The Market:
Catalyst for Rationality and Filter of Irrationality[*]

John A. List
University of Maryland and NBER

Daniel L. Millimet
Southern Methodist University

17 December 2004

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 perhaps the central issue facing behavioral economics today: to what extent does market experience mitigate various forms of individual irrationality? We find considerable evidence that the market is a catalyst for rationality. The study then focuses on aggregate market outcomes by examining empirically whether individual rationality is a prerequisite for market efficiency. Using field data gathered from more than 380 subjects of age 6-18 in multi-lateral bargaining markets at a shopping mall, we find that the market is a filter of irrationality—even when markets are populated solely by irrational buyers, aggregate market outcomes quickly converge to neoclassical predictions.

JEL: C93 (Field Experiments), D4 (Market Structure and Pricing)
Key words: rationality, market learning, field experiment, competitive market theory

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

---

[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.

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*.

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. 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? The latter question addresses one of the most fundamental issues facing behavioral economics today.

To answer these questions, while allowing for the fact that preferences may evolve over time, we conduct 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. This treatment represents a particularly demanding test of the market's role in 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 or "rules of thumb" that can be applied in future transactions in the *same* market over similar tasks.[4]

Empirical results indicate that: (i) not all agents exhibit rational choice behavior, and (ii) market experience facilitates the development of such behavior. Specifically, we find that subjects participating in the sportscard market had, on average, nearly three fewer GARP violations; and while 66% of sportscard market participants exhibited behavior consistent with utility maximization, only 34% of non-participants exhibited such behavior. Moreover, the panel aspect of our experimental design permits tests of the impact of market participation on the learning of rational choice behavior. We find that the number of GARP violations decreased by two, on average, over the sample period for all market participants, while it remained virtually unchanged for non-participants. Finally, we document diminishing returns to market experience, as *new* entrants into the market experienced a reduction of nearly four GARP violations on average. 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, and several econometric modeling assumptions.

While these findings are important 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

---

[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).

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 a prerequisite for market efficiency in multi-lateral decentralized markets?

Empirical results indicate that even in decentralized bargaining markets populated entirely by irrational children buyers, market outcomes are efficient and prices follow neoclassical 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 does not unduly influence aggregate efficiency in bilateral bargaining markets.

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

II. Experimental Design

The experimental treatments were run at various sportscard shows at a shopping mall in a large southern city during the period November 2000-June 2001. In the first set of sportscard 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

4

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 young individuals 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.[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. 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

---

[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] We should highlight that while we believe GARP is the most fundamental definition of rationality, one could use several other measures of rationality. In this spirit, it is possible that subjects who have few (many) violations in our setting have many (few) violations of rationality in other less complicated settings.

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.

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 that have previous market experience and those that (may) have 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 may prohibit one from drawing firm 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 given). 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.

The top panel of 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*. Of course, over the six-month period between the first
and second treatments subjects in this treatment could have learned to choose more rationally 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 market experience.

III. Results

The Appendix provides descriptive statistics of the variables of interest. 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).

Result 1: Not all individuals behave rationally.

Table 2 provides detailed summary characteristics of the GARP results for the 381
observations. On average, subjects had over 3.5 GARP violations, with 55% of subjects having
at least one violation.[8] 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. Figure 2 shows the distribution of

---

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

GARP violations. Restricting attention to the sub-sample with at least one GARP violation, the distribution is fairly uniform from two to eleven GARP violations.

Result 2: Market experience and rational choice behavior are directly related.

Table 2 provides summary characteristics of the GARP results for the 381 subjects disaggregated by treatment assignment as well. In round one of the experiment (which took place in November), non-participants in the sportscard market who were given the $25 "parting gift" (*MALLgift*) had roughly 4.5 GARP violations on average; nearly 70% 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 30% 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 ($t = 5.53$; to be conservative, we provide all test results assuming a two-sided alternative, even when theory provides strong guidance that a one-sided alternative is appropriate). While interesting, the positive correlation between market participation and rational choice behavior *may* be spurious, reflecting unobservables associated with the self-selection of certain individual types into the sportscard market.

Examination of the round two results (conducted the following June) 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, two subjects in *MALLgift* never took part in the sportscard market (they neither traded or 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, 21 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, 19 *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 November-June period. Finally, eight 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 23 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 2.85 GARP violations on average, while *MALLnogift* individuals had 3.78 violations on average (although the difference is not statistically significant: $t = 1.01$, $p = 0.32$ at conventional levels). The fact that round two subjects in the *SPORT* treatment have fewer GARP violations on average (1.32) than the *MALLgift* group of market participants either

reflects selection effects (e.g., children participating in the sportscard market with no inducement tend to be of higher ability) or indicates that duration of market experience matters.

To assess the robustness of the summary results discussed heretofore, as well as control for possible sources of bias such as non-random attrition and the endogeneity of "compliance," we use a variety of econometric methods to analyze the effect of market participation on rational choice. The first set of specifications is nested in the following estimating equation:

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

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 Prob($GARP_{it}$ is observed) = $\Phi(z_{it}\gamma)$; and, $\eta_{it} = \mu_i + e_{it}$ is the error term, where μ_i is an individual-specific random effect and e_{it} represents idiosyncratic shocks. The vector x includes controls for age and gender, while z includes controls for age, gender, and mall visits per month. $\phi(\theta_{it})/\Phi(\theta_{it})$ is the inverse Mills' ratio, and is used to control for possible self-selection bias due to non-random attrition (mall visits per month represents the exclusion restriction used for identification).[9]

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 model assumes that the number of violations for

---

[9] Imposing the restrictions $\pi = 0$ 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 $\pi = 0$ reduces (1) to a Generalized Least Squares (GLS) random effects model. Finally, restricting only $\sigma_\mu^2 = 0$ reduces (1) to the standard Heckman selection model.

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 1 reveals an interesting statistic: over 40% 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 has been termed a variety of names, most notably zero-inflated Poisson (ZIP), and is a natural extension of the Poisson formulation in equation (2):

$$g_{it} = 0 \qquad\qquad \text{with probability } P_{it} \tag{3}$$

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

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

$$\Pr(g_{it} = 0) = P_{it} + (1 - P_{it})R_{it}(0) \tag{5}$$

12

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

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

$$\Psi_{it} = w_{it}\omega. \tag{8}$$

When estimating equation (7) it is important to include variables in the vector w_{it} that signal whether individual i will clear the initial "hurdle" and register any GARP violations. As our data set includes no additional controls likely to influence the clearance of this hurdle, we use the same variables in w as in x, as well as the dummy variable, *PART*.

The final estimation technique – the propensity score matching (PSM) estimator developed in Rosenbaum and Rubin [1983] – relaxes the parametric 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 by pairing each market participant with the experimental subject not participating in the market who most resembles the market participant

in terms of observable characteristics (hence, PSM is a "selection on observables" estimation technique).   To determine the appropriate match, a logit equation is estimated where the dependent variable is a binary variable indicating whether the subject is a market participant. After estimation of this first-stage regression, the predicted probability of each subject participating in the market is constructed and each market participant is paired with the non-participant in the market whose predicted probability of participating in the market (known as the propensity score) is most similar to his or her own.   Upon matching each market participant to the appropriate non-participant, the estimated impact of market participation on the number of GARP violations is given by the mean difference across all matched pairs.[10]

Empirical results from the various econometric specifications are presented in Tables 3 and 4.   Table 3 utilizes the 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, we re-estimate the various specifications utilizing only the sub-sample of non-participants in the market at the time of the initial experiment (*MALLgift* and *MALLnogift*), excluding non-compliers in round two (see Table 4). 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 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 [2003]), and is commonly referred to as "intent

---

[10] For further details, the reader is referred to Dehejia and Wahba [1999, 2002], Heckman et al. [1997], or Smith and Todd [2004].

to treat"). Comparison of the IV and non-IV results informs us about the randomness of the experimental noncompliers.

Four primary findings emerge from the empirical results in Tables 3 and 4. First, 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, Heckman selection, Poisson, ZIP, and PSM); the point estimates are statistically significant in every case in Table 3 and virtually every case in Table 4. 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). The effect is reduced by approximately one-half, however. This is consonant with roughly 50% of the effect in Table 3 being attributable to unobservable attributes positively associated with both rational behavior and participation in the sportscard market (i.e., selection effects), and the remaining 50% indicating a sizeable effect of market participation on rational choice behavior.

Third, the IV point estimates in Table 4 continue to be negative, but are smaller in magnitude and statistically indistinguishable from zero.[11] A likelihood ratio (LR) test of the exogeneity of *actual* market participation fails to reject the null, however; thus, the IV results are less efficient than the other models.[12] Finally, in addition to being statistically significant, the results confirm our hypothesis that market participation is "economically" meaningful, as the (OLS, GLS, and Heckman selection-corrected) point estimates in Table 3 (Table 4) imply that market participation reduces the number of GARP violations by roughly 2.3 (1.3). Given that

---

[11] The first-stage regression indicates that intent to treat, D, is a highly significant determinant of actual participation ($p = 0.00$).

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

the mean number of violations in the full sample is 3.6, this represents a sizeable reduction. Coefficient estimates from the Poisson and ZIP models, while interpreted differently, suggest effects of similar magnitudes.

Result 3: Age and rational choice behavior are directly related.

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 in the (OLS, GLS, and Heckman selection-corrected) econometric results in Tables 3 and 4, where the number of GARP violations decreases by one-half per year *ceteris paribus*.[13]

Result 4: The intensity of market participation and rational choice behavior are directly related.

The analysis of the experimental sub-sample in Table 4 pooled compliers and partial compliers in the *MALLgift* group, explicitly treating them 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). To examine the effects of the intensity of market participation on rational choice behavior, we first note that mean GARP violations are lower for compliers in *MALLgift* (2.16) than partial compliers (3.48), although this difference is not statistically significant at

---

[13] 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.

conventional levels ($p = 0.22$). As before, however, the difference between compliers and partial

compliers may (at least partially) reflect selection effects.

To assess the role of market intensity more formally, Table 5 estimates several variants of

the previous models reported in Table 4, differentiating between compliers and partial compliers.

Specifically, models 1-6 in Table 5 mimic the corresponding models in Table 4, except partial

compliers are excluded from the sample.[14] 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 ten noncompliers in the *MALLgift* and

MALLnogift groups and instruments for participation with the "intent to treat" variable.[15]

Finally, models 7-12 in Table 5 are identical to those in models 1-6 except in these models

compliers in the *MALLgift* group are excluded from the sample, but partial compliers are

retained. Consequently, the effect of participation is identified from the variation in GARP

violations across *MALLgift* partial compliers and *MALLnogift* compliers, where the only subjects

with *PART* = 1 are subjects in the *MALLgift* group who entered the market to merely

buy/sell/trade to maximize profits from the contents of their gift bag. The IV results (model 8)

are obtained from the sample that still excludes *MALLgift* compliers, but utilizes the ten

noncompliers in the *MALLgift* and *MALLnogift* groups and instruments for participation with the

"intent to treat" variable.[16]

---

[14] In the interest of brevity, we only report results conditioning on a linear function of age.

[15] The coefficient on the instrument in the first-stage regression is highly significant ($p = 0.02$).

[16] The coefficient on the instrument in the first-stage regression is highly significant ($p = 0.01$).

Given the difference in intensity of market participation between the compliers and partial compliers in the *MALLgift* group, comparison of the results in the upper and lower panels provides empirical evidence of the impact of market intensity. Such a comparison indicates that intensity matters a great deal. Whereas the coefficients in the upper panel (excluding partial compliers) are negative, large in magnitude, and statistically significant in the majority of cases, the coefficients in the lower panel are (for the most part) of the correct sign, but smaller in magnitude and statistically insignificant at conventional levels. Thus, one time participation in the market conveys some benefit, but repeated participation is associated with more rational behavior. 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.[17]

> *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. Table 6 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 slightly more than one from the first to the second round, with roughly 32% (41%) 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 tremendously across subject groups,

---

[17] We also estimated the models including separate dummy variables for "One-Time Participation" (equal to one for *MALLgift* partial compliers in round two, zero otherwise) and "Full Market Participation" (equal to one for *MALLgift* compliers in round two, zero otherwise), both including and excluding noncompliers. Consonant with the results in Table 5, we obtain a coefficient of -0.48 (*s.e.* = 0.78) for "One-Time Participation" and a coefficient of -2.28 (*s.e.* = 0.83) for "Full Market Participation" using OLS and excluding noncompliers, and virtually identical results using noncompliers as well—IV estimation of such a model is not possible, as there is only one instrument.

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 nearly one-half fewer GARP violations, on average, in round two. Since these respondents did not change their market participation, and any selection bias involved from these individuals self-selecting into the market remains constant temporally, it is not surprising that the decline in violations by almost one-half is close to the predicted decline in GARP violations from an increase in age by more than seven months, as detailed in Tables 3 and 4.

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 60% 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 two. GARP violations for noncompliers in the *MALLgift* treatment increased by five 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 $p < 0.05$ 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 four 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.

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 1.93 (0.09) fewer (more) GARP violations on average in round two. Thus, the estimated treatment effect of market participation is -2.02; a standard *t*-test rejects the null that the treatment effect equals zero at the $p = 0.08$ level. This is a particularly strong 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 nonrandomness is time varying, will the previous comparisons confound market and selection effects. Yet there may still be a bias imparted due to nonrandom 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).

The beneficial effect of market participation on the development of rational choice behavior is further evidenced 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.[18] The top panel of Table 7 uses the full sample and provides a robustness test of the learning results documented above: estimates contained in rows 1-4 shows that market participants have a larger decrease in GARP violations than do non-participants, and the latter

_____

[18] We still include age and gender in the differenced regression as these variables may affect the degree of learning.

two estimates are statistically significant at the $p < 0.10$ level. Focusing on the experimental sub-sample, 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. The IV results, however, now reject the exogeneity of actual participation at the $p < 0.10$ level, and the point estimates are not surprisingly extremely imprecise.[19]

While the experimental results of the impact of the market on learning in Table 7 are weaker than the results pertaining to the effects of participation on the level of rational behavior, the learning results may be misleading if intensity of market participation matters. 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* (-3.58) than partial compliers (-0.43), and this difference is statistically significant at the $p < 0.02$ level.

To examine the role of market participation intensity more formally, Table 8 summarizes estimates of several variants of the previous models that differentiate between compliers and partial compliers. As in Table 5, models 1-3 in Table 8 mimic the corresponding models in Table 7, except partial compliers are excluded from the sample. The IV results (model 2) are obtained from the sample that excludes partial compliers, but uses the ten noncompliers in the *MALLgift* and *MALLnogift* groups and instruments for participation with the "intent to treat" variable.[20] Finally, models 4-6 in Table 8 are identical to those in models 1-3 except now compliers in the *MALLgift* group are excluded from the sample, but partial compliers are retained. The IV results (model 5) are obtained from the sample that still excludes *MALLgift*

---

[19] The first-stage model indicates that intent to treat is a significant determinant of actual participation ($p = 0.00$).
[20] The coefficient on the instrument in the first-stage regression is highly significant ($p = 0.00$).

compliers, but utilizes the ten noncompliers in the *MALLgift* and *MALLnogift* groups and instruments for participation with the "intent to treat" variable.[21]

Comparison of the empirical results in the upper and lower panels in Table 8 indicates that intensity matters for the learning process. Whereas the coefficients in the upper panel (excluding partial compliers) are negative, large in magnitude, and statistically significant except for the IV model (although exogeneity is not rejected), the coefficients in the lower panel are generally of the correct sign, but smaller in magnitude and not statistically significant at conventional levels. Thus, as in the levels specifications in Table 5, one time participation in the market conveys some benefits, but regular market participation is associated with economically significant changes in the development of rational behavior.[22]

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

---

[21] The coefficient on the instrument in the first-stage regression is highly significant ($p = 0.00$).

[22] For completeness, we also estimated the models using the change in GARP violations as the dependent variable including separate dummy variables for "One-Time Participation" (equal to one for *MALLgift* partial compliers in round two, zero otherwise) and "Full Market Participation" (equal to one for *MALLgift* compliers in round two, zero otherwise), both including and excluding noncompliers. Consonant with the results in Table 8, we obtain a coefficient of -0.62 (*s.e.* = 1.27) for "One-Time Participation" and a coefficient of -3.08 (*s.e.* = 1.37) for "Full Market Participation" using OLS and excluding noncompliers, and virtually identical results using noncompliers.

substantively rational choices." And continues, remarking (p. 218) that "Camerer (1990a) 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."[23]

IV. Market Experimental Design and Results

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)).

---

[23] The interested reader should also see Camerer (1990b).

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 <u>one</u> 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 market periods, each lasting ten minutes. 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.

We follow this simple procedure in each of four market treatments, which are summarized in column 1 of Table 1B and can be read as follows: row 1, column 1 of Table 1B 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 3 presents buyer and seller induced values, which are taken from Davis and Holt [1995, pp. 14-15]. In Figure 3, 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 3.

Experimental Results

Table 9 contains summary statistics for the experimental data. Entries in Table 9 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 9 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 6: Competitive theory adequately organizes the data, even in markets populated entirely by irrational buying agents.*

By observing average trading prices across the five market sessions in Table 9, one can see Result 6 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. These results highlight the strength of the invisible hand, as competitive conditions prevail even when irrational agents are the sole buyers in the marketplace. This insight is important in light of recent empirical findings of individual irrationality.[24]

---

[24] 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.

Another interesting insight gained from Table 9 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 systematic data collection must be carried out to make this a formal result.

V. Conclusion

A central issue facing behavioral 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 a straightforward test for rationality relying on the theory of revealed preference, we find that the market is a catalyst for rationality. This result is surprising in that our experimental design constitutes a particularly demanding test of learning, as our exploration measures participation in one 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.

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

Camerer, C. 1990b. "Do Markets Correct Biases in Probability Judgment? Evidence from Market Experiments," in J. Kagel and L. Green (Eds.), Advances in Behavioral Economics, 2, Northwood, NJ: Ablex Publishers, 125-172.

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. 2003. "Teachers, Race, and Student Achievement in a Randomized Experiment," *Review of Economics and Statistics*, forthcoming.

Dehejia, R.H. and S. Wahba. 1999. "Casual Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 94, 1053-62.

Dehejia, R.H. and S. Wahba. 2002. "Propensity Score Matching for Nonexperimental Causal Studies," *Review of Economics and Statistics*, 84, 151-61.

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

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., 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.

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*, forthcoming.

List, J.A. 2004b. "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field," *Quarterly Journal of Economics*, 119(1): pp. 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.

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. 2000. "Does Matching Address Lalonde's Critique of Nonexperimental Estimators," *Journal of Econometrics*, forthcoming.

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.
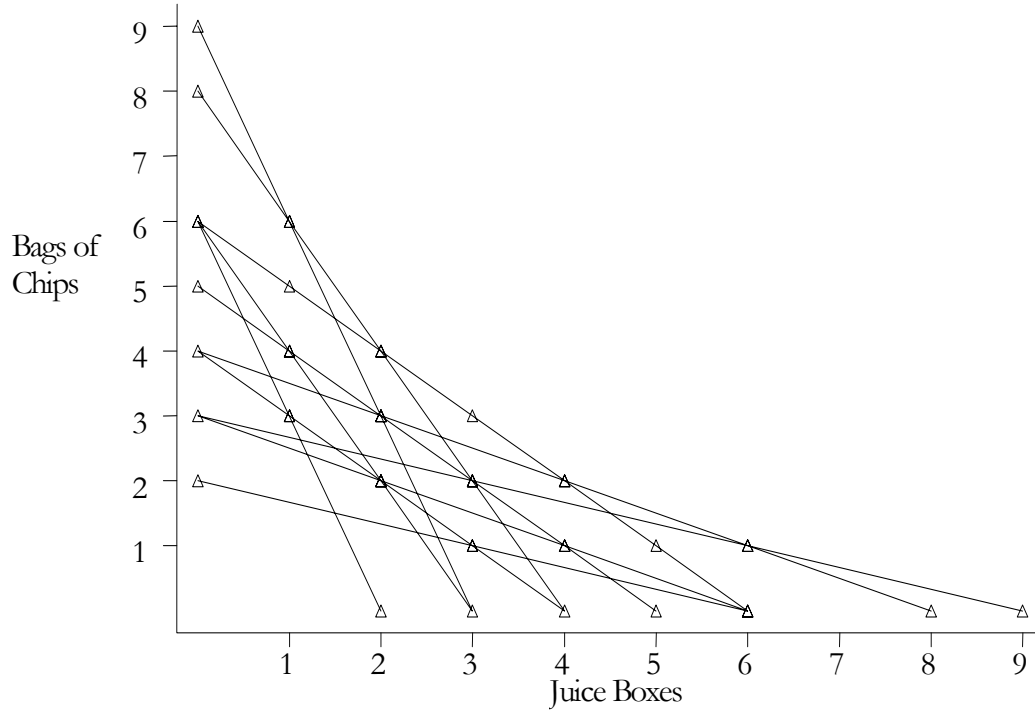
Figure 1. Choice Sets
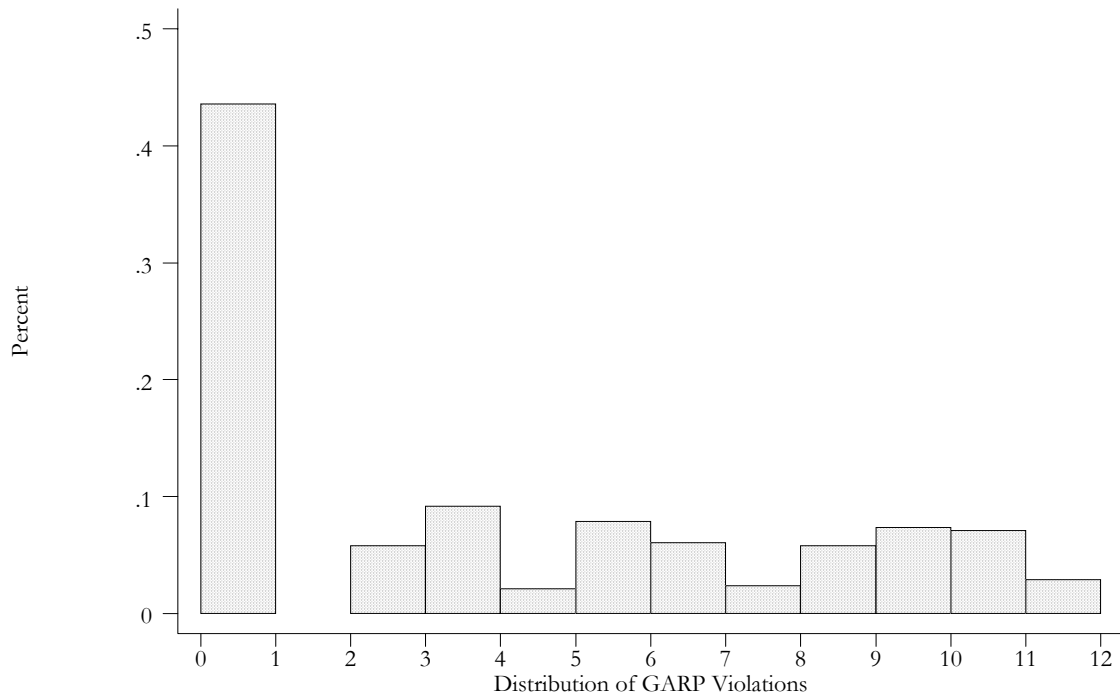
Figure 2. GARP Violations

Table 1A Individual Choice (GARP) Experimental Design

	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.

Table 1B Market Experimental Design

	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.

Table 2 Summary of GARP Violations

Group	GARP Violations	H$_O$: Equality of Means Across Participants & Non-Participants	Observations
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 3 Determinants of GARP Violations: Full Sample

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Age Dummies Included	LM Test Random Effects	Hausman Test (FE v. RE)	N
			Independent Variable					
(1) OLS w/ robust standard errors	-0.47† (0.05)	-0.04 (0.41)	-2.35† (0.37)		No			381
(2) OLS w/ robust standard errors		-0.06 (0.41)	-2.32† (0.38)		Yes			381
(3) GLS-RE	-0.48† (0.06)	-0.06 (0.40)	-2.33† (0.37)		No	$\chi^2(1)=2.97$ [p=0.08]	$\chi^2(3)=1.34$ [p=0.72]	381
(4) GLS-RE		-0.06 (0.39)	-2.27† (0.37)		Yes	$\chi^2(1)=3.31$ [p=0.07]	$\chi^2(14)=12.95$ [p=0.53]	381
(5) Heckman Selection	-0.48† (0.08)	-0.05 (0.40)	-2.35† (0.37)	0.06 (1.65)	No			554
(6) Heckman Selection		-0.03 (0.39)	-2.32† (0.37)	-0.38 (1.57)	Yes			554
(7) Poisson w/ robust standard errors	-0.15† (0.02)	-0.01 (0.10)	-0.77† (0.14)		No			381
(8) Poisson w/ robust standard errors		-0.02 (0.10)	-0.76† (0.14)		Yes			381
(9) Zero-Inflated Poisson w/ robust std. errors					No			381
(a) Violations = 0	0.22† (0.04)	-0.07 (0.27)	1.39† (0.26)					
(b) # Vltns. ≥ 0	-0.07† (0.01)	-0.03 (0.06)	-0.18† (0.07)					
(10) Zero-Inflated Poisson w/ robust std. errors					Yes			381
(a) Violations = 0		-0.10 (0.27)	1.38† (0.27)					
(b) # Vltns. ≥ 0		-0.03 (0.06)	-0.19† (0.07)					
(11) Propensity Score Matching			-1.76† (0.40)		No			274

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 mall trips per month as an exclusion restriction in the first-stage selection equation (the coefficient is statistically significant at the $p < 0.01$ level in each specification). Propensity score matching restricts matches to be from the same experiment round. The matching algorithm balances age, gender, mall trips per month, and lagged GARP violations (for round two matched pairs). † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Table 4 Determinants of GARP Violations: Experimental Sub-Sample

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Age Dummies Included	LR Test of Exogeneity	LM Test Random Effects	Hausman Test (FE v. RE)	N
(1) OLS w/ robust standard errors	-0.55† (0.06)	-0.12 (0.43)	-1.33† (0.60)		No				282
(2) OLS w/ robust standard errors		-0.14 (0.44)	-1.27† (0.64)		Yes				282
(3) IV	-0.55† (0.07)	-0.04 (0.42)	-0.59 (2.02)		No	$\chi^2(1)$=0.21 [p=0.65]			292
(4) IV		-0.09 (0.44)	-0.69 (2.14)		Yes	$\chi^2(1)$=0.12 [p=0.73]			292
(5) GLS-RE	-0.56† (0.07)	-0.16 (0.43)	-1.40† (0.57)		No		$\chi^2(1)$=3.02 [p=0.08]	$\chi^2(2)$=0.27 [p=0.87]	282
(6) GLS-RE		-0.19 (0.44)	-1.38† (0.58)		Yes		$\chi^2(1)$=3.03 [p=0.08]	$\chi^2(12)$=15.67 [p=0.21]	282
(7) Heckman Selection	-0.56† (0.08)	-0.13 (0.42)	-1.30† (0.60)	0.36 (1.50)	No				428
(8) Heckman Selection		-0.16 (0.42)	-1.24† (0.60)	0.42 (1.48)	Yes				428
(9) Poisson w/ robust standard errors	-0.15† (0.02)	-0.04 (0.10)	-0.37‡ (0.20)		No				282
(10) Poisson w/ robust standard errors		-0.04 (0.10)	-0.36† (0.20)		Yes				282
(11) Zero-Inflated Poisson w/ robust std. errors					No				282
(a) Violations = 0	0.23† (0.04)	0.01 (0.28)	0.85‡ (0.38)						
(b) # Vltns. $\geq$ 0	-0.07† (0.01)	-0.03 (0.06)	-0.03 (0.10)						
(12) Zero-Inflated Poisson w/ robust std. errors					Yes				282
(a) Violations = 0		-0.02 (0.28)	0.77‡ (0.40)						
(b) # Vltns. $\geq$ 0		-0.04 (0.06)	-0.06 (0.10)						
(13) Propensity Score Matching			-1.06 (0.78)		No				80

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 mall trips per month as an exclusion restriction in the first-stage selection equation (the coefficient is statistically significant at the $p < 0.01$ level in each specification). For further details, see Table 3 and text. † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level.

Table 5 Determinants of GARP Violations: Experimental Sub-Sample Sensitivity Analysis

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	LR Test of Exogeneity	LM Test Random Effects	Hausman Test (FE v. RE)	N
Independent Variable								

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	LR Test of Exogeneity	LM Test Random Effects	Hausman Test (FE v. RE)	N
Excluding MALLgift Partial Compliers								
(1) OLS w/ robust standard errors	$-0.56^{\dagger}$ (0.07)	-0.32 (0.46)	$-2.32^{\dagger}$ (0.84)					261
(2) IV	$-0.57^{\dagger}$ (0.07)	-0.26 (0.42)	-3.75 (5.21)		$\chi^2(1)=0.05$ [p=0.83]			271
(3) GLS-RE	$-0.57^{\dagger}$ (0.07)	-0.36 (0.45)	$-2.53^{\dagger}$ (0.80)			$\chi^2(1)=1.59$ [p=0.21]	$\chi^2(2)=1.58$ [p=0.45]	261
(4) Heckman Selection	$-0.57^{\dagger}$ (0.09)	-0.33 (0.44)	$-2.30^{\dagger}$ (0.83)	0.28 (1.62)				407
(5) Poisson w/ robust standard errors	$-0.15^{\dagger}$ (0.02)	-0.08 (0.11)	$-0.72^{\dagger}$ (0.33)					261
(6) Zero-Inflated Poisson w/ robust std. errors								261
(a) Violations = 0	$0.22^{\dagger}$ (0.05)	0.13 (0.30)	$1.21^{\dagger}$ (0.59)					
(b) # Vltns. ≥ 0	$-0.07^{\dagger}$ (0.01)	-0.04 (0.07)	-0.17 (0.16)					
Excluding MALLgift Compliers								
(7) OLS w/ robust standard errors	$-0.58^{\dagger}$ (0.06)	-0.38 (0.43)	-0.46 (0.78)					263
(8) IV	$-0.58^{\dagger}$ (0.07)	-0.22 (0.42)	3.08 (2.16)		$\chi^2(1)=1.25$ [p=0.26]			273
(9) GLS-RE	$-0.59^{\dagger}$ (0.07)	-0.37 (0.44)	-0.39 (0.75)			$\chi^2(1)=2.68$ [p=0.10]	$\chi^2(2)=0.53$ [p=0.77]	263
(10) Heckman Selection	$-0.62^{\dagger}$ (0.09)	-0.41 (0.43)	-0.38 (0.78)	1.28 (1.67)				409
(11) Poisson w/ robust standard errors	$-0.15^{\dagger}$ (0.02)	-0.10 (0.10)	-0.11 (0.22)					263
(12) Zero-Inflated Poisson w/ robust std. errors								263
(a) Violations = 0	$0.25^{\dagger}$ (0.05)	0.15 (0.29)	0.05 (0.59)					
(b) # Vltns. ≥ 0	$-0.07^{\dagger}$ (0.01)	-0.06 (0.06)	0.05 (0.11)					

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 mall trips per month as an exclusion restriction in the first-stage selection equation (the coefficient is statistically significant at the $p < 0.01$ level in each specification). For further details, see Table 4 and text. $\dagger$ indicates significance at the $p < 0.05$ level; $\ddagger$ at the $p < 0.10$ level.

Table 6 Summary of Learning Rates

Group	Change in GARP Violations	H$_O$: Equality of Means Across Participants & Non-Participants	Observations
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

Notes:

1. 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.

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.

Table 7 Determinants of Learning Rates

Model	Age	Gender	Market Participant	Inverse Mills' Ratio	Age Dummies Included	LR Test of Exogeneity	N
Independent Variable							
Full Sample							
(1) OLS	0.22 (0.14)	1.28 (1.00)	-1.64 (1.01)		No		104
(2) OLS		1.09 (1.08)	-1.46 (1.06)		Yes		104
(3) Heckman Selection	0.34† (0.16)	1.92‡ (1.04)	-1.70‡ (0.98)	-3.21 (2.00)	No		277
(4) Heckman Selection		1.85 (1.15)	-1.54‡ (0.97)	-3.92‡ (2.08)	Yes		277
(5) Propensity Score Matching			0.02 (0.73)		No		158
Experimental Sub-Sample							
(6) OLS	0.23 (0.19)	2.18‡ (1.23)	-1.67 (1.15)		No		63
(7) OLS		1.65 (1.32)	-1.41 (1.19)		Yes		63
(8) IV	0.15 (0.18)	2.24† (1.08)	-0.83 (1.36)		No	$\chi^2(1)=3.00$ [p=0.08]	73
(9) IV		1.93‡ (1.05)	0.45 (1.64)		Yes	$\chi^2(1)=5.56$ [p=0.02]	73
(10) Heckman Selection	0.29 (0.20)	2.76† (1.26)	-2.07‡ (1.10)	-3.18‡ (1.69)	No		209
(11) Heckman Selection		2.23‡ (1.24)	-1.76‡ (1.04)	-3.33† (1.63)	Yes		209
(12) Propensity Score Matching			-1.39 (1.07)		No		80

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 mall trips per month as an exclusion restriction in the first-stage selection equation (the coefficient is statistically significant at the $p < 0.01$ level in each specification). † indicates significance at the $p < 0.05$ level; ‡ at the $p < 0.10$ level. For further details, see Tables 3 and 4.

Table 8 Determinants of Learning Rates: Experimental Sub-Sample Sensitivity Analysis

| | **Independent Variable** | | | | | |
Model	**Age**	**Gender**	**Market Participant**	**Inverse Mills' Ratio**	**LR Test of Exogeneity**	**N**
Excluding MALLgift Partial Compliers						
(1) OLS	0.38	0.97	-3.33[†]			42
	(0.24)	(1.68)	(1.48)			
(2) IV	0.22	1.65	-2.31		$\chi^2(1)=1.07$	52
	(0.21)	(1.34)	(1.81)		[p=0.30]	
(3) Heckman Selection	0.49[†]	2.67	-3.68[†]	-4.50[†]		188
	(0.25)	(1.78)	(1.34)	(2.25)		
Excluding MALLgift Compliers						
(4) OLS	0.08	2.20	-0.48			44
	(0.22)	(1.38)	(1.30)			
(5) IV	0.01	2.25[‡]	0.66		$\chi^2(1)=4.72$	54
	(0.20)	(1.19)	(1.59)		[p=0.03]	
(6) Heckman Selection	0.20	2.13	-0.80	-3.54[‡]		190
	(0.23)	(1.37)	(1.21)	(2.13)		

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 mall trips per month as an exclusion restriction in the first-stage selection equation (the coefficient is statistically significant at the $p < 0.01$ level in each specification). [†] indicates significance at the $p < 0.05$ level; [‡] at the $p < 0.10$ level. For further details, see Tables 3 and 4.
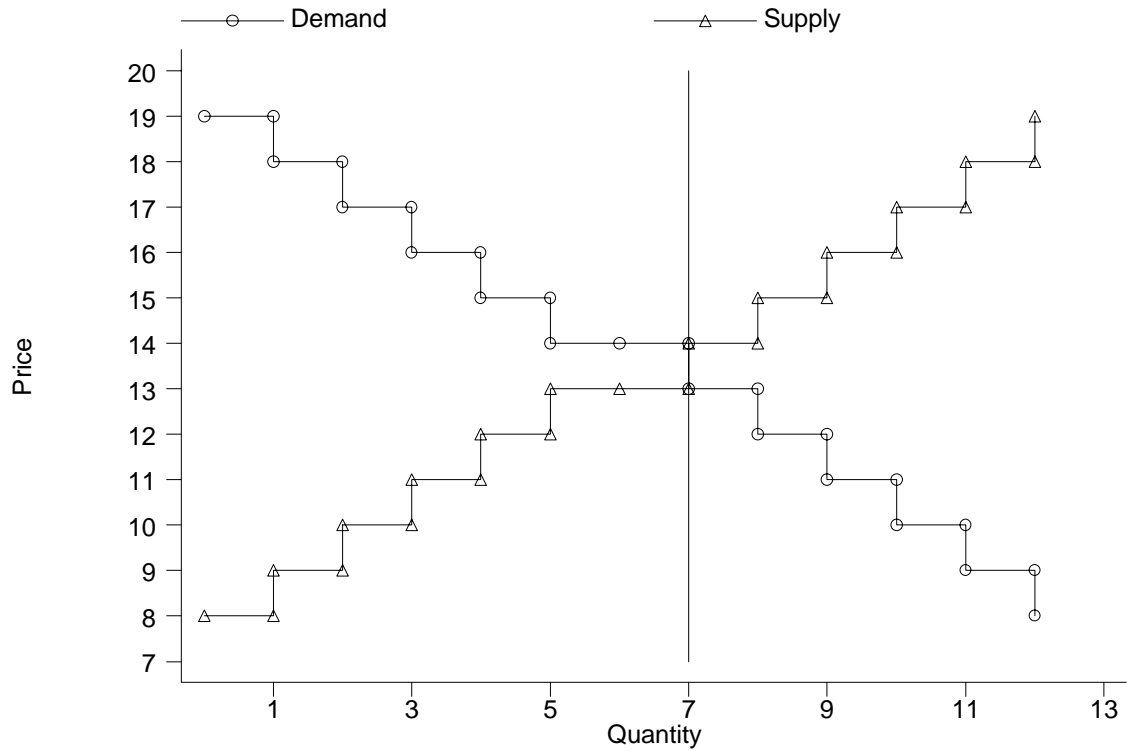
Figure 3. Buyer and seller induced values for market experiment

Table 9 Experimental Results—Chamberlin Market

Treatment	Market Period				
	(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.

Appendix: Summary Statistics

Sample	GARP Violations		Age		Gender		Mall Trips Per Month	
	Round 1	**Round 2**	**Round 1**	**Round 2**	**Round 1**	**Round 2**	**Round 1**	**Round 2**
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]

The header "Variable" spans across the Age and Gender columns.

Notes:

1. Numbers represent sample means. Standard deviations are in parentheses. Number of observations is in brackets.

2. Gender: 0 = male; 1 = female.

3. *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.