Draft: Comments Welcome

Laboratory Experiments

John H. Kagel

Department of Economics

Ohio State University

June 8, 2009

To be published in *Methods of Modern Experimental Economics*, Guillaume Frechette and Andrew Schotter (eds). Special thanks to my long-time co-authors Dan Levin and David Cooper as the ideas expressed here have been largely developed in working with them. Research has been partially supported by National Science Foundation grants SES-0452911, 0851674, and 0924764. Any opinions, findings and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The stated goal of this book, and the series it is part of, is to "formally open a dialogue on methodology," in this case for modern experimental economics. The particular topic I was asked to write about is lab experiments within the subheading "The Lab and the Field." In what follows I discuss laboratory experiments in relationship to field experiments, and more generally, in relationship to empirical research in economics and economic theory. All three empirical techniques, lab experiments, field experiments, and field data, are in principle fully complimentary to each other, and are informed by, and in turn inform, economic theory. Further, evaluating the results of laboratory experiments in relationship to results from other lab experiments, field experiments, field data and theory are all critical in determining the external validity of lab results. This discussion will take place within the context of two case studies I am reasonably familiar with, one dealing with auctions, the other dealing with gift exchange in labor markets.

In writing about these relationships, it is important to keep in mind two things in terms of the context, or background, I bring to the discussion. First, behavior in experiments is often marked by a considerable learning period, which is no doubt true in field settings as well. In this context one of the most important insights I have gained over the years is that whatever learning there is tends to be context dependent and difficult to generalize to new situations that are moderately, no less very, different from the learning context itself. This is not only my impression from the experimental results reported on here (which is rather overwhelming in the auction data I am most familiar with), but from reading results from the psychology explicitly designed to study learning and learning generalizability.[1] A second important insight is that oftentimes the same data says different things to different readers. That is, there is the experiment and then there is the interpretation of the results of the experiment – what the results mean in terms of the question(s) that motivated the study. It is not necessary to "buy" the authors interpretation of the results of an experiment in order to "buy" the experimental results. In fact, it is often conflicting interpretations of experimental results that leads to new experiments and better understanding of the phenomena in question, either by the initial

[1] There is a whole body of psychological literature indicating the difficulty of learning generalizing across different contexts (see, for example, Gick and Holyoak, 1980; Perkins and Salomon, 1988; Salomon and Perkins, 1989).

experimenters or competing groups. It is part of the healthy dialogue that is inherent to the research process.

I. The Winner's Curse

a. An initial set of results and how they relate to theory and experimental methodology

The winner's curse story begins with Capen, Clapp, and Campbell (1971), three petroleum engineers who claimed that oil companies suffered unexpectedly low returns "year after year" in early Outer Continental Shelf (OCS) oil lease auctions. OCS auctions are common value auctions, where the value of the oil in the ground is essentially the same to all bidders. Each bidder has their own estimate of the (unknown) value at the time that they bid. Even if these estimates are unbiased, bidders must account for the informational content inherent in winning the auction: the winner's estimate of value is (one of) the highest estimates. If bidders ignore this *adverse selection effect* inherent in winning the auction, it will result in below normal or even negative profits. The systematic failure to account for this adverse selection effect is referred to as the winner's curse: you win, you lose money, and you curse.[2]

Similar claims regarding a winner's curse have been made in a variety of other contexts: book publication rights (Dessauer, 1981), professional baseball's free agency market (Cassing and Douglas, 1980; Blecherman and Camerer, 1998), corporate takeover battles (Roll, 1986), and in real estate auctions (Ashenfelter and Genesore, 1992). These claims have traditionally been greeted with a good deal of skepticism by economists as they imply that bidders repeatedly err, violating basic notions of rationality which are unsustainable in the longer run. It is exceedingly difficult to support claims of a winner's curse with field data because of data reliability problems and plausible alternative explanations.[3]

The ambiguity inherent in interpreting field data, and the controversial nature of the winner's curse, provided the motivation for laboratory studies of the winner's curse. Bazerman

---

[2] Unfortunately, many economists, particularly theorists, characterize the winner's curse as the difference between the expected value of the item conditional on winning and the unconditional, naive expectation, using the term to refer to bidders fully accounting for this difference, rather than failing to do so and losing money as a consequence. This can make for some confusion.

[3] For example, Hendricks, Porter, and Boudreau (1987) found that in early OCS lease sales, average profits were negative in auctions with seven or more bidders. Hendricks et al. note that one possible explanation for this outcome is the increased severity of the adverse selection problem associated with more bidders. However, they note that the data could also be explained by bidder uncertainty regarding the number of firms competing on a given tract (their preferred explanation) so that bidders discount by the usual number of rivals, which is less than seven, earning negative profits as a result of the failure to accurately anticipate the number of rivals.

and Samuelson (1983) conducted the first laboratory experiment demonstrating a winner's curse. Using M.B.A. students at Boston University the experiment was conducted in class, with students participating in a series of first-price sealed-bid auctions in which bidders formed their own estimates of the value of each of four commodities - jars containing 800 pennies, 160 nickels, 200 large paper clips each worth four cents, and 400 small paper clips each worth two cents. Unknown to subjects, each jar had a value of $8.00. (Subjects bid on the value of the commodity, not the commodity itself.) In addition to their bids, subjects provided their best estimate of the value of the commodities and a 90% confidence bound around these estimates. A prize of $2.00 was given for the closest estimate to the true value in each auction. Auction group sizes varied between 4 and 26.

The average value estimate across all four commodities was $5.13 ($2.87 below the true value). In contrast, the average winning bid was $10.01 resulting in an average loss to the winner of $2.01, with losses occurring in over half of all the auctions.[4] Further analysis showed that winning bidders were substantially more aggressive than other bidders, so that average winning bids were sensitive to a handful of grossly inflated bids.

The results of this experiment show that the winner's curse is easy to observe. However, many economists would object to the fact that subjects had no prior experience with the problem and no feedback regarding the outcomes of their decisions between auctions, so that the results could be attributed to the mistakes of totally inexperienced bidders. The robustness of these results is even more suspect given their sensitivity to a handful of grossly inflated bids, which one might suppose would be eliminated as a result of bankruptcies or learning in response to losses incurred in earlier auctions. In fact, it was just these objections that motivated my initial common value auctions experiments along with my colleague Dan Levin (Kagel and Levin, 1986; Kagel et al., 1989).

Our initial experiments were designed to correct for these problems: First, we conducted a *series* of experiments with cash payouts, preceded by two or three dry runs intended to deal with the experience issue. Second, we provided bidders with unbiased estimates of the value of the (fictitious) commodity they were bidding on, so as to insure that the affiliated values

[4] Winning bidders paid these losses out of their own pockets, or from earnings in the other auctions (Max Bazerman, personal communication).

assumption underlying bidding in common value auctions would be satisfied, and that winning bids yielding negative profits could not be attributed to biased estimates of the (unconditional) expected value of the item (as distinct from a biased estimated conditional on the event of winning). Third, following each auction we provided bidders with the full set of bids along with the corresponding signal values and the true value of the item, as well as the winning bidders profits, designed to speed up the learning process and to help subjects recognize the adverse selection problem (i.e., that winning bidders tend to have the highest, or one of the highest, estimates of the value of the item).

Our working hypothesis was that after a few rounds, bidders would settle down to earning a substantial share of the expected risk neutral Nash equilibrium (RNNE) profits in these auctions. Boy, were we wrong. Auctions with inexperienced bidders showed a pervasive winner's curse (Kagel, et al., 1989): Profits in the first nine auctions averaged -$2.57 compared to the RNNE prediction of $1.90, with only 17% of all auctions having positive profits. Note, this is after bidders had participated in 2-3 dry runs, with full feedback, so that the results cannot be attributed to a total lack of experience. These negative profits are not a simple matter of bad luck either, or a handful of grossly inflated bids, as 59% of all bids and 82% of the high bids were above the expected value conditional on winning the item. Further, 40% of all subjects starting these auctions went bankrupt. In short, the winner's curse was alive and well even after our best efforts to establish an environment that would correct for the objections of most mainstream economists. These results, "the usual disaster", have since been replicated by others under similar conditions (Lind and Plott, 1991; Cox, Dinkin, and Swarthout, 2001).[5]

Not to be deterred, we brought back experienced subjects who had participated in this initial experiment and began to look at some of the comparative static predictions of the theory (Kagel and Levin, 1986). [6] This was motivated by two thoughts: First, maybe subjects just needed more experience. Second, perhaps subjects were just miscalibrated, so that the comparative static predictions of the theory would be satisfied in spite of the overbidding, a far more damaging mark against the underlying theory (in our minds at least) than simply overbidding. The results looked pretty good, at least for small numbers of bidders, as auctions

[5] At the same time there were a parallel series of laboratory experiments demonstrating a robust winner's curse in the corporate takeover game first reported in Samuelson and Bazerman (1985). Results of which are reviewed in Kagel (1995) and Kagel and Levin (2008); also see Charness and Levin (2009).

[6] The curious juxtaposition of the paper with experienced bidders being published *before* the paper with inexperienced bidders is explained by the vagaries of the publication process.

with 3 or 4 bidders earned substantial positive profits averaging $4.32 per auction (significantly greater than zero, but still well below the RNNE prediction of $7.48 per auction). However, for these same bidders, bidding in larger groups (auctions with 6-7 bidders) profits averaged -$0.54 per auction, compared to the RNNE prediction of $4.82. Note that predicted profits decreased by $2.66 per auction while actual profits decreased by almost twice as much ($4.86), along with a sizable frequency of the winner's curse as measured by the high bid exceeding the expected value conditional on winning.

At the same time we observed significant violations of two key comparative static predictions of the underlying theory. First, regressions showed that, other things equal, individual bids increased in response to increased numbers of rivals, a characteristic considered to be a telltale sign of a winner's curse in field data. Second, public information in the form of announcing the lowest private estimate of the value of the item, which in equilibrium, should in theory at least increase sellers' profits (Milgrom and Weber, 1982) did so in actions with 3-4 bidders. But this same public information had just the opposite effect in auctions with 6-7 bidders, reducing sellers' profits.

We reached several conclusions based on these results. First, the overbidding identified as a winner's curse, was not just a matter of subjects being miscalibrated, but extended to violating key comparative static predictions of the theory, including a key public policy prescription that public information, if available, should be used to raise revenue in the sale of government assets with a significant common value element. Second, the reemergence of a winner's curse in going from experienced bidders in auctions with 3-4 bidders to one with 6-7 bidders with its heightened adverse selection effect was the first indication that the underlying learning processes resulting in elimination of the worst effects of the winner's curse was context specific rather than involving some sort of "theory absorption" that readily generalized to new environments. There have since been other manifestations of this, most notably in almost common value auctions (Rose and Kagel, 2008).

Third, we showed a striking similarity between our experimental results and an important anomaly in the field data on common value auctions, thereby addressing the ever present question in laboratory experiments of external validity: Rates of return calculated for drainage leases in OCS sales showed that *both* neighbors and non-neighbors earned higher rates of return

on drainage compared to wildcat leases Mead et al. (1983, 1984).[7] A wildcat leases is one for which no drilling data are available, so that bidders have symmetric information based on seismic readings. A drainage lease is one in which hydrocarbons have been located on an adjacent tract so that there is asymmetric information, with companies who lease the adjacent tracts (neighbors) having superior information to other companies (non-neighbors). Theory predicts that in equilibrium, with an asymmetric information structure of this sort, neighbors will earn substantially more, on average, than on wildcat leases, and non-neighbors will earn substantially less (essentially zero profits according to the leading model of common value auctions with "insider" information at the time).[8] Obviously, the field data does not square with this theoretical result. However, it is consistent with the laboratory results once one recognizes that there is a considerable amount of public information about the oil in the ground associated with drainage tracts.[9] Similar to the laboratory results, this public information may have corrected for a winner's curse that depressed rates of return on wildcat tracts. Now clearly, this is not the only possible explanation for the field data: The leading alternative explanation is that the lower rate of return on wildcat leases reflects the option value of the proprietary information that will be realized on neighbor tracts if hydrocarbons are found. However, what the explanation based on public information increasing bidders' profits in the presence of a winner's curse has going for it is its consistency with the original claims of Capen, Clapp and Campbell (1971), the low rates of return reported on wildcat leases (Meade et al., 1983; Hendricks et al., 1987) along with the higher rates of return for non-neighbors on drainage leases than the wildcat leases, and parsimony.

b. Further questions related to external validity – field experiments of a sort.

Conventional laboratory experiments in economics typically use financially motivated students as subjects. One ongoing issue with this subject population of convenience is that perhaps this ignores important selection effects that jeopardize the external validity of the results reported. There are several dimensions to this argument: Students might be self-selected in some way so as to exclude individuals with characteristics that are important determinants of the behavior in the underlying subject population of interest. This could consist of any number of

[7] Hendricks and Porter (1992) obtained net rate of return estimates quite similar to Mead et al. on this score as well.
[8] Wilson (1967), Weverbergh (1979), Engelbrecht-Wiggans, Milgrom and Weber (1983), Hendricks, Porter and Wilson (1994).
[9] See Cooper (1998) for discussion of the extensive spying that goes on between rival companies once drilling starts on a tract and the difficulties involved in keeping drilling results out of the hands of competitors.

things including inexperience with the task at hand, general immaturity, and the failure to select out individuals with the relevant "smarts" to thrive in markets where poorly performing individuals are eliminated from the market as part of the competitive process.

One possible solution to this issue is to compare the behavior of professionals who have experience with the institution of interest to students. Dyer, Kagel and Levin (1989) report one such experiment comparing bids of seasoned executives in the commercial construction industry to students. To enhance the ecological validity of the experiment, we employed a first-price low bid wins auction, the same as the competitive bidding environment the executive operated in. Other than that, the experimental procedures were the same as in the demand side auctions already discussed.[10] Similar results were found between the construction executives and the students: Both suffered from a winner's curse earning negative average profits, with winning bidders consistently bidding below the expected value conditional on winning. Further, announcing the highest private information signal for the executives raised the average offer price by a statistically significant amount, consistent with the demand side results in the presence of a winner's curse (Kagel and Levin, 1986). Regression results showed that the executives bid slightly higher in auctions with larger numbers of competitors, qualitatively in the direction predicted by the RNNE, but the magnitude of the response was not large enough to avoid even larger average losses with more bidders. This response to changing numbers of bidders is qualitatively different from the students in both low and high price auctions, and may well represent a difference between the two subject pools as a result of the executives past experience. However, there is an alternative explanation based on the losses the executives experienced in the small numbers market (as opposed to the positive profits the students had in the small markets), so that with increased numbers of rivals simple survival pressures required bidding less aggressively for the executives to avoid bankruptcy and exit from the market.[11]

One can, of course, dismiss the executives suffering from a winner's curse on the grounds that they failed to take the experimental incentives seriously. But the data suggests otherwise as

[10] In terms of Harrison and List's (2004) proposed taxonomy of field experiments, this is an "artifactual field experiment", the same as a "conventional" laboratory experiment, employing abstract framing and an imposed set of rules but using a nonstandard subject pool, professionals who presumably have some experience with the relevant institution, thereby alleviating some of the potential selection effects associated with student subjects.

[11] Notice that we come back again to different interpretations of the same set of results. We have results which, hopefully, are reasonably non-controversial, and then the question of what they mean, where reasonable people (actually, in this case, the same people) can reach different conclusions regarding their broader implications.

they were carefully attentive throughout and estimates of bid functions were qualitatively similar to the students as bids were monotonically increasing in signal values.

Further investigation (Dyer and Kagel, 1996) begins to resolve the apparent discrepancy between the executives suffering a winner's curse in the lab with their apparent success in field settings. Two possibilities, which are not necessarily mutually exclusive, were identified. One is that the executives had learned a set of situation-specific rules of thumb which enabled them to avoid the winner's curse in the field, but which could not be applied in the laboratory. For example, an important determinant of the risk associated with bidding a job, which impacts the cost estimates assigned to job components and the markup assigned to the contract, involves the architect/owner's reputation for admitting mistakes, or ambiguities, in construction plans, and their willingness to implement functionally equivalent construction outcomes using alternative, and cheaper, construction techniques than originally specified.[12] Needless to say, the contractors did not have any of these situation-specific rules to rely on when bidding within the relatively abstract context of the laboratory experiment. This factor is fully consistent with the context specific nature of learning that psychologists have identified and that we have seen in our auction experiments.

The second factor at play is that the bidding environment created in the experiment, which is based on theoretical work, is not fully representative of the environment encountered in the construction industry. These involve two factors: (i) repeated play elements which at times permit low bidders to withdraw their bids *without penalty* following a "reasonable" mistake in calculating the bid estimate and (ii) differences in bidder estimates of construction costs are a much smaller factor in the construction industry compared to OCS auctions, with overhead and the amount of idle resources anticipated often playing the decisive role in determining the low bid, so that the executives were not prepared for the large bid factor needed to avoid the winner's curse in the lab. As for this second factor, looking at the results of other experiments comparing students with professionals (for example, Burns, 1985; Garret, Walker and Wooders, 2004) I am pretty well convinced that the major insight gained from such studies is the ability to identify unanticipated differences between field settings and how we model these situations.

[12] This is important enough to the point that in the experiment, at least one of the executives jokingly inquired, "Who is the architect associated with this job?"

More recently Harrison and List (2004, 2008) present results which, on some dimensions at least, appear to show how practicing professionals are able to avoid the winners curse in an "artifactual field experiment" as they have *"... developed a heuristic that "travels" from problem domain to problem domain."*[13] In what follows I offer an alternative explanation for this result that is more in line with my argument regarding context specific learning. Harrison and List also report results from a second experiment in which the same professionals dealing with a commodity they are familiar with, sports cards, provides "… *tentative support for the hypothesis that naturally occurring markets are efficient because certain traders use heuristics to avoid the inferential error that underlies the winner's curse.*" With regards to this second experiment I think there is some misunderstanding of the nature of the winner's curse. More importantly, I think there is some fundamental misunderstanding of how we view the results of our experiments identifying a winner's curse as they relate to behavior in "mature" markets.

In one experiment, Harrison and List (2008) compare bids of sports card dealers with non-dealers using the same abstract, induced valuation procedures reported on above. They employ both a symmetric information procedure in which all bidders have the same level of information about the common value (a single draw from a uniform distribution whose mid point is the common value) as well as an asymmetric information treatment where one bidder, the insider, knows the true value with certainty while all other bidders draw an information signal from a symmetric distribution around the true value. In all cases subjects bid in a single auction after having participated in a minimum of 10 practice auctions with either 4 or 7 bidders.[14] Dealers rarely suffer from a winner's curse in the symmetric information treatment as they typically bid below the expected value conditional on winning. In contrast, non-dealers suffered from the typical winner's curse, biding above the expected value conditional on winning and earning negative average profits. There are statistically significant differences in the estimated bid functions between dealers and non-dealers, mostly having to do with bidder sensitivity to the degree of uncertainty inherent in bidders' private information signals. In the asymmetric treatment they are unable to reject a null hypothesis of any difference in the bid functions for dealers and non-dealers in their role as "outsiders." The frequency of the winner's curse

[13] See Harrison and List (2004, p. 1027) for both quotes (italics in the original).

[14] Subjects were not provided with any starting capital balances or participation fees to cover potential losses. However, a second experiment that was not announced until after the first was completed was used to insure that everyone went home with positive profits.

increases for dealers and non-dealers (it's still higher for non-dealers), to the point that there is a 25-30% rate of the winner's curse for dealers.

Harrison and List argue that the superior performance of dealers with symmetric information reflects field experience in comparable settings, and that, unlike the construction contractors, "… context-specific experience does appear to carry over to comparable settings, at least with these types of auctions." They attribute the substantial increase in the frequency of the winner's curse for dealers in the asymmetric information treatment to the fact that "*when dealers are placed in an unfamiliar role they perform relatively poorly.*" (Harrison and List, 2008, p. 835, italics in the original) So at best, their results provide an argument for very limited learning generalizability: dealers having adapted to adverse selection effects in field settings with symmetric information do not recognize the heightened adverse selection effect when an insider is present, succumbing to the winner's curse.

While this result is consistent with my argument for context specific learning, I believe there is an alternative explanation for the symmetric information results. Dealers in buying trading cards must purchase them at low enough prices to be able to sell them at a profit and would most certainly be in the habit of doing so; e.g., List and Lucking-Reiley (2000) show that dealers bid just under $50 for cards with a retail value of $70 in a Vickrey auction. So applying such large discount factors, which non-dealers in their typical role as buyers would not be in the habit of doing, could very well protect them from a winner's curse.[15] There is independent support for this interpretation of their results from an experiment by Garrat et al. (2004) who report similar results for eBay buyers and sellers in a second-price sealed bid auction with induced values. On this interpretation, the "*heuristic that "travels" from problem domain to problem domain"* is that dealers have learned to buy low and sell high which adventitiously, in this case, protects them from the winner's curse.[16]

In their second experiment, the professional card dealers are bidding on a commodity they are familiar with, an *unopened* package of L*eaf* sports cards (packages containing 10 cards of

[15] In that experiment List and Lucking-Reiley assume that the trading card market is best approximated by a private value auction.

[16] One should, of course, ask why this same heuristic does not help the construction contractors? The answer is simple. General contractors do not buy and sell in anything approaching the same way that card dealers do. Rather they solicit bids from large numbers of subcontractors who are responsible for fulfilling their commitments, and then add in their own estimated general contactor's costs.

unknown value, and an *established* retail price of between $9-$10). In comparing bids of dealers with non-dealers there is not a single bid above $10 for dealers and only a handful of bids above $10 for non-dealers. These dealer bids underlie the claim cited above "… *that naturally occurring markets are efficient because certain traders use heuristics to avoid the inferential error that underlies the winner's curse.*" While we agree the L*eaf* sports card auctions constitute a common value auction, it is one in which there is no scope for a winner's curse. The cards have a well known market value, which precludes any adverse selection effect based on different estimates of their value, particularly on the part of professional card dealers. As such, for the dealers at least, this experiment is comparable to auctioning off a $9 or $10 bill.

But there is a broader claim here regarding the efficiency of bidding in established markets and the absence of a winner's curse that requires some clarification. Finding a winner's curse with inexperienced bidders in the laboratory, in conjunction with field reports of a winner's curse (e.g., Capen, Clapp and Campbell, 1971), along with the parallels between field data and laboratory data reported on above, leave us reasonably well convinced that a winner's curse was present in early OCS auctions and is likely to exist *at least in the start up phase* of auction markets with a strong common value element. Further, if our interviews with the construction contractors are at all representative, it's very likely to be present for new entrants into those markets, or even experienced agents when entering a new segment of the market (Dyer and Kagel, 1996). But as the laboratory experiments have also shown, people do learn and there are market selection effects so that one would *not* expect long established players in any given market to suffer from a winner's curse on a *regular* basis.

II. Gift Exchange in Experimental Labor Markets

Laboratory auction experiments are generally considered to be a reasonably good model for their field counterparts in that underlying structure of the lab and target environments are similar; e. g., a first-price sealed-bid auction in the lab has the same rules as its field counterpart. That is, the empirical interpretation of the concepts and models underlying auctions are quite similar between laboratory and field settings. To be sure there are still important differences. For example, laboratory investigations of the winner's curse typically employ a single-unit pure common value auction environment with a known number of bidders, whereas field settings often involve a mix of common and private value elements, with an unknown number of actual bidders, frequently demanding multiple units. But to the extent that there is a strong common

value element in the field setting with a number of competing bidders there is a strong adverse selection effect that bidders' must be contend with in both cases. Thus, the insights about behavior gained in the laboratory can, and have, provided insights into field settings: One of the arguments in favor of open, ascending price auctions for the Federal Communication Commission's sale of spectrum (air-wave) rights was that "… allowing bidders to respond to each other's bids diminishes the winner's curse: that is, the tendency for naive bidders to bid up the price beyond the licenses' actual value, or for shrewd bidders to bid cautiously to avoid over paying." McAfee and McMillan (1996, p. 161). Further, to the extent that laboratory experiments are used as a testbed for designing auctions to be implemented in field settings (e.g., Plott, 1997; Ledyard, et al., 1997) there are special reasons to think that the experiment provides a good working model for the target environment, and hence for the external validity of the results reported.

a. *An initial set of results comparing laboratory and field experiments:* There is considerably greater controversy regarding the correspondence between the empirical concepts underlying laboratory studies of gift exchange in labor markets relative to the target environment of labor relationships in firms (see Bardsley et al., 2009, for example). Akerlof (1982) was the first to introduce the idea of gift exchange in labor markets; firms paying above market wages because of the negative effects of low wages on worker morale, resulting in less shirking as a reciprocal response on the part of workers. Among other things, the model has been used to explain the reluctance of firms to cut nominal wages as unemployment increases in an economic downturn. Of course, there are many competing models designed to explain this same phenomena (see Bewely, 2004, for a review in relationship to empirical findings from field studies). The idea of gift exchange as a reciprocal response of employees to higher wages is one that has been extensively studied, originally in a series of laboratory experiments pioneered by Fehr and his colleagues (e.g., Fehr, Kirchsteiger, and Reidl, 1993). These results have recently been challenged on the basis of several field experiments (e.g., Gneezy and List, 2006).

The typical gift exchange game is a two stage game. In stage 1 employers' make costly wage offers to potential employees. In stage 2 employees decide to accept or reject the proposed offer and then provide a costly "effort level" to employers, with more effort being more costly. The higher the effort level provided the greater the employer's profits. In most early lab experiments matching of firms and workers was anonymous, taking place over a finite number of

trials announced in advance, all taking place well within two hours, so that there is no opportunity for workers to develop individual reputations, or for other repeated play game effects. "Effort" is defined in terms of "employees" choosing higher numbers which reduce their experimental earnings while increasing the earnings of their "employer."

Since there are a finite number of plays of the game with a known end point, standard unraveling arguments predict wage offers at the bottom end of the range permitted in all periods, accompanied by minimum effort levels. Yet there is typically a clear positive relationship between wages and effort levels resulting in a Pareto improving outcome with higher earnings for both firms and workers than under the competitive equilibrium outcome. These results do not appear to depend on the fine points of the market institutions (e. g., posted offer vs. double oral auction or one-to-one matching) and are robust to introducing a real effort task into environment (e.g., Gneezy, 2004, who used solving mazes as his real effort task).

The role of positive and negative reciprocity between workers and firms in this setting is reasonably non-controversial. For example, Fehr and Falk (1999) look at gift exchange in very demanding double auction markets. In one treatment, with opportunities for gift exchange which could, potentially, improve outcomes for both "employers" and "employees," employers offer "wages" above the competitive equilibrium level, and spurn efforts of employees to undercut these wages, as they anticipate lower effort levels. However, with exogenously determined effort levels, employers accept only the lowest wage offers, with wages forced down to close to the competitive equilibrium level. [17] This suggests that employers' behavior is driven mainly by the expectation of positive reciprocity on the part of workers to higher wages, as opposed a simple desire to raise workers earnings out of "fairness" considerations. To take another example, Charness (2004) compares effort levels in a typical two stage gift exchange experiment with one-to-one matching of employers with employees. In one treatment "wages" are determined exogenously, either randomly or by the experimenter, with the distribution of these exogenous wage offers designed to match the distribution observed with employer determined wages. There are higher effort levels with exogenously determined as opposed to employer determined wage offers at lower wage rates, with no difference at higher wage rates. He interprets this as evidence for negative reciprocity on the part of workers as they respond with

[17] There is zero worker surplus/profits at the competitive equilibrium so that wages are a bit higher than the competitive equilibrium itself.

the minimum possible effort to low employer determined wages. He interprets the absence of higher effort levels in response to higher employer determined wages as evidence against positive reciprocity. However, this is not the only possible interpretation, with the Fehr and Falk (1999) experiment providing evidence for positive reciprocity at higher wage rates.[18]

To be sure there are variations in effort levels reported in different experiments. For example, Hannan et al. (2002) report marked differences in reciprocity between undergraduate students and MBAs under the same experimental conditions, with MBAs providing significantly higher effort levels at comparable wage rates, as well as providing higher average wages. They conjecture that the difference between MBAs and undergraduate results from MBAs greater experience in jobs where gift exchange plays an important role, thereby sensitizing the MBAs to the reciprocity considerations inherent in higher wage offers. In contrast, most undergraduate work in the US is associated with minimum wage jobs where there is no, or minimal gift exchange. There is some support for this conjecture in a subsequent study with the same subject population of MBAs as subjects with above-median professional work experience provided higher mean effort compared to subjects with below-median professional work experience (Hannan, 2005). Hannan et al. also report no significant difference in effort levels of workers in response to comparable wage offers from high versus low productivity firms (where high productivity firms find it less costly to provide higher wages than do low productivity firms).[19] They conjecture that this lack of responsiveness, which was true for both MBAs and undergraduates, is due to the fact that the relationship between firm profits and productivity is an indirect one, hence not as likely to be as salient to subjects as differences in wage offers. In a subsequent experiment, Charness et al. (2004) found that the degree of gift exchange is surprisingly sensitive to how payoffs are framed: they find substantially less effort provided (among undergraduates) when a comprehensive payoff table relating wages and effort levels to workers' payoffs and employers' earnings was provided as part of the instructions as opposed to the usual format of providing payoffs functions and a set of examples for subjects to work

[18] Also see Owens and Kagel (2009) for evidence of positive reciprocity at higher wage rates. The issue of whether positive or negative reciprocity plays a stronger role in subject behavior is still under debate (see, for example, Offerman, 2002 and Cox, Sadiraj, and Sadiraj (2008) for contrasting points of view). Hannan (2005) provides evidence that employees reduce effort levels more in response to a reduction in wage rates followng a negative profit shock than the increase in effort in response to an increase in wage rates following a positive profit shock.

[19] Wage offers were tagged with the firm's productivity level in a posted offer labor market.

through.[20] However, similar payoff tables had no apparent deleterious effect on MBAs effort levels compared to results typically reported (Hannan et al., 2002).

In a field experiment Gneezy and List (2006) question the staying power of positive reciprocity in response to higher than anticipated wage rates. They looked at two tasks – computerizing library holdings over a 6 hour period and a door-to-door fundraising effort over a single weekend day. The gift exchange treatment was operationalized by advertising a given wage rate, with higher than advertised wages paid to one of two treatment groups; e.g., an advertised wage of $12 per hour for the library task, with half the subjects given the "surprise" wage of $20 an hour upon showing up. They break their data in half reporting significantly higher effort levels for the high wage group in the first half of the day, but no significant differences in the last half of the day. They interpret these results in terms of the psychology literature on reference point effects, arguing that after a while workers reference point shifts so that the new higher wage serves as the fair wage reference point, with a resulting drop in effort. Clearly, this is not the only possible interpretation of their results: Perhaps the higher wage workers became fatigued from working harder and/or the higher wage workers provided the gift level they thought appropriate to the higher than advertised wage in the first half of the day and slacked off after that.[21]

Subsequent field experiments report positive gift exchange between firms and workers that in a number of cases is *increasing* over time. Kube et al. (2006) look at gift exchange in a library cataloguing task, focusing on both positive and negative reciprocity. Students were hired for a six hour shift to catalogue books, with the recruitment e-mail announcing a *presumptive* salary of 15 Euros per hour. Upon arrival one-third of the subjects were told the wage would be 20 Euros per hour (the "Kind" treatment), with another third told that the wage was actually 10 Euros per hour (the "Unkind" treatment), with the last group getting the 15 Euro per hour wage (the "Neutral" treatment). The number of books catalogued increased over time for all three treatments. The Unkind treatment starts out at a much lower rate of cataloguing than the Neutral

[20] Note there was a positive relationship between effort levels and wages as typically reported in laboratory gift exchange experiments; just that average wages and effort levels were significantly lower with the payoff table than without it.

[21] A number of writers have cited Hennig-Schmidt et al. (2008) as support for the Gneezy-List results. What Hennig-Schmidt et al. report is an absence of positive reciprocity to higher than anticipated wages in a lab experiment with real work effort and a corresponding field experiment. An additional real-effort lab treatment supports their hypothesis that the absence of explicit cost and surplus information that enables employees to calculate employer's surplus eliminates the negative results reported.

treatment and remains below it throughout, with the differences statistically significant in each 90 minute interval. The Kind treatment starts at the same rate as the Neutral treatment but catalogues at higher rates for each 15 minute interval after that. However, with the exception of the middle interval these differences are not statistically significant at conventional levels. Kube et al. conclude that negative reciprocity is a stronger force than positive reciprocity, which may well be the case. But the small sample sizes involved (9 and 10 subjects in each treatment), in conjunction with the time pattern observed in the data between the Kind and Unkind treatments, would seem to be a shaky basis to reject the existence of positive reciprocity in field experiments of this sort. Indeed Al-Ubaybli et al. (2006) provide evidence in favor of positive reciprocity in a field experiment with temporary workers in which effort levels are growing over time in both the baseline treatment and the positive gift exchange treatment, with the differences in favor of the gift exchange treatment increasing over time, to the point that they appear to be statistically significant at the end of their two day trial.[22]

The discussion so far illustrates both the strengths and weaknesses of laboratory and field experiments on this topic. Subjects in field experiments do not know they are in an experiment, the tasks they are being asked to perform parallel those they would normally perform, and, in some cases, the experiment takes place in the context of a longer term relationship. As such, these experiments should have a closer relationship to the field setting that authors like Akerlof had in mind. However, the cost of this verisimilitude is high. There is considerable loss of control in these experiments, as we neither know the cost of effort, the perceived benefits of effort to the employer, nor the game that the employees think they are playing. Measurement is a problem in many of these studies as workers in field settings can respond to incentives along multiple dimensions, so that the experimenter may miss important elements of employees' responses to higher wage rates. Also one must account for the level of baseline wages relative to market wages for comparable work, as higher than normal baseline wages may already elicit a strong gift response. As Cohn et al. (2007) note, this may impose a ceiling effect resulting in a downward bias in the response to the gift wage treatment, and baseline wages in relationship to market wages are not always reported.

[22] Unfortunately, the statistical specification omits any interaction term for the time trend variable with the treatment effects which would reveal whether or not effort in the gift exchange treatment grew significantly faster over time than the controls, nor is the data split into early versus late responses as in Gneezy and List (2006).

b. *So what do any of these experiments have to tell us about the gift exchange and sticky downward wages in the workplace?*

None of the experiments reported on so far directly address the question of sticky downward wages to clear unemployment in a recession. Further, the very structure of these experiments, both laboratory and field studies, cannot approach the target environment of ongoing labor relations within firms, no less one in which there is a downturn in economic activity for the economy. In addition, at least one prominent student of survey research on why firms typically do not cut wages during recessions (Bewley, 2004), seriously doubts the empirical relevance of the morale theory of wage rigidity proposed by Akerlof (1982) and others (Solow, 1979; Akerlof and Yellen, 1988, 1990). Bewely does so on the grounds that "…. employers say that they do not see much connection between effort or morale and wage levels; effort and morale do not increase with pay levels…" (Bewely, 2004, p. 7). He goes on to note that what is accurate in the theory is that employers avoid cutting wages because to do so would hurt morale. In contrast, higher wages per se have little effect because employees usually have little notion of a fair market value for their services, but quickly come to believe that they are entitled to their existing pay, no matter how high it may be.

Nevertheless Bewley sees considerable relevance for the gift exchange experiments described here to the issue at hand – sticky downward wages. He does so on the grounds that in lab experiments:

> "The most important finding is the prevalence of reciprocity. ….The general willingness to reciprocate good to good is the essence of good morale. Negative reciprocity is what underlies the insult effect of pay cuts, which is resentment caused by the firm's perceived breach of positive reciprocity; workers expect employers to offer pay increases, not cuts, in exchange for loyalty and effort." (Bewley, 2004, p. 16).

Thus, although there is a certain unreality associated with the structure of both laboratory and field experiments dealing with gift exchange in labor markets in relationship to the field settings they are designed to provide insight into, what the experiments isolate and validate is the prevalence of reciprocity. That is, what the experiments have done is to isolate and study the core principle underlying sticky downward wages. It is on the basis of this, rather than a fully faithful empirical model of the target environment (an impossibility under any circumstance), that makes for the relevance of the experiments for the target environment.

Summary and Conclusions

I have discussed the use of laboratory experiments and their relationship to field data, field experiments, and economic theory. I have done this in the context of two examples I am reasonably familiar with – common value auctions and the winner's curse and gift exchange in experimental labor markets. I have tried to make three overarching substantive points.

First, learning, which is endemic to most experimental studies, tends to be context specific and difficult to generalize from one environment to another. This is totally consistent with the psychological literature on learning which distinguishes between near transfer (e.g., someone who knows how to drive a car can typically handle driving a light truck) and far transfer (e.g., knowing how to drive a car does not transfer immediately to driving a semi-trailer no less an airplane or speedboat). This has a number of implications. Among other things, one should not automatically expect economic agents who are fine tuned to the field settings they typically operate in to be able to adjust to experimental environments that deprive them of their typical contextual cues or that vary even in small, but important, ways from their natural habitat. And although agents in an experiment have converged on an equilibrium in a given economic setting, this does not necessarily imply that they will respond in the way that theory suggests, without a new round of learning, when the environment changes.

Second, although hopefully we can all agree on the "facts" of a particular economic experiment, there is typically wide room for disagreement on the interpretation of those facts as they apply to the broader issues at hand. I've supplied two examples on this point, one for each of the two cases I've considered. (I could provide many more as well in any number of areas of active experimental research past and present.) I don't expect everyone to agree with my alternative interpretations, no less the investigators that reached the original conclusions. In fact it's these differences in how one interprets the facts that leads to new and interesting experiments that further narrow what we do and don't know on a particular topic.

Third, I have tried to show that even in cases where the laboratory setting seems rather removed and abstract relative to field setting one has in mind, that the experimental results may be quite relevant to that field setting. The trick here is that the experiment, as abstract and as seemingly far removed from the target environment in question, needs to capture the essential behavioral issue(s) at hand – do real agents behave in the ways our theories predict and for the

reasons the theory postulates? Of course, coming back to my second point, there will no doubt be some who will disagree with my interpretation for the particular case discussed.

References

Akerlof, Goerge A. 1982. Labor contracts as partial gift exchange. *Quarterly Journal of Economics,* 97, 543-569.

Akerlof, George A. and Janet Yellen. 1988. Fairness and unemployment. *American Economic Association, Papers and Proceedings*, 78, 44-49.

Akerlof, George, A. and Janet Yellen. 1990. The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics,*105, 255-283.

Al-Ubaydli, O., S. Andersen, U. Gneezy, and John A. List. 2006. For love or money? Testing non-pecuniary and pecuniary incentive schemes in a field experiment. Working Paper, University of Chicago.

Ashenfelter, O. and D. Genesore. 1992. Testing for price anomalies in real estate auctions. *American Economic Review: Papers and Proceedings* 82:501-505.

Bardsley, Nicholas, Robin Cubitt, Garham Loomes, Peter Moffatt, Chris Starmer, and Robert Sugden. (2009) Experimental economics: Rethinking the rules. Princeton University Press (in press).

Bazerman, M. H. and W. F. Samuelson. 1983. I won the auction but don't want the prize. *Journal of Conflict Resolution* 27:618-634.

Bewley, Truman. 2004. Fairness, reciprocity, and wage rigidity. *IZA* Discussion Paper # 1137,

Blecherman, B. and C. F. Camerer. 1998. Is there a winner's curse in the market for Baseball Players? Mimeograph. Brooklyn Polytechnic University**.**

Burns, P. 1985. Experience and decision making: a comparison of students and businessmen in a simulated progressive auction," in Vernon L. Smith (ed.) *Research in Experimental Economics, Vol 3*, Greenwich: JAI Press, 139-53.

Capen, E. C., R. V. Clapp, and W. M. Campbell. 1971. Competitive bidding in high-risk situations. *Journal of Petroleum Technology* 23: 641-653.

Cassing, J. and R.W. Douglas. 1980. Implications of the auction mechanism in baseball's free agent draft. *Southern Economic Journal* 47:110-121.

Charness, G. 2004. Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics*, 22, 665-88.

Charness, G., G. Frechette, and J. H. Kagel. 2004. How robust is laboratory gift exchange. *Experimental Economics*, 7: 189-205.

Charness, G. and D. Levin. 2009. The origin of the winner's curse: A laboratory study. *American Economic Journal: Microeconomics.*

Cohn, A., Ernst Fehr, and L. Goette. 2007. Gift exchange and effort: Evidence from a field experiment. Working Paper. Institute for Empirical Research in Economics, University of Zurich.

Cooper, Christopher. 1998. Oil firms still rely on corporate spies to be well-informed. *Wall Street Journal*, Dec 7:1&23.

Cox, James C, Sam Dinkin, and James T. Swarthout. 2001. Endogenous entry and exit in common value auctions. *Experimental Economics*, 4, ….

Cox, James C., Klarita Sadiraj, and Vjolica Sadiraj. 2008. Implications of trust, fear and reciprocity for modeling economic behavior. *Experimental Economics*, 11, 1-24.

Dessauer, J. P. 1981. *Book publishing*. New York: Bowker.

Dyer, D., J. H. Kagel, and D. Levin. 1989. A comparison of naive and experienced bidders in common value offer auctions: A laboratory analysis. *Economic Journal* 99:108-115.

Dyer, D. and J. H. Kagel. 1996. Bidding in common value auctions: How the commercial construction industry corrects for the winner's curse. *Management Science* 42:1463-1475.

Engelbrecht-Wiggans, R., P.R. Milgrom, and R.J. Weber. 1983. Competitive bidding and proprietary information. *Journal of Mathematical Economics*, 11:161-69.

Fehr, Ernst and Armin Falk. 1999. Wage rigidity in a competitive incomplete contract market. *Journal of Political Economy*, 107, 106-134.

Fehr, Ernst, Georg Kirchsteiger, and Arno Reidl. 1993. Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics*, 108, 437-60.

Garratt, R., M. Walker, and J. Wooders. 2004. Behavior in second-price auctions by highly experienced eBay buyers and sellers. Working paper, No. 1181., Department of Economics, University of California Santa Barbara.

Gick, M.L. and K.J. Holyoak. 1980. "Analogical Problem Solving." Cognitive Psychology, 12: 306-355.

Gneezy, Uri. 2004. Do high wages lead to high profits? An experimental study of reciprocity using real effort. University of Chicago, Graduate School of Business, Working Paper.

Gneezy, Uri and J. List. 2006. Putting behavioral economics to work: Field evidence of gift exchange. *Econometrica*, 74, 1365-84.

Hannan, R. Lynn. 2005. The combined effect of wages and firm profit on employee effort. *The Accounting Review*, 80, 167-188.

Hannan, R. Lynn, John Kagel, and Donald Moser. 2002. Partial gift exchange in experimental labor markets: Impact of subject population differences, productivity differences, and effort requests on behavior. *Journal of Labor Economics*, 20, 923-51.

Harrison, G. W. and J. A. List. 2004. Field Experiments. *Journal of Economic Literature*. 42: 1013-59.

Harrison, G. W. and J. A. List. 2008. Naturally occurring markets and exogenous laboratory experiments: A case study of the winner's curse," *Economic Journal*. 118: 822-843.

Hendricks, K. and R.H. Porter. 1992. Bidding Behavior in OCS Drainage Auctions: Theory and Evidence, paper presented at the 1992 European Economics Association meetings.

Hendricks, K. and R.H. Porter and B. Boudreau. 1987. Information, returns, and bidding behavior in OCS auctions: 1954-1969. *The Journal of Industrial Economics*, 35:517-542.

Hendricks, K., R. H. Porter, and C. A. Wilson. 1994. Auctions for oil and gas leases with an informed bidder and a random reservation price. *Econometrica*. 62:1415-1444.

Hennig-Schmidt, Heike, Bettina Rockenbach, and Abdolkarim Sadrieh. 2008. In search of workers real effort reciprocity – a field and laboratory experiment. *Journal of the European Economic Association*, in press.

Kagel, J. H. 1995. Auction: Survey of experimental research. In Alvin E. Roth and John H. Kagel, Editors, Princeton University Press.

Kagel, J. H. and D. Levin. 1986. The winner's curse and public information in common value auctions. *American Economic Review* 76:894-920.

_____ and _____. 2008. Auctions: A survey of experimental research, 1995 – 2008. Working paper, Ohio State University.

_____, _____, R. Battalio, and D.J. Meyer.1989. First-price common value auctions: bidder behavior and the winner's curse. *Economic Inquiry*, 27:241-58.

Kube, S., M.A. Maréchal, C. Puppe. 2006. Putting reciprocity to work – positive versus negative responses in the field. Working paper.

Ledyard, John O., David P. Porter, and Antonio Rangel. 1997. Experiments testing multi-object allocation mechanisms. *Journal of Economics and Management Strategy* 6:639-75.

Lind, B. and C. R. Plott. 1991. The winner's curse: Experiments with buyers and with sellers. *American Economic Review* 81:335-346.

List, J. A. and D. Lucking-Reiley. 2000. Demand reduction in multi-unit auctions: Evidence from a sportscard field experiment. *American Economic Review* 90:961-972.

McAfee, R.P., and J. McMillan. 1996. Analyzing the airwaves auction." *Journal of Economic Perspectives,* 10: 159-176.

Mead, W.J., A. Moseidjord, and P.E. Sorensen. 1983. The rate of return earned by leases Under cash bonus bidding in OCS oil and gas leases. *Energy Journal*, 4:37-52.

_____, _____, and _____. 1984. Competitive bidding under asymmetrical information: behavior and performance in Gulf of Mexico drainage lease sales, 1954-1969. *Review of Economics and Statistics*, 66:505-08.

Milgrom P. R. and R. J. Weber. 1982. A theory of auctions and competitive bidding. *Econometrica* 50:1089-1112.

Offerman, Theo. 2002. Hurting hurts more than helping helps. *European Economic Review*, 46, 1423-37.

Owens, Mark F. and John H. Kagel, 2009, Minimum Wage Restrictions and Employee Effort in Labor Markets with Gift Exchange Present. Working paper, Ohio State University.

Perkins, D.N. and G. Salomon. 1988. Teaching for transfer. *Educational Leadership*, 46: 22-32.

Plott, C. 1997. Laboratory experimental test beds: application to the PCS auction. *Journal of Economics and Management Strategy,* 6: 605-638.

Roll, R. 1986. The hubris hypothesis of corporate takeovers. *Journal of Business*, 59:197-216.

Rose, S. L. and J. H. Kagel. 2008. Bidding in almost common value auctions: An experiment. *Journal of Economic Management Strategy……*

Salomon, G. and D.N. Perkins. 1989. Rocky roads to transfer: rethinking mechanisms of a neglected phenomenon. *Education Psychologist*, 24: 113-142.

Samuelson, W.F. and M.H. Bazerman. 1985. The winner's curse in bilateral negotiations, in V.L. Smith (ed.), *Research in Experimental Economics,* 3. Greenwich, CT.: JAI Press.

Solow, Robert M. 1979. Another possible source of wage stickiness. *Journal of Macroeconomics,* 1, 79-82.

Weverbergh, M. 1979. Competitive bidding with asymmetric information reanalyzed," *Management Science*, 25, 291-94.

Wilson, B. R. 1967. Insider competitive bidding with asymmetric information. *Management Science* 13:816-820.