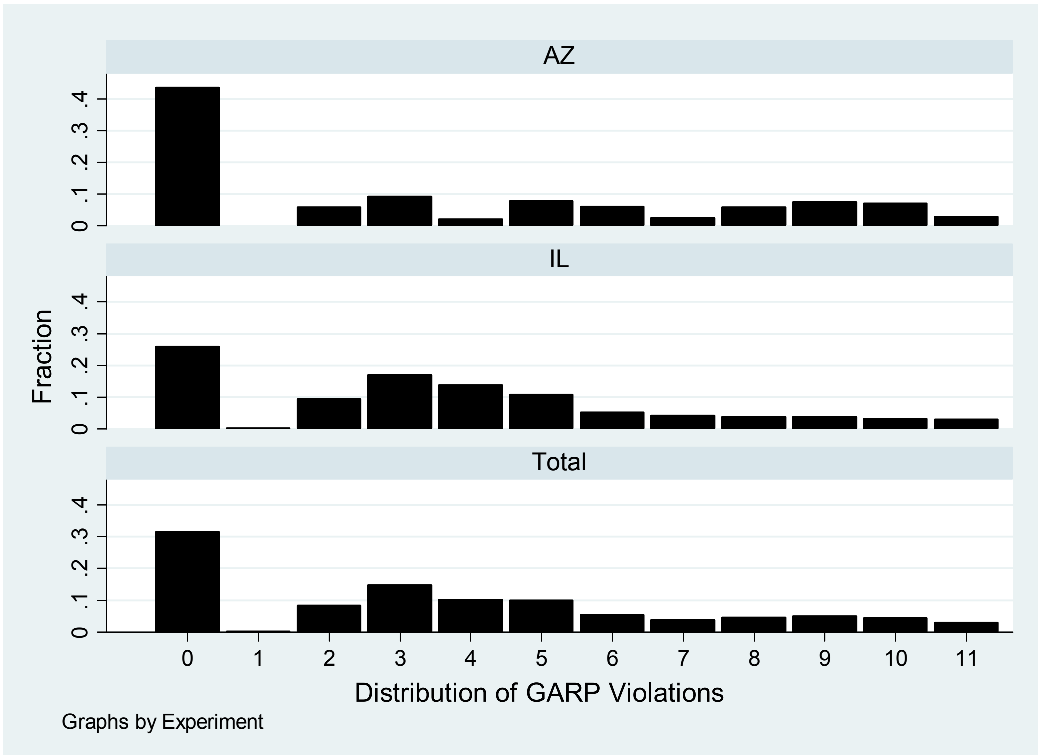

Figure 2. GARP Violations

A second result naturally follows:

Result 2: Market experience and rational choice behavior are directly related: market participation reduces the number of GARP violations by 0.4 to 0.8.

Table 2 provides preliminary support for this result by including summary characteristics of the GARP results for the 1239 subjects disaggregated by treatment assignment. In round one of the experiment, non-participants in the sportscard market who were given the $25 "parting gift" (*MALLgift*) had roughly 4.2 GARP violations on average; nearly 72% had at least one violation. Non-participants in the sportscard market who were not given the $25 "parting gift"

(*MALLnogift*) displayed virtually identical levels of rational behavior. This is critical to our ability to circumvent the self-selection issue as it verifies that the "parting gifts" were exogenously assigned.

Subjects participating in the sportscard market at the time of round one (*SPORTS*) registered fewer than two GARP violations on average, and less than 38% of those in the *SPORTS* treatment had non-zero GARP violations. The difference in mean GARP violations across market participants (*SPORTS*) and non-participants (*MALLgift* and *MALLnogift*) is statistically significant at the $p < 0.01$ level ($t = 5.27$; to be conservative, we provide all test results assuming a two-sided alternative). While interesting, the positive correlation between market participation and rational choice behavior *may* be spurious, reflecting observables (such as age or gender) or unobservables (such as innate ability) associated with the self-selection into the sportscard market.

Examination of the round two results (conducted seven months later) allows us to disentangle selection effects from market participation effects through the comparison of experimental subjects exogenously induced into the market through the "parting gift" given after round one (*MALLgift*) with non-participants not given this gift (*MALLnogift*). Estimating the exogenous effect of market participation is not straightforward, however, as some subjects failed to "comply" with the experiment, and those who did comply, did so in varying degrees. Specifically, five subjects in *MALLgift* never took part in the sportscard market (they neither traded nor sold items in their gift bag nor entered the market at a later date). Utilizing terminology from the treatment effects literature, we shall henceforth refer to these two subjects as "noncompliers." Of the remaining *MALLgift* subjects observed in round two, 99 entered the sportscard market to buy/sell/trade items, but they did not subsequently engage in any market activities after their gift bag items were exhausted (referred to hereafter as "partial compliers"). And, 95 *MALLgift* subjects entered the sportscard market to buy/sell/trade items and subsequently became regular participants in the market (referred to hereafter as "compliers"). These agents reported that they had attended sportscard shows to buy/sell/trade items not related to their gifts within the seven month intervening period. Finally, 18 subjects in *MALLnogift* entered the sportscard market on their own volition between rounds one and two (referred to hereafter as "noncompliers" as well), while the remaining 159 subjects in *MALLnogift* observed in round two remained out of the market (referred to hereafter as "compliers").

Ignoring the possible nonrandomness of compliance for the moment and focusing only on the sub-set of experimental compliers (including partial compliers), Table 2 reveals that *MALLgift* subjects had 3.21 GARP violations on average, while *MALLnogift* individuals had 3.62 violations on average (although the difference is not statistically significant at conventional levels: $t = 1.28$, $p =$

0.20). The fact that round two subjects in the *SPORT* treatment have fewer GARP violations on average (1.64) than the *MALLgift* group of market participants reflects selection effects (e.g., children participating in the sportscard market with no inducement tend to be older or of higher ability) and/or indicates that duration of market experience matters.

Although interesting, the summary results discussed heretofore may not reflect the *causal* effect of market experience on rational choice behavior due to bias arising from non-random attrition and the endogeneity of compliance. In addition, even though assignment to the *MALLgift* and *MALLnogift* treatments is random, there are efficiency gains from controlling for observable attributes of subjects (see, e.g., Hirano et al. [2003] and Millimet and Tchernis [2008]). Fortunately, the econometric methods needed to address these potential sources of bias and inefficiency are well-developed. To proceed, we begin by first estimating various specifications nested in the following equation:

$$GARP_{it} = x_{it}\beta + \delta PART_{it} + \pi[\phi(\theta_{it})/\Phi(\theta_{it})] + \eta_{it} \qquad i = 1,\ldots,N;\ t = 1,2 \qquad (1)$$

where $GARP_{it}$ is the number of GARP violations for subject i in round t; x_{it} is a vector of individual attributes; $PART_{it}$ is a dummy variable equal to unity if subject i has participated in the sportscard market as of round t; β, δ, and π are parameters to be estimated; ϕ (Φ) represents the standard normal density (cumulative density) function; $\theta_{it} = z_{it}\hat{\gamma}$, where Pr($GARP_{it}$ is observed) = $\Phi(z_{it}\gamma)$; and, $\eta_{it} = \mu_i + e_{it}$ is the error term, where μ_i is an individual-specific effect and e_{it} represents idiosyncratic shocks. The vector x includes controls for age, gender, and state, while z includes controls for age, gender, state, a quadratic in mall visits per month, and state interacted with a quadratic in mall visits per month.[9] $\phi(\theta_{it})/\Phi(\theta_{it})$ is the inverse Mills' ratio, and is used to control for possible bias due to non-random attrition (Heckman [1976]). The exclusion of mall visits per month from x ensures that the model is nonparametrically identified.[10]

The second set of empirical specifications utilizes the fact that our dependent variable is a count measure of the number of GARP violations; such a regressand is typically analyzed using a Poisson regression model. The Poisson

[9] We also replaced the linear age term with a full set of age dummies in all specifications. The estimated coefficient on *PART* is qualitatively unaffected; these results are available upon request.

[10] Imposing the restrictions π = 0 and $\sigma_\mu^2 = 0$, where σ_μ^2 is the variance of the random effects, (1) reduces to a simple pooled Ordinary Least Squares (OLS) regression. Restricting only π = 0 reduces (1) to a Generalized Least Squares (GLS) random effects model, or a fixed effects (FE) model, depending on whether μ_i is assumed to be independent of the included regressors. Finally, restricting $\sigma_\mu^2 = 0$ reduces (1) to the standard Heckman selection model.

model assumes that the number of violations for individual *i* at time *t* is drawn from a Poisson distribution with parameter λ_{it}. Consequently, the probability of observing a given number of violations is given by:

$$\Pr(GARP_{it} = g_{it}) = \frac{\exp\{-\lambda_{it}\}\lambda_{it}^{g_{it}}}{g_{it}!}, \qquad g_{it} = 0,1,2,\ldots \tag{2}$$

where $\ln(\lambda_{it}) = x_{it}\beta+\delta PART_{it}$, x_{it} and $PART_{it}$ are defined above, β is a vector of unknown parameters, and δ is a single, unknown parameter.

Although estimation of equation (2) is straightforward via maximum likelihood, Figure 2 reveals an interesting statistic: over 30% of observations had zero GARP violations. This finding represents a potential problem in estimation of equation (2) since the number of observations with zero violations exceeds that predicted by the standard Poisson model. To circumvent this problem, we consider the underlying data generation process more closely. Conditional on the presence of certain attributes, some individuals may never display any irrational behavior; such individuals would always register zero violations, independent of the data generation process. Moreover, there are other individuals for whom the number of violations conceivably follows a Poisson process, implying that the number of violations may again be zero, in part due to the data-generating process.

A technique to account for this two-step process is discussed in Greene (2003), among others. The procedure, commonly referred to as a zero-inflated Poisson (ZIP), is a natural extension of the Poisson formulation in equation (2):

$$g_{it} = 0 \qquad \text{with probability } P_{it} \tag{3}$$
$$g_{it} \sim \text{Poisson}(\lambda_{it}) \qquad \text{with probability } 1 - P_{it}, \tag{4}$$

where $\ln(\lambda_{it}) = x_{it}\beta$, and therefore

$$\Pr(g_{it} = 0) = P_{it} + (1-P_{it})R_{it}(0) \tag{5}$$
$$\Pr[g_{it} > 0] = [1-P_{it}]R_{i}(\text{not } 0). \tag{6}$$

P_{it} represents the state probability and R_{it} is the Poisson distribution for g_{it}. To model the state probability, P_{it}, we use the logistic specification

$$P_{it} \sim \text{Logistic}(\Psi_{it}), \tag{7}$$

where $\Psi_{it} = x_{it}\tilde{\beta}+\tilde{\delta} PART_{it}$, with $\tilde{\beta}$ and $\tilde{\delta}$ representing unknown parameters.

Our final estimation technique – the propensity score matching (PSM) estimator developed in Rosenbaum and Rubin [1983] – relaxes the distributional and functional form assumptions utilized in the estimations of (1) and (2), as well as the zero-inflated Poisson model. The goal of the matching method is to identify the effect of a binary treatment on an outcome of interest. Blundell and Costa-Dias [2002] provide an excellent introduction to the matching method, concluding that "matching methods have been extensively refined in the recent evaluation literature and are now a valuable part of the evaluation toolbox."

Briefly, the goal of the matching method in the present context is to estimate the effect of market participation on the number of GARP violations of market participants (the so-called "treatment effect on the treated"). To do so, one must estimate the missing counterfactual: the number of GARP violations a market participant *would have committed* had the subject not participated in the market. This is accomplished for each market participant by weighting the outcomes of experimental subjects not participating in the market, where the weights reflect the "similarity" of the market participant and non-participants in terms of observable characteristics (hence, PSM is a "selection on observables" estimation technique). To measure similarity, we use the estimated propensity score (i.e., the predicted probability of each individual receiving the treatment conditional on the observable attributes), estimated using a probit model where the dependent variable is a binary variable indicating whether the subject is a market participant and the covariates include those in x plus additional higher order and interaction terms following the guidance provided in Millimet and Tchernis [2008]. Upon estimating the propensity score, we utilize kernel matching (using the Epanechnikov kernel with a bandwidth of 0.05) and obtain standard errors via bootstrap.[11]

Empirical results from the various econometric specifications are presented in Tables 3 and 4. Table 3 utilizes the pooled, full sample of subjects; thus, *PART* = 1 for subjects participating in the sportscard market prior to the experiment (*SPORTS*), as well as those that entered the market over the course of the experiment (even if only to sell/trade one's parting gifts). To circumvent the potential endogeneity of market participation, Table 4 displays the results obtained utilizing only the sub-sample of non-participants in the market at the time of the initial experiment (*MALLgift* and *MALLnogift*), excluding the 23 non-compliers in round two. Furthermore, Table 4 includes an additional model which uses all observations in the *MALLgift* and *MALLnogift* treatment groups (compliers, partial compliers, and noncompliers), and instruments for *actual* market participation with *intended* participation. In other words, we use D as an

[11] For further details, the reader is referred to Heckman et al. [1997] or Smith and Todd [2005].

instrumental variable (IV) for *PART*, where *D* is defined as unity if the subject belongs to the *MALLgift* group, and zero if the subject is a member of the *MALLnogift*. Such an instrument has been utilized in other contexts (see, e.g., Dee [2004]), and is commonly referred to as the "intent to treat". Comparison of the IV and non-IV results informs us about the randomness of the 23 experimental noncompliers.

Table 3. Determinants of GARP Violations: Pooled Sample

	Independent Variable						
Model	**Age**	**Gender**	**Market Participant**	**Inverse Mills' Ratio**	**LM Test RE**	**Hausman Test (FE v. RE)**	***N***
(1) OLS w/ robust standard errors	-0.52† (0.03)	0.20 (0.20)	-1.05† (0.18)				1223
(2) GLS-RE	-0.49† (0.03)	0.19 (0.23)	-0.91† (0.17)		$\chi^2(1)$=47.96 [p=0.00]	$\chi^2(2)$=11.51 [p=0.00]	1223
(3) FE	-0.05 (0.18)		-1.02† (0.26)				1223
(4) Heckman Selection	-0.54† (0.05)	0.19 (0.20)	-1.05† (0.20)	0.70 (1.37)			1560
(5) Poisson w/ robust standard errors	-0.15† (0.01)	0.06 (0.05)	-0.32† (0.06)				1223
(6) Zero-Inflated Poisson w/ robust std. errors							1223
(a) Violations = 0	0.32† (0.03)	-0.25 (0.17)	0.29‡ (0.15)				
(b) # Vltns. ≥ 0	-0.06† (0.01)	0.002 (0.04)	-0.20† (0.04)				
(7) PSM			-1.23† (0.24)				1223

Notes: Dependent variable is the number of GARP violations. OLS = Ordinary Least Squares; GLS = Generalized Least Squares; RE = random effects; FE = fixed effects; *N* = number of observations. Robust standard errors are heteroskedasticity-consistent. The Heckman selection model uses a quadratic in mall trips per month and state/experiment interacted with a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically significant at the $p < 0.01$ level in each specification). Models (1) – (6) also include a dummy variable for state/experiment. Propensity score matching (PSM) imposes the common support; standard errors obtained by bootstrap (200 repetitions) using kernel matching (Epanechnikov kernel with a bandwidth of 0.05). The propensity score is estimated via probit and includes a cubic in age, gender, state/experiment, gender interacted with a cubic in age, and state/experiment interacted with a cubic in age. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Five conclusions emerge from the empirical results in Tables 3 and 4. First, utilizing the pooled, full sample (Table 3) indicates that participants in the sportscard market exhibit considerably more rational behavior as measured by the number of GARP violations. This finding is robust across a wide range of econometric methodologies (OLS, GLS, FE, Heckman selection, Poisson, ZIP, and PSM); the point estimates are statistically significant in every case.

Beyond the estimated effect of market participation, the fact that Heckman selection model yields *no statistically meaningful evidence* of non-random attrition is particularly noteworthy; the coefficient on the inverse Mills' ratio is statistically insignificant at conventional levels ($\hat{\pi} = 0.70$; s.e. = 1.37). Note, even if the coefficient on the selection correction term were statistically significant, inclusion of this term eliminates the bias arising from non-random attrition assuming the model is well-specified. This is consistent with Figure 3, which reveals little difference in the distribution of GARP violations in round one by attrition status.

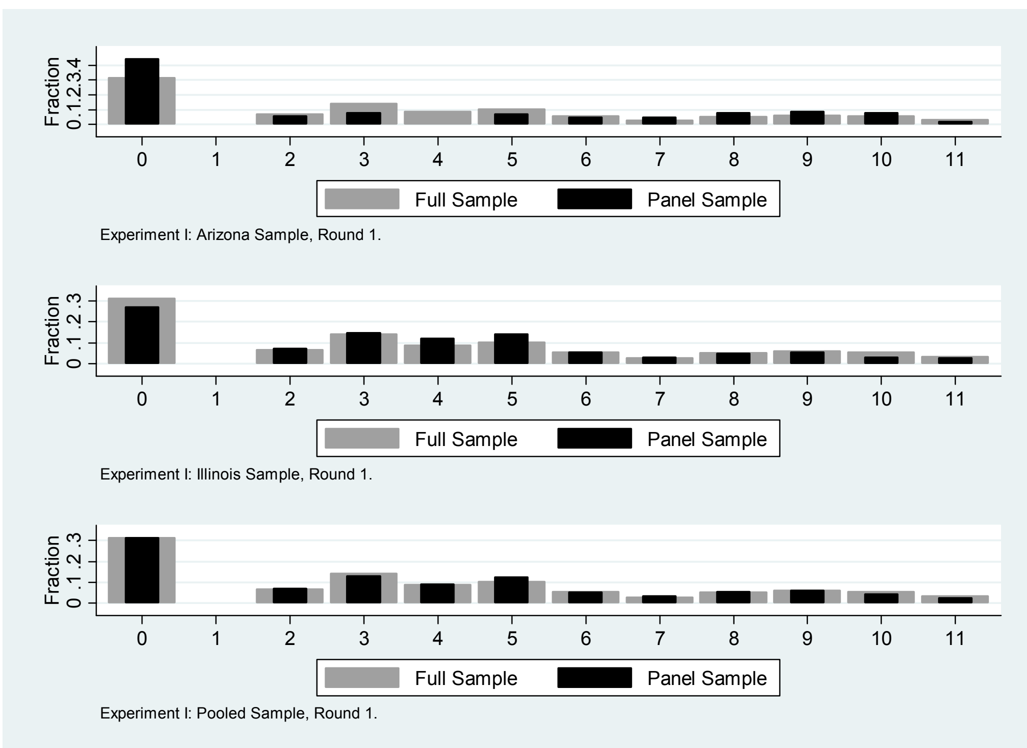

Figure 3. GARP Violations in Round 1 by Attrition Status

Given the critical importance of this result, we undertake two additional tests of non-random attrition. First, we estimate semiparametric versions of the selection model. Specifically, rather than including the inverse Mills' ratio, which relies on the assumption of joint normality, we estimate $\theta_{it} = z_{it}\hat{\gamma}$ using a first-stage probit model, but then include a fifth order polynomial of θ_{it} in the second-stage regression model as a flexible approximation to an unknown correction term. The results are virtually identical, and we fail to reject the null that the coefficients on the correction terms are jointly zero at conventional levels of statistical significance. Second, we implement a falsification test utilizing only the 420 subjects observed in both rounds. Specifically, we regress the number of GARP violations *in round one* on age, gender, state, and market participation *in round two*. Since market participation in round two cannot influence rational behavior in the preceding round, the coefficient on participation should be statistically insignificant. A statistically significant coefficient would suggest bias arising from non-random attrition and/or endogenous market participation in round two. However, the coefficient is *positive* and statistically insignificant (0.17, s.e. = 0.29). As a result, not only does our analysis *yield* statistically significant effects where one might expect, it also *fails to yield* statistically significant effects where none should exist.

Second, the impact of market participation on the development of rational behavior holds even when we restrict attention to the sub-sample exogenously induced into such participation (Table 4).[12] The effect is reduced by more than one-half, however. This is consonant with roughly 60% of the effect in Table 3 being attributable to unobservable attributes positively associated with rational behavior and participation in the sportscard market (i.e., selection effects) and/or significant effects of the duration of market participation, and the remaining 40% indicating a still sizeable effect of market participation on rational choice behavior.[13] Furthermore, the Heckman selection model continues to yield no statistically meaningful evidence of non-random attrition; the coefficient on the inverse Mills' ratio is statistically insignificant at conventional levels ($\hat{\pi}$ = 0.87; s.e. = 1.25).

[12] Note, Table 4 and subsequent tables do not present the FE estimates when a Hausman test fails to reject the random effects assumption. FE results are available upon request.

[13] The fact that our falsification test described above *and* the IV results discussed below provide little evidence of non-random selection into market participation, but we do find significant effects of the duration of market participation (also discussed below) suggests that the decline in magnitude is due to the lower intensity of market participation among the subjects included in the sample in Table 4, and not selection effects.

Table 4. Determinants of GARP Violations: Experimental Sub-Sample (Pooled Sample)

Model	Independent Variable							
	Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	LM Test RE	Hausman Test (FE v. RE)	*N*
(1) OLS w/ robust	-0.55†	0.05	-0.41†					1078
standard errors	(0.03)	(0.21)	(0.21)					
(2) IV w/ robust	-0.56†	0.08	-0.15		$\chi^2(1)=0.39$			1101
standard errors	(0.03)	(0.21)	(0.55)		[p=0.53]			
(3) GLS-RE	-0.53†	0.03	-0.44†			$\chi^2(1)=44.02$	$\chi^2(2)=2.95$	1078
	(0.04)	(0.24)	(0.20)			[p=0.00]	[p=0.23]	
(4) Heckman	-0.57†	0.05	-0.39	0.87				1383
Selection	(0.05)	(0.21)	(0.24)	(1.25)				
(5) Poisson w/ robust	-0.15†	0.02	-0.12‡					1078
standard errors	(0.01)	(0.05)	(0.06)					
(6) Zero-Inflated Poisson w/ robust std. errors								1078
(a) Violations = 0	0.35†	-0.11	-0.39‡					
(0.03)	(0.17)	(0.21)						
(b) # Vltns. ≥ 0	-0.06†	-0.002	-0.18†					
	(0.01)	(0.04)	(0.04)					
(7) PSM			-0.42†					1078
			(0.21)					

Notes: Dependent variable is the number of GARP violations. Experimental sub-sample includes non-participants in the market at the time of round one (*MALLgift* and *MALLnogift*). All models also include a dummy variable for state. Model (2) utilizes the entire sub-sample; all remaining models use only experimental compliers and partial compliers. The Heckman selection model uses a quadratic in mall trips per month and state/experiment interacted with a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically significant at the $p < 0.01$ level in each specification). Models (1) – (6) also include a dummy variable for state/experiment. For further details, see Table 3 and text. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Third, the IV point estimate in Table 4 remains negative, but is smaller in magnitude and statistically indistinguishable from zero due to a noticeable loss in precision.[14] However, we fail to reject the null that *actual* market participation is exogenous; thus, the IV results are less efficient than the other models.[15]

[14] Given the relatively low rate of non-compliance, it is not surprising that the first-stage regressions indicate that the intent to treat, *D*, is a highly significant determinant of actual participation ($p = 0.00$).

[15] Given the finding that noncompliers are random, we re-estimated the models in Table 4 using the total experimental sub-sample (compliers and noncompliers). Empirical results are

Nonetheless, we also re-do our falsification test from above; we estimate by IV the effect of market participation *in round two* on the number of GARP violations *in round one* using the 376 subjects in the experimental sub-sample observed in both rounds. Again, the coefficient on participation should be statistically insignificant, and it is (0.06, s.e. = 0.17).

Finally, in addition to being statistically significant, the results confirm our hypothesis that market participation is "economically" meaningful, as the (OLS, GLS, Heckman selection-corrected, and PSM) point estimates in Table 3 (Table 4) imply that market participation reduces the number of GARP violations by roughly one (0.4). Given that the mean number of violations in the full sample is 3.6, even a reduction of 0.4 is sizeable given the nature and duration of the market experience. Coefficient estimates from the Poisson and ZIP models, while interpreted differently, suggest effects of similar magnitudes.

A related third result relates to age and rational choice behavior:

> *Result 3: Age and rational choice behavior are directly related: the number of GARP violations decreases by roughly one-half per year, ceteris paribus.*

Harbaugh et al. [2001] document that behavior of seven-year-olds is less consistent with utility maximization than behavior of eleven-year-olds: 75% (40%) of seven- (eleven-) year-olds incur at least one GARP violation. Empirical results contained herein extend this finding by showing that (i) older subjects tend to exhibit more rational behavior on average over a wider range (ages 6-18), and (ii) market participation offsets the effect of youth on rational choice behavior. Specifically, the beneficial aspects of the aging process are documented across all econometric specifications in Tables 3 and 4, indicating that the number of GARP violations decreases by one-half per year *ceteris paribus*.[16] As a result, even when focusing on the experimental sub-group (Table 4), we find that market experience of the type measured herein has an effect on rational behavior equivalent to that of one year of aging.

Given these first three insights, a natural query revolves around the intensity of market experience. A further insight results:

> *Result 4: The intensity of market participation and rational choice behavior are directly related: one-time experience in the market has no impact on GARP violations, while repeated engagement reduces the number of GARP violations by close to one.*

---

qualitatively similar, suggesting if anything, larger impacts of market participation on rational choice.

[16] As stated previously, the Poisson and ZIP models also yield statistically significant estimates of the effects of age on GARP violations, but the coefficients are interpreted differently.

Table 5. Determinants of GARP Violations: Experimental Sub-Sample Excluding *MALLgift* Partial Compliers (Pooled Sample)

Model	Independent Variable: Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	LM Test RE	Hausman Test (FE v. RE)	*N*
(1) OLS w/ robust standard errors	-0.56† (0.04)	-0.01 (0.23)	-0.84† (0.27)					983
(2) IV w/ robust standard errors	-0.57† (0.04)	0.05 (0.24)	-0.45 (1.03)		$\chi^2(1)=0.17$ [p=0.68]			1006
(3) GLS-RE	-0.54† (0.04)	-0.07 (0.25)	-0.96† (0.27)			$\chi^2(1)=36.29$ [p=0.00]	$\chi^2(2)=3.47$ [p=0.18]	983
(4) Heckman Selection	-0.59† (0.05)	-0.01 (0.23)	-0.79† (0.33)	0.75 (1.39)				1291
(5) Poisson w/ robust standard errors	-0.15† (0.01)	0.001 (0.06)	-0.25† (0.09)					983
(6) Zero-Inflated Poisson w/ robust std. errors								983
(a) Violations = 0	0.36† (0.03)	0.002 (0.18)	0.05 (0.26)					
(b) # Vltns. ≥ 0	-0.06† (0.01)	0.004 (0.04)	-0.22† (0.06)					
(7) PSM			-0.85† (0.27)					983

Notes: Dependent variable is the number of GARP violations. Model (2) utilizes noncompliers; all remaining models exclude experimental noncompliers. All models also include a dummy variable for state. The Heckman selection model uses a quadratic in mall trips per month and state/experiment interacted with a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically significant at the $p < 0.01$ level in each specification). Models (1) – (6) also include a dummy variable for state/experiment. For further details, see Tables 3 and 4 and text. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

To provide evidence of this result, consider the analysis of the experimental sub-sample in Table 4 that explicitly treated compliers and partial compliers in the *MALLgift* group identically. Clearly, though, they are not identical, as the extent of market participation is much greater for compliers (those who became regular market participants during the course of the experiment) than for partial compliers (those whose only experience in the market is via buying/selling/trading the contents of their gift bag). Of initial interest in this case is that average GARP violations are lower for compliers in *MALLgift*

(2.82) than partial compliers (3.62); the difference is statistically significant at conventional levels (t = 2.15, p = 0.03). However, the difference between compliers and partial compliers may (at least partially) reflect selection effects.

To assess the role of market intensity more formally, we include Table 5, which is identical to Table 4 except partial compliers are excluded from the sample.[17] As a result, the effect of participation is identified from the variation in GARP violations across *MALLgift* compliers and *MALLnogift* compliers, where now the only subjects with *PART* = 1 are subjects in the *MALLgift* group who became regular participants in the sportscard market. The IV results (model 2) are obtained from the sample that excludes partial compliers, but utilizes the 23 noncompliers in the *MALLgift* and *MALLnogift* groups and instruments for participation with the "intent to treat" variable.[18]

Given the difference in intensity of market participation between the compliers and partial compliers in the *MALLgift* group, comparison of the results to those in Table 4 provide empirical evidence of the impact of market intensity. Such a comparison indicates that intensity matters a great deal; the coefficients in Table 5 are 2-3 times larger in magnitude than the corresponding coefficient in Table 4. Thus, the results found in Table 4 are driven by the effect of market participation on subjects who were exogenously induced into repeatedly entered the sportscard market. In light of these findings, one may interpret the effects of participation documented in Table 4 as a weighted average of the effects on compliers and partial compliers.

Combining Results 2 and 4 leads to a fifth insight:

Result 5: The market and the intensity of market participation play an important role in the development of rational choice behavior.

While the preceding empirical analysis demonstrates the impact of market participation on the *level* of rational behavior, market interaction also plays a crucial role in the *learning* behavior of individuals. For evidence of this result, we present Table 6, which displays the breakdown of the *change* in GARP violations between the two rounds of the experiment by treatment group. Overall, the mean number of GARP violations declined by more than one-half from the first to the second round, with roughly 30% (42%) of respondents registering no change (a decline) in the number of violations.

While the sample as a whole exhibited behavior consonant with the learning of rational behavior over time, the magnitude of the learning rates vary

[17] In the interest of brevity, we do not report results retaining partial compliers but excluding compliers. Results are available upon request, but fail to yield any statistically significant effects of market experience.

[18] The coefficient on the instrument in the first-stage regression are highly significant ($p < 0.01$).

tremendously across subject groups, and this variation allows one to identify the impact of market participation on the learning process. First, subjects participating in the market prior to round one of the experiment (*SPORTS*) experienced one-half fewer GARP violations, on average, in round two. Since these respondents did not change their market participation, and any bias involved from these individuals self-selecting into the market likely remains constant temporally, it is not surprising that the decline in violations is close to the predicted decline in GARP violations from an increase in age by more than seven months, as detailed in Tables 3-5.

Table 6. Summary of Learning Rates: Pooled Sample

Group	Change in GARP Violations	H_O: Equality of Means Across Participants & Non-Participants	*N*
Full Sample	-0.58 (3.10) [30.24%, 41.67%]		420
Group 1 (*MALLgift*):			
Non-Participants	3.40 (4.10) [40.00%, 0%]	*t* = 3.05 [p = 0.00]	5
Participants	-0.94 (3.12) [26.80%, 47.94%]		194
Group 2 (*SPORT*):			
Participants	-0.50 (2.80) [47.72%, 34.09%]		44
Group 3 (*MALLnogift*):			
Non-Participants	-0.10 (2.91) [28.93%, 35.85%]	*t* = 3.19 [p = 0.00]	159
Participants	-2.44 (3.28) [33.33%, 55.56%]		18

Notes: Figures represent mean change in GARP violations. Standard deviations are in parentheses. Percentage of observations with no change in GARP violations, followed by percentage of observations with a decrease in GARP violations, is in brackets. *P*-values are in brackets beneath *t*-tests. Group 1 includes non-participants in the market at the time of round one who were randomly given $25 worth of free cards after round one was completed. Group 2 includes subjects who were already in the market at the time of round one. Group 3 includes non-participants in the market at the time of round one who were not given any free cards after round one was completed.

Because the *SPORTS* group participated in the market prior to round one, however, they provide no information about the effect of participation on the development of rational choice, which is a primary goal of the experiment. Examination of the *MALLgift* and *MALLnogift* groups enables us to estimate the impact of market participation. In the *MALLgift* group, subjects exogenously induced into market participation between the two experimental rounds exhibited behavior consistent with the learning of rational behavior, with almost 50% of compliers and partial compliers in the *MALLgift* group registering a decrease in the number of GARP violations and the mean number of violations declining by nearly one. GARP violations for noncompliers in the *MALLgift* treatment increased by 3.4 on average; a *t*-test rejects the null hypothesis that the change in GARP violations is equal across non-participants and new market participants at the $p < 0.01$ level.

We also observe a statistically significant difference between compliers (non-participants) and noncompliers (market participants) in the *MALLnogift* group, with noncompliers who entered the market without inducement incurring almost 2.5 fewer GARP violations in round two compared to their round one performance, whereas compliers that remained out of the sportscard market exhibited almost no change on average in their level of rational choice. Again, a *t*-test rejects the null hypothesis that the change in GARP violations is equal across non-participants and new market participants at the $p < 0.01$ level.

More relevant, however, for deducing the effect of exogenous market participation is the comparison between compliers and partial compliers in the experimental group *MALLgift* and compliers in the *MALLnogift* group, as these observations most closely conform to the "ideal" treatment and control groups. Here, the treatment (control) group registered 0.94 (0.10) fewer GARP violations on average in round two. Thus, the estimated treatment effect of market participation is -0.84; a standard *t*-test rejects the null that the treatment effect equals zero at the $p = 0.01$ level.

This is an important finding, as the differencing over time removes any time-invariant unobservables related to the decision to (partially) comply. Only if compliers or partial compliers are a nonrandom sub-sample of the *MALLgift* and *MALLnogift* treatment groups, and the source of this non-randomness is time-varying, will the previous comparisons confound market and selection effects. Yet, there may still be a bias imparted due to non-random attrition. Moreover, even though the IV results in Table 4 failed to reject exogeneity of actual market participation in the experimental sub-sample (using compliers, partial compliers, and noncompliers) with respect to the *level* of rational behavior, this does not necessarily imply that compliance is exogenous with respect to *learning* (we return to this below).

Table 7. Determinants of Learning Rates by Experiment: Pooled Sample

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	*N*
	Independent Variable					
Full Sample						
(1) OLS	0.10‡	0.39	-0.84†			419
	(0.06)	(0.35)	(0.31)			
(2) Heckman Selection	0.16†	0.56	-0.95†	-2.90†		780
	(0.07)	(0.41)	(0.31)	(1.44)		
(3) PSM			-0.63			419
			(0.40)			
Experimental Sub-Sample						
(4) OLS	0.09	0.59	-0.74†			352
	(0.07)	(0.38)	(0.32)			
(5) IV	0.09	0.61‡	-0.79‡		$\chi^2(1)=7.86$	375
	(0.06)	(0.37)	(0.42)		[p=0.01]	
(6) Heckman Selection	0.12	0.74‡	-0.92†	-3.15†		680
	(0.07)	(0.43)	(0.32)	(1.25)		
(7) PSM			-0.34			352
			(0.35)			

Notes: Dependent variable is the change in GARP violations from round one to round two. Experimental sub-sample includes non-participants in the market at the time of round one (*MALLgift* and *MALLnogift*). Models (8) and (9) utilize the entire sub-sample; all remaining models using the experimental sub-sample use only experimental compliers and partial compliers. The Heckman selection model uses a quadratic in mall trips per month and state/experiment interacted with a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically significant at the $p < 0.01$ level in each specification). Models (1) – (6) also include a dummy variable for state/experiment. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level. For further details, see Tables 3 and 4 and text.

To address these concerns, we turn to the econometric analysis, with the results provided in Table 7. Table 7 presents some straightforward econometric results, estimating equation (1) via OLS, IV, the Heckman selection model, and PSM, except that now the dependent variable is the *change* in the number of GARP violations from round one to round two.[19] The top panel of Table 7 uses the full sample and provides confirmation of the learning results documented above; estimates contained in rows 1-3 shows that market participants have a larger decrease in GARP violations than do non-participants, and first two estimates are statistically significant at the $p < 0.05$ level. Interestingly, there is

[19] We still include age, gender, and state in the differenced regression.

now evidence of non-random attrition; subjects possessing unobservables associated with greater learning of rational behavior are more likely to remain in the sample. However, correcting this has little impact on the estimated effect of market participation on learning; the point estimate actually *increases* in magnitude.

Focusing on the experimental sub-sample in the bottom panel of Table 7, we find similar results in terms of magnitude and statistical significance when actual participation is treated as exogenous and we utilize only compliers and partial compliers. Moreover, we continue to find evidence of positive selection into the sample; subjects with unobservables associated with greater learning of rational behavior are less likely to drop out of the sample. Again, though, correcting this actually *increases* the magnitude of the estimated effect of market participation. Lastly, the IV results now reject the exogeneity of actual participation at the $p = 0.01$ level, but the estimated coefficient is unaffected (in magnitude) and remains statistically significant at the $p < 0.10$ confidence level.[20] In sum, not only are the experimental estimates of the impact of the market on *learning* in Table 7 statistically significant, the magnitudes are roughly double the magnitudes of the experimental estimates of the impact of the market on the *level* of rational behavior in Table 4.

For our final exercise, we assess whether the intensity of market participation matters on learning. Thus, we proceed as before and differentiate between compliers and partial compliers, first noting that the mean change in GARP violations is greater for compliers in *MALLgift* (-1.54) than partial compliers (-0.32), and this difference is statistically significant at the $p < 0.01$ level.

To examine the role of market participation intensity more formally, Table 8 mimics the final four rows of Table 7 except partial compliers are excluded from the sample.[21] The IV results (row 2) are obtained from the sample that excludes partial compliers, but uses the 23 noncompliers in the *MALLgift* and *MALLnogift* groups and instruments for participation with the "intent to treat" variable.[22]

Comparison of the empirical results in Table 7 and Table 8 indicates that intensity matters for the learning process. The coefficients in Table 8 are negative, near or above one in absolute value, and statistically significant at the $p < 0.05$ confidence level. Thus, as in the levels specifications in Table 5, one time

[20] The first-stage models indicate that intent to treat is a significant determinant of actual participation ($p < 0.01$).

[21] In the interest of brevity, we do not report results retaining partial compliers but excluding compliers. Results are available upon request, but fail to yield any statistically significant effects of market experience.

[22] The coefficient on the instrument in the first-stage regressions is highly significant ($p < 0.01$).

participation in the market conveys no statistically meaningful benefit, but regular market participation is associated with economically significant changes in the development of rational behavior.

Table 8. Determinants of Learning Rates: Experimental Sub-Sample Excluding *MALLgift* Partial Compliers (Pooled Sample)

	Independent Variable					
Model	**Age**	**Gender**	**Market Participant**	**Inverse Mills' Ratio**	**Test of Exogeneity**	***N***
(1) OLS	0.06 (0.08)	0.20 (0.47)	-1.33† (0.39)			257
(2) IV	0.06 (0.07)	0.35 (0.45)	-1.00† (0.45)		$\chi^2(1)$=4.55 [p=0.03]	280
(3) Heckman Selection	0.11 (0.09)	0.78 (0.62)	-1.51† (0.39)	-3.79† (1.64)		588
(4) PSM			-0.96† (0.41)			257

Notes: Dependent variable is the change in GARP violations from round one to round two. Model (2) utilizes noncompliers; all remaining models exclude noncompliers. All models also include a dummy variable for state. The Heckman selection model uses a quadratic in mall trips per month and state/experiment interacted with a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically significant at the $p < 0.01$ level in each specification). Models (1) – (6) also include a dummy variable for state/experiment. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level. For further details, see Tables 3 and 4 and text.

The underlying learning mechanism that leads to the mediation of irrationality as agents gain market experience is an important normative issue that merits serious consideration. In this case, the market experience we examine includes buying, trading, and selling goods in an active marketplace. Hence, it might be such that what agents learned in the marketplace, which includes many arbitrageurs, was to formulate rational and consistent valuations and this behavior transfers to the experimental choice context. The learning occurs because the market reveals to a person that her instinctual behavior is costly, inducing her to behave in a more systematized manner. In a related spirit, Slonim (1999, p. 217) notes that such learning could result because "people may imitate others, people may form habits based on past successful trials or people may employ best reply or stimulus response rules such that the dynamics may lead to more substantively rational choices." And continues, remarking (p. 218) that "Camerer (1990) hypothesizes that people may learn to (properly) think through the task once they have experience with paths of the task (what he calls experiential learning) rather than hypothetically using backward induction to determine (substantively) rational behavior; people may learn to use a more rational rule by learning to incorporate more paths of play."

B. Heterogeneity Across Experiments

To this point, our analysis has pooled the data from the two experiments. An issue that is becoming more salient in the literature relates to whether the response to a particular treatment is heterogeneous across populations. To explore this issue, we re-conduct much of the previous statistical analysis separately by experiment. In the interest of brevity, we place the results in Appendix A, Tables A2-A8, and only provide an abbreviated discussion. Nonetheless, our analysis provides the following insight.

> *Result 6: Although the qualitative impact of market participation on the change in rational behavior is similar across both experiments, the magnitude of the impact is heterogeneous.*

Tables A2-A5 present the results assessing the effect of market participation on the level of rational choice. Using the full sample (subjects in the *MALL* and *SPORTS* treatment groups) in Table A3 indicates an effect of market participation that is roughly eight times larger in magnitude in the Arizona sample. When we focus only on the experimental sub-sample (compliers, partial compliers, and non-compliers) in Table A4, the point estimates continue to be about six to seven times larger in the Arizona sample; in fact, the point estimates are typically not statistically significant at conventional levels in the Illinois sample. However, Table A5 indicates that market intensity matters in both experiments; the effect of market participation is statistically significant in most cases in both experiments when experimental partial compliers are excluded from the analysis. Again, though, the effect is much larger – four to five times as large – in the Arizona sample. Finally, it is noteworthy that there is no evidence of non-random attrition, nor evidence of non-random compliance, in either sample.

Tables A6-A8 present the results assessing the effect of market participation on learning behavior. Using the full sample (subjects in the *MALL* and *SPORTS* treatment groups) in Table A7 indicates an effect of market participation that is roughly two to three times larger in magnitude in the Arizona sample. When we focus only on the experimental sub-sample (compliers, partial compliers, and non-compliers) in Table A7, the point estimates do not change much in terms of magnitude in either sample, but fewer estimates are statistically significant at conventional levels. In addition, there is now evidence of non-random attrition and compliance in the Arizona sample. Finally, Table A8 indicates that market intensity matters in both experiments; the effect of market participation is statistically significant in all but one case across both experiments when experimental partial compliers are excluded from the analysis. In addition, the effect of market participation continues to be about three times larger in the Arizona sample.

What might account for the heterogeneous effects found across the two experiments? One possible explanation may lie in the difference in the age structure of the two samples. As seen in Table A1, subjects in the Illinois sample are roughly one year older on average across the two rounds. Because older subjects are likely to have acquired more experience in various markets during their lifetime, the effects of exogenously induced participation into the sportscard market may have a smaller effect if there are diminishing returns. To assess this, we re-estimated the Heckman selection model in Table 4 (row 4) and Table 7 (row 6) using the pooled, experimental sub-sample. However, we now also include an interaction term between market participation and age. In both cases, the interaction term is positive and statistically significant at the $p < 0.05$ confidence level (indicating a smaller, negative effect of market participation on the level of or change in the number of GARP violations); the point estimate is 0.22 in Table 4 and 0.28 in Table 7. Thus, the age difference across the experiments is one contributing factor to smaller effects found in the Illinois sample.

C. Robustness to Additional Covariates

As a final sensitivity check, we re-estimate the models using data on the experimental sub-sample from the Illinois experiment, but make use of some the additional covariates available. Tables A9-A10 examine the level of rational behavior; Tables A11-A12 examine learning. All models now include a dummy variable indicating if math or science is the individual's favorite subject, a dummy variable indicating attendance at a private school, an index reflecting self-reported perception of the individual by one's teachers, three race dummies, and a quadratic for hours spent watching television per week in addition to the covariates previously included.

When assessing the level of rational behavior (Tables A9-A10), the results are qualitatively unchanged. When assessing the impact of market participation on learning, the point estimates actually increase in absolute value. For example, the point estimate in the Heckman selection model changes from -0.64 in Table A7 (Panel II, row 6) to -0.82 in Table A11 (row 3), and -1.05 in Table A8 (Panel II, row 3) to -1.37 in Table A12 (row 3).

IV. Chamberlin Market Experiment

A. Experimental Design

Whether, and to what extent, individual-level choice violations – such as those observed above – impact the operation of markets is an open issue that undoubtedly depends critically on the market institution. For example, making use of the Walrasian tatonnement mechanism, Becker [1962] proves that several fundamental features of economics, such as correctly sloped supply and demand schedules, may result even when agents are irrational, serving to sufficiently relax the utility-maximizing assumption inherent in economic modeling. Similarly, Smith [1962] showed that the Walrasian tatonnement mechanism was not necessary for market outcomes to approximate equilibrium predictions if agents were given a chance to learn. In such cases, a double oral auction could produce outcomes that were highly efficient.

In this section, rather than examine behavior in oral double-auctions, we explore market outcomes in multi-lateral bargaining contexts. Our market treatments are similar in spirit to Chamberlin [1948], who over a half-century ago executed what is believed to be the first experiment to test neoclassical competitive market theory. Rather than examining behavior of undergraduate students milling around the classroom, however, we examine behavior in an actual marketplace: the sportscard market (see List [2004a; 2004b] for related experiments; the experimental design description below follows List [2004a]).

In these market sessions, each participant's experience typically followed four steps: (1) consideration of the invitation to participate in an experiment, (2) learning the market rules, (3) actual market participation, and (4) conclusion of the experiment and exit interview. In Step 1, before the market opened, a monitor approached dealers at the sportscard show and inquired about their interest in participating in an experiment that would take about 60 minutes during the sportscard show. Since most dealers are accompanied by at least one other employee, it was not difficult to obtain their agreement after it was explained to them that they could earn money during the experiment.

The novelty of the experimental design in this section is that the non-dealer subject pool included children aged 12-14 who participated in the GARP exercise discussed above. We are therefore able to link actual behavior in one experimental environment to behavior in another market-based environment to examine the effects of rationality on market outcomes.

Upon agreement of the prerequisite number of dealers (sellers) and non-dealers (children buyers) to participate, monitors thoroughly explained the experiment rules in Step 2. The experimental instructions were standard, and borrowed from Davis and Holt [1995, pp. 47-55] with the necessary adjustments.

Before continuing, a few key aspects of the experimental design should be highlighted. First, all individuals were informed that they would receive a $10 participation fee upon completion of the experiment. In addition, following Smith [1965], to ensure that marketers would engage in a transaction at their reservation prices, we provided a $0.05 commission for each executed trade for both buyers and sellers.

Second, the child participants (non-dealers) were informed that the experiment consisted of five periods and that they would be buyers in the experiment. In each of the five periods, we used Smith's [1976] induced value mechanism by providing each buyer with a "buyer's card" containing a number – known only to that buyer – representing the maximum price that he or she would be willing to pay for *one* unit of the commodity. Dealers were informed that they would be sellers in the market and, in each of the five periods, that each would be given a "seller's card" containing a number – known only to that seller – representing the minimum price that he or she would be willing to sell <u>one</u> unit. Importantly, both buyers and sellers were informed that this information was strictly private and that reservation values would change each period. They were also informed about the number of buyers and sellers in the market (explained more fully below) and informed that agents may have different values.

Third, the monitor explained how market earnings would be determined. The difference between the contract price and the maximum reservation price determined the market earnings of buyers; the difference between the contract price and the minimum reservation price determined sellers' earnings. Several examples illustrated the irrationality associated with buying (selling) the commodity above (below) the induced value.

Fourth, the homogeneous commodities used in the experiment were 1982 *Topps* Ben Oglivie baseball cards, upon which moustaches had been drawn, thereby rendering the cards valueless outside of the experiment. Consequently, the assignment given to buyers was clear: enter the marketplace and purchase the "moustache" card for as little as possible. Likewise, the task confronting sellers was equally as clear, and an everyday occurrence: sell the Oglivie "moustache" card for as much as possible. The cards and participating dealers were clearly marked to ensure buyers had no trouble finding the commodity of interest. Finally, buyers and sellers engaged in two five-minute practice periods to gain experience with the market.

In Step 3, subjects participated in the market. Each market session consisted of five, ten minute market periods. After each ten-minute period, a monitor privately gathered the buyers and gave each a new buyer's card; a different monitor privately gave each seller a new seller's card. Throughout the market process careful attention was paid to prohibit discussions between agents that could induce collusive outcomes. Much like the early writers in this area, we wanted to give neoclassical theory its best chance to succeed. Step 4 concluded the experiment: after subjects completed a confidential survey, they were paid their earnings in private.

Table 9. Market Experimental Design (Arizona Sample)

	Low GARP Violations	High GARP violations
No market experience	*NElowGARP* 12 buyers, 12 sellers *n* = 24	*NEhighGARP* 12 buyers, 12 sellers *n* = 24
High market experience	*ElowGARP* 12 buyers, 12 sellers *n* = 24	*EhighGARP* 12 buyers, 12 sellers *n* = 24

Notes: Each cell represents one unique treatment. For example, "*NElowGARP*" in row 1, column 1, denotes that 12 "non-experienced low GARP violation" buyers and 12 sellers competed in markets for five periods. No subject participated in more than one treatment.

We follow this simple procedure in each of four market treatments, which are summarized in column 1 of Table 9 and can be read as follows: row 1, column 1 of Table 9 contains treatment *NElowGARP*, denoting that the 12 buyers had no sportscard market experience prior to our experiments and had a low number of GARP violations (fewer than 3.5) in round 2 of the GARP treatment described above. Treatment *NEhighGARP* included 12 buyers who also had no sportscard market experience prior to our experiments, but had a high number of GARP violations (more than 3.5) in round 2 of the GARP treatment described above. The two experienced subject treatment groups, *ElowGARP* and *EhighGARP*, were chosen analogously to the two previous treatment groups.

In all four treatments, the market is composed of twelve buyers (sellers), each with unit demand (supply). Figure 4 presents buyer and seller induced values, which are taken from Davis and Holt [1995, pp. 14-15]. In Figure 4, each step represents a distinct induced value that was given to buyers (demand curve) and sellers (supply curve). The efficient, perfectly competitive outcome in this treatment yields \$37 in rents per period, which occurs at the static price/quantity equilibrium of *Price* = \$13–\$14 and *Quantity* = 7 as predicted by competitive price theory. This is the extreme point of the intersection of the buyer and supplier rent areas in Figure 4.

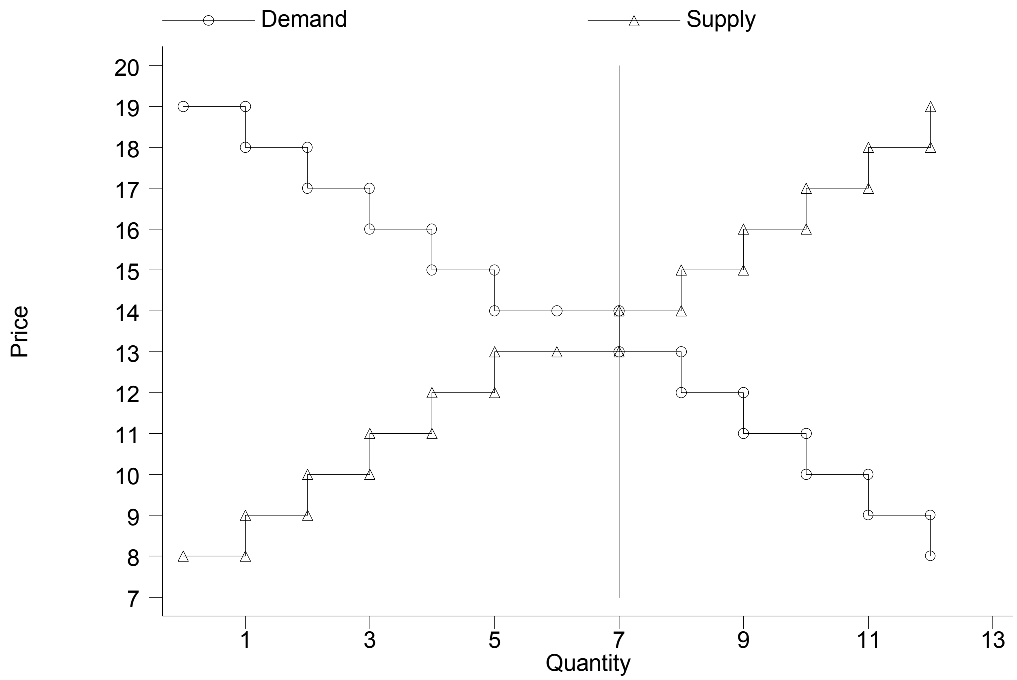

Figure 4. Buyer and Seller Induced Values for Market Experiment

B. Experimental Results

Table 10 contains summary statistics for the experimental data. Entries in Table 10 are at the period level and include the average price and its standard deviation, quantity traded, total buyer and seller per-period profits, and a measure of efficiency (total rents captured divided by available rents). Table 10 can be read as follows: on average, in period 1 of the *NElowGARP* session 7.3 cards were purchased at a trading price of $13.53 (std. dev. = 1.9). Total buyer and seller profit was $15.75 and $17.25, respectively, and traders captured 89% of the available rents for the period. Perusal of the data summary for the various treatment groups yields our major finding in this section:

> *Result 7: Even in markets populated entirely by irrational buying agents, prices and quantities tend toward the intersection of the supply and demand functionals.*

Table 10. Experimental Results: Chamberlin Market (Arizona Sample)

	Market Period				
Treatment	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
NElowGARP					
Average price	16.75	15.30	14.20	13.71	14.00
Std. deviation	(0.50)	(0.45)	(0.45)	(0.76)	(0.89)
Quantity	(Q=4)	(Q=5)	(Q=5)	(Q=7)	(Q=6)
Profits					
Buyers	3.00	8.50	10.00	11.00	12.00
Sellers	25.00	21.50	14.00	20.00	19.00
Efficiency	76%	81%	65%	84%	84%
NEhighGARP					
Average price	18.00	15.75	15.00	14.91	13.71
Std. deviation	(0.0)	(1.5)	(0.00)	(0.66)	(0.76)
Quantity	(Q=2)	(Q=4)	(Q=5)	(Q=6)	(Q=7)
Profits					
Buyers	1.00	3.00	10.00	9.50	15.00
Sellers	18.00	17.00	20.00	20.50	16.00
Efficiency	51%	54%	81%	81%	84%
ElowGARP					
Average price	11.50	12.00	13.57	13.86	13.13
Std. deviation	(0.71)	(1.4)	(0.97)	(1.5)	(0.84)
Quantity	(Q=2)	(Q=2)	(Q=7)	(Q=7)	(Q=8)
Profits					
Buyers	5.00	4.00	18.00	12.00	21.00
Sellers	4.00	4.00	17.00	20.00	14.00
Efficiency	24%	21%	94%	86%	95%
EhighGARP					
Average price	14.75	13.50	13.64	13.67	13.07
Std. deviation	(3.4)	(1.7)	(1.9)	(1.2)	(1.3)
Quantity	(Q=4)	(Q=6)	(Q=7)	(Q=6)	(Q=7)
Profits					
Buyers	7.00	17.00	14.50	19.00	20.50
Sellers	8.00	14.00	14.50	15.00	15.50
Efficiency	41%	84%	78%	92%	97%

Note: Figures in table represent averages in each treatment. Summary statistics are provided for period price, standard deviation of period price, quantity traded in period, and total buyer and seller profits in each period. For example, in the "*NElowGARP*" treatment, period 1 had an average trading price of $16.75 with a standard deviation of $0.50. Four trades were made, and total buyer (seller) profit was $3.00 ($25.00) for the period. Market efficiency was 76 percent.

By observing average trading prices across the five market sessions in Table 10, one can see Result 7 most vividly. For each of the four treatment groups, the average price is within the neoclassical range ($13–$14) and quantity levels are very close to neoclassical expectations (Q = 7) in many periods. Average efficiency rates are also quite high, reaching 97% in the final period of the *ElowGARP* treatment, as traders learn particulars about the market. Moreover, the data reveal that after a few periods, prices rapidly converge to the competitive level, and in *each* of the four treatments, prices "settle down" to approximately neoclassical expectations in the final two periods. Akin to the results in Forsythe et al. [1992] and Gode and Sunder [1993], who explore related issues in much different environments, these results highlight the strength of the invisible hand; competitive conditions prevail even when irrational agents are the sole buyers in the marketplace.[23]

Another insight gained from Table 10 concerns the allocation of rents. First, there is a tendency for experienced buyers to wrest more than sellers in the final market period: $21.00 (20.50) versus $14.00 ($15.50) (see ElowGARP and EhighGARP entries). Second, comparing outcomes across buyer types, an interesting observation is that GARP violations influence the allocation of rents among the inexperienced buyer group, but only in the early periods of the market. This result suggests that for one-shot games, or markets that allow minimal experience, GARP violations *might* be an important determinant of rent allocation among inexperienced buyers, but much more data collection must be carried out to make this a formal result.

V. Conclusion

An important issue facing behavioral and experimental economics today is the role of the market in shaping and influencing behavior. Yet, the extent to which experience in naturally-occurring markets influences rational choice behavior has heretofore been unexplored empirically. This study begins by presenting evidence from a set of experimental treatments that exogenously induce market participation by individuals, finding that the market is more powerful than most surmise: using one straightforward test for rationality relying on the theory of revealed preference, we find that the market is a catalyst for this type of rationality. This result is surprising in that our experimental design constitutes a

[23] This result contrasts with List [2004a], who finds that in markets populated entirely by inexperienced sellers and inexperienced buyers, market efficiency is frustrated. Combining these insights suggests that a necessary condition for market outcomes to be efficient in bilateral markets is that either (i) one side of the market – buyers or sellers – must be experienced/rational agents or more stringently that (ii) sellers need to be experienced/rational. This represents an area of interesting future research.

particularly demanding test of learning, as our exploration measures participation in one particular market and rational choice behavior in a *separate*, quite distinct, "market." Insights gained from these treatments suggest that economic rationality is a social, not an individual construct.

These findings naturally lead to an examination of market outcomes in multi-lateral decentralized markets that are populated by irrational and rational buying agents. Such an analysis lends insights into whether rationality influences rent allocation and whether equilibrating properties of markets are affected by the presence of a significant number of irrational agents. Our findings suggest that even in markets populated entirely by irrational actors, several fundamental features of markets, such as price and quantity realizations, meet neoclassical predictions after a few rounds of market experience. In light of recent findings of individual irrationality, this empirical result should have significance to scholars interested in both positive and normative economic modeling and aggregate market outcomes.

Given that the questions posed herein pertain to basic underpinnings of economic modeling, additional research is warranted. As in all experiments, one issue that arises is the generalizability of the results. We explore this along geographic lines in the current study, finding that our qualitative results hold in different regions in the U.S. However, other dimensions of generalizability may also prove profitable for future study. In particular, one may wish to examine if the effects of market participation are heterogeneous across markets and/or agents. If participation in different markets yields different responses in terms of the development of rational behavior, what are the salient features of markets that are associated with the learning of rational behavior? If participation varies across individuals – adults versus children, male versus female, etc. – what accounts for such differences? Exploring different types of rationality and how they change with market experience is also of utmost importance. We reserve these discussions for another occasion.

Appendix A

Table A1. Summary Statistics

Sample	GARP Violations		Age		Gender		Mall Trips Per Month	
	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2
I. Arizona Sample								
MALLgift	4.47	2.95	10.39	11.02	0.35	0.31	6.54	8.69
	(3.93)	(3.51)	(3.12)	(2.93)	(0.48)	(0.47)	(6.25)	(8.18)
	[110]	[42]	[110]	[42]	[110]	[42]	[110]	[42]
MALLnogift	4.46	3.45	10.67	11.39	0.52	0.48	4.18	6.65
	(3.90)	(3.59)	(3.14)	(3.04)	(0.50)	(0.51)	(4.60)	(6.79)
	[109]	[31]	[109]	[31]	[109]	[31]	[109]	[31]
SPORTS	1.74	1.32	10.53	10.65	0.05	0.00	5.40	4.68
	(3.17)	(2.14)	(3.39)	(3.30)	(0.22)	(0.00)	(5.15)	(3.21)
	[58]	[31]	[58]	[31]	[58]	[31]	[58]	[31]
Full Sample	3.90	2.62	10.53	10.72	0.35	0.27	5.37	6.88
	(3.92)	(3.28)	(3.18)	(3.06)	(0.48)	(0.45)	(5.50)	(6.77)
	[277]	[104]	[277]	[104]	[277]	[104]	[277]	[104]
II. Illinois Sample								
MALLgift	4.05	3.29	11.21	12.27	0.23	0.22	4.52	4.26
	(3.30)	(2.33)	(2.59)	(2.49)	(0.42)	(0.41)	(4.84)	(4.26)
	[255]	[157]	[255]	[157]	[241]	[157]	[233]	[151]
MALLnogift	3.75	3.59	11.43	12.45	0.22	0.24	4.61	4.35
	(3.17)	(3.26)	(2.36)	(2.33)	(0.41)	(0.43)	(4.76)	(4.60)
	[267]	[146]	[267]	[146]	[266]	[145]	[253]	[141]
SPORTS	2.45	2.38	12.05	12.54	0.05	0.08	4.95	5.46
	(2.35)	(2.40)	(2.19)	(2.47)	(0.22)	(0.28)	(5.50)	(6.35)
	[20]	[13]	[20]	[13]	[20]	[13]	[19]	[13]
Full Sample	3.84	3.39	11.35	12.36	0.22	0.22	4.58	4.36
	(3.22)	(2.81)	(2.47)	(2.41)	(0.41)	(0.42)	(4.81)	(4.51)
	[542]	[316]	[542]	[316]	[542]	[316]	[505]	[305]

Notes: Numbers represent sample means. Standard deviation is in parentheses. Number of observations is in brackets. Gender: 0 = male; 1 = female. Age was not collected in period 2 for the Illinois sample. Figures based on the age in period plus one. *MALLgift* includes non-participants in the market at the time of round one who were randomly given $25 worth of free cards after round one was completed. *MALLnogift* includes non-participants in the market at the time of round one who were not given any free cards after round one was completed. *SPORTS* includes subjects who were already in the market at the time of round one.

Table A2. Summary of GARP Violations by Experiment

Group	GARP Violations	H_O: Equality of Means Across Participants & Non-Participants	Observations
I. Arizona Sample			
Full Sample	3.55 (3.79) [55.41%]		381 [Round 1: *n* = 277] [Round 2: *n* = 104]
Group 1 (*MALLgift*):			
Round 1			
Non-Participants	4.47 (3.93) [69.09%]		110
Round 2			
Non-Participants	5.00 (7.07) [50.00%]	$t = 0.84$ [$p = 0.40$]	2
Participants	2.85 (3.39) [47.50%]		40
Group 2 (*SPORT*):			
Round 1			
Participants	1.74 (3.17) [29.31%]		58
Round 2			
Participants	1.32 (2.14) [32.26%]		31
Group 3 (*MALLnogift*):			
Round 1			
Non-Participants	4.46 (3.90) [66.97%]		109
Round 2			
Non-Participants	3.78 (3.80) [65.22%]	$t = 0.87$ [$p = 0.39$]	23
Participants	2.50 (2.93) [50.00%]		8

Notes:
1. Figures represent mean GARP violations. Standard deviations are in parentheses. Percentage of observations with non-zero GARP violations or *p*-values is in brackets.
2. Group 1 includes non-participants in the market at the time of round one who were randomly given $25 worth of free cards after round one was completed. Group 2 includes subjects who were already in the market at the time of round one. Group 3 includes non-participants in the market at the time of round one who were not given any free cards after round one was completed.
3. Participants include both regular market participants and one-time participants (partial compliers).

Table A2 (Cont.). Summary of GARP Violations by Experiment

Group	GARP Violations	H_O: Equality of Means Across Participants & Non-Participants	Observations
II. Illinois Sample			
Full Sample	3.68 (3.08) [74.01%]		858 [Round 1: n = 542] [Round 2: n = 316]
<u>Group 1 (*MALLgift*)</u>:			
Round 1			
Non-Participants	4.05 (3.30) [72.94%]		255
Round 2			
Non-Participants	2.33 (2.08) [66.67%]	t = -0.72 [p = 0.47]	3
Participants	3.31 (2.34) [81.82%]		154
<u>Group 2 (*SPORT*)</u>:			
Round 1			
Participants	2.45 (2.35) [60.00%]		20
Round 2			
Participants	2.38 (2.40) [61.54%]		13
<u>Group 3 (*MALLnogift*)</u>:			
Round 1			
Non-Participants	3.75 (3.17) [74.53%]		267
Round 2			
Non-Participants	3.60 (3.28) [69.85%]	t = 0.09 [p = 0.93]	136
Participants	3.50 (3.14) [70.00%]		10

Notes:
1. Figures represent mean GARP violations. Standard deviations are in parentheses. Percentage of observations with non-zero GARP violations or p-values is in brackets.
2. Group 1 includes non-participants in the market at the time of round one who were randomly given $25 worth of free cards after round one was completed. Group 2 includes subjects who were already in the market at the time of round one. Group 3 includes non-participants in the market at the time of round one who were not given any free cards after round one was completed.
3. Participants include both regular market participants and one-time participants (partial compliers).

Table A3. Determinants of GARP Violations: Full Sample by Experiment

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	LM Test RE	Hausman Test (FE v. RE)	N
	Independent Variable						
I. Arizona Sample							
(1) OLS w/ robust	-0.47†	-0.08	-2.36†				381
standard errors	(0.05)	(0.40)	(0.37)				
(2) GLS-RE	-0.48†	-0.10	-2.34†		$\chi^2(1)$=2.98	$\chi^2(2)$=0.88	381
	(0.06)	(0.40)	(0.37)		[p=0.08]	[p=0.64]	
(3) Heckman	-0.48†	-0.09	-2.36†	0.24			554
Selection	(0.07)	(0.39)	(0.37)	(1.52)			
(4) Poisson w/ robust	-0.15†	-0.02	-0.77†				381
standard errors	(0.02)	(0.10)	(0.14)				
(5) Zero-Inflated Poisson w/ robust std. errors							381
(a) Violations = 0	0.22†	-0.04	1.40†				
	(0.04)	(0.26)	(0.26)				
(b) # Vltns. ≥ 0	-0.07†	-0.03	-0.18†				
	(0.01)	(0.06)	(0.07)				
(6) PSM			-2.46†				381
			(0.39)				
II. Illinois Sample							
(1) OLS w/ robust	-0.56†	0.13	-0.31				842
standard errors	(0.04)	(0.23)	(0.19)				
(2) GLS-RE	-0.51†	0.14	-0.27		$\chi^2(1)$=59.92	$\chi^2(2)$=3.91	842
	(0.05)	(0.28)	(0.18)		[p=0.00]	[p=0.14]	
(3) Heckman	-0.60†	0.16	-0.29	1.99			1006
Selection	(0.11)	(0.25)	(0.23)	(4.80)			
(4) Poisson w/ robust	-0.15†	0.05	-0.09‡				842
standard errors	(0.01)	(0.06)	(0.06)				
(5) Zero-Inflated Poisson w/ robust std. errors							842
(a) Violations = 0	0.49†	-0.08	-0.69†				
	(0.05)	(0.21)	(0.23)				
(b) # Vltns. ≥ 0	-0.06†	0.02	-0.22†				
	(0.01)	(0.05)	(0.05)				
(6) PSM			-0.33‡				842
			(0.20)				

Notes: Dependent variable is the number of GARP violations. OLS = Ordinary Least Squares; GLS = Generalized Least Squares; RE = random effects; FE = fixed effects; N = number of observations. Robust standard errors are heteroskedasticity-consistent. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are statistically significant at the $p < 0.01$ level in the Arizona sample only). Propensity score matching (PSM) imposes the common support; standard errors obtained by bootstrap (200 repetitions) using kernel matching (Epanechnikov kernel with a bandwidth of 0.05). The propensity score is estimated via probit and includes a cubic in age, gender, state/experiment, gender interacted with a cubic in age, and state/experiment interacted with a cubic in age. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Table A4. Determinants of GARP Violations: Experimental Sub-Sample

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	LM Test RE	Hausman Test (FE v. RE)	*N*
I. Arizona Sample							
(1) OLS w/ robust standard errors	-0.55†	-0.12	-1.33†				282
	(0.06)	(0.43)	(0.60)				
(2) IV w/ robust standard errors	-0.56†	-0.04	-1.04		$\chi^2(1)$=0.05		292
	(0.06)	(0.44)	(2.04)		[p=0.82]		
(3) GLS-RE	-0.56†	-0.16	-1.40†		$\chi^2(1)$=3.02	$\chi^2(2)$=0.27	282
	(0.07)	(0.43)	(0.57)		[p=0.08]	[p=0.87]	
(4) Heckman Selection	-0.56†	-0.13	-1.30†	0.29			428
	(0.08)	(0.42)	(0.60)	(1.47)			
(5) Poisson w/ robust standard errors	-0.15†	-0.04	-0.37‡				282
	(0.02)	(0.10)	(0.20)				
(6) Zero-Inflated Poisson w/ robust std. errors							282
(a) Violations = 0	0.23†	0.01	0.85‡				
	(0.04)	(0.28)	(0.38)				
(b) # Vltns. ≥ 0	-0.07†	-0.03	-0.03				
	(0.01)	(0.06)	(0.10)				
(7) PSM			-1.21‡				282
			(0.73)				
II. Illinois Sample							
(1) OLS w/ robust standard errors	-0.55†	0.09	-0.17				796
	(0.04)	(0.24)	(0.21)				
(2) IV w/ robust standard errors	-0.56†	0.11	0.01		$\chi^2(1)$=0.13		809
	(0.04)	(0.23)	(0.53)		[p=0.72]		
(3) GLS-RE	-0.51†	0.13	-0.18		$\chi^2(1)$=53.10	$\chi^2(2)$=2.72	796
	(0.05)	(0.28)	(0.20)		[p=0.00]	[p=0.26]	
(4) Heckman Selection	-0.60†	0.13	-0.15	2.40			955
	(0.11)	(0.26)	(0.25)	(4.64)			
(5) Poisson w/ robust standard errors	-0.15†	0.04	-0.05				796
	(0.01)	(0.06)	(0.06)				
(6) Zero-Inflated Poisson w/ robust std. errors							796
(a) Violations = 0	0.48†	-0.05	-1.03†				
	(0.05)	(0.22)	(0.27)				
(b) # Vltns. ≥ 0	-0.06†	0.02	-0.22†				
	(0.01)	(0.05)	(0.05)				
(7) PSM			-0.18				796
			(0.22)				

Notes: Dependent variable is the number of GARP violations. Experimental sub-sample includes non-participants in the market at the time of round one (*MALLgift* and *MALLnogift*). Models (3) and (4) utilize the entire sub-sample; all remaining models use only experimental compliers and partial compliers. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are statistically significant at the $p < 0.01$ level in the Arizona sample only). For further details, see Table A3 and text. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Table A5. Determinants of GARP Violations: Experimental Sub-Sample Excluding Partial Compliers by Experiment

Model	Independent Variable: Age	Independent Variable: Gender	Independent Variable: Market Participant	Inverse Mills' Ratio	Test of Exogeneity	LM Test RE	Hausman Test (FE v. RE)	*N*
I. Arizona Sample								
(1) OLS w/ robust standard errors	-0.56†	-0.32	-2.32†					261
	(0.07)	(0.46)	(0.84)					
(2) IV w/ robust standard errors	-0.57†	-0.27	-2.69		$\chi^2(1)$=0.06			271
	(0.06)	(0.64)	(5.64)		[p=0.94]			
(3) GLS-RE	-0.57†	-0.36	-2.53†			$\chi^2(1)$=1.59	$\chi^2(2)$=1.58	261
	(0.07)	(0.45)	(0.80)			[p=0.21]	[p=0.45]	
(4) Heckman Selection	-0.57†	-0.33	-2.30†	0.26				407
	(0.09)	(0.44)	(0.83)	(1.60)				
(5) Poisson w/ robust standard errors	-0.15†	-0.08	-0.72†					261
	(0.02)	(0.11)	(0.33)					
(6) ZIP w/ robust std. errors								261
(a) Violations = 0	0.22†	0.13	1.21†					
	(0.05)	(0.30)	(0.59)					
(b) # Vltns. ≥ 0	-0.07†	-0.04	-0.17					
	(0.01)	(0.07)	(0.16)					
(7) PSM			-2.29‡					261
			(0.77)					
II. Illinois Sample								
(1) OLS w/ robust standard errors	-0.57†	0.10	-0.48‡					722
	(0.04)	(0.27)	(0.27)					
(2) IV w/ robust standard errors	-0.58†	0.14	-0.22		$\chi^2(1)$=0.06			735
	(0.04)	(0.27)	(0.93)		[p=0.81]			
(3) GLS-RE	-0.52†	0.09	-0.52†			$\chi^2(1)$=47.50	$\chi^2(2)$=3.12	722
	(0.05)	(0.30)	(0.26)			[p=0.00]	[p=0.21]	
(4) Heckman Selection	-0.62†	0.10	-0.42	1.80				884
	(0.14)	(0.29)	(0.34)	(5.26)				
(5) Poisson w/ robust standard errors	-0.15†	0.04	-0.14‡					722
	(0.01)	(0.07)	(0.09)					
(6) ZIP w/ robust std. errors								722
(a) Violations = 0	0.49†	0.02	-0.40					
	(0.05)	(0.23)	(0.30)					
(b) # Vltns. ≥ 0	-0.06†	0.03	-0.23†					
	(0.01)	(0.05)	(0.07)					
(7) PSM			-0.45‡					722
			(0.25)					

Notes: Dependent variable is the number of GARP violations. Models (2) and (8) utilize noncompliers; all remaining models exclude experimental noncompliers. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are statistically significant at the $p < 0.01$ level in the Arizona sample only). ZIP = zero-inflated Poisson. For further details, see Table A4 and text. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Table A6. Summary of Learning Rates by Experiment

Group	Change in GARP Violations	H_O: Equality of Means Across Participants & Non-Participants	Observations
I. Arizona Sample			
Full Sample	-1.07 (4.28) [31.73%, 41.35%]		104
Group 1 (*MALLgift*):			
Non-Participants	5.00 (7.07) [50.00%, 0%]	*t* = 2.22 [p = 0.03]	2
Participants	-1.93 (4.21) [22.50%, 57.50%]		40
Group 2 (*SPORT*):			
Participants	-0.45 (3.30) [51.61%, 25.81%]		31
Group 3 (*MALLnogift*):			
Non-Participants	0.09(4.72) [17.39%, 30.43%]	*t* = 2.18 [p = 0.04]	23
Participants	-4.00 (4.04) [37.50%, 62.50%]		8
II. Illinois Sample			
Full Sample	-0.43 (2.58) [29.75%, 41.77%]		316
Group 1 (*MALLgift*):			
Non-Participants	2.33 (2.08) [33.33%, 0%]	*t* = 1.90 [p = 0.06]	3
Participants	-0.68 (2.73) [27.92%, 45.45%]		154
Group 2 (*SPORT*):			
Participants	-0.62 (0.87) [38.46%, 53.85%]		13
Group 3 (*MALLnogift*):			
Non-Participants	-0.13 (2.51) [30.88%, 36.76%]	*t* = 1.31 [p = 0.19]	136
Participants	-1.20 (1.93) [30.00%, 50.00%]		10

Notes: Figures represent mean change in GARP violations. Standard deviations are in parentheses. Percentage of observations with no change in GARP violations, followed by percentage of observations with a decrease in GARP violations, is in brackets. *P*-values are in brackets beneath *t*-tests. Group 1 includes non-participants in the market at the time of round one who were randomly given $25 worth of free cards after round one was completed. Group 2 includes subjects who were already in the market at the time of round one. Group 3 includes non-participants in the market at the time of round one who were not given any free cards after round one was completed.

Table A7. Determinants of Learning Rates by Experiment

	Independent Variable					
Model	**Age**	**Gender**	**Market Participant**	**Inverse Mills' Ratio**	**Test of Exogeneity**	***N***
I. Arizona Sample						
Full Sample						
(1) OLS	0.22	1.17	-1.69‡			104
	(0.14)	(0.98)	(1.00)			
(2) Heckman Selection	0.30‡	1.52	-1.74‡	-2.00		277
	(0.15)	(1.02)	(0.98)	(1.75)		
(3) PSM			-1.45			104
			(0.97)			
Experimental Sub-Sample						
(4) OLS	0.23	2.18‡	-1.67			63
	(0.19)	(1.23)	(1.15)			
(5) IV	0.16	2.58†	-0.40		$\chi^2(1)=3.37$	73
	(0.18)	(1.15)	(1.53)		[p=0.07]	
(6) Heckman Selection	0.28	2.69†	-2.07‡	-2.91‡		209
	(0.19)	(1.26)	(1.11)	(1.68)		
(7) PSM			-1.69			63
			(1.73)			
II. Illinois Sample						
Full Sample						
(1) OLS	0.04	-0.05	-0.60†			315
	(0.06)	(0.35)	(0.29)			
(2) Heckman Selection	0.03	-0.06	-0.70†	-4.06		503
	(0.08)	(0.48)	(0.31)	(4.75)		
(3) PSM			-0.58‡			315
			(0.30)			
Experimental Sub-Sample						
(4) OLS	0.02	0.03	-0.53‡			289
	(0.07)	(0.37)	(0.31)			
(5) IV	0.03	-0.06	-0.44		$\chi^2(1)=1.68$	302
	(0.06)	(0.36)	(0.33)		[p=0.20]	
(6) Heckman Selection	0.01	0.03	-0.64†	-3.83		471
	(0.09)	(0.48)	(0.31)	(4.35)		
(7) PSM			-0.52			289
			(0.34)			

Notes: Dependent variable is the change in GARP violations from round one to round two. Experimental sub-sample includes non-participants in the market at the time of round one (*MALLgift* and *MALLnogift*). Models (8) and (9) utilize the entire sub-sample; all remaining models using the experimental sub-sample use only experimental compliers and partial compliers. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are statistically significant at the $p < 0.01$ level in the Arizona samples only). † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level. For further details, see Tables A3 and A4.

Table A8. Determinants of Learning Rates: Experimental Sub-Sample Excluding *MALLgift* Partial Compliers by Experiment

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	*N*
	Independent Variable					
I. Arizona Sample						
(1) OLS	0.38 (0.24)	0.97 (1.68)	-3.33† (1.48)			42
(2) IV	0.24 (0.22)	2.03 (1.48)	-1.98 (2.10)		$\chi^2(1)$=1.08 [p=0.30]	52
(3) Heckman Selection	0.48‡ (0.25)	2.58 (1.79)	-3.70† (1.35)	-4.31† (2.24)		188
(4) PSM		-0.78 (1.78)				42
II. Illinois Sample						
(1) OLS	-0.04 (0.08)	-0.29 (0.45)	-0.95† (0.36)			215
(2) IV	-0.02 (0.07)	-0.38 (0.43)	-0.85† (0.39)		$\chi^2(1)$=0.84 [p=0.36]	228
(3) Heckman Selection	-0.05 (0.11)	-0.01 (0.86)	-1.05† (0.37)	-3.12 (5.69)		400
(4) PSM		-0.85† (0.38)				215

Notes: Dependent variable is the change in GARP violations from round one to round two. Models (2) and (5) utilize noncompliers; all remaining models exclude noncompliers. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are statistically significant at the $p < 0.01$ level in the Arizona sample only). † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level. For further details, see Tables A3 and A4.

Table A9. Determinants of GARP Violations: Experimental Sub-Sample (Illinois Sample with Additional Controls)

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	LM Test RE	Hausman Test (FE v. RE)	*N*
	Independent Variable							
(1) OLS w/ robust	-0.55†	0.10	-0.25					687
standard errors	(0.04)	(0.26)	(0.22)					
(2) IV w/ robust	-0.56†	0.11	-0.14		$\chi^2(1)$=0.03			696
standard errors	(0.04)	(0.26)	(0.56)		[p=0.85]			
(3) GLS-RE	-0.51†	0.13	-0.27			$\chi^2(1)$=41.77	$\chi^2(2)$=2.68	687
	(0.05)	(0.30)	(0.21)			[p=0.00]	[p=0.26]	
(4) Heckman	-0.68†	0.14	-0.20	4.17				837
Selection	(0.15)	(0.35)	(0.31)	(4.94)				
(5) Poisson w/ robust	-0.15†	0.04	-0.08					687
standard errors	(0.01)	(0.07)	(0.07)					
(6) Zero-Inflated Poisson w/ robust std. errors								687
(a) Violations = 0	0.51†	-0.05	-1.01‡					
	(0.06)	(0.24)	(0.28)					
(b) # Vltns. ≥ 0	-0.06†	0.02	-0.24†					
	(0.01)	(0.05)	(0.05)					
(7) PSM			-0.24					687
			(0.21)					

Notes: Dependent variable is the number of GARP violations. Experimental sub-sample includes non-participants in the market at the time of round one (*MALLgift* and *MALLnogift*). All models also include a dummy variable indicating if math or science is the individual's favorite subject, a dummy variable indicating attendance at a private school, an index reflecting self-reported perception of the individual by one's teachers, three race dummies, and a quadratic for hours spent watching television per week. Model (2) utilizes the entire sub-sample; all remaining models use only experimental compliers and partial compliers. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically insignificant at conventional levels). The propensity score is estimated via probit and includes a cubic in age, cubic in hours spent watching television per week, gender, state/experiment, gender interacted with a cubic in age, , gender interacted with a cubic in hours spent watching television per week, a dummy variable indicating if math or science is the individual's favorite subject, a dummy variable indicating attendance at a private school, an index reflecting self-reported perception of the individual by one's teachers, and three race dummies,. For further details, see Table A4 and text. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Table A10. Determinants of GARP Violations: Experimental Sub-Sample Excluding *MALLgift* Partial Compliers (Illinois Sample with Additional Controls)

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	LM Test RE	Hausman Test (FE v. RE)	*N*
	Independent Variable							
(1) OLS w/ robust standard errors	-0.57† (0.05)	0.12 (0.30)	-0.53‡ (0.29)					620
(2) IV w/ robust standard errors	-0.58† (0.05)	0.14 (0.30)	-0.37 (0.98)		$\chi^2(1)$=0.01 [p=0.92]			629
(3) GLS-RE	-0.52† (0.05)	0.08 (0.32)	-0.65† (0.28)			$\chi^2(1)$=38.10 [p=0.00]	$\chi^2(2)$=3.25 [p=0.20]	620
(4) Heckman Selection	-0.71† (0.17)	0.09 (0.36)	-0.46 (0.38)	3.85 (5.09)				770
(5) Poisson w/ robust standard errors	-0.15† (0.01)	0.04 (0.08)	-0.16‡ (0.10)					620
(6) Zero-Inflated Poisson w/ robust std. errors								620
(a) Violations = 0	0.52† (0.06)	-0.03 (0.25)	-0.35 (0.31)					
(b) # Vltns. ≥ 0	-0.06† (0.01)	0.03 (0.06)	-0.24† (0.08)					
(7) PSM			-0.53‡ (0.32)					620

Notes: Dependent variable is the number of GARP violations. Model (2) utilizes noncompliers; all remaining models exclude experimental noncompliers. All models also include a dummy variable indicating if math or science is the individual's favorite subject, a dummy variable indicating attendance at a private school, an index reflecting self-reported perception of the individual by one's teachers, and a quadratic for hours spent watching television per week. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically insignificant at conventional levels). For further details, see Table A5 and A9. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Table A11. Determinants of Learning Rates: Experimental Sub-Sample (Illinois Sample with Additional Controls)

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	*N*
			Independent Variable			
(1) OLS	0.04 (0.07)	0.15 (0.41)	-0.74† (0.35)			252
(2) IV	0.06 (0.07)	0.01 (0.39)	-0.62‡ (0.36)		$\chi^2(1)=2.57$ [p=0.11]	261
(3) Heckman Selection	0.07 (0.09)	0.20 (0.51)	-0.82† (0.34)	-3.70 (3.51)		414
(4) PSM			-0.94‡ (0.51)			252

Notes: Dependent variable is the change in GARP violations from round one to round two. Experimental sub-sample includes non-participants in the market at the time of round one (*MALLgift* and *MALLnogift*). Model (2) utilizes the entire sub-sample; all remaining models using the experimental sub-sample use only experimental compliers and partial compliers. All models also include a dummy variable indicating if math or science is the individual's favorite subject, a dummy variable indicating attendance at a private school, an index reflecting self-reported perception of the individual by one's teachers, and a quadratic for hours spent watching television per week. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically insignificant at conventional levels). † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level. For further details, see Tables A4, A5, and A10.

Table A12. Determinants of Learning Rates: Experimental Sub-Sample Excluding *MALLgift* Partial Compliers (Illinois Sample with Additional Controls)

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Test of Exogeneity	*N*
			Independent Variable			
(1) OLS	-0.02	-0.35	-1.26†			185
	(0.09)	(0.51)	(0.40)			
(2) IV	0.003	-0.53	-1.18†		$\chi^2(1)$=0.96	194
	(0.08)	(0.47)	(0.41)		[p=0.33]	
(3) Heckman Selection	-0.02	-0.23	-1.37†	-1.28		347
	(0.09)	(0.73)	(0.39)	(3.70)		
(4) PSM			-0.97			185
			(0.60)			

Notes: Dependent variable is the change in GARP violations from round one to round two. Model (2) utilizes noncompliers; all remaining models exclude noncompliers. All models also include a dummy variable indicating if math or science is the individual's favorite subject, a dummy variable indicating attendance at a private school, an index reflecting self-reported perception of the individual by one's teachers, and a quadratic for hours spent watching television per week. The Heckman selection model uses a quadratic in mall trips per month as exclusion restrictions in the first-stage selection equation (the coefficients are jointly statistically insignificant at conventional levels). † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level. For further details, see Table A11.

Appendix B

Confidential Survey Summary

These questions will be used for statistical purposes only. THIS INFORMATION WILL BE KEPT STRICTLY CONFIDENTIAL AND WILL BE DESTROYED UPON COMPLETION OF THE STUDY.

1. How long have you been active in the sportscard and memorabilia market? _______yrs
2. How much money do you spend per month at sportscard shows?________
3. How many trades do you make in a typical month?_______________
4. Gender: 1) Male 2) Female
5. Age ______ Date of Birth ____________
6. Race______________
7. Grade in school______________
8. Do you attend private or public school?____________
9. How often do you visit the mall?_______________ times per month.
10. How many brothers and sisters do you have?__________
11. Do you think it will rain next week?_____________
12. Do you think your teacher/teacher(s) at school like you?__________
13. Hours spent watching TV per week?____________
14. Favorite subject in school?__________

Name________________________ Email____________________telephone number___________

Parents:

Relationship between adult respondent and subject_______________

1. How long have you been active in the sportscard and memorabilia market? ______yrs
2. How many sportscard or memorabilia shows do you attend in a typical year? _______
3. Gender: 1) Male 2) Female
4. Age ______ Date of Birth ____________
5. What is the highest grade of education that you have completed. (Circle one)
   1) Eighth grade 3) 2-Year College 5) 4-Year College
   2) High School 4) Other Post-High School 6) Graduate School Education
6. What is your approximate yearly income from all sources, before taxes? (Circle one)
   1) Less than $10,000 5) $40,000 to $49,999
   2) $10,000 to $19,999 6) $50,000 to $74,999
   3) $20,000 to $29,999 7) $75,000 to $99,999
   4) $30,000 to $39,999 8) $100,000 or over

References

Andreoni, J. and J. Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 70, 737-53.

Battalio, R.C., J.H. Kagel, R.C. Winkler, E.B. Fisher, R.L. Basmann, and L. Krasner. 1973. "A Test of Consumer Demand Theory Using Observations of Individual Consumer Purchases," *Western Economic Journal*, 11, 421-8.

Becker, G.S. 1962. "Irrational Behavior and Economic Theory," *Journal of Political Economy*, 70, 1-13.

Blundell, R. and M. Costa-Dias. 2002. "Alternative Approaches to Evaluation in Empirical Microeconomics," *Portuguese Economic Journal*, 1, 91-115.

Bowles, S. 1998. "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions," *Journal of Economic Literature*, 36, 75-111.

Brown, M. and D. Heien. 1972. "The S-Branch Utility Tree: A Generalization of the Linear Expenditure System," *Econometrica*, 40, 737-47.

Camerer, C. 1990. "Behavioral Game Theory," in R. Hogarth (Ed.), *Insights in Decision Making*, 311-336.

Camerer, C. 2002. *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton University Press.

Chamberlin E.H. 1948. "An Experimental Imperfect Market," *Journal of Political Economy*, 56, 95-108.

Davis, D. and C. Holt. 1995. *Experimental Economics,* Princeton: Princeton University Press.

Dee, T. 2004. "Teachers, Race, and Student Achievement in a Randomized Experiment," *Review of Economics and Statistics*, 86, 195-210.

Forsythe Robert, Forrest Nelson, George R. Neumann, Jack Wright. 1992. **"**Anatomy of an Experimental Political Stock Market**,"** *American Economic Review*, Vol. 82(5), pp. 1142-1161.

Georgescu-Roegen, N. 1950. "The theory of Choice and the Constancy of Economic Laws," *Quarterly Journal of Economics*, 44, 125-38.

Gode, Dhananjay K. and Shyam Sunder, "Allocative Efficiency of Markets with Zero-Intelligence Traders: Markets as a Partial Substitute for Individual Rationality," *Journal of Political Economy*, 101:1 (1993), pp. 119-137.

Greene, William H. 2003. *Econometric Analysis*, 5th Edition, Upper Saddle River, NJ: Prentice Hall.

Harbaugh, W.T., K. Krause, and T.R. Berry. 2001. "GARP for Kids: On the Development of Rational Choice Behavior," *American Economic Review*, 91, 1539-45.

Heckman, J.J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement*, 5, 475-492.

Heckman, J.J., H. Ichimura, and P.E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 64, 605-654.

Hirano, K., G.W. Imbens, and G. Ridder 2003. "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score," *Econometrica*, 71, 1161-1189.

Houttaker, H.S. and L.D. Taylor. 1966. *Consumer Demand in the United States: Analyses and Projections*, Cambridge, MA: Cambridge.

Koopmans, T.C. 1964. "On Flexibility of Future Preferences," in S. Maynard and B. Glenn (eds.) *Human Judgments and Optimality*, New York: Wiley.

List, J.A. 2004a. "Testing Neoclassical Competitive Theory in Multi-Lateral Decentralized Markets," *Journal of Political Economy*, 112, 1131-56.

List, J.A. 2004b. "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field," *Quarterly Journal of Economics*, 119, 49-89.

MacCrimmon, K.R. and M. Toda. 1969. "The Experimental Determination of Indifference Curves," *Review of Economic Studies*, 36, 433-51.

Manser, M.E. 1976. "Elasticities of Demand for Food: An Analysis Using Non-Additive Utility Functions Allowing for Habit Formation," *Southern Economic Journal*, 43, 879-91.

May, K.O. 1954. "Intransitivity, Utility, and the Aggregation of Preference Patterns," *Econometrica*, 22, 1-13.

Myerson, R.B. and M. Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265-81.

Millimet, D.L. and R. Tchernis. 2008. "On the Specification of Propensity Scores: with Applications to the Analysis of Trade Policies," *Journal of Business & Economic Statistics*, forthcoming.

Philips, L. 1972. "A Dynamic Version of the Linear Expenditure Model," *Review of Economics and Statistics*, 54, 450-98.

Pollak, R.A. 1978. "Endogenous Tastes in Demand and Welfare Analysis," *American Economic Review*, 68, 374-9.

Rosenbaum, P. and D. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41-55.

Roth, A.E. 1995. "Introduction to Experimental Economics," in J.H. Kagel and A.E. Roth (eds.) *The Handbook of Experimental Economics*, Princeton University Press, 1995, pp. 3-98.

Sippel, R. 1997. "An Experiment on the Pure Theory of Consumers' Behavior," *Economic Journal*, 107, 1431-44.

Slonim, R.L. 1999. "Learning Rules of Thumb or Learning More Rational Rules," *Journal of Economic Behavior and Organization*, 38, 217-36.
Smith, J. and P. Todd. 2005. "Does Matching Address Lalonde's Critique of Nonexperimental Estimators," *Journal of Econometrics*, 125, 305-53.
Smith, V.L. 1962. "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy*, 70, 111-37.
Smith, V.L. 1965. "Experimental Auction Markets and the Walrasian Hypothesis," *Journal of Political Economy*, 73, 387-93.
Smith, V.L. 1976. "Experimental Economics: Induced Value Theory," *American Economic Review*, 66, 274-9.
Stone, R. 1954. "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," *Economic Journal*, 64, 511-27.
Varian, H. 1982. "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50, 945-72.
Wales, T.J. 1971. "A Generalised Linear Expenditure Model of the Demand for Non-Durable Goods in Canada," *Canadian Journal of Economics*, 4, 471-84.