Social Preferences: Some Thoughts from the Field

John A. List

Department of Economics, University of Chicago, Chicago, Illinois 60637; and National Bureau of Economics Research, Cambridge, Massachusetts, 02138-5398; email: jlist@uchicago.edu

Annu. Rev. Econ. 2009. 1:563–79

First published online as a Review in Advance on June 17, 2009

The *Annual Review of Economics* is online at econ.annualreviews.org

This article's doi: 10.1146/annurev.economics.050708.142958

1941-1383/09/0904-0563$20.00

Key Words

nonstandard preferences, lab and field experiments

Abstract

This review steps back from the burgeoning economics literature on measuring social preferences and considers more carefully the empirical evidence from the lab and the field. I place the claims from the ardent supporters of the literature into three bins: one for claims that are supported by the data upon closer scrutiny, one for claims that are not supported by the data upon closer scrutiny, and one for claims that may or may not be true. The third set of claims highlights important theoretical and empirical investigations that need to be done to further our understanding of the nature and import of social preferences.

> [W]hen self-interest and ethical values with wide verbal allegiance are in conflict, much of the time, most of the time in fact, self-interest theory. . .will win.
>
> Stigler (1981)

1. INTRODUCTION

Consonant with the spirit of the words of Stigler in the epigraph, one of the traditional assumptions in economic modeling is that agents are purely self-interested. An influential collection of laboratory experiments measuring social preferences has called into question the validity of this position (see, e.g., Camerer & Weigelt 1988, Fehr et al. 1993, Berg et al. 1995, Charness 2004, Fehr et al. 1997, Charness & Rabin 2002). This literature is complemented by an entire body of theoretical research exploring the economic consequences of social preferences, wherein agents have preferences that are measured over their own and others' material payoffs (Rabin 1993, Levine 1998; Fehr & Schmidt 1999; Bolton & Ockenfels 2000; Andreoni & Miller 2002; Charness & Rabin 2002). In this review, I explore social preferences under this broad definition and am not interested in pinpointing whether the behavior consistent with social preferences is due to altruism, reciprocity, fairness, inequality aversion, or another motive.

The social preference results have attracted much attention, as some have argued that they are relevant beyond the context inherent in the laboratory. Indeed, Fehr et al. (1993, p. 437) note that their gift-exchange results "provide. . .experimental support for the fair wage-effort theory of involuntary unemployment." In more recent work, Fehr et al. (2007) push the results further by arguing how labor market models should be changed to reflect the manner in which fairness perceptions affect the labor market.

Overall, this research agenda has taught us much about social preferences and some of the important economic consequences of including social preferences in our models. This is an invaluable step in the discovery process, and I laud the literature for this progress. I find this research agenda fascinating and firmly believe in its potential. Yet, whether the findings in this literature should lead to a fundamental reformulation of our theoretical models—from changing models of contracting and optimal wage theories to revising our expectations of the various equilibria that we expect markets to approach—remains an open empirical question.

To lend insight into such issues, I take a step back from the burgeoning literature and consider more carefully the empirical evidence—from both the lab and the field—that has caused this most recent ruckus in the economics community. What is open for debate is whether the nature and extent of these preferences should induce us to rethink our overall economic modeling approach. A useful appraisal of the social preferences literature can be done by placing the claims from the ardent supporters into three bins: (*a*) claims that are supported by the data upon closer scrutiny, (*b*) claims that are not supported by the data upon closer scrutiny, and (*c*) claims that may or may not be true; further theoretical and empirical investigation must be conducted to move such claims from the third bin into the first or second bin.

Bin *a* includes claims such as "there is individual heterogeneity, but a portion of economic agents has social preferences." Indeed, a collection of agents certainly has social preferences, and substantial lab and field data support this claim (examples in my own field experimental data include List & Lucking-Reiley 2002, List 2004, Landry et al. 2006, Karlan & List 2007). Bin *b* includes arguments such as "in one-shot environments,

the effect of social preferences on aggregate market efficiency is large, and trusting actions are profitable for principals in sequential prisoner's dilemma games." Data from both the lab and field refute such claims.[1] Bin *c* contains claims such as "repeated interactions are a powerful multiplier of the effect of fairness concerns." In addition, claims concerning issues about generalizability of results from the lab to the field should be housed in bin *c*. I provide some evidence below concerning such generalizability, but any strong general conclusions concerning generalizability would be premature, since the literature is in its infancy.

The remainder of my review proceeds as follows. I begin by presenting a framework in which to interpret the received experimental results. I then briefly summarize the findings within this framework, focusing on individual heterogeneities, as well as the effects of stakes, the nature and extent of scrutiny, context, and the experimental duration on the received results. I conclude with some remarks on future work.

2. A FRAMEWORK

Recently, with Levitt (Levitt & List 2007), I have argued that human decisions are influenced not only by monetary calculations but also by at least five other factors: (*a*) the presence of moral and ethical considerations, (*b*) the extent to which one's actions are scrutinized by others and the nature of that scrutiny, (*c*) the particular context in which the decision is embedded, (*d*) the subject pool of individuals making the decisions, and (*e*) the stakes of the game. Here I reconsider briefly the framework we introduced (see also List 2008).

A utility-maximizing individual i is faced with a choice regarding a single action $a \in (0,1)$. The choice of action affects the agent's utility through two channels. The first effect is on the individual's wealth (denoted W_i). The higher the stakes or monetary value of the game, denoted v, the greater the decision's impact on W_i. The second effect is the nonpecuniary moral cost or benefit associated with action i, denoted as M_i. If, for instance, an individual has strong social preferences, he will derive utility from giving pecuniary and nonpecuniary gifts to others, for example.

In practice, many factors influence the moral costs associated with an action, but for modeling purposes, we focused on just three aspects of the moral determinant: (*a*) the greater the negative impact of an action on others, the more negative the moral payoff M_i; (*b*) the strength of social norms or legal rules (n) that govern behavior in a particular society influences behavior, and (*c*) moral concerns depend on the nature and extent of scrutiny of an individual's actions. Scrutiny is inherently multidimensional, but for simplicity we considered only its nature and extent. An example of the nature of scrutiny is the presence of an experimenter, who potentially alters the subject's perception of the situation. More broadly, the experimental environment itself might draw upon a different set of expectations than markets do. The extent of scrutiny relates to the anonymity of the subject's decision.

In certain cases, the nature of scrutiny in the field closely resembles that in the lab. Perhaps doctor/patient or drill sergeant/trainee settings are the closest parallels to the experimenter/subject relationship. Clearly, the nature of scrutiny varies considerably across field settings, as does the extent of anonymity (e.g., purchasing art in an Internet

[1]Yet, I am sure that a vector of laboratory parameters exists that will yield a measurable effect of social preferences on market efficiency that is also profitable for principals in the lab. My intuition is that if one simply increases the multiplier, that will do the trick.

auction versus at a Sotheby's auction). In the model below, I denote the effect of scrutiny as s, with higher levels of s associated with greater moral costs.

Focusing on the case in which utility is additively separable in the moral and wealth arguments, I make the utility function when an individual i takes action a as

$$U_i(a, v, n, s) = M_i(a, v, n, s) + W_i(a, v). \quad (1)$$

Solving this simple decision problem yields several predictions (Levitt & List 2007). For example, the greater the social norm against the wealth-maximizing choice or the greater the degree of scrutiny, the larger the deviation from that choice. Further, as the stakes of the game rise, wealth concerns will increase in importance relative to fairness concerns; that is, $|\partial M/\partial v| < |\partial W/\partial v|$. Such a framework makes it clear that the more the lab environment mirrors the naturally occurring setting that it is modeling, the more confident one can be that the lab results will be generalizable. If the lab diverges from the environment of interest, the model provides a framework for predicting in what direction behavior in the lab will deviate from that outside the lab.

3. EMPIRICAL EVIDENCE

The model can speak to a wide range of experimental results, but its bite is likely to be greatest for those games in which there is the potential for a strong moral component to behavior. Research on social preferences, one of the most influential areas in experimental economics in recent years, fits this bill. This broad class of games includes dictator and ultimatum bargaining games, public goods games, trust games, and gift exchange games. **Table 1** describes each game, the overall pattern of results, and the social preference interpretation. Results from these types of experiments have been used to argue that prosocial preferences are important in a wide range of real world settings, based on the assumption that the experimental findings are descriptive of the world at large.

In what follows, I use the model as a guide to examine the empirical evidence regarding possible complications to extrapolating the experimental findings outside the lab. I am certainly not denying that individuals have social preferences; indeed, my own results mentioned above and the framework summarized in Section 2 assume that moral costs can be influenced by a concern for others as well as a concern for one's own appearance. Rather, I am interested in the extent to which the lab provides reasonable guidance as to the importance of such behavior in a wide range of naturally occurring settings. One important assumption maintained in the literature is that laboratory experiments are useful in this regard because the experimenter precludes that reciprocal responses (W) will lead to future material rewards, effectively isolating social preferences (M). The key behind this approach is that the analyst creates an environment that the subjects truly view as one-shot and act accordingly.

3.1. Subject Pool

According to the simple framework in Section 2, several assumptions must be made if we are to conclude strongly that behavior in the lab is a good indicator of behavior in the field. In regard to the subject pool, I demonstrate this point with a simple graphical interpretation of this model (see also List 2006a, which this discussion closely follows). For ease of exposition, assume that the action represents how much money to send to an

Table 1 Summary of experimental games used to measure social preferences*

Name of game	Brief description	Typical finding	Social preference interpretation
Ultimatum game[a]	A two-stage game where two people, a proposer and a responder, bargain over a fixed amount of money. First stage: proposer offers a split of the money. Second stage: responder decides to accept or reject the offer. If accepted, each player receives money according to the offer; if rejected, each player receives nothing	Proposers: majority of offers in the range of 25%–50% of fixed amount. Few offers below 5%. Responders: frequently reject offers below 20% of fixed amount	Proposers: fairness. Responders: punish unfair offers: negative reciprocity, fairness preferences (such as inequity aversion)
Dictator game[b]	A simple variant of the ultimatum game. Strategic concerns are absent as the proposer simply states what the split will be and the proposer has no veto power, rendering the proposed split as effective	Usually >60% of subjects pass a positive amount of money, with the mean transfer roughly 20% of the endowment	Altruism; fairness preferences, such as inequity aversion
Trust game[c]	A sequential prisoner's dilemma game. The first mover decides how much money to pass to the second mover. All money passed is increased by a factor greater than 1. The second mover then decides how much money to return to the first mover. The second mover is a dictator who has been given his endowment by the first mover	Proposers: average transfer of roughly 50% of endowment. Responders: repayment is increasing in transfer. Average repayment rate is nearly 50% of transfer	Proposers: trust; foresee positive reciprocity. Responders: trustworthiness, positive reciprocity
Gift exchange game[d]	Similar to the trust game, but the first mover's money passed (or "wage" or "price" offer) is not increased by a factor. Rather, it represents a pure lump-sum transfer. Also, the first mover requests a desired effort, or quality, level in return for the "wage" or "price" offer. The second mover then chooses an effort or quality level that is costly to provide but increases the first mover's payoff	Proposers: "wage" or "price" offer is typically greater than the minimum allowed. Responders: effort, or quality increases in "wage" or "price" offer	Proposers: trust; foresee positive reciprocity. Responders: trustworthiness, positive reciprocity

(*Continued*)

Table 1 (*Continued*)

Name of game	Brief description	Typical finding	Social preference interpretation
Public goods game[e]	Generalization of the prisoner's dilemma game in that n group members decide simultaneously how much to invest in the public good. The payoff function is given by $P_i = e - g_i + \beta\sum_n g_j$, where e represents initial endowment, g_i is the number of tokens that subject i places in the group account, β is the marginal payoff of the public good, and $\sum_n g_j$ is the sum of the n individual contributions to the public good. The dilemma arises when $0 < \beta < 1 < n\beta$	Players' contribution to public good is roughly 50% of endowment in one-shot games. Many players' contributions "unravel" to approach 0% in latter rounds of multiperiod games	Altruism; fairness preferences, conditional reciprocity

*This table is taken from Levitt & List (2007).

[a]See Roth (1995) for a discussion of ultimatum and dictator games. This game was first proposed in the economics literature by Guth et al. (1982).

[b]This game was first proposed in the economics literature by Kahneman et al. (1986). A related game is the punishment game, whereby an observer can, for a cost, punish the first mover by subtracting a portion of the first mover's payoff.

[c]This game was first proposed in the economics literature by Berg et al. (1995).

[d]This game was first proposed in the economics literature by Fehr et al. (1993), and a related game is described by Camerer & Weigelt (1988). The payoff function description for the buyer is similar to Fehr et al.'s (1997) S13–S16 treatments. In this case, the price represents a pure lump-sum transfer, which differs from the earlier joint profit equation (Fehr et al. 1993), which was characterized by price increases leading to an increase in the sum of payoffs when $q < 1$.

[e]This is a generalization of the famous prisoner's dilemma game.

anonymous stranger in another room (i.e., a dictator game). Assume that $\tau_i = y_{i1} - y_{i0}$ is the treatment effect for individual *i*, which in this case represents what some have argued is the degree of social preferences of agent *i*. **Figure 1** shows a hypothetical density of τ_i in the population, a density assumed to have mean τ^*. In this case, the parameter τ^* is equivalent to the average treatment effect; this is the treatment effect of interest if the analyst is pursuing an estimate of the average social preferences in this population.

Of first concern is that selection into the lab experiment is not random, but might occur with a probability related to τ. One example of such an argument can be found in the social psychology literature, where it has been asserted that "scientific do-gooders" interested in the research, or students who readily cooperate with the experimenter (a professor) and seek social approval, are those students who select into the lab (Orne 1962). Using this notion to formulate the selection rule leads to positive selection: When students are offered the option to volunteer as subjects in an experiment, those with higher τ values are more likely to participate. In **Figure 1**, I denote the cutoff value of τ_i as τ^+: students above τ^+ participate, those below do not.

In this example, the treatment effect on the treated, τ_{TT}, is what is measured in the lab experiment. τ_{TT} is equal to $E(\tau_i|\tau_i > \tau^+)$, which represents the estimate of social preferences for those who participate. A lack of recognition of selection causes the analyst to mismeasure the treatment effect for the population of interest. **Figure 1** also shows the treatment effect on the untreated, τ_{TU}. This τ_{TU} is equal to $E(\tau_i|\tau_i < \tau^+)$, which represents the unobserved estimate of social preferences for those who chose not to participate. Therefore, the population parameter of interest, τ^*, is a mixture of these two effects: $\tau^* = Pr^*\tau_{TT} + (1-Pr)^*\tau_{TU}$, where Pr represents the probability of $\tau_i > \tau^+$. Even if one assumes that the population density of τ_i among students is isomorphic to the density of the population of interest, such selection frustrates proper inference.

Even with no selection effects, a related question is whether the density of τ_i in the student population is similar to the density of the population of interest. Under the framework of Equation 1, this revolves around the question of whether population distributions have similar structures (that is, $M_i \neq M_j$ for individuals *i* and *j*, for example). One approach to investigating this question is to run experiments with professionals, or other representative agents [artifactual field experiments as per Harrison & List (2004)], and compare the results to those obtained with students in similar laboratory experiments.

Fehr & List (2004) provide one example in a game that trades off morality and wealth. They examine experimentally how chief executive officers (CEOs) in Costa Rica

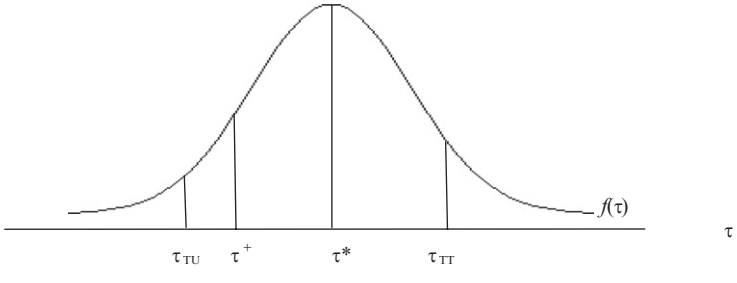

Figure 1

Simple illustration of heterogeneous treatment effects (List 2006a).

behave in trust games and compare their behavior with that of Costa Rican students. They find that CEOs are considerably more trusting and exhibit more trustworthiness than students. These differences in behavior may mean that CEOs are more trusting in everyday life, or it may be that CEOs are more sensitive to the lab and nonanonymity effects, or that the stakes are so low for the CEOs that the sacrifice to wealth of making the moral choice is infinitesimal. Despite these ambiguities, the experiment does show lab evidence consonant with social preferences from a much different subject pool than what is usually studied.

Subsequent results have lent similar insights. For example, I have used a set of field experiments to explore behavior in linear public goods games (List 2004). I find that the level and temporal nature of the contributions are consistent with data reported in lab experiments. Furthermore, I find that whereas younger and middle-aged subjects tend to contribute at rates consistent with extant laboratory data, more mature subjects contribute larger amounts of their endowment to the public good—in both multiple-shot and one-shot games. This reported correlation of age and prosocial preferences has since been found to be quite robust (see, e.g., Carpenter et al. 2005b, 2008; Holm & Nystedt 2005; Bellemare & Kröger 2007; Sutter 2007; Sutter & Kocher 2007; Gächter & Herrmann 2007; Guth et al. 2007). Although this evidence broadly paints a similar picture, as the model in the previous section makes clear, in order for these laboratory findings to be interpreted as deep structural parameters, it must be the case that the scope of lab and nonanonymity effects is similar across experimental populations.

Perhaps making this point most vividly is the fascinating work of Henrich et al. (2005), who conducted ultimatum, dictator, and public goods games in 15 different small-scale communities in developing countries. They report enormous variation in behavior across communities that correlate with patterns of everyday life and the social norms operating in these various communities (the following passage closely follows Levitt & List 2007). For instance, as Henrich et al. (p. 31) note, the Orma readily recognized "that the public goods game was similar to the *harambee*, a locally-initiated contribution that Orma households make when a community decides to construct a public good such as a road or school," and consequently gave quite generously. Likewise, when the ultimatum game is conducted among the whale-hunting Lamalera of Indonesia and the Ache in Paraguay, societies with strong norms of sharing, very generous offers are observed and very few offers are rejected. Alternatively, in small-scale foraging societies, such as the Hadza of Tanzania, low offers and high rejection rates are observed in ultimatum games.

In all of the experiments Henrich et al. (2005) conducted, the context that the experimenter can control—the payoffs, the description of the way the game is played, and so on—was almost identical. But the context that actors themselves brought to the game and experimenters cannot control—like past experiences and internalized social norms—proved centrally important in the outcome of play. Indeed, as Henrich et al. note (p. 33), these "contrasting behaviors seem to reflect their differing patterns of life, not any underlying logic of hunter-gatherer life ways." One must consider whether the experimenter can ever induce the true one-shot settings necessary to ascribe behaviors consonant with one-shot theories.

3.2. Stakes

The model also predicts that in games that have both a morality and a wealth component, financial concerns will take on increasing prominence as the stakes rise. The evidence in the literature is only partially consistent with this view. In dictator games, a large increase

in stakes generally leads to a less-than-proportionate increase in money transferred. For example, Carpenter et al. (2005a) report an increase in stakes from $10 to $100 caused the median offer to drop from 40% to 20% of the endowment. This result is much weaker for smaller changes in stakes: Cherry et al. (2002) find no perceptible differences in offers between a $10 and $40 dictator game. Stakes effects have also been found in second-mover play in ultimatum games, in which the acceptance rate is generally increasing in the amount offered, conditional on the share offered—that is, a $1 offer in a $5 game is rejected more often than a $100 offer in a $500 game. Slonim & Roth (1998) find that in each range of offers below 50%, the acceptance rate goes up as the level of stakes increases (from 60 to 1500 Slovak koruna, the latter of which represents eight days of wages for the typical worker). In another type of game that involves some form of trust, the centipede game, Parco et al. (2002) similarly find that raising financial incentives causes a breakdown in mutual trust. Fehr et al. (2002), however, report fairness concerns play an important role in both low and high stakes in trust and gift exchange games.

The point of the model is not to say that low-stakes games in the lab have no market parallels—we take part in such transactions in markets every day. The point is that if the analyst does not account properly for the differences in stakes across settings, inaccurate inference concerning the importance of prosocial preferences might result. The magnitude of such mismeasurement is a rich area for future research, and it would be interesting to compare such effects to the size of the effects of other factors, such as context and selection.

3.3. Context: Time Horizons

One important aspect of the context of lab experiments that is not oft-discussed is duration. Laboratory experiments usually consist of at most a few hours of activities. For example, in gift exchange games in the laboratory, student subjects typically play several rounds of the game by choosing an effort or wage level (by circling or jotting down a number) in response to pecuniary incentive structures. The experiment usually lasts about an hour and a result often observed is that effort levels and wages are positively correlated. Such results are often interpreted as providing support for the received labor market predictions of Akerlof (1982) on gift exchange.

Yet, standard economic theory reminds us that a short-run change in wages will only create substitution effects, whereas long-run changes also include income effects, which tend to cancel out substitution effects. Insights from the psychology literature also point to potentially important behavioral differences between short-run ("hot") and long-run ("cold") decision making (see, e.g., Loewenstein 2005).

Accordingly, at least two relevant issues arise. First, is real-world on-the-job effort different in nature from that required in lab tasks? Second, does the effect that we observe in the lab manifest itself over longer time periods? The evidence is sparse within the experimental economics literature on these issues, but studies are beginning to emerge.

One early study is the work of Gneezy & List (2006). To provide a first test of the gift exchange hypothesis in an actual labor market, Gneezy invited people to take part in an effort to computerize the holdings of a small library at a large university in the Midwest. Recruitment was done via posters that promised participants *one-time* work that would last six hours and that would pay $12 per hour, or $72. Participants were not informed that they were taking part in an experiment. The first treatment paid a flat wage of $12 per hour (denoted "Non-Gift"), as promised. In the second treatment (denoted "Gift"),

once the task was explained to the participants, they were told that they would be paid $20 per hour, not the $12 that had been promised.

Figure 2 (see color insert) depicts the quantity of work done by the two groups. In line with the gift exchange hypothesis, participants in the $20-per-hour wage treatment provided significantly higher effort in the first 90 min than participants in the $12-per-hour wage treatment. After 90 min on the job, however, effort levels were indistinguishable across the two treatments.

In a second field experiment, List invited students to work in a door-to-door fundraising drive to support the Natural Hazards Mitigation Research Center at a large university in the Southeast (see Landry et al. 2006 for full results). Similar to the library task, participants were told that this was one-time work for which they would be paid $10 per hour. Yet an important difference in this case is that workers have a better idea about the surplus and how much the employer valued the task. If workers know only the promised wage and not the surplus, as in our library task, only the promised wage can serve as a reference point. If the surplus is known, the share of the surplus that the workers receive will determine whether they perceive their wage as fair.

The first treatment was a flat wage of $10 per hour, as promised. In the second treatment, once the solicitors were trained, they were told that they would be paid $20 per hour, not the $10 that had been advertised. Empirical results differ slightly from those of the library treatment but tell a similar story: Whereas the number of doors knocked on in any of the hours worked is not different across treatments, solicitors in the $20-per-hour wage treatment raised significantly more money in the first few hours of the task than solicitors in the $10-per-hour wage treatment. Yet, after a few hours the observed outcomes were indistinguishable. I provide **Figure 3** (see color insert) as visual evidence.

The notion that positive wage shocks do not invoke long-run effects in effort levels is also consistent with data reported by Al-Ubaydli et al. (2008), Hennig-Schmidt et al.'s. (2006) field experiment, and Kube et al. (2006). These results suggest that great care should be taken before extrapolating from short-run laboratory experiments to long-run field environments.

Naturally occurring data concerning the effects of pay shocks on work effort are mixed. Chen (2005), who uses a large data set drawn from the Australian Workplace Industrial Relations Survey to explore reciprocity in the workplace, finds little evidence consistent with reciprocity. Lee & Rupp (2007) examine the effort responses of U.S. commercial airline pilots following recent pay cuts and find that such effects are very short-lived, consistent with Gneezy & List (2006). In the first week after a pay cut, frequent and longer flight delays are observed, but after the first week, airline flight performance reverts to previous levels. In contrast, Krueger & Mas (2004) provide evidence *consistent with*[2] negative reciprocity on the part of disgruntled Firestone employees, and Mas (2006) documents persistent adverse effects on police performance following arbitration decisions in favor of the municipality. In addition, Al-Ubaydli et al. (2008) and Kube et al. (2006) do find some evidence of short- and long-run effects of negative gifts, though in Kube et al. (2006) there is some evidence of convergence, consonant with Gneezy & List (2006).

As the preceding discussion suggests, the evidence is generally mixed or negative on the gift exchange relationship in the long run when the repeated-game incentives are sup-

[2]I stress *consistent with* because, of course, such results are consonant with other models as well. For example, simple efficiency wage models also predict such data patterns.

pressed. Thus, a useful take-away point is that the impact of gift exchange in one-shot interactions on aggregate market efficiency is small. Equally important, engaging in gift exchange is unprofitable for principals in these settings. This interpretation of the lab and field results is a general departure from the earlier literature that argued ferociously about the empirical importance of such preferences in one-shot settings (see Fehr et al. 2007 for more on this departure).

3.4. Context: Choice Set

Another contextual issue in experiments relates to the set of rules governing the interactions, the wording of instructions, and the set of actions the subject is allowed to take. In contrast to the lab, where the researcher designs these factors, in naturally occurring environments the choice set often is almost limitless and institutions arise endogenously. Importantly, agents with market power are typically those who are responsible for constructing the context of the situation.

Even among those who choose to participate in lab experiments, restrictions on the available choice set can affect observed behavior. For example, prosocial behavior might be observed less frequently in markets merely because people can avoid situations where they must make costly contributions to signal their generosity. This idea is illustrated by Lazear et al. (2004), who allowed agents an opportunity to pay to opt out of playing the dictator game experiment. They find that "the majority of subjects share without really wanting to, as evidenced by their willingness to avoid the dictator game and to even pay for avoiding it." Such forces are readily observable in the field as well—panhandlers receive less in gifts if passersby can easily "sort" themselves to the other side of the road to avoid interaction.

Another example of how the available choice set influences play in the dictator game is provided by List (2007) and Bardsley (2008). In the typical dictator game, the subject is given, say, $10 and asked what portion the subject would like to share with the other player, who received less than $10. The experiment is framed such that "giving nothing" is the least generous act, and substantial sums of money are given away. If instead the subject is given $10 and is told that the rules allow giving any portion of this money away to the second player, or confiscating up to an additional $10 from the other player, little is known about how the results change. This is exactly what List (2007) and Bardsley (2008) offer.

Figures 4 and **5** (see color insert) summarize two of the treatments in List (2007). In total, I had 24 dictators in the baseline treatment (**Figure 4**) and 50 dictators in the "take" treatments (**Figure 5**). Of first note is the finding that the baseline data are qualitatively similar to results reported in other dictator games: 17 of 24 dictators gave a nonzero amount, and the mean amount given was roughly 25% of the endowment ($32 of $120 was given). Simply manipulating the choice set has considerable effects, however. A comparison of **Figures 4** and **5** shows that allowing taking significantly shifts the distribution leftward: In the take treatment few agents gave a positive amount—5 of 50—a proportion that is significantly lower than the proportion of givers in the baseline treatment. This result shows that by simply making the action set symmetric, nearly all of the giving vanishes. If true social preferences underlie the rates of giving observed in the standard dictator game, then similar numbers of subjects should give across the baseline and the take treatments. Clearly this does not occur, calling into question the underlying

mechanism at work when agents are observed giving in dictator games. The results in **Figures 4** and **5** are more in line with the model described in Section 2.

Other empirical results yield similar insights. In an earlier study, Andreoni et al. (2002) change the payoff space in a sequential laboratory public goods game with an asymmetric equilibrium and report results consistent with the data in List and Bardsley. Perhaps establishing the effect of context most vividly, Ellingsen et al. (2008) cleverly show how behavior changes drastically when labels are changed. Further, they show that people respond to labels that others may be using to interpret and evaluate their behavior.

These examples also highlight that laboratory experiments often restrict the response mode to a single dimension, whereas real-world settings almost always involve multiple response modes. Consider again the act of giving within a dictator game. An agent who is inclined to help others might give money in the dictator game in the lab. In the field, this same agent might give nothing, instead using other more efficient means to express generosity, such as volunteering time to help others. In this example, the laboratory evidence is consistent with some type of broader preference, but that preference might be expressed through a different activity in the field.

3.5. Nature and Extent of Scrutiny

Because Levitt & List (2007) expand on the effects of anonymity, I focus here on the effect of the nature of scrutiny. In the typical lab experiment, subjects enter an environment in which they are keenly aware that their behavior is being monitored, recorded, and subsequently scrutinized. Decades of research within psychology highlight the power of the role obligations of being an experimental subject, the power of the experimenter herself, and the experimental situation (see, e.g., Pierce 1908, Orne 1962). Clearly, the nature of scrutiny inherent in the lab is rarely encountered in the field and represents an important aspect of the situation that needs to be accounted for when generalizing laboratory results.

One example of the effect of scrutiny is given by List (2006b). In that study, I carried out gift exchange experiments in which buyers make price offers to sellers, and in return sellers select the quality level of the good provided to the buyer. Higher quality goods are costlier for sellers to produce than lower quality goods but are more highly valued by buyers. I began by running a standard gift exchange game in a laboratory context, but using experienced sportscard traders as subjects. The results, shown in **Figure 6** (see color insert), mirror the typical findings with other subject pools: strong evidence for social preferences was observed, in the sense that buyers offered higher quality levels to sellers who offered higher prices—although the buyers were not obligated by the rules of the game to do so.

I then carried out a second lab experiment that maintained the central elements of the gift exchange game, but the goods exchanged in this lab treatment were actual baseball cards whose market values are heavily influenced by minor differences in condition that are difficult for untrained consumers to detect. If social preferences are present on the part of card sellers, then buyers who offer more money should be rewarded with higher quality cards. When card sellers were brought into the lab to sell their cards, which were subsequently professionally graded, the results paralleled those obtained in **Figure 6** and in the standard gift exchange game with student subjects.

I then moved from a lab environment in which sellers knew their behavior was being scrutinized to their natural environment. Importantly, dealers in this treatment were unaware that their behavior was being recorded as part of an experiment. Confederates were

sent as buying agents to approach dealers on the floor of a sportscard show. Confederates had been instructed to offer different prices in return for sportscards of varying quality, just as in the lab treatment described above. When the dealers believed that consumers could not have the card graded, or when there was likely to be little future interaction, a minor statistical relationship between price and quality emerged. Only when there were reputational consequences to a dealer (i.e., when quality was verifiable and the potential for a long-term relationship existed) was high quality provided. The social preferences so routinely observed in the lab—even for this very same group of traders—were attenuated in the field.

Figure 7 (see color insert) summarizes these data from the local dealer group (those who would have reputational concerns). In all of the treatments except for the Floor-NoGrading treatment, both reputation concerns and social preferences are "turned on;" thus, a positive relationship between price and quality can be caused by one or both factors. When I "turn off" reputational concerns in the Floor-NoGrading treatment, the relationship between price and quality is positive but significantly dampened.

Other field-generated data yield similar conclusions. For example, making use of personnel data from a leading United Kingdom–based fruit farm, Bandiera et al. (2005) find that behavior is consistent with a model of social preferences when workers can be monitored. Under a relative compensation scheme, workers internalize the negative externality that they impose on others when other workers can observe their productivity. Yet, this effect disappears when workers cannot monitor each other, which rules out pure altruism as the underlying cause of workers' behavior. Being monitored proves to be the critical factor influencing behavior in this study.

Relatedly, Benz & Meier (2008) compare how individuals behave in laboratory experiments involving charitable donation and how the same individuals behave in the field. They find some evidence of correlation across situations but find that subjects who have never contributed in the past to the charities gave 75% of their endowment to the charities in the lab experiment. Similarly, those who never gave to the charities subsequent to the lab experiment gave more than 50% of their experimental endowment to the charities in the lab experiment. Similar insights are reported by Laury & Taylor (2007), who compared an "altruism parameter" estimated from a public goods lab experiment and actual contributions to a real public good (urban tree planting nonprofit organization).

In a "dining game," Gneezy et al. (2004) find that behavior in a social dilemma game in the laboratory exhibits considerable cooperative behavior—i.e., in the lab, students showed great reluctance to impose negative externalities. Yet, in a framed field experiment that resembles the laboratory game—diners were taken to eat at a restaurant—the authors find no evidence of cooperative play, even though both experimental samples are drawn from the same student population. They speculate that unfamiliarity with the task and confusion are two reasons why negative externalities are influential in the lab but not in the field. Such results are consistent with the simple model detailed above.

Overall, these results are consistent with the wealth of psychological literature that suggests there is only weak evidence of cross-situational consistency of behavior (e.g., Mischel 1968, Ross & Nisbett 1991). Long ago, Hartshorne & May (1928) discovered that the people who cheat in one situation are not the people who cheat in another. If this result spills over to measurement of prosocial preferences, it means that (*a*) there is not a general cross-situational trait called "social preferences," and/or (*b*) the subjects view one situation as relevant to social preferences and the other as irrelevant.

4. DISCUSSION

The literature has now evolved to the point of arguing that "Repeated interactions are a. . .powerful multiplier of the effect of fairness concerns" (Fehr et al. 2007). Although this sounds plausible, I know of no empirical evidence that unambiguously shows this result. More specifically, I am unaware of data that suggest reputational concerns by themselves do not yield the data relationships that are consistent with gift exchange—that is, reputational concerns are able to explain the results in exclusion of fairness concerns.

The main message of this review is that the evidence of social preferences from laboratory games is more mixed than the strong advocates conclude. Interpretation from the field studies that are argued to provide strong results in the gift exchange game is confounded—both reputational concerns and social preferences are likely at work. And in those cases when in fact the field data are able to provide a clean measure of social preferences, the effect is found to be small. Lab experiments potentially avoid some of the confounds in field studies, but slight perturbations of experimental conditions can dramatically alter behavior, and the important properties of the lab situation are not conducive to fluid generalizability to the extralab world. In these ways, estimation of deep preference parameters is lacking, especially in light of the fact that we have no theory to generalize such parameters.

I view the lab as having a comparative advantage at providing unique qualitative insights. In addition, the lab is able to shed light on what can happen, rather than pinpoint what will happen in a certain field situation. For the social preference literature, the lab evidence has certainly highlighted an interesting phenomenon. Because the lab systematically differs from most naturally occurring environments on certain dimensions, experiments may not always yield results that are readily generalizable. This point, of course, applies with equal force to data generated from naturally occurring environments. In the end, the properties of the situation should be compared and contrasted, and theory and empirical evidence should be guiding forces to inform us of generalizability across domains.

DISCLOSURE STATEMENT

The author is not aware of any biases that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This paper draws heavily upon work I have published with numerous coauthors, most importantly Gneezy & List (2006) and Levitt & List (2007), and solo work [especially List (2006a, 2008)]. I thank Daron Acemoglu for sharp comments that improved the content tremendously.

LITERATURE CITED

Akerlof GA. 1982. Labor contracts as partial gift exchange. *Q. J. Econ.* 97(4):543–69

Al-Ubaydli O, Andersen S, Gneezy U, List JA. 2008. *Incentive schemes to promote optimal work performance: evidence from a multi-tasking field experiment*. Work. Pap., Univ. Chicago

Andreoni J, Brown P, Vesterlund L. 2002. What makes an allocation fair? Some experimental evidence. *Games Econ. Behav.* 40(1):1–24
Andreoni J, Miller J. 2003. Giving according to GARP: an experimental test of the consistency of preferences for altruism. *Econometrica*. 70:737–53
Bandiera O, Rasul I, Barankay I. 2005. Social preferences and the response to incentives: evidence from personnel data. *Q. J. Econ.* 120(3):917–62
Bardsley N. 2008. Dictator game giving: altruism or artifact? *Exp. Econ.* 11(2):122–33
Bellemare C, Kröger S. 2007. On representative social capital. *Eur. Econ. Rev.* 51(1):183–202
Benz M, Meier S. 2008. Do people behave in experiments as in real life? Evidence from donations. *Exp. Econ.* 11(3):299–321
Berg J, Dickhaut JW, McCabe KA. 1995. Trust, reciprocity, and social history. *Games Econ. Behav.* 90(1):166–93
Bolton GE, Ockenfels A. 2000. ERC-A Theory of equity, reciprocity, and competition. *Am. Econ. Rev.* 10(1):122–42
Camerer CF, Weigelt K. 1988. Experimental tests of a sequential equilibrium reputation model. *Econometrica* 56(1):1–36
Carpenter J, Connolly C, Myers K, Caitlin M. 2008. Altruistic behavior in a representative dictator experiment. *Exp. Econ.* 11(3):282–98
Carpenter J, Verhoogen E, Burks S. 2005a. The effect of stakes in distribution experiments. *Econ. Lett.* 86(3):393–98
Carpenter J, Verhoogen E, Burks S. 2005b. Comparing students to workers: the effects of stakes, social framing, and demographics on bargaining outcomes. In *Field Experiments in Economics*, ed. J Carpenter, G Harrison, JA List, pp. 261–90. New York: JAI
Charness G. 2004. Attribution and reciprocity in an experimental labor market. *J. Labor Econ.* 22(3):665–88
Charness G, Rabin M. 2002. Understanding social preferences with simple tests. *Q. J. Econ.* 117(3):817–69
Chen P. 2005. *Reciprocity at the workplace: Do fair wages lead to higher effort, productivity, and profitability*? Work. Pap., Dep. Econ., Australian Natl. Univ.
Cherry T, Frykblom P, Shogren JF. 2002. Hardnose the dictator. *Am. Econ. Rev.* 92(4):1218–21
Ellingsen T, Johannesson M, Munkhammar S, Möllerström J. 2008. *Why labels affect cooperation.* Work. Pap., Dep. Econ., Stockholm School Econ.
Fehr E, Fischbacher U, Tougareva E. 2002. *Do high stakes and competition undermine fairness—evidence from Russia*. Work. Pap. No. 120, Inst. Emp. Res. Econ., Univ. Zürich
Fehr E, Gachter S, Kirchsteiger G. 1997. Reciprocity as a contact enforcement device: experimental evidence. *Econometrica* 65(4):833–60
Fehr E, Goette L, Zehnder C. 2007. *The behavioral economics of the labor market: central findings and their policy implications*. Unpublished manuscript, Boston Federal Reserve
Fehr E, Kirchsteiger G, Riedl A. 1993. Does fairness prevent market clearing? An experimental investigation. *Q. J. Econ.* 108(2):437–59
Fehr E, List JA. 2004. The hidden costs and returns of incentives—trust and trustworthiness among CEOs. *J. Eur. Econ. Assoc.* 2(5):743–71
Fehr E, Schmidt K. 1999. A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114:8173–68
Gächter S, Herrmann B. 2007. *The limits of self-governance when cooperators get punished: experimental evidence from urban and rural Russia*. CeDEx Disc. Pap. No. 2007–11
Gneezy U, Haruvy E, Yafe H. 2004. The inefficiency of splitting the bill: a lesson in institution design. *Econ. J.* 114:265–80
Gneezy U, List JA. 2006. Putting behavioral economics to work: testing for gift exchange in labor markets using field experiments. *Econometrica* 74(5):1365–84

Guth W, Schmidt C, Sutter M. 2007. Bargaining outside the lab: a newspaper experiment of a three-person ultimatum game. *Econ. J.* 117:449–69
Guth W, Schmittberger R, Schwarze B. 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3:367–88
Harrison GW, List JA. 2004. Field experiments. *J. Econ. Lit.* 42(4):1009–55
Hartshorne H, May MA. 1928. *Studies in Deceit*. New York: Macmillan
Henrich J, Boyd R, Bowles S, Gintis H, Fehr E, et al. 2005. 'Economic man' in cross-cultural perspective: behavioral experiments from 15 small-scale societies. *Behav. Brain Sci.* 28(6):795–855
Holm HJ, Nystedt P. 2005. Intra-generational trust—a semi-experimental study of trust among different generations. *J. Econ. Behav. Organ.* 58(3):403–19
Kahneman D, Knetsch JL, Thaler RH. 1986. Fairness as a constraint on profit seeking: entitlements in the market. *Am. Econ. Rev.* 76(4):728–41
Karlan D, List JA. 2007. Does price matter in charitable giving? Evidence from a large-scale natural field experiment. *Am. Econ. Rev.* 97(5):1774–93
Krueger AB, Mas A. 2004. Strikes, scabs, and tread separations: labor strife and the production of defective Bridgestone/Firestone tires. *J. Polit. Econ.* 112(2):253–89
Kube S, Maréchal MA, Puppe C. 2006. *Putting reciprocity to work: positive versus negative responses in the field*. Work. Pap. No. 2006-07, Univ. St. Gallen
Landry C, Lange A, List JA, Price MK, Rupp N. 2006. Toward an understanding of the economics of charity: evidence from a field experiment. *Q. J. Econ.* 121(2):747–82
Laury S, Taylor L. 2007. *Altruism spillovers: Are behaviors in context-free experiments predictive of altruism toward a naturally occurring public good*? Work. Pap., Georgia State Univ.
Lazear EP, Malmendier U, Weber RA. 2004. *Sorting in experiments*. Work. Pap., Stanford Univ.
Lee D, Rupp NG. 2007. Retracting a gift: How does employee effort respond to wage reductions? *J. Labor Econ.* 25(4):725–62
Levine DK. 1998. Modeling altruism and spitefulness in experiments. *Rev. Econ. Dynam.* 1(3):593–622
Levitt SD, List JA. 2007. What do laboratory experiments measuring social preferences tell us about the real world? *J. Econ. Perspect.* 21(2):153–74
List JA. 2004. Young, selfish, and male: field evidence of social preferences. *Econ. J.* 114(492):121–49
List JA. 2006a. Field experiments: a bridge between lab and naturally occurring data. *Adv. Econ. Anal. Policy.* 6(2):8
List JA. 2006b. The behavioralist meets the market: measuring social preferences and reputation effects in actual transactions. *J. Polit. Econ.* 114(1):1–37
List JA. 2007. On the interpretation of giving in dictator games. *J. Polit. Econ.* 115(3):482–94
List JA. 2008. *Experimenting with Fish Has Some Advantages*. Boston: Boston Federal Reserve. In press
List JA, Lucking-Reiley D. 2002. Effects of seed money and refunds on charitable giving: experimental evidence from a university capital campaign. *J. Polit. Econ.* 110(1):215–33
Loewenstein G. 2005. Hot-cold empathy gaps and medical decision-making. *Health Psychol.* 32:151–65
Mas A. 2006. Pay, reference points, and police performance. *Q. J. Econ.* 121(3):183–221
Mischel W. 1968. *Personality and Assessment*. New York: Wiley
Orne MT. 1962. On the social psychological experiment: with particular reference to demand characteristics and their implications. *Am. Psychol.* 17(10):776–83
Parco JE, Rapoport A, Stein WE. 2002. Effects of financial incentives on the breakdown of mutual trust. *Psychol. Sci.* 13:292–97
Pierce AH. 1908. The subconscious again. *J. Philos. Psychol. Sci. Methods* 5:264–71
Rabin M. 1993. Incorporating fairness into game theory and economics. *Am. Econ. Rev.* 83(5):1281–302

Ross L, Nisbett RE. 1991. *The Person and the Situation; Perspectives of Social Psychology.* New York: McGraw-Hill

Roth AE. 1995. Bargaining experiments. In *The Handbook of Experimental Economics*, ed. JH Kagel, AE Roth, 4:253–342. Princeton, NJ: Princeton Univ. Press

Slonim R, Roth AE. 1998. Learning in high stakes ultimatum games: an experiment in the Slovak Republic. *Econometrica* 66(3):569–96

Stigler G. 1981. Economics or ethics? In *Tanner Lectures on Human Values*, ed. S McMurrin, pp. 1–31. Cambridge, UK: Cambridge Univ. Press

Sutter M. 2007. Outcomes versus intentions: on the nature of fair behavior and its development with age. *J. Econ. Psychol.* 28(1):69–78

Sutter M, Kocher M. 2007. Trust and trustworthiness across different age groups. *Games Econ. Behav.* 59(2):364–82

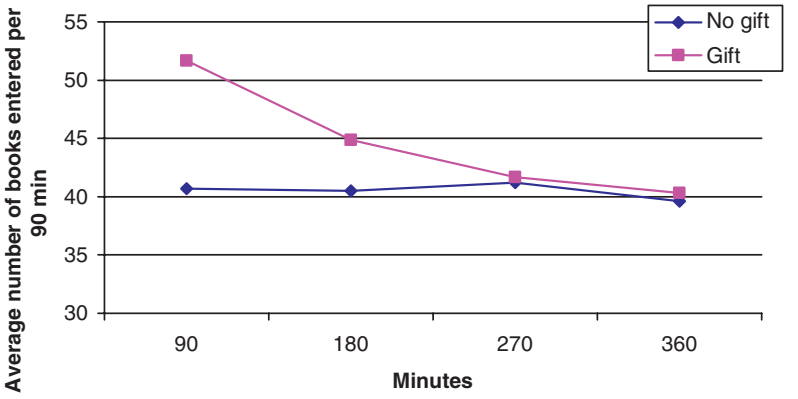

Figure 2

Average number of books logged per time period (Gneezy & List 2006).

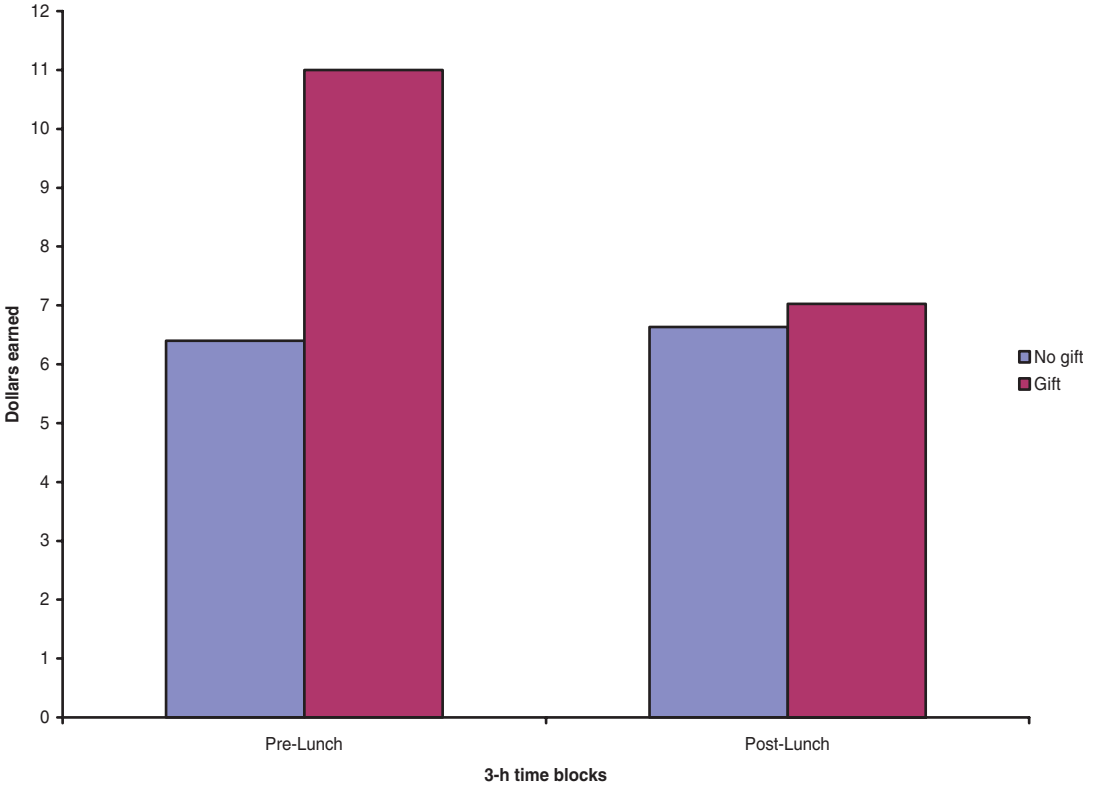

Figure 3

Average earnings by three-hour block (Gneezy & List 2006).

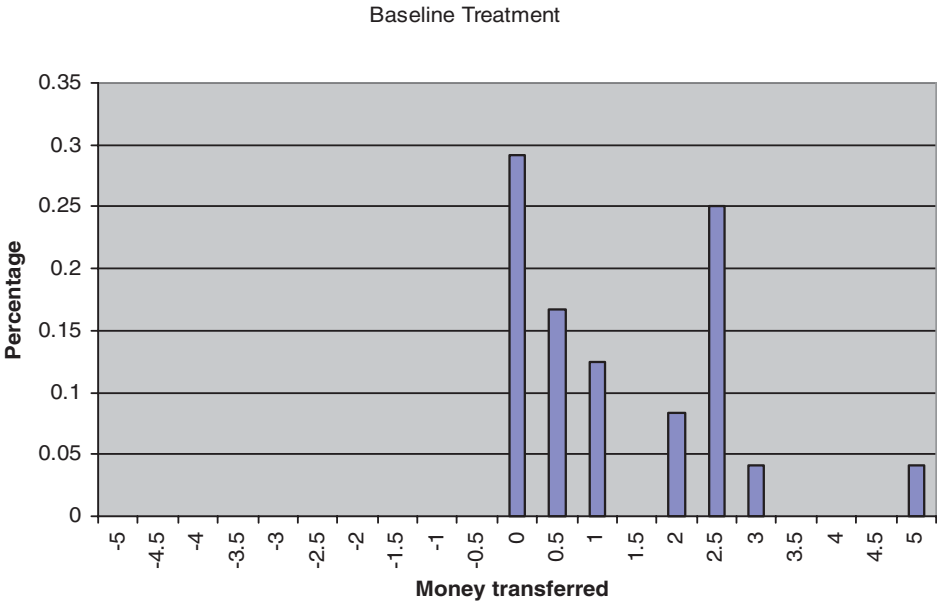

Figure 4

Dictator game giving—baseline treatment (List 2007).

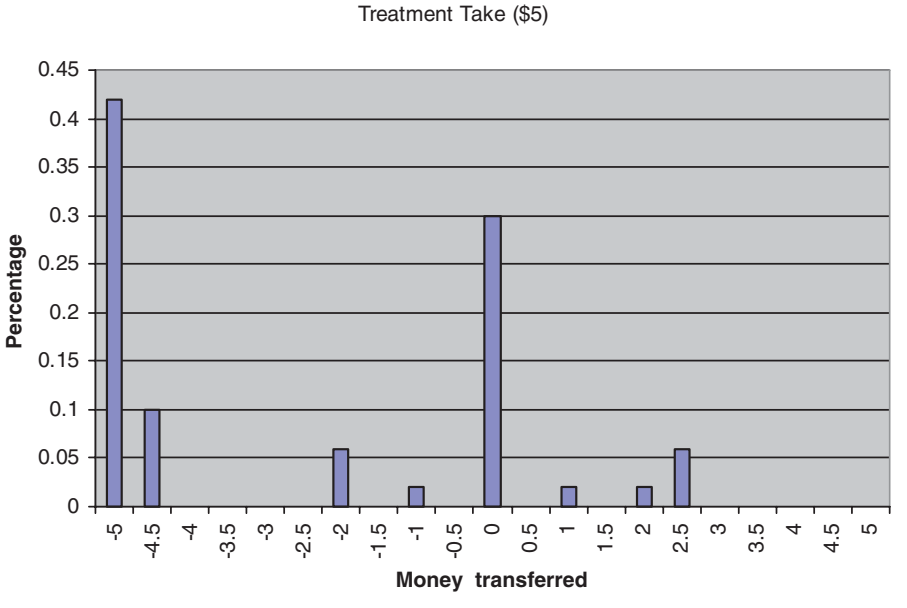

Figure 5

Dictator game giving—baseline treatment with "take" option (List 2008).

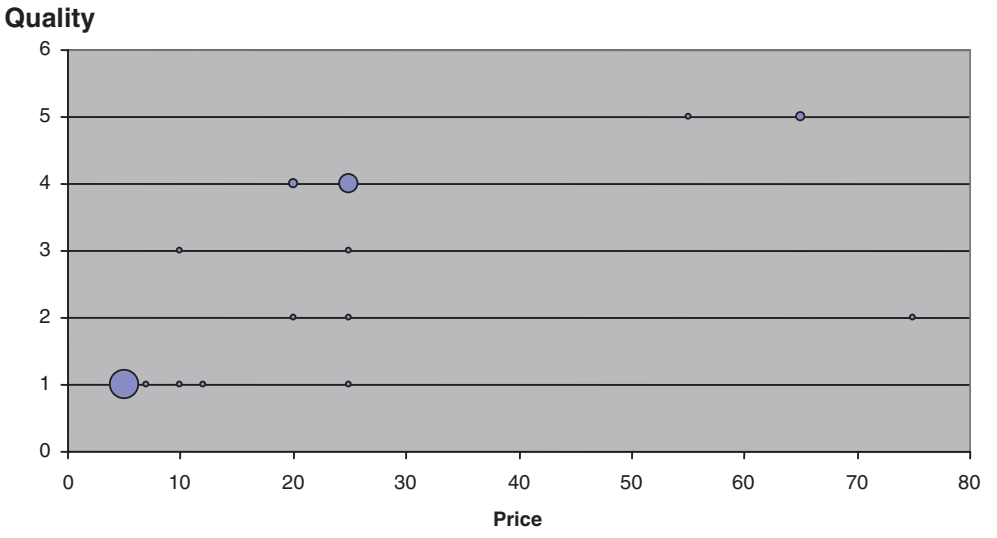

Figure 6

Gift exchange (List 2006b). Larger circles indicate that a greater number of observations occur at that point.

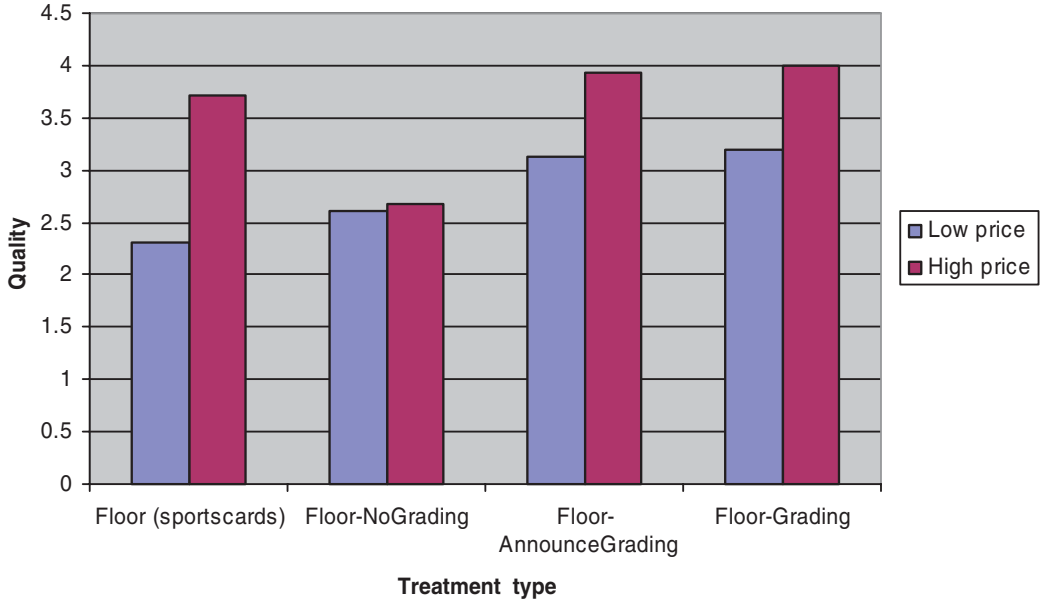

Figure 7

Price/quality relationship for local dealers.

Annual Review of Economics

Volume 1, 2009

Contents

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://econ.annualreviews.org