Experimental Methods and Elicitation of Values

by

Glenn W. Harrison, Ronald M. Harstad and E. Elisabet Rutström[†]

September 2003

ABSTRACT

Experimental methods are currently being extensively used to elicit subjective values for commodities and projects. Three methodological problems are not systematically addressed in this emerging literature. The first is the potential for laboratory responses to be censored by field opportunities, so that lab responses can be confounded by uncontrolled knowledge of the field; the second is the potential for subjective perceptions about field opportunities, and hence valuation responses, to be affected by the institution used to elicit values; and the third is the potential for some elicitation institutions to influence subjective perceptions of characteristics of the commodity or project being valued, and hence change the very commodity being valued. All three problems result in potential loss of control over the value elicitation process. For example, we show that censoring affects conclusions drawn in a major study of beef packaging valuation. We derive implications for experimental designs that minimize the potential effect of these methodological problems.

JEL Classification: B4, C9, D8
Keywords: experiments, valuation, auctions, incentive-compatibility

[†] Department of Economics, College of Business Administration, University of Central Florida (Harrison and Rutström), and Rutgers Business School, Rutgers University (Harstad). E-mail: GHARRISON@BUS.UCF.EDU, ERUTSTROM@BUS.UCF.EDU and HARSTAD@RUTGERS.EDU. Rutström thanks the U.S. National Science Foundation for research support under grants NSF/IIS 9817518, NSF/MRI 9871019 and NSF/POWRE 9973669. We are grateful for comments from two referees.

Correspondence: Professor Glenn W. Harrison, Department of Economics, College of Business Administration, University of Central Florida, Orlando, FL 32816, USA. E-mail: GHARRISON@BUS.UCF.EDU.

The literature in experimental economics has been dominated by the use of one important design feature: experimenter-induced values. The explosion of experimental applications in recent decades testifies to the power that controlling via induced values creates. Book-length surveys by Davis and Holt [1993], Friedman and Sunder [1994] and Kagel and Roth [1995] predominately report induced-values experiments.

More recently, however, experimental methods have been used to *elicit the homegrown values* of individuals or groups for commodities or projects that exist outside the laboratory. The objective of eliciting values is quite different from the objectives of applications of the induced-values technique. The adjective "homegrown" simply means "not induced," and refers to values that are neither controlled nor known *a priori* by the experimenter. Though more widely applicable, the elicitation of homegrown values is particularly central to the fields of marketing,[1] environmental damage assessment,[2] and the general estimation of individual preferences.[3]

We argue that three fundamental methodological issues must be addressed when designing experiments intended to elicit homegrown values. These issues, almost without exception, have been neglected in the literature purporting to elicit values truthfully and accurately.[4] Moreover, they relate directly to the value-elicitation objective, so that they naturally have not arisen in the literature on induced-values experiments.[5] These issues arise within the mainstream microeconomic context in which individuals behave as constrained expected-utility maximizers, and hence are assumed to hold well-behaved reservation values for traded commodities. Even under ideal circumstances, we

---

[1] For example, see Menkhaus, Borden, Whipple, Hoffman and Field [1992] and Hoffman, Menkhaus, Chakravarti, Field and Whipple [1993].

[2] See Blackburn, Harrison, and Rutström [1994], Cummings and Harrison [1994], Neill et al. [1994] and Cummings, Harrison and Rutström [1995]. In this literature the elicitation of homegrown values for a lab commodity is used either as a proxy for the field environmental good or as a way to calibrate field survey responses.

[3] For example, see Coller and Williams [1999], Harrison, Lau and Williams [2002] and Holt and Laury [2002].

[4] The experimental literature eliciting homegrown values has grown in recent years. Much of the early literature is reviewed in Cummings and Harrison [1994].

[5] The concern with whether a subject in an induced-values experiment is operating with thoroughly induced values, as when subjects in ultimatum bargaining experiments might be supposed to value fairness or its appearance, or to assume that other subjects do so, is the closest that literature comes to the concerns developed here.

argue, serious qualifications remain as to the ability of received experimental procedures to accurately elicit values from individuals. No doubt violations of the conventional assumptions would compound the difficulties discussed here.

First, the problem of *field-price censoring* arises from the availability of immediate substitutes to the chosen laboratory commodity. A rational subject will not agree to obtain the same commodity in an experiment at a price that he perceives can be beaten outside the lab with sufficiently high probability. This implies that elicited values will be censored at the perceived extra-laboratory price of the good.

The second concern is related to field-price censoring and applies particularly to situations where subjects are uncertain of the commodity's extra-laboratory price. The concern arises when *beliefs about field prices are affiliated*; that is, when it is rational for one subject's beliefs to be positively responsive to the beliefs of subjects whose information differs from his. In such cases respondents may revise their valuations after observing the stated values of other respondents. This possibility implies that informational aspects of the institutions used matter in ways that are irrelevant in induced value settings. For example, an institution with repetition of choices or sequential revelation of values could give a very different answer than an institution with one-time, simultaneous, or privately communicated, value statements. Although use of the latter type of institution should avoid this problem, variants of the first type of institution are widely employed in homegrown value settings.

The third concern addresses the possibility that subjects might also have *affiliated beliefs about the quality of the laboratory commodity* itself. Subjects who are uncertain about the characteristics of the commodity might rationally infer information about those characteristics from observing other respondents' stated values. This has the same implication as the problem with affiliated beliefs about field prices: institutions with repetition of choices or sequential revelation of values could alter values in the process of eliciting them in ways that never occur for induced values.

The next three sections describe these problems in turn. Subsequently, section 4 shows the potential impact of field-price censoring, re-analyzing observations in a seminal elicitation study

(Hoffman, Menkhaus, Chakravarti, Field and Whipple [1993]). Section 5, finally, draws implications for experimental designs to mitigate the effects of the problems we identify. We do not claim that these problems will always affect the conclusions of studies eliciting values in a laboratory, or that the effects will always be quantitatively dramatic. But we do claim that it makes no sense to knowingly design an experiment that could fall prey to these issues when alternative designs exist.

1. Field-Price Censoring

We maintain the assumption that a subject has a (homegrown) value for a commodity,[6] and denote this value V. Of course, V is known to the subject and must be discerned by the experimenter. The subject is asked the classic, simple *dichotomous choice* question, "Would you be willing to buy one unit of the commodity at a price X?" The hope is that the subject will respond "yes" if $V > X$ and "no" if $V < X$.

Initially, consider the case of a commodity that is transacted on an idealized frictionless market outside the laboratory. That is, with no transactions costs, subjects can buy or sell this commodity at a going market price P. An example approximating frictionless conditions might be if the subjects were Pennsylvania residents working in Trenton, NJ, and the commodity was a bridge-toll token.[7] In this case, it is irrational for the subject to base his response on any comparison between his value V and the laboratory elicitation price X. His sole rational response is "no" if $X > P$ and "yes" if $X < P$. If for some reason $V > X > P$ (on this frictionless market, he would have to be away from his consumption optimum), the subject nonetheless should refuse to buy the commodity in the laboratory on worse terms of trade than he could acquire the identical commodity in the field. At $X = P$ the subject faces only the trivial choice of exchanging one coin for another.

---

[6] We are not concerned here with the process by which a subject determines his preferences. We do, however, consider the possibility that subjects are uncertain about the quality or other characteristics of a commodity. A subject may be able to respond consistently to a few dichotomous choice questions even if he cannot be said to have a firm, noncapricious value for a commodity. Laboratory observations of consistent responses do nothing to prove that the subject in fact possesses a well-defined value.

[7] The example assumes, as we once observed, that bridge tokens can be obtained as change or used to pay for purchases at stores and restaurants in the vicinity, and virtually amount to an additional denomination of coinage.

Figure 1 illustrates our point: the commodity is shown horizontally, with a composite commodity on the vertical axis, and AB the field budget constraint. The subject comes into the experiment at endowment point E. If the subject has not consumed the commodity in the field the endowment point E would be at point A. This possibility is quite likely in practice, particularly if one defines the commodity narrowly. The circumstance in which $X < P$ is illustrated by offering laboratory purchase opportunities along the dashed line EF. He rationally must accept the arbitrage opportunity presented: buy from the experimenter at X in lieu of some planned later purchase at P, or buy from the experimenter for resale at P. Since he should accept the offer independent of his private valuation, the presence of the field price acts as a censoring device on the experimenter's ability to observe his private valuation. When $X > P$, the experimenter is simply asking the subject if he wishes to move inside his budget set, from E to G, by offering a price $X > P$. In this case, a rational subject should refuse the offer even when V exceeds X. Again, the field price acts to censor the private values. Hence, in this ideal frictionless market a dichotomous choice question *elicits no information whatsoever* about a rational subject's value.

Next consider introducing realistic market frictions. Before doing so, expand the classic dichotomous choice question to allow quantity responses in a range from 0 to M units. Transactions costs are still assumed away, and the subject can obtain the commodity outside the laboratory at price P. Now, however, any unit of the commodity, once obtained in the laboratory or outside, becomes "used" and hence can only be resold in the field for a price $R < P$.

As before, eliciting at a laboratory price $X > P$ is simply a movement inside the budget constraint, and is thus censored. The subject's feasible consumption opportunities, when responses are elicited using a laboratory price $X < P$, are illustrated in Figure 2. The axes and the field budget constraint (AB) are as in Figure 1. The solid line EF illustrates eliciting at a laboratory price X for which $R < X < P$: the subject can move outside the budget constraint from E to F by purchasing M units of the commodity. Since $R < X$, reselling the commodity in the field would only move back along the dotted line FH (whose slope is R), inside EF. The subject optimally responds by choosing the point along EF where the highest indifference curve is reached. Thus, for values of X

that are less than *P*, but greater than *R*, the experimenter can elicit uncensored responses.

The situation when the laboratory price *X* is chosen such that is $X < R$ is illustrated by the dashed line EG. By purchasing all *M* units allowed in the laboratory, and then possibly reselling some, the subject can move along the dotted line GJ (slope *R*). Buying only some of the *M* units from the experimenter would place the subject between E and G, strictly worse off than along GJ. Hence, in this case, we again have censoring since subjects should always buy all *M* units, independent of their private valuation. To conclude on the case where the field resale value is less than the field purchase price, then, a rational subject's response reflects *his private valuation accurately* only when laboratory prices are set *below* the field purchase price yet *above* the field resale price.[8]

The case typically used in laboratory studies is slightly more complex, but still exhibits field-price censoring. Allowing for the possibility of transactions costs in the field would imply a wider range of lab prices for which responses would be uncensored, but the logic is otherwise the same.

Censored observations will only take the values 0 and *M*. However, either of these could also be an uncensored response: at some price between the upper and the lower censoring points, demand could be less than one unit, or more than *M*.[9] Hence only an intermediate elicited response, $0 < Q_e < M$, is transparently identified as uncensored. In this sense, then, unless the possibility of field-price censoring can be ruled out *a fortiori*, the classical dichotomous choice mechanism is an inefficient laboratory elicitation design. If 0 and 1 units purchased are the only feasible options, neither is *per se* a necessarily uncensored observation.[10] Nevertheless, even if all observations were

---

[8] Coller and Williams [1999] pioneer explicit recognition of this concern in their elicitation of individual discount rates. They recognize that censoring from above can occur by subjects' perceptions of field borrowing rates, and from below by perceived field lending rates. They query subjects for these perceptions, and incorporate them in censored-observations statistical analysis. Cherry et al. [2001] undertake direct tests of the effects of field price censoring in an induced-value setting, and report evidence confirming the hypothesis that it affects bidding behavior in the qualitative manner expected.

[9] For some commodities, such as cappuccino machines, resale and gift-giving are likely the only explanations for a subject buying more than one.

[10] This concern also applies to open-ended mechanisms which in effect only elicit binary information at the price the subject states. For example, a subject who states a value by choosing how high to bid in an English auction is simply indicating a unit demand at that price and a zero demand at higher prices. The zero demand at higher prices could result from field-price censoring, so the stated exit price could be censored. In contrast, if the experimental design used the Ausubel [2002] auction, an ascending-bid design which incorporates the payment rule of the Vickrey auction, then subjects would state a demand curve by decreasing units demanded at various points as prices rose. If allowed to demand at least three units, the prices at which

known to be in the uncensored range, censoring would still create truncated error terms, so standard statistical techniques for dealing with censored observations would still be appropriate.

To see the importance of designs allowing for elicitations that are transparently uncensored, consider the limitations of an experimenter's information about the parameters determining the censoring bounds in the presence of transactions costs. Even if the transaction cost is generally nonnegative, it may be zero if a subject already planned other purchases from a retailer selling this commodity, or it may be very high if a subject perceived his time to be quite valuable, or his mobility limited, or his best extra-laboratory purchase opportunity to entail visiting an unsafe location after dark. By its very nature, the relevant transactions cost is clearly a subjective calculation, likely not under experimenter's control, nor readily observed.

The earlier discussion is also misleading for most commodities in its casual application of the "Law of One Price." Nearby retailers may post differing prices, and the experimenter may not know which price the subject considers the relevant basis for comparison. Moreover, subjects are unlikely to have much experience with the resale market, and thus are likely to have quite diffuse subjective expectations of both resale prices and transactions costs. The experimenter cannot know this boundary more precisely than the subject perceives it.

Methods for dealing with censored observations are, of course, well established in the econometrician's toolkit. Those methods, however, take the censoring boundaries as part of the input. Therefore, we are not sanguine about the prospects for simply passing off field-price censoring as a concern in data analysis alone, rather than as a broader experimental design issue.

The problem of field-censored laboratory responses is methodological in nature: it cannot simply be addressed by avoiding trivial or familiar commodities. A subject who has diffuse beliefs about the field price of an unfamiliar commodity nonetheless ought not accept a real economic commitment and obtain the commodity in the laboratory if he perceives that the price can be beaten in the field with high enough probability. The existence of imperfect substitutes makes the problem of establishing the censoring bounds even more difficult. A subject's response could depend on his

---

the subject decreased demand from three to two, and from two to one, would be uncensored.

perception of which field substitute corresponds to the laboratory commodity, what the price of the field substitute is, and how the relative quality of the field substitute compares to the laboratory commodity. Without careful design the experimenter cannot hope to observe the price at which censoring arises.

Field censoring does not cause lab responses to be incorrect, or untruthful. They are correct in the sense that they are the observed responses of the subject in the context in which the subject is placed. The issue is rather that the experimenter may have incompletely controlled the context, which in this case is the degree of substitution between the laboratory and field commodities. Given that context, the respondent has behaved rationally. Censored responses provide observations answering a different question than do uncensored responses.

2. Affiliated Beliefs About Field Substitutes

The problem of field-price censoring could arise in an experiment simply designed to elicit a homegrown value from a *single* subject. The next two issues are related, but arise when a subject might be influenced in his decision about what value to state through his observations of other subjects' behavior. First we consider the possibility that observing others might influence the price at which a subject censors stated values, due to beliefs about field prices. Institutions that were designed for use with induced values do not necessarily control for the information flow between subjects, since that is not an issue with induced values (unless the subjects mis-perceive the task). When eliciting homegrown values, however, the design of the institution must build in additional controls, or else the bids will be affected by these information flows and will not accurately reflect the underlying subjective valuations.

Suppose an experimenter decides to elicit values by having subjects compete in an open ascending-bid ("English") auction to obtain the same commodity. In particular, suppose, as in Rutström [1998], that prices are called out ascendingly, and subjects cross from one side of the room to the other when the price reaches their willingness-to-pay. To work with a concrete example, let us assume that the commodity is a jar of some brand of spaghetti sauce acquired by the

experimenter in some store in a large metropolitan area, and that all subjects have just tasted a sample of pasta covered with this sauce.

Consider a subject with a true willingness-to-pay of V and diffuse (prior) beliefs about the market price P, beliefs which allow for plausible prices both well below and well above V. Suppose this subject has good reason to believe that other subjects may have better information about the market price than he has. For example, his family ordinarily makes spaghetti sauce from scratch, or some other family member does the grocery shopping.

Consider the inferences this subject should draw if he notices several other subjects, apparently interested in the product, cross the room at prices well below V. He ought to allow for two possibilities: that other subjects' exit prices reflect much lower willingness-to-pay, and/or that their exit prices are field-censored as the result of beliefs which are divergent from (and quite possibly better-informed than) his beliefs about the price of this spaghetti sauce and perhaps prices of close substitutes. To the extent that the latter possibility may be true, he ought rationally to update his beliefs about P. Extrapolating from the standard terminology of auction theory, we refer to this rational updating phenomenon as *affiliated beliefs*. With enough low-price exits by others, or sufficient allowance for better-informed censoring, this subject may rationally exit well below V.[11] Notice that his value has not changed, but that it is rational to censor his stated value, and the price at which he ought to censor may be influenced by the procedure used to elicit values.

As stressed by Rutström [1998], valuations that are elicited by lab experiments may be influenced by the institution that is used to elicit them. This point is to some extent self-evident, since one of the hallmarks of experimental methods has been the demonstration that "institutions matter." What gives it more force is that the choice of institution can matter *even when choosing from institutions that are all behaviorally demand-revealing in induced-value settings*.

---

[11] The affiliated-beliefs effects also potentially contaminates institutions that are *not* behaviorally demand-revealing. For example, consider an institution like the Vickrey auction, which is demand-revealing in theory, but yields bids exceeding values in the laboratory (Kagel, Harstad and Levin [1987], Harstad [2000]). To the extent that overbidding relative to values exhibits some statistical pattern across subjects, one would still expect the Vickrey auction to elicit a different bid from a subject who got to observe the bids of some of his rivals before choosing his own bid.

Rationality requires that subjects update their beliefs about prices of field substitutes whenever useful inferences can be drawn from information revealed to the subject in the elicitation process. But subjects may update their beliefs even when the information revealed is not useful in an objective sense, since it is sufficient that the subject believes it to be useful. For example, a subject bidding in repeated Vickrey auctions for irradiated sandwiches, as in Hayes et al. [1995], may adjust his bid in the direction of the bids he learns that others made in earlier repetitions, in the *belief* that he knows so little about the value of irradiating a sandwich that the others must be better informed. In fact, they may be just as ignorant on the matter as he is, but nonetheless the feedback from prior rounds could influence all their bids. This type of bidding behavior may then erroneously be interpreted as a convergence of the valuations over time, when the true underlying process is really a convergence in the beliefs about field prices.

The general experimental design question becomes one of recognizing that perceptions of the likely price of field substitutes may vary across subjects, as may the confidence with which perceptions are held. To some extent these perceptions may be proxied by information about the subject, such as experience in purchasing this commodity in recent times. Experimental institutions vary in their propensity to transmit information about these perceptions between participants. Thus, the problem presented in this section further limits the information an experimenter is likely to obtain about the particular price level at which a subject begins to censor his stated value. As before, the problem is not confined to commodities with perfect field substitutes: subjects observing others' valuations should rationally draw inferences about others' beliefs about prices of imperfect substitutes as well.

We conclude that experimental control requires attention to the role that the information revelation properties of institutions can play in affecting the elicited values. We do not claim that values influenced by elicitation institutions in the way mentioned above are "wrong." Instead, we claim that observers who ignore this interaction might draw unwarranted conclusions about the true valuations of subjects.

3. Affiliated Beliefs About Characteristics

Subjects may not only draw inferences about prices of field substitutes from observing other subjects, but they may also draw inferences about the quality or other characteristics of the commodity itself. These two kinds of affiliated belief effects would usually occur together, although experiments could be designed that distinguish between them. To simplify, in this section we suppose that there is common knowledge that all subjects have identical, possibly imprecise information about prices of field substitutes.

Again, consider a concrete example. Suppose valuations are being elicited for dancing lessons at a long-established local dancing school. Subjects in the experiment are aware that anyone who has had dancing lessons at this school have been eliminated from the subject pool, and that anyone with absolutely no interest in dancing lessons has no incentive to show up for the experiment. Subjects are given a brochure describing the school's six-session dancing class, and asked to bid for an admission ticket for that class using an English auction as the one described above. Obviously, a subject's value for this instruction is a function of the perceived frequency with which past attendees of these classes have gained some dancing competence.

Consider a subject who realizes that none of his friends have ever mentioned anything positive or negative about the school's classes. Based upon his dim perception of the quality of the instruction, he formulates a value V for the class. If it were a single-shot Vickrey auction, he could do no better than to exit from the bidding at price V. However, in an English auction, he may have more information which he rationally must take into account. Since this dancing school has had enough customers to stay in business, there may well be subjects in the room who do have second-hand knowledge about the commodity's quality, through their different networks of friends that might include graduates of the school.

Suppose this subject observes several *potentially* well-informed rival bidders cross the room well before the price in the English auction reaches V. Since we have assumed that everyone has the same price information, he ought rationally to allow for two possibilities: that these other subjects value dancing lessons of a given quality less than he does, which ought not influence his

bid; and/or that they may be exiting on the basis of second-hand information suggesting that the dancing lessons being offered are of low quality. He should rationally attach some positive probability to the latter possibility. This could lead him to exit the auction well below V. In contrast, if the price has risen near V without much early exit by rival bidders, he ought to allow for the *possibility* that several other bidders may not value dancing lessons of known quality as highly as he, and must therefore have continued competing because they have positive information about quality. This is not the only reason they might still be competing, but allowing for it could rationally lead him to compete past V.

Consequently, as in the previous section, institutions that *perform identically* in induced-values settings can nonetheless *elicit different values* from the same subject due to affiliated beliefs about commodity characteristics.[12]

4. An Example

A. The Field Experiment

Hoffman et al. [1993] conducted a seminal field experiment in which they elicited valuations for beef steaks in two alternative retail packages.[13] In many respects their experiment provides an ideal testing ground for the issues we raise. They were eliciting homegrown values for commodities that had relatively close substitutes in the field, and they did indeed inform subjects of the approximate field prices for these substitutes. Their auctions were conducted in a series, such that all subjects in later periods had been able to observe earlier prices, and possibly draw some inferences from those prices about likely market prices for the goods being valued. Finally, the two products differed in terms of several characteristics, which were explained in detail to subjects. In fact, one of their treatments was to provide extra information on the less familiar product and a demonstration

[12] This is quite separate from any claims that the subject is "researching preferences" or changing preferences because of the value elicitation process. Such claims rest on a misunderstanding of the essence of (von Neumann-Morganstern) expected-utility theory, which is to allow subjects to revise their subjective probabilities of the states of nature while maintaining their underlying preferences over final outcomes.

[13] Menkhaus et al. [1992] and Schmitz et al. [1993] provide complementary analyses of the same field experiment. All data and statistical code used in our analysis is available for replication at HTTP://WWW.BUS.UCF.EDU/GHARRISON/DATA/EE/BEEF/.

of one of it's characteristics.

The two products in their experiment were beef steaks: in one case packaged with the customary Styrofoam tray wrapped in see-through plastic, and in the other case wrapped in a "vacuum skin." The latter product was regarded as the new, less familiar product, and the objective was to determine how consumers would value the new product relative to the old product. The new, vacuum packed product had a longer shelf life, but did not have the red color normally associated with fresh beef until it had been opened for several minutes.

Three information treatments were employed. In the "no information" case all subjects were told what the products were, and virtually nothing else. In the "information" treatment all subjects were told some of the virtues of the vacuum skin package. Finally, in the "demonstration" treatment all subjects were given information about the virtues of the vacuum skin package *and* were allowed to see how it returns to a regular beef color after the package is opened and allowed to sit for some minutes. The sample of 765 subjects was assigned to each of these treatments roughly equally.

All valuations were undertaken using a uniform price sealed bid auction in which 8 bidders competed to purchase 4 steaks. It was common knowledge that there were 8 bidders at all times. Each bidder could purchase only one steak per auction period, so no complications from multiple-unit bidding by the same subject would arise. The selling price was determined as the fifth-highest bid, with all winners paying that price rather than their bid. So this auction was a multiple-unit analogue of the single-unit Vickrey auction, which provides an incentive for truthful revelation. Indeed, subjects were instructed at some length that it was in their best interest to bid exactly what they thought the steak was worth. The instructions went further than simply asking subjects to bid truthfully: they explained, as well as one can to ordinary consumers, the logic behind this being the best strategy.

Each subject participated in 4 trial auctions, in which bids were hypothetical and no products were actually purchased. Then there were 6 auctions in which the bids were real and the purchase consequences binding. Each subject was informed of the market price at the end of each period, but not informed of what the bids of the other participants were. So we know exactly what information

was provided to subjects as the experiment proceeded.

In one treatment the old product was auctioned first, after which the new product was auctioned; this order was reversed in the other treatment. Thus the subject in the first treatment would have had 2 trial periods valuing the old product, followed by 2 trial periods valuing the new product, 3 purchase periods valuing the old product, and then 3 purchase periods valuing the new product.

All subjects were given a questionnaire asking a range of standard socio-demographic questions, including gender, age, education level, family income, employment status, marital status, and household size. They were also asked a series of questions about their concerns with beef steaks, such as costliness, fat content, and taste. Additionally, they were asked if their meat usage habits had changed in the past 3 years, and how important they considered the packaging of beef to be when making a purchase decision. All of these questions had been asked before the sample was recruited for the valuation experiment, by telephone. Only subjects who identified themselves as the "primary meat purchaser" for their household and who had actually purchased beef steaks in the past month were recruited for the experiments. The sample design called for there to be an equivalent sample by income and gender in each group of 8 bidders; with respect to gender, each session had 6 women and 2 men.

The experiments were conducted in 1989 in Denver and Los Angeles, with roughly equal samples in each location. The experimenters also recorded which "monitor" conducted the experiment, so that one can control for experimenter effects in valuations.

B. Analysis

Statistical analysis of these data require that one account for the panel structure of the data. Each group of 8 bidders participated in a time series of 10 periods, the first 4 being trial periods. There are many hypotheses one could evaluate with this rich design, but we focus on evaluating the theoretical concerns we raised.

Censoring

Censoring is an issue in this experiment because subjects were asked to value a product which had a clear counterpart in regular supermarkets. Indeed, the experimental instructions clearly explained that the beef steak within the package was just the same as one would find in a supermarket:

> These steaks have been obtained from a major meat wholesaler and are similar to the very best steaks you could purchase in your local grocery store. In fact, they are USDA choice steaks and are as fine as any you would find in your area. *They currently retail for about $6.00 a pound at major supermarkets.* The meat is identical except for the packaging. One is a vacuum skin package; the other is an overwrapped styrofoam tray. Steaks in each package type were cut from the same piece of meat and were packaged within the last few days. (p.335; our emphasis)

We assume that subjects took this information from the experimenter at face value, although concerns that this may not be the case apply in varying degrees to any experiment. For our purposes it is noteworthy that subjects were given a clear and uniform signal as to the market price of the commodity: $6 per pound. Since the old product being valued was the same as the meat packaged in the supermarkets referred to here, this $6 price can be taken as the price of perfect substitutes for that *old* product. It may not be the price of a perfect substitute for the *new* product, since subjects may not view the vacuum skin as comparable to the product they could buy in supermarkets for $6. In either case, bids should be affected by the knowledge that there is a $6 alternative available.

The effect of field censoring can be captured in an econometric model by recognizing that no individual would rationally bid higher than $6 if the product being valued were considered a perfect substitute. Some individuals did bid greater than $6, although very few and only in the earlier periods. Inspection of histograms of bids in each period for the two products suggests an eyeball-apparent trend for the mode to move down as time passes. The distribution of bids for the new product also appears to be slightly more skewed to the right than the distribution of bids for the old product. In (trial) period 1 some 3.5% of all bids were above $6, and this fraction was slightly higher for the new product compared to the old product (4.4% vs. 2.6%). This overall fraction drops to just 1.7% in (trial) period 2, and is less than 1% for each of the binding purchase periods.

Nevertheless, it is not necessary that observed bids be "spiked" at $6 or near it for there to be an effect from censoring. The key issue statistically is whether the *error* term for an individual is "allowed" to include values above $6. Thus somebody that had a bid in one period of $5.75 may have an effect from censoring if the standard error on that bid was large enough. If that standard error was $1, for example, then a statistical analysis that ignored censoring would implicitly attach a significant probability to the subject bidding in excess of $6. The appropriate estimation technique for this issue is a generalized Tobit regression, in which an upper bound is set for each individual.[14] For logical consistency we also apply a lower bound at $0.

Affiliated Beliefs

This concern arises because the subjects might try to learn from others about the appropriate field price for these goods, or infer from market valuations the value to them of the characteristics of the good. This concern is particularly likely to be an issue for the new product, since it is the one which is less familiar to consumers. The implication is that subjects' bids in later periods could be affected by the level of earlier selling prices.

Is it possible in this design to tease apart the two ways in which affiliated beliefs could affect valuations? One pathway is via their influence on beliefs about field prices, and the other pathway is via their influence on beliefs about the quality of the good. Since subjects were directly told that the field price of the ordinary beef was $6, in this design it would be unlikely that the first pathway would be significant. Thus we conclude that any effects from affiliation in this setting are almost certainly due to the second pathway.

[14] This upper bound is $6 unless the individual bid greater than $6, in which case it is set equal to the observed bid. Allowance for bids above $6 is based on the possible presence of transactions costs for obtaining the field substitute that may differ across individuals.

Statistical Model

A regression model is used to identify these possible effects in these data. The model takes into account the panel structure of the data, and includes variables to control for several factors that might otherwise influence bids: age; gender; family income; marital status; employment status; education; household size; and the stated importance of packaging, the "familiar bright red color," and whether the beef "looks appetizing and attractive" in the decision to purchase steaks. Dummy variables to capture treatments include effects for the order in which the packages were valued, the location, the period, the experimenter, whether the auction was for the vacuum skin package,[15] whether information was provided,[16] whether the new product was demonstrated,[17] and pairwise interactions between the last three factors.[18]

In comparison to the findings we report below, Hoffman et al. [1993] report that the average winning bids are higher for the vacuum packed product than for the styrofoam packed product. They also find that providing information increases bids for both products, but that the demonstration had no significant effect. We test for the effect of explicitly allowing for the possibility of censoring by comparing a Panel Generalized Least Squares model (GLS) with a Panel Tobit model. Additionally, we test for the effect of explicitly allowing for the possibility of affiliated beliefs by including variables capturing lagged values of the selling prices. We include all bids, not just the winning ones, in our analyses.

Table 1 reports the results of estimating this statistical model, focusing on the effects of the main experimental treatments on bids and dropping the estimation results for all other controls. Panel A reports the basic regression that ignores the possible effects of market prices and censoring. This panel is therefore the one that is most comparable to the analysis undertaken in Hoffman et al. [1993]. Panel B includes lagged effects from market prices, as well as average trial period prices.[19]

---

[15] Variable `Vacuum` captured this treatment, taking a value 1 if the package was the vacuum skin variant.

[16] Variable `Info` captured this treatment.

[17] Variable `Demo` captured this treatment.

[18] Variables `VacInfo` and `VacDemo` captured these interactions.

[19] Although lagged market prices are not a lagged dependent variable, since the dependant variable is the individual's bid and not the market price, it is possible that they are endogenous with respect to bids. We

Panel C augments the specification in Panel B with allowance for censoring. We regard Panel C as the appropriate estimator for this experimental design.

From Panel C we infer that the vacuum skin packaging is *not* valued more than the styrofoam packaging *unless information is provided and/or the product is demonstrated.* The effect of providing information is to increase valuations by roughly 42 cents per pound, and this is a statistically significant effect. Adding a product demonstration increases valuations of the new product by roughly 7 cents per pound, and this is not as statistically significant an effect as providing information. Simply providing the product itself, with no information or demonstration, would not generate greater valuations. Ignoring the statistical effects of censoring and market prices, Panel A would lead to *overestimates* of the direct effect of information on the valuation of both products.[20]

The results in Panels B and C are very similar, implying that the influence from affiliation on the treatment effects is independent of censoring. Table 2 provides an alternative, and direct, test of the effects of affiliation.[21] Using estimation results from the appropriate statistical specification of Panel C of Table 1, the coefficient of lagged market prices is positive and statistically significant. Thus we infer that higher trial period prices, or prior auction prices, were associated with higher valuations and bids, and that the effect was systematic.[22] In summary, allowing for the possibility that affiliated beliefs and field price censoring contaminate the bids significantly influences the conclusions.

---

allowed for this possibility with a panel instrumental variables estimator, and obtained virtually identical results.

[20] The coefficient on `Info` in Panel A is 0.18, implying that valuations for the styrofoam packed meat (the reference product) increases by 18 cents per pound on average with the provision of information. This result is statistically significant at the 7% level, using a two-tailed test. The coefficient on `VacInfo` is 0.45, implying an *additional* 45 cents increase for the vacuum packed meat over the basic 18 cents.

[21] Variables `tbid_vs` and `tbid_st` are the average trial bid prices for the two commodities, and `priceL1` is one-period lagged price.

[22] It might be argued that the effects of previous market prices on valuations could just be due to individual "learning" about the optimal bidding strategy. We agree. However, without a formal model of a specific learning process, it is impossible to reject such explanations for the observed data (indeed, *any* data). If we posit that the learning behavior is individual-specific, and takes the empirical reduced form of an auto-regressive process of order 1 that varies for each individual, we can allow for learning and still test if market prices have some effect on valuations. They do, and the effect is still statistically significant. Constraining the auto-regressive process to be the same for all individuals results in the same conclusion. However, our tests are conditional on these particular specifications of a process that we readily concede could take many forms.

5. Implications for Experimental Practice

The concerns we raise have implications for best practice in experimental design if value elicitation is the objective, where "best" is defined with respect to the ability to claim control. Certain uncontrolled for factors *might* not be a problem, but in the absence of controls for their presence or mitigation we simply do not know and must rely on assertion. To paraphrase Smith [1982; p.938], it is precisely this type of frustration with the need to rely on assertion which prompted so many researchers to turn to using experimental methods to elicit homegrown values.

First, to elicit from subjects those valuations for a commodity which they had upon arriving at the experiment, one should use one-shot institutions rather than repeated institutions with the same good.[23] In our example, based on Hoffman et al. [1993], we found significant effects of the lagged selling prices, consistent with the presence of affiliated beliefs. Specifically, if a Vickrey auction is to be employed it is better to use a one-shot Vickrey auction than to have many repetitions of a Vickrey auction with the same good. We concede that there is some advantage in repetitions of a Vickrey auction in terms of training subjects in the logistical procedures of the experiment (the evidence from induced-value experiments is that they do not learn the dominant strategy properties of the Vickrey auction from experience and feedback; cf. Harstad [2000]). However, it is possible to train subjects up in another good altogether such as the tasty choice of (Hershey's) Kisses by Coller and Williams [1999]. Similarly, it is possible to simply inform subjects of the dominant strategy property of the institution they are using, such as in Neill et al. [1994] and Rutström [1998]. Such information is appropriate when the objective of the experimental design is to elicit homegrown values *assuming* that subjects know the dominant strategy property, rather than experiments whose objective is to *test* whether subjects understand the dominant strategy property.[24]

---

[23] It is conceivable that a research question could call for eliciting the values of subjects after they have made adjustments for the values stated by others, and/or adjustments for observed transaction prices. Such a context might call for a repeated Vickrey auction design. However, these considerations should not justify the use of repeated Vickrey auctions on the vague grounds that they elicit "better informed" valuations.

[24] Hey [1991] provides a valuable methodological contrast between foundational experimental tests of theories of individual decision-making (his part II) and non-foundational tests of those theories in conjunction with further structural assumptions (his part III). Although the latter are conditional on the former, there is no reason to forego the latter if one is willing to risk the possibility that the foundations need revision.

Can one just do repeated Vickrey auctions and report the results of the first period for those skeptics, like us, who dismiss the later periods as potentially contaminated? Unfortunately not. The problem is that the subjects typically are told that the experiment will last for several periods, and we do not then know how they decide to bid in the first round. It is perfectly plausible that a subject might understand the logic of a dominant strategy for a one-shot auction but not be able to see that this logic applies equally to each of the stage games of a repeated (experimental) game.[25]

The second implication for best practice experimental design is the importance of having simultaneous bid submission rather than having real-time bid submission or real-time sequential bid submission such as one would find in an open-bid English auction, again to avoid the confounding effects that arise due to affiliated beliefs. The implication here is that one might best use sealed-bid institutions rather than standard forms of sequential-bid, or real-time, bid institutions.[26]

A third implication for experimental design is to try to build in some controls for field substitutes. It is important to distinguish between the subjective transactions costs that individuals might have in going from the laboratory to the field and the subjective beliefs that individuals might have with respect to the price of field substitutes. Each of these are important dimensions of the problem of controlling for field substitutes, and each is to some extent amenable to experimental control. Attempts to do that by way of eliciting information on subject characteristics, and appropriate statistical methods that allow for censoring, are provided by Coller and Williams [1999], Harrison, Lau and Williams [2002], and our re-analysis of the Hoffman et al. [1993] data.

---

[25] Indeed, if the stage game Nash Equilibria are not unique, one runs into further possible problems if the subjects do not believe that the horizon for the repeated game is finite and certain. Experimenters that run the risk of computer failure or fire alarms always have subjects who rationally entertain some expectation that the experiment will end unexpectedly. Hence one could have a repeated (experimental) game with a finite but unknown horizon.

[26] Non-standard, "silent clock" institutions could be easily devised, in which subjects could revise their bids in real-time and yet not be informed of the prior bids of other agents. One mitigating factor in favor of real-time English auctions is that subjects appear to understand their dominant strategy better in that setting than in one-shot sealed-bid auctions, as demonstrated by Rutström [1998] and Harstad [2000]. In an asymmetric auction situation which was not incentive-compatible, Kirchkamp and Moldovanu [2001] found substantial differences between dynamic auction formats in which bidders' exits were publicly communicated to the auctioneer and formats in which their exits were privately (silently) communicated to the auctioneer. This setting gave bidders strong reasons to respond to the observed behavior of rival subjects. Further study of the behavioral properties of silent clock auctions in incentive-compatible settings seems particularly worthwhile to us.

6. Conclusion

The basic attraction of experimental methods is the control that they afford the student of human behavior. The issues we raise are "problems" only for particular experiments attempting to elicit *homegrown values* if allowance for censoring or affiliation is neglected: the experiment *or* inferences drawn from it may be contaminated. In other words, they are only problems when the experimenter overstates the degree to which the elicitation procedure was controlled. It is an open question whether valuations based on affiliated beliefs about the commodity or field prices are more "reliable" from a positive perspective, or more "relevant" from a normative perspective, than those valuations which are not so affiliated. We have opinions on these matters, but our central argument is that they ought to be explicit aspects of experimental design.

References

Ausubel, Lawrence M., "An Efficient Ascending-Bid Auction for Multiple Objects," *Working Paper No. 97-06*, University of Maryland, August 2002; forthcoming, *American Economic Review*.

Blackburn, McKinley; Harrison, Glenn W., and Rutström, E. Elisabet, "Statistical Bias Functions and Informative Hypothetical Surveys", *American Journal of Agricultural Economics*, 76, December 1994, pp. 1084-1088.

Cherry, Todd L.; Frykblom, Peter; List, John A.; Shogren, Jason F., and Williams, Melonie B., "Laboratory Testbeds and Nonmarket Valuation: The Case of Bidding Behavior in a Second-price Auction with an Outside Option," *Unpublished Manuscript*, Department of Economics, Appalachian State University, January 2001.

Coller, Maribeth, and Williams, Melonie B., "Eliciting Individual Discount Rates," *Experimental Economics*, 2, 1999, 107-127.

Cummings, Ronald G., and Harrison, Glenn W., "Was the Ohio Court Well Informed in Their Assessment of the Accuracy of the Contingent Valuation Method?", *Natural Resources Journal*, 34(1), Winter 1994, 1-36.

Cummings, Ronald G.; Harrison, Glenn W., and Rutström, E. Elisabet, "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?", *American Economic Review*, 85(1), March 1995, 260-266.

Davis, Douglas D., and D, Charles A., *Experimental Economics* (Princeton, NJ: Princeton University Press, 1993).

Friedman, Daniel and Sunder, Shyam, *Experimental Methods: A Primer for Economists* (New York:

Cambridge University Press, 1994).

Harrison, Glenn W.; Lau, Morten Igel, and Williams, Melonie B., "Estimating Individual Discount Rates for Denmark: A Field Experiment," *American Economic Review*, 92(5), December 2002, 1606-1617.

Harstad, Ronald M., "Dominant Strategy Adoption and Bidders' Experience with Pricing Rules," *Experimental Economics*, 3(3), December 2000, 261-280.

Hayes, Dermot J.; Shogren, Jason; Shin, Seung Y., and Kliebenstein, James B., "Valuing Food Safety in Experimental Auction Markets," *American Journal of Agricultural Economics*, 77, February 1995, 40-53.

Hey, John D., *Experiments in Economics* (Cambridge, UK: Blackwell, 1991).

Hoffman, Elizabeth; Menkhaus, Dale J.; Chakravarti, Dipinkar; Field, Ray A., and Whipple, Glen D., "Using Laboratory Experimental Auctions in Marketing Research: A Case Study of New Packaging for Fresh Beef," *Marketing Science*, 12(3), Summer 1993, 318-338.

Holt, Charles A., and Laury, Susan K., "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), December 2002, 1644-1655.

Kagel, John H.; Harstad, Ronald M., and Levin, Dan, "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study," *Econometrica*, 55, November 1987, 1275-1304.

Kagel, John, H., and Roth, Alvin E. (eds.), *The Handbook of Experimental Economics* (Princeton: Princeton University Press, 1995).

Kirchkamp, Oliver, and Moldovanu, Benny, "An Experimental Analysis of Auctions with Interdependent Valuations," *Working Paper*, Department of Economics, Mannheim University, 2001; forthcoming, *Games and Economic Behavior*.

Menkhaus, Dale J.; Borden, George W.; Whipple, Glen D.; Hoffman, Elizabeth; and Field, Ray A., "An Empirical Application of Laboratory Experimental Auctions in Marketing Research," *Journal of Agricultural and Resource Economics*, 17(1), 1992, 44-55.

Neill, Helen R.; Cummings, Ronald G.; Ganderton, Philip T.; Harrison, Glenn W., and McGuckin, Thomas, "Hypothetical Surveys and Real Economic Commitments," *Land Economics*, 70(2), May 1994, 145-154.

Rutström, E. Elisabet, "Home-Grown Values and the Design of Incentive Compatible Auctions," *International Journal of Game Theory*, 27(3), 1998, 427-441.

Schmitz, John D.; Menkhaus, Dale J.; Whipple, Glen D.; Hoffman, Elizabeth, and Field, Ray A., "Impact of Changing Consumer Preferences On Willingness-to-Pay for Beef Steaks In Alternative Retail Packaging," *Journal of Food Distribution Research*, 24, September 1993, 23-35.

Smith, Vernon L., "Microeconomic Systems as an Experimental Science," *American Economic Review*, 72(5), December 1982, 923-955.

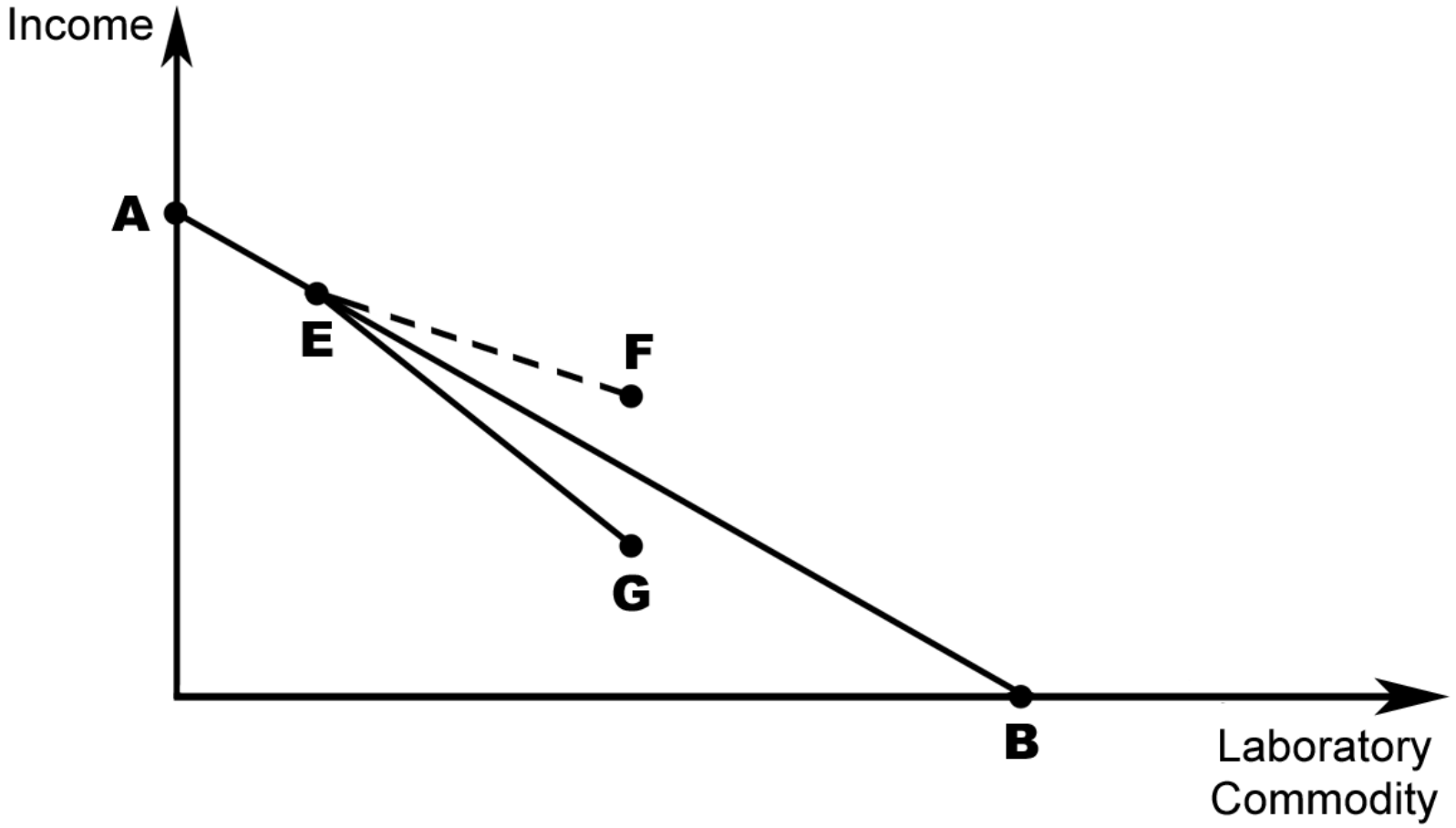

Figure 1: Frictionless Field Market

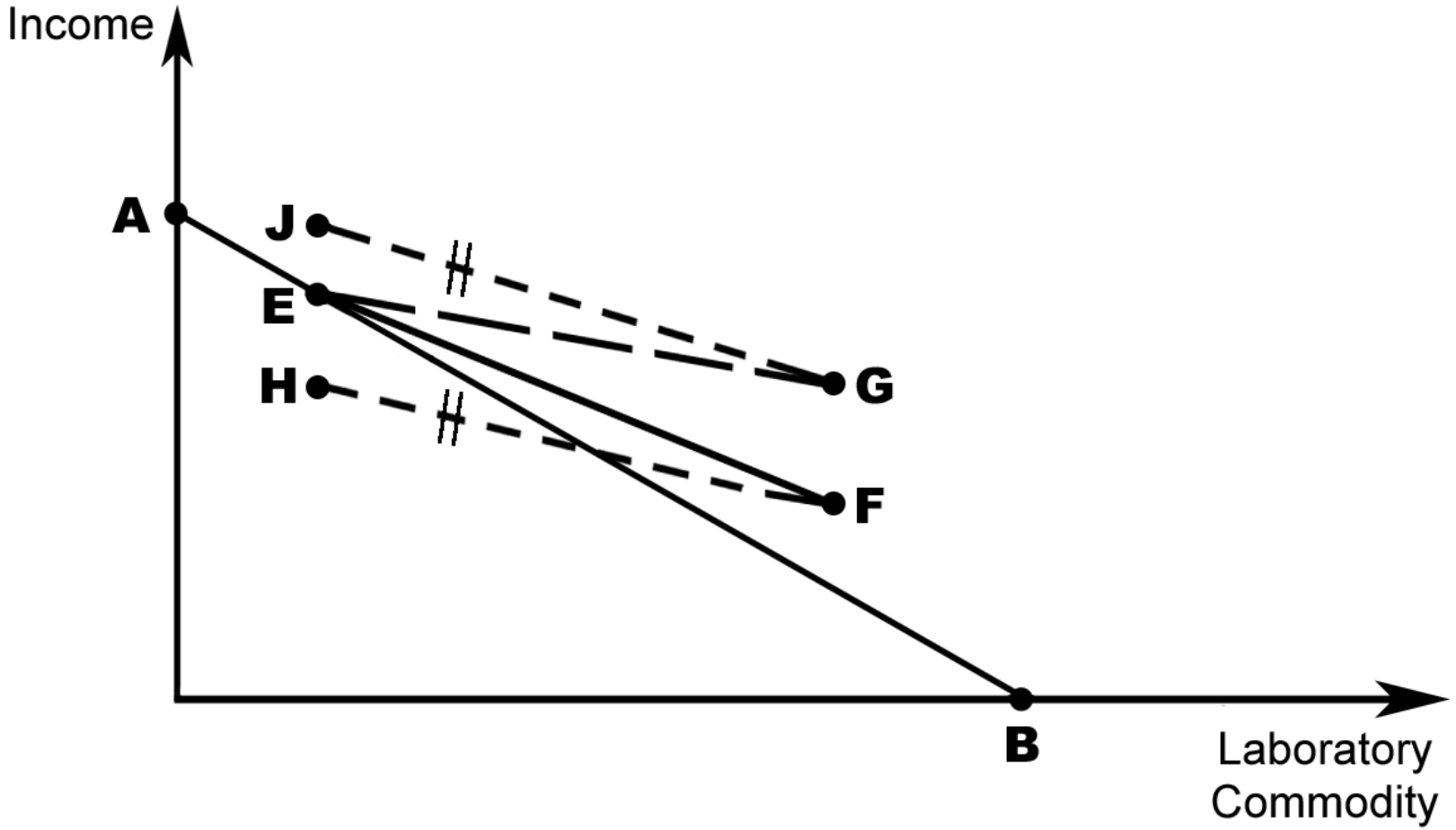

Figure 2: Resale Below Purchase Price

Table 1: Estimated Effects of Experimental Treatments on Bids for Meat

Estimates obtained using random effects specification

Treatment	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
				A. Panel GLS		
Vacuum	-.0549387	.0333573	-1.65	0.100	-.1203179	.0104404
Info	.1824116	.0987915	1.85	0.065	-.0112161	.3760393
Demo	-.1472108	.0986941	-1.49	0.136	-.3406477	.0462262
VacInfo	.449179	.0470355	9.55	0.000	.3569912	.5413669
VacDemo	.0739316	.0468498	1.58	0.115	-.0178924	.1657555
				B. Panel GLS With Market Price Effects		
Vacuum	-.0411943	.0324661	-1.27	0.204	-.1048268	.0224381
Info	-.1916986	.0571364	-3.36	0.001	-.3036839	-.0797133
Demo	.0095262	.0552915	0.17	0.863	-.0988431	.1178955
VacInfo	.3984435	.0458179	8.70	0.000	.3086422	.4882449
VacDemo	.0736634	.0455684	1.62	0.106	-.015649	.1629758
				C. Panel Tobit Model With Market Price Effects		
Vacuum	-.0440431	.0346063	-1.27	0.203	-.1118703	.023784
Info	-.2454266	.0623487	-3.94	0.000	-.3676278	-.1232254
Demo	.0468671	.0617209	0.76	0.448	-.0741037	.1678379
VacInfo	.4198155	.0488044	8.60	0.000	.3241607	.5154703
VacDemo	.0686025	.0484393	1.42	0.157	-.0263367	.1635418

Table 2: Estimated Effects of Market Prices on Bids for Meat

Estimates obtained using Panel Tobit specification with random effects

Variable	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
tbid_vs	.3996299	.0247609	16.14	0.000	.3510994	.4481605
tbid_st	.2883476	.0249211	11.57	0.000	.2395032	.337192
priceL1	.3555019	.0172282	20.63	0.000	.3217353	.3892686

Note: Variables `tbid_vs` and `tbid_st` are the average trial bid prices for the two commodities, and `priceL1` is one-period lagged price.