Laboratory experiments employing an induced-values methodology often report on allocative efficiencies observed. That methodology requires experimenters know subjects’ motivations precisely, questionable in labs, impossible in field experiments. Allocative efficiency implies a hypothetical costless aftermarket would be inactive. An allocation mechanism’s outcome is defined to be behaviorally efficient if an appropriate aftermarket is actually appended to the mechanism and measures at most a negligible size of remaining mutually beneficial gains. Methodological requirements for an appropriate aftermarket are specified. A first demonstration observes more frequent and ex-ante larger behavioral inefficiencies in second- than in first-price auctions. A simple field demonstration indicates when a public-good increase can be observed to cover marginal cost to subjects’ mutual benefit, without knowing valuations. A wide variety of empirical economic-policy studies can utilize this methodology to observe comparative evidence of alternative policies’ allocative-efficiency shortfalls.

C9; C93; D01; D61; D03; D46; Keywords: revealed thresholds, behavioral efficiency, aftermarket, field experiment methodology, allocative efficiency, empirical political economy, valuation revelation
1. INTRODUCTORY MOTIVATION

Recommendations for policy adoption or alteration are more valuable if evidence of the size of shortfalls from allocative efficiency can be provided for the allocation mechanisms or policy instruments under consideration. Such evidence has so far come from laboratory experiments using an induced-values methodology for an abstract commodity.¹ That methodology requires experimental subjects’ motivations be known to the experimenter; it may fall short when subjects have interpersonal utilities or when induced values are otherwise incomplete; it is wholly unable to assess field experiment efficiency.

Field experiment transactions might involve provision or allocation of: irrigation water, adult education, pollution permits, medical services, regulated financial transactions, microfinance, or similar “naturally occurring” goods, services and contracts.² The experimenter can observe transacting behavior but not valuations or motivations, and so cannot calculate the Pareto set nor measure shortfalls from it.

This observability constraint appears to have seriously stunted field-experiment investigations. Other than induced-values experiments conducted using as subjects Peruvian farmers, single mothers in Malawi, et al.,³ the bulk of field experiments have been single-subject studies, examining risk aversion, time preference, likelihood of loan repayment, willingness to trust, preference for more-egalitarian or more-

---

¹ A subject might, for example, be told that she is one of $N$ buyers or $M$ sellers of an abstract good called $X$ (not “tennis lessons” or “coupons for video downloads”) that will be traded, with her payoff being the difference between trading prices and induced values, as in “the first unit of $X$ you buy can be resold to the experimenter for $8.75, the second for $6.80, the third for $5.10” to a potential buyer, or “the first unit of $X$ you sell can be obtained from the experimenter for $3.10, etc.” Cf. e.g., Smith (1962,1976,1982), Davis and Holt (1993), Kagel and Roth (1995), Plott and Smith (2008), and original sources cited in the latter two works.

² In “conducted” field experiments, these transactions occur on a market constructed and controlled by the experimenter; in “unconducted” field experiments, the experimenter observes but cannot control a naturally occurring market. Harrison and List (2004) define “natural” field experiments to add to unconducted experiments a less critical requirement that subjects are unaware they are in an experiment. They define “framed” field experiments just as conducted experiments.

³ It may be methodologically clearer to call these “specified-subjects” experiments than to lump them in with field experiments.
meritocratic compensation schemes, and other aspects of individual preferences, individual consumer behavior or labor supply. These studies could appropriately be described as decision-science experiments; few directly study market outcomes or any issues of allocative or even interactive (i.e., game-theoretic) behavior. One more set of field experiments observes and perhaps controls allocations, but reports only positive (not normative) observations: project cost or subsidization fraction, number of dropouts or take-ups, post-project survey responses, fraction of adult-education graduates employed one year later, and the like.

This paper proposes a definition and a specific but broadly usable methodology to allow observations relevant to allocative efficiency in field experiments, despite observing only revealed thresholds, i.e., only behaviors that stem from unobserved motivations and preferences (sometimes even incompletely observed feasibility constraints). An appropriately designed aftermarket is appended to an experiment, not to improve, but solely to assess the outcome. The aftermarket attains revealed-threshold evaluations without either the ability or the need to know their basis.

I cover theoretical issues arising with this methodology. Following that, sections 3-5 provide a first demonstration of the concept: a simple but concrete example of the appropriate usage of a properly constructed aftermarket to observe allocative-efficiency shortfalls without relying on knowing subjects’ motivations. It finds first-price auctions less behaviorally inefficient than second-price auctions; in this context, subjects’ bidding in these auctions was unaffected by knowing there would be an aftermarket.

Then, a first small field aftermarket is reported, observing whether an increase in output of a public good from an ad hoc starting point can be achieved as a mutually beneficial reallocation. I describe how these demonstrations relate to broader field usage.
The antecedent methodology is Bohm (1984). Bohm suggested an “interval methodology” for field study of excludable binary (0 or 1 units) public-good decisions. A public project’s potential beneficiaries are randomly divided into two groups. Each subject in Group U (my labels, mnemonic for “Understaters”) is asked to state a willingness to pay (WTP) for the project’s provision, with each committed to pay, annually, a to-be-specified common fraction of their stated WTP. Each subject in Group O (“Overstaters”) is asked to state a WTP for provision, with the understanding that any subject stating at least \( W_{\text{min}} \) would be permitted to benefit from the project indefinitely for an annual total payment of \( W_{\text{min}} \) (a parameter set well below per-capita annual project cost). The project is to be undertaken only if the sum across groups of stated WTPs provides (statistically) sufficient confidence that the sum of actual WTPs exceeds project cost, in which case the payment fraction for Group U members will be the ratio: \{\text{half project cost/sum of Group U WTPs}\}.

The assumptions underlying Bohm’s method are [i] each Group U respondent has a transparent incentive to understate WTP, [ii] each Group O respondent has a transparent incentive to overstate WTP, and [iii] the confidence intervals cover any asymmetry in how large misstatements these incentives generated. Notice, no subject

4 The stated WTPs form an Evaluation Interval, which is the convex hull of \{twice the 95% confidence interval for the sum of Group U WTPs\} and \{twice the 95% confidence interval for the sum of Group O WTPs\}. If the lower end of the Evaluation Interval exceeds project cost, then that Group U’s (presumably understated) average stated WTP exceeds per-capita project cost is considered to imply that it is efficient to undertake the project. If the upper end falls short of project cost, then that per-capita cost exceeds Group O’s (putably overstated) average stated WTP is considered to imply that it is inefficient. Should a narrow Evaluation Interval include project cost, indifference over the efficiency of project undertaking is reported. Should project cost fall within a wide Evaluation Interval, then WTP under- and over-statements are deemed too large for a definitive conclusion.

5 Bohm (1984) reports a demonstration for an actual project in Sweden, with 136 respondents in each group (nearly the totality of beneficiaries). Project cost was not far above the Evaluation Interval’s lower end; the project was done, as outcomes near the ends of the Evaluation Interval are viewed as less likely to represent the sum of actual WTPs than outcomes nearer the point estimate. The methodology’s payment rules left a deficit of 20% of project cost. The Evaluation Interval’s width was just over 38% of the point estimate. \( W_{\text{min}} \) was \{project cost/400\}; 36% of Group O respondents stated WTPs exactly equaling \( W_{\text{min}} \), a potential anchoring effect. This large potential effect casts some doubt on [ii]. Clearly a Group U respondent had an incentive to state a WTP below (actual
had an incentive to state WTP correctly, and the methodology does not yield estimates of the size of efficiency gains or shortfalls therefrom.⁶

Treating the allocation to be evaluated as 0 units of public good, the exact aftermarket design demonstrated in section 6.2 could be used to observe whether the efficiency gain from 1 unit was positive or negative, and by how much.

2. Methodological Steps

2.1. Pareto Efficiency

In the 1950’s it became commonplace among several developers of general equilibrium theory to add imagery in a reinterpretation of the usual maximization-problem definition of Pareto efficiency: Suppose an allocation were to be Pareto-efficient. Then a hypothetical costless aftermarket would be inactive: upon reaching a Pareto-efficient allocation, there would be no remaining mutually beneficial transactions to exploit.

The simple starting point of this research project is to replace the hypothetical with the actual: to append a carefully designed aftermarket to an experiment reaching an allocation decision, so as to measure the shortfall from allocative efficiency of the allocation reached.

2.2. Behavioral efficiency

Behavioral efficiency: an outcome of an allocation mechanism is said to be behaviorally efficient if an appropriate (incentive-compatible, suitably transparent, and approximately costless) aftermarket is actually (and immediately) appended to the

---

⁶ Bohm believed the binary nature of the public good decision he studied was a large advantage: “The case of divisible public goods, requiring the revelation of WTP functions, or at least WTP for several alternative quantities, is referred to the science fiction department for the time being.” (pp. 138-9) This paper, decades later, accepts that challenge.
allocation mechanism\textsuperscript{7} and at most a negligible aggregate size of mutually beneficial gains is observed on the aftermarket. Natural extensions include at least the following: [a] an allocation mechanism is said to be behaviorally efficient in a particular context if it reliably yields behaviorally efficient outcomes; [b] a social or economic policy \(Y\) is said to be \textit{behaviorally less inefficient} in a particular context than an alternative policy \(Z\) if the shortfall from behaviorally efficient outcomes under policy \(Y\) is robustly observed in such aftermarkets to be significantly smaller than observed under \(Z\).

2.3. \textit{Aftermarket Methodology}

The aftermarket of this definition must be designed and implemented so as to support the intended normative interpretation. It likely aids first to set aside three straightforward disqualifications: [i] In general, simply repeating an allocation mechanism does not suffice to draw meaningful conclusions about efficiency of the initial application of the mechanism.\textsuperscript{8} [ii] Whatever its structure, a \textit{resale} market (Zheng [2002]) does not suffice. A key terminological distinction: unlike resale markets (e.g., for US Treasury debt), an \textit{aftermarket} necessarily involves \textit{the same} economic actors as the original market (original allocation mechanism), none added and none absent.\textsuperscript{9} Resale of notes purchased at a Treasury auction last month neither implies nor denies an inefficiency in that auction’s allocation. [iii] Later transactions involving an informationally distinct commodity cannot support interpretation of an earlier allocation as inefficient.\textsuperscript{10}

\textsuperscript{7} Any allocation-reaching method, even informal, I harmlessly mislabel an allocation mechanism.

\textsuperscript{8} Were repeating the same mechanism to suffice in some special circumstance, likely it would create needless confusion for subjects, e.g., lead them to infer they did something wrong.

\textsuperscript{9} This definition clearly fits the three-score-old reinterpretation of Pareto efficiency in section 2.1.

\textsuperscript{10} For example, suppose one of a group of competing used-car dealers obtains a particular car at an auction of cars whose leases have ended. Some days later, the consigner of this particular car allows the winning bidder to return it and gives a full refund; that winning bidder continues to be considered a financially reliable bidder by the auctioneer. Even if exactly the same set of bidders are competing when the car is re-auctioned, a different bidder winning the re-auction does not imply any inefficiency of the original auction. The knowledge that the car was returned, inferred from the fact of its re-
If shortfalls from allocative efficiency are to be discerned, the aftermarket must be constructed so as to identify and assess any and all remaining mutually beneficial transactions involving the same set of traders, under the same information as occurs when the original market (or other mechanism) reaches an allocation. This requires subjects’ behavior in the aftermarket be interpretable as revealing the border between potential transactions they prefer to make and prefer not to make, thus allowing the experimenter to infer from observations all relevant willingnesses-to-pay and willingnesses-to-accept. In straightforward situations, this can be accomplished via a typical incentive-compatibility characterization: that the price to any partner in any transaction be independent of his or her own behavior, with the impact of the behavior limited to affecting whether (to be precise, the probability with which) an aftermarket transaction occurs. It is to this feature that the revealed thresholds in the title refers.

Formally (cf. section 2.5), the aftermarket must be designed in such a way that a supposed equilibrium behavior in the original allocation mechanism, together with truthful revelation in the aftermarket, constitute an equilibrium of the mechanism-cum-aftermarket game. Importantly, this is not because equilibrium behavior is assumed or even expected either in the original allocation mechanism or in the aftermarket. Rather, it is because an aftermarket less stringently designed could not possibly be informative as to the original outcome’s efficiency.

If the aftermarket had to be an allocation mechanism in its own right, this revelation requirement would typically be impossible (Myerson and Satterthwaite, 1983, provide an impossibility theorem for perhaps the simplest case). However, the aftermarket is to be appended to a mechanism, not to obtain an allocation, but merely

---

11 I know of antecedent reports of only two experiments employing aftermarkets, Grether, Isaac and Plott (1979,1989) and Rassenti, Smith and Bulfin (1982). Both experiments are fine laboratory studies of airport landing rights; in neither is the aftermarket designed to identify all possible remaining gains from trade, and in both the induced values are utilized to analyze pre- and post-aftermarket efficiency.
to evaluate normatively the allocation that was reached before the aftermarket was used. The simplification thus obtained is characterized in section 2.5.

This bears emphasis: for the purpose of evaluating efficiency, an aftermarket does not have to be an allocation mechanism, and the behavior preceding the aftermarket does not have to be equilibrium or even rational behavior. Moreover, as envisioned in this research methodology, an aftermarket is conducted not to improve an outcome, but solely to measure the allocative efficiency shortfall attained in the preceding market outcome.

That the aftermarket be approximately free of transactions costs, and that it be suitably transparent, necessarily have less exacting interpretations; these are dimensions where designing aftermarkets cannot be as scientific. Negligibility of transactions costs is most usefully evaluated relative to the size of potential mutual gains from further transacting. Indeed, transactions costs yield a calibration: an appropriate aftermarket identifies and quantifies all mutually beneficial transactions for which the perceived gains from trade exceed the perceived transactions costs.12

Experimental psychology and laboratory economic literatures yield insights into transparency that are extensive, though mainly anecdotal and subjective.13

A frequent query is whether a second aftermarket might uncover an inefficiency in the first aftermarket. The question is pointless, because an aftermarket is a measurement tool, not a palliative—it does not try to reach an efficient outcome.

---

12 When subjects have already been congregated, either physically or via simultaneous interaction on the Internet, an aftermarket run fairly quickly and with simple, transparent tasks for subjects is likely to sluff off transactions-costs concerns. When congregating is only required for the aftermarket, it may be that an appropriate aftermarket design compensates subjects for the costs of congregating, being careful to compensate in a manner unrelated to observed aftermarket activity.

13 That price-clock-based ascending-price mechanisms are notably more transparent than sealed-bid mechanisms seems a solid inference to draw from the laboratory experiments reported in Harstad (2000). For this reason, ascending-price mechanisms are the presumptive technology, used in both the lab demonstration of sections 3-5 and section 6’s field demonstration.
Moreover, the (first) aftermarket may reveal yet not implement a gain from reallocation, which information could affect behavior in a second aftermarket.  

2.4. Aftermarket Transactions

A meaningful aftermarket needs to observe thresholds with respect to any potential transaction that may be mutually beneficial; this need will vary with the topics and mechanisms of field studies. When allocation of identical units of a single private good is the concern (or when identical units of multiple goods are at stake, but issues of complementarities or income effects can reasonably be assumed absent), observations of behavioral valuations of bilateral trades suffice.

Suppose the original allocation mechanism determines a public-good quantity and its production-cost allocation. Now an aftermarket merely observing potential bilateral transactions is insufficient; transactions which alter public good output and attain some adjustment of cost shares in accordance with increased or reduced public good production cost must be considered. Some field studies may require the aftermarket be designed to observe valuations of external effects or of dynamically structured contracts; e.g., allocations of common-pool resources.

The usual maximization characterization of efficiency determines marginal conditions and assumes the appropriate convexities to imply that a local optimum is a global optimum. A similar limitation may be needed in many cases to keep aftermarkets sufficiently simple and straightforward. For example, consider examining in an aftermarket both an increase and a decrease in public good output, each by some more-than-differential amount that is in context small. Concluding a behaviorally efficient outcome from an observed inability to find any mutually

---

14 Examples include version 2 in section 2.5, and the first demonstration’s aftermarket (section 4).
15 When goods $L$ and $R$ are complements for some subjects, the possibility that subject 1 might value an additional set $\{L, R\}$ by an amount sufficient to compensate both subject 2 for forgoing one unit of $L$ and subject 3 for forgoing one unit of $R$ must be observable by design.
16 Public-good examples are in section 6.2 and Appendix A.3, an externality example in Appendix A.2.
beneficial increase or decrease by that given amount assumes the unobserved motivations were consistent with marginal valuations decreasing more rapidly than marginal production cost. Should a behavioral inefficiency be found, the study would indicate in which direction public good output could be altered so as to obtain a perceived mutual gain, but not how far such a movement could continue. In most contexts, the imaginable alternative of checking several possible increases in public-good output of varying sizes, and corresponding decreases, is likely to rob an aftermarket of a required low-transaction-costs character.  

2.5. Theoretical Issues

Suppose the allocation mechanism being studied is sufficiently formal to permit analysis as a game $G$. Let $G = \{ G, G+ \}$ be the strategic-form game with nature that implements a properly constructed aftermarket $G+$, where nature is modeled in $G$ as making all choices that nature or any other player made in $G$. Then an appropriate necessary condition for construction of the aftermarket to yield behavioral-efficiency observations is that truthful revelation for all real players be a Nash equilibrium in $G+$ for any play in $G$ (including any equilibrium $E$ of $G$). When possible, it is desirable to increase this necessary condition from Nash equilibrium to dominant strategy.

Otherwise, a behavioral efficiency (or inefficiency) conclusion is unwarranted. It bears repeating: this is not because equilibrium behavior in $G$ is assumed or even expected. If the design of $G+$ distorted incentives so that $E$ was no longer an

---

17 Correspondingly, suppose an aftermarket appended to a mechanism allocating a given quantity of identical units of a private good were to observe that a potential buyer willing to pay the most for an additional unit could not cover the lowest price at which some potential seller was willing to provide the additional unit. Assuming there was no mutually beneficial trade in which this buyer would acquire two units of the good (thus, assuming unobserved motivations included diminishing marginal utility) might be preferable to running an aftermarket that also priced 2-unit (and larger?) trades, and allowed for multiunit trades to have multiple parties on the same side of the trade.

18 This supposition is not trivial: some of the cultural incentive schemes discussed in Ostrom (1998) may be difficult to formalize as allocation mechanisms.
equilibrium of $G$ when $G$ appended $G^+$, then even rational behavior following which the aftermarket measures inefficiencies (behavior under different incentives than those generated by $G$ itself) cannot possibly be the behavior we desire to measure (that of $G$ itself). [ii] If truthful revelation in $G^+$ were not an equilibrium in $G^+$, following any behavior in $G$, whether an equilibrium of $G$ or not, then nothing could be inferred from the observed behavior in $G^+$ about the efficiency of behavior in $G$. Only then can aftermarket behavior possibly measure $G$’s inefficiencies.

Notice these are necessary conditions. As all that is observed is behavior, to expect any sufficient conditions is illogical.19

Aftermarket construction can thus be viewed as a particular type of mechanism design problem. While formal constraints of mechanism design are often limiting, the mechanism design challenge posed here should always be attainable. An aftermarket constructor has three critical dimensions of flexibility generally unavailable in mechanism design. The experiment’s aftermarket: [a] does not have to balance the budget, though hopefully limiting the size of any deficit; [b] does not have to implement any transactions observed to be mutually beneficial with probability one, but merely with positive probability (and thus does not have to be an allocation mechanism);20 [c] can implement one observed mutually beneficial transaction $\alpha$ at the expense of not implementing a different observed mutually beneficial transaction $\beta$ even though the observed efficiency gain from $\beta$ is larger, provided the ex ante probability of implementing a given observed mutually beneficial transaction is increasing in the size of the observed efficiency gain.

---

19 Valuations are not observed, merely inferred from aftermarket behavior. Feasibility conditions may be incompletely observed; subjects’ perceptions of feasibility conditions are not observed.
20 It may be worth noting that, where the implementation is financial, this implies a positive probability that the commitments made are financially incurred. I see no opportunity for surveys about whether subjects wished to reallocate, or about hypothetical terms under which subjects wished to reallocate, to substitute for an aftermarket.
Appendix A.2 provides an example of an allocation situation in which an appropriate aftermarket must take advantage of dimension [c].

As a simple illustration, suppose a field experiment has observed a failure to reach an agreeable transaction in a bilateral-bargaining situation. Motivations are unobserved, so it is unknown whether the failure was an efficient outcome, or a mutually beneficial bargain was possible but not reached.\(^21\) However, at least three distinct example aftermarket constructions can observe whether the outcome already observed was behaviorally efficient.\(^22\)

Each construction asks the seller to state the lowest price that he is willing to accept, and the buyer to state the highest price that she is willing to pay.\(^23\) The experimenter has carefully explained to the subjects, in advance, what use will be made of their responses. Aftermarket construction 1 will implement the transaction whenever \(B\), her stated willingness-to-pay, exceeds \(S\), his stated willingness-to-accept, with the pre-announced rule that she will pay \(S\), and he will receive \(B\). Construction 1 is incentive-compatible, implements any transaction observed to be mutually beneficial, and requires the experimenter to cover the deficit \(B - S\). Aftermarket construction 2 draws a random variable \(R\) from a distribution exogenous to all information provided by this pair of subjects, and transacts at random price \(R\) if \(B \geq R \geq S\). Construction 2 is incentive-compatible, implements any transaction observed to be (strictly) mutually beneficial with positive probability, and balances the budget. Aftermarket construction 3 also draws a random variable \(R\) from an exogenous distribution, transacts if \(B \geq R \geq S\), but she pays \(R\) and he receives \(1.05R\), achieving incentive compatibility, implementing with positive probability any transaction.

\(^{21}\) Myerson and Satterthwaite (1983) show no bilateral-bargaining allocation mechanism can insure efficient outcomes when the potential seller’s and potential buyer’s valuations are private information.

\(^{22}\) If the failure to reach a transaction occurred in an unconducted field experiment, the aftermarket would require a transition to a conducted field experiment.

\(^{23}\) Depending on context and the background and culture of the subjects, this may well not be the language in which the experimental instructions state the request. “To state” may be implemented via clicking on a price clock, as in sections 4 and 6.2.
observed to be mutually beneficial, but requiring the experimenter cover a deficit of \( R/20 \) when transactions occur.

Note that mechanism design requirements can still impinge on aftermarket design desiderata. In particular, consider an aftermarket construction 4 which attempted to alter aftermarket 1 above by only transacting when the deficit \( B - S \) did not exceed a maximum desired experimenter cash infusion \( M \). This construction would no longer suffice for incentive compatibility.\(^{24}\)

3. LABORATORY SETTING

Five-bidder sealed-bid auctions of a single abstract asset were conducted, in seven sessions (110 subjects) via first-price rules, and in six sessions (85 subjects) via second-price rules.\(^{25}\) There were 10, 15, 20, 25 or 30 subjects in the laboratory during a session, with random reassignments into groups of five each period.\(^{26}\)

Affiliated asset valuations (Milgrom and Weber, 1982) for subjects were determined as follows.\(^{27}\) In each period, first a random number \( C \), called a central tendency, was drawn uniformly from \([50, 1000]\) (all random variables are multiples of \(0.01\)). Then, given a realization \( c \) of \( C \), for each subject \( j \) an estimate \( X_j \) was drawn uniformly from \([c-10, c+10]\), conditionally independent. Finally, asset value to subject \( j \) was \( V_j = \frac{3}{4} X_j + \frac{1}{4} C \); this system incorporates private values (the first \( X_j \) is a random variable which is independent of \( C \)).

\(^{24}\) The seller might attain the outcome of no transaction and no gain if he truthfully stated \( S \), as \( B - S \) might exceed \( M \), while some overstatement \( X > S \) might yield a gain of \( B - X > 0 \) should \( B - X \) be less than \( M \). Thus, instead of an incentive to truthfully reveal his willingness-to-accept, the seller (and perhaps the buyer) would optimally trade off a greater gain in the event of transaction against a higher probability of a gain by some degree of misstatement. (Even if, instead of announcing \( M \), the experimenter merely stated that the transaction occurs “unless the deficit is too large,” the construction would still be insufficient to warrant conclusions as to behavioral efficiency.)

\(^{25}\) Thus, contrasts across pricing rules are between-subject comparisons.

\(^{26}\) Subjects were University of Arizona undergraduates, recruited campuswide via website, sat at visually isolated computers, and had no experience in similar experiments. A second-price auction session for which less than ten subjects showed was eliminated from data analysis. The experiments were conducted in October and November 2009, using the Z-Tree programming environment (Fischbacher, 2007). Instructions are here and here.

\(^{27}\) With respect to the experiment’s primary purpose as a demonstration of measuring auctions’ shortfalls from efficiency via aftermarkets, that asset values were induced is an irrelevant detail.
term, $X_j$) to introduce efficiency issues, as well as a natural, small common-value component ($C$). These rules were carefully explained and examples given. $C$ was not revealed to subjects until end-of-period feedback, which gave a complete, anonymous report of $C$, $X_j$'s, $V_j$'s, all behavior, and profit calculations. The session protocol is in Appendix B. This setting offers a contest between theory and prior experiments, cf. Appendix E.

4. Aftermarket Implementation

To discern from subjects’ behavior whether an auction attained an efficient outcome, the experiment appended an aftermarket designed as follows. Once all subjects had typed and submitted their bids, the acquiring bidder was determined (by fair, random tie-breaking if necessary). Each bidder was informed of the auction price and whether his bid acquired the asset (no other feedback). Upon a few seconds’ warning, the aftermarket began; subjects had been told before bidding that this aftermarket would follow the auction.

A price clock ticked up on all subjects’ screens, rising by $0.50 every two seconds (though more slowly in the first period with an aftermarket), beginning at a random price calculated to be acceptable to all potential purchasers but noisy enough to

---

28 A principal reason for including a common-value component was to avoid a throw-away bid problem: with independent private values, most values will yield so low a chance of winning as to make serious consideration of what to bid not worthwhile. Unfortunately, how high a value a subject has to draw before he or she chooses to pay attention is unobservable. In the current design, all estimates between $60 and $990 have the same expected profitability, removing throwaway bid concerns. (Data analysis only includes the ~98% of cases where subjects’ estimates were in the [$60, $990] range.)

29 The instructions stated that a reserve price, below which the asset would not be sold, was drawn anew before each auction, uniformly from $[(c – 10), (c – 6)]$, and would not be revealed until end-of-period feedback. As planned, the reserve price was never binding.

30 Subjects began the experiment with a bank balance of $12, with profits added and losses subtracted during the session, and final balances paid in cash. These valuation procedures call for a small winner’s curse correction; the 90% confidence interval for the loss in the event a winning bid exceeded the symmetric, risk-neutral equilibrium bid by exactly the winner’s curse correction (were all rival bids in equilibrium) is about $[1.50, 4.25]$. Thus, three to five such losses could likely be handled without the balance going negative. If a subject’s balance became negative, he was given a $20 loan to be repaid out of his final bank balance. Two of 195 subjects could not quite repay the loan; it was of course forgiven and they were paid only the usual $5 show-up fee.
avoid revealing information about the still-unknown $C$. The bidder who acquired the asset was labeled the acquirer, and asked to click the “Accept” button on the screen as soon as the price reached the lowest price at which he was willing to sell the asset just acquired in the auction to one of the four rival bidders. Each of the four bidders that did not acquire the asset was asked to do nothing so long as the prices being shown were prices at which he would be willing to buy the asset from the acquirer, and then to click “Accept” at the highest such price.$^{31}$

Instructions had carefully described the rules relating these Rebids (prices the four nonacquiring bidders accepted) and the Offer (price where the acquirer clicked Accept) to possible aftermarket transactions. [1] If the Offer exceeds all four Rebids, there is no aftermarket transaction. [2] If at least two Rebids are no lower than the Offer, the asset is transferred from the acquirer to the highest Rebid’s bidder, at a price set by the second-highest Rebid. [3] If the highest Rebid exceeds the Offer which exceeds all other Rebids, a random number $R$, drawn before the auction from multiples of $0.01$ in [$c$, $(c + 15]$] equiprobably, determines the aftermarket outcome. If $R$ falls between the highest Rebid and the Offer, the asset is transferred from the acquirer to the highest Rebidder, at a price set equal to $R$; otherwise, there is no aftermarket transaction.

This aftermarket design justifies the inferences about behavioral efficiency to be drawn from observations of aftermarket behavior: any mutually beneficial trade revealed is transacted with positive probability, and in no aftermarket transaction is the price determined by the behavior of either transacting party. Among possible aftermarket designs, prior experimental evidence (Harstad, 2000) suggests the use of the price clock makes aftermarket incentives as transparent as possible. Whenever at least one nonacquiring bidder selects a Rebid above the acquirer’s Offer, a mutually

$^{31}$ No subject observed any information about other subjects’ behavior in the aftermarket until all five had clicked on a price.
beneficial trade that the auction did not achieve has been identified and the gain measured (whether or not the aftermarket actually transacts that trade).

5. Aftermarket Observations

First-price [second-price] auctions were observed to be behaviorally efficient in 72% [57%] of observations (cf. Table 1). In 28% of 203 first-price auctions, and 43% of 142 second-price auctions, at least one bidder who was outbid was observed to be willing to buy the asset from the high bidder for mutual gain. This difference is significant at the 1% level in a Pearson test.  

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Observations</th>
<th>First-Price Auctions</th>
<th>Second-Price Auctions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Aftermarkets Observed</td>
<td>203</td>
<td>142</td>
<td></td>
</tr>
<tr>
<td>2 Behaviorally Efficient</td>
<td>72%</td>
<td>57%</td>
<td></td>
</tr>
<tr>
<td>3 Mean of Observed Shortfalls</td>
<td>$4.45</td>
<td>$4.16</td>
<td></td>
</tr>
<tr>
<td>4 Shortfall Capacity</td>
<td></td>
<td>$10.26</td>
<td></td>
</tr>
<tr>
<td>5 Aftermarket Fraction</td>
<td></td>
<td>41%</td>
<td></td>
</tr>
<tr>
<td>6 Expected Behavioral Shortfall</td>
<td>$1.246</td>
<td>$1.788</td>
<td></td>
</tr>
</tbody>
</table>

Where aftermarket behavior exhibited such gains, the difference between the most an outbid bidder will pay and the least the acquiring bidder will accept is a behavioral measure of the shortfall from efficiency, averaged in row 3. While shortfalls when observed were larger in first-price auctions, with zero shortfalls

---

32 These percentages sum occurrences where the aftermarket transacted with those where the random price fell below the Offer or above the one Rebid exceeding that Offer.
averaged in for the behaviorally efficient outcomes, the expected shortfall in row 6 is smaller for first-price auctions.

It bears emphasis that, while these auctions sold induced-value assets, the behavioral efficiency and shortfall measures make no use of any information contained in the induced values. These reports stem solely from subjects’ behaviors: their sealed bids, Accept Bids and Accept Asks, and in no way depend on any information about subjects’ motivations, nor on completeness of values induction.33

Reports of allocations reached in induced-values experiments typically provide efficiency measures in percentages, because the dollar value of total gains from trade in Pareto-efficient allocations can be calculated from the induced values. This methodology cannot be used absent experimenter knowledge of valuations.

In some situations, behavior in the original allocation mechanism can offer a benchmark for the economic significance of the size of shortfalls from efficiency, which I call the shortfall capacity. This experiment demonstrates both the possibility and its limitations.

Second-price auctions are incentive compatible, in that the bid selected determines only whether the bidder wins or not; the price is solely determined by the highest rival bid. In particular, the risk-neutral symmetric equilibrium bid equals expected asset value conditioned on an assumption that the bid is pivotal.34 For the distributions of section 3, this implies that bids differ from bidders’ expected values by a constant.35 Hence, differences between two bids submitted in second-price auctions equal (in symmetric equilibrium) differences between the two bidders’ willingnesses-to-pay and therefore measure the gain from trade if the asset were

33 That is, if a subject had an unusual degree of innate aggressiveness, this was not observed, but it potentially affected auction bids, frequency that subject won, and that subject’s aftermarket behavior.
34 For private values, this is the dominant strategy discovered by Vickrey (1961). This feature of second-price, common-value auctions was first found by Matthews (1977); the intuition is developed in Harstad and Bordley (1996).
35 This neglects estimates in the ranges [50, 60] and [990, 1000], for which the difference is not constant. Data for such estimates was excluded from data analysis.
hypothetically to be transferred from the lower bidder to the higher bidder. These second-price observations include 588 differences between a losing bid and the high bid in the same auction. The average of these bid differences is the $10.26 reported in row 4 above. This is the capacity for shortfalls from efficiency in the sense that, had the original auctions allocated the asset equiprobably among inefficient acquirers, aftermarkets that reallocated to the efficient acquirer could average $10.26 in gains from trade unattained by such a complete misallocation to a random inefficient acquirer.

The best measure I can envision to attach an economic significance to the $4.16 mean of shortfalls revealed by aftermarkets following second-price auctions is that it is 41% of the shortfall capacity.

The absence in Table 1 of a comparable first-price benchmark is not an oversight. Differences between a losing first-price bid and the high bid in the same auction could be averaged. However, without the incentive compatibility of second-price auctions, these first-price bid differences have no similarly strong argument to measure gains from hypothetical transfers between the bidders.\(^36\) In each circumstance where I have envisioned an outline of an appropriate aftermarket design, any potential transaction perceived to be mutually beneficial that is observed provides an absolute magnitude of the perceived gain.\(^37\)

5.1. Submitted-bid Impact

The unique risk-neutral symmetric Bayesian equilibrium of the game consisting of the sealed-bid auction (either pricing rule) followed by the aftermarket is for each

\(^{36}\) Optimizing the \{profitability given winning/probability of winning\} tradeoff in the risk-neutral symmetric equilibrium of the first-price auction yields a nonlinear term in the bid function (corresponding to the term that is a constant in second-price auctions). In that equilibrium, if bidder \(A\) outbids bidder \(B\) by $8, \(A\)'s willingness-to-pay exceeds \(B\)'s, but the $8 bid difference is not a measure of the willingness-to-pay difference.

\(^{37}\) Comparison of these measures with the Pareto-efficiency measures obtained from induced-values data is considered in Appendix C.
bidder to make the same equilibrium bid as if there were no aftermarket, and then truthfully reveal in the aftermarket. Such theoretical separability is essential to an appropriate aftermarket. It remains an empirical question whether subjects bid the same way when they know an aftermarket will follow; there might be reasons subjects would find for bidding less, or for bidding more, in an auction when knowing there will be an aftermarket.\textsuperscript{38} The protocol in Appendix B assists addressing this question.

The following linear bid function was estimated separately from the first-price and second-price data:

\[ M_{st} = \text{const} + \beta_x X_{st} + \beta_a After + \text{error}_{st}, \]

where the markup \( M_{st} \) was the observed bid minus the asset value estimate \( X_{st} \) for subject \( s \) in period \( t \); \( X_{st} \) was a control for possible learning effects, the number of periods of experience in the affiliated-values auctions; \( After \) was a dummy variable taking the value 1 if the subject knew the auction in period \( t \) would be followed with an aftermarket, 0 if the subject knew the auction would not be followed with an aftermarket.

Estimates obtained from OLS linear regressions with clustering by subject are presented in Table 2. For both types of auction rules, a null hypothesis that subjects bid no differently when knowing there would be an aftermarket as when there would not cannot be rejected at anything vaguely approaching conventional levels of significance.\textsuperscript{39}

\textsuperscript{38} Among possibilities are that a subject might perceive an opportunity to win the auction profitably and then profit further by selling in the aftermarket, which could be perceived as suggesting more aggressive bidding than if there were to be no aftermarket; or a subject might perceive the aftermarket as a second chance to obtain the asset, which could suggest less aggressive bidding than without an aftermarket. Both of these possibilities are in equilibrium illusory.

\textsuperscript{39} Failure to reject this null was found in alternatives that did not cluster or that added subject fixed effects; alternatives where the bid was the dependent variable and asset value estimate an independent variable were insignificantly different.
Table 2

<table>
<thead>
<tr>
<th></th>
<th>First-Price Data</th>
<th>Second-Price Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-2.346</td>
<td>-1.435</td>
</tr>
<tr>
<td>Std. error</td>
<td>(0.567)</td>
<td>(1.629)</td>
</tr>
<tr>
<td>Significance</td>
<td>0.001</td>
<td>0.38</td>
</tr>
<tr>
<td>$\beta_x$</td>
<td>0.045</td>
<td>0.393</td>
</tr>
<tr>
<td>Std. error</td>
<td>(0.075)</td>
<td>(0.370)</td>
</tr>
<tr>
<td>Significance</td>
<td>0.55</td>
<td>0.29</td>
</tr>
<tr>
<td>$\beta_y$</td>
<td>-1.242</td>
<td>-5.063</td>
</tr>
<tr>
<td>Std. error</td>
<td>(1.431)</td>
<td>(4.738)</td>
</tr>
<tr>
<td>Significance</td>
<td>0.39</td>
<td>0.29</td>
</tr>
<tr>
<td># Observations</td>
<td>2215</td>
<td>1075</td>
</tr>
<tr>
<td>F test</td>
<td>1.24</td>
<td>0.58</td>
</tr>
<tr>
<td>Significance</td>
<td>0.293</td>
<td>0.562</td>
</tr>
</tbody>
</table>

That subjects did not significantly alter their bidding behavior between auctions known to be followed by an aftermarket and auctions known to be final determiners of payoffs is also suggestive of suitable transparency of the aftermarket structure. It would be possible in most field settings to surprise subjects with an aftermarket that they almost surely did not anticipate; this will not always be best practice.

6. **Field Demonstration**

6.1. **Field Context**

To focus this demonstration on the aftermarket, the initial allocation mechanism is submersed via an exogenous outcome: production of one unit of public good, with the experimenter’s budget covering costs of that production. The aftermarket then considers whether a second unit’s production cost can be allocated to the
behaviorally perceived mutual benefit of all members of the economy. Aftermarket design contrasts with the auctions aftermarket. Here it institutes with probability one an allocation found mutually favorable, but does not balance the budget: it has the experimenter cover a deficit if the sum of stated marginal benefits exceeds unit production cost.

The public good studied is a uniform distribution of small packets of Haribo candy, a product in international distribution and prominent on the shelves of local grocery and convenience stores. It is natural to think of candy as a private good, but in this experiment it was allocated under strict adherence to the definition of a pure public good: [a] there was group exclusion but no individual exclusion in consumption, and [b] there was no rivalry in consumption. Either all subjects in the economy received one unit of candy apiece, or all subjects received two units of candy apiece, depending on whether stated willingnesses-to-pay summed to at least the production cost. All groups (economies) studied consisted of six subjects.  

Subjects were seated at visually isolated computers. Instructions were passed out and read aloud, questions encouraged and answered. The initial unit of Haribo candy was given to each subject; they were allowed to consume it immediately if they were uncertain of the quality of the candy or for any other reason. They were informed that a second unit would be provided to every member of the group if the most each group member was willing to pay summed to at least €1. These were elicited, the second unit provided or not, and subjects paid to the experimenter their cost share for the additional unit (which was necessarily less than the show-up fee).

---

40 18 subjects (in one session, 12, due to no-shows) were in the room during a session; no subject knew which others were in the same group. Subjects were students recruited by website signup from several campuses of the University of Montpellier. (Show-up fees ranged from €3-8, depending on the distance from their home campus to the Experimental Economics Lab at the Richter campus.)

41 Instructions (English translation) are available at http://harstad.missouri.edu/Instructs/.

42 Haribo candy was of course available for purchase outside the lab, and a subject’s transactions costs of doing so were unknown. Hence, it was important to keep the per-capita threshold for public good production below extra-laboratory prices, so that censoring stated valuations by extra-laboratory availability could not be an issue; cf. Harrison, Harstad and Rutstrom (2004).
6.2. Methodology Implementation

Recall the allocation of 1 unit of public good is treated as if it arose via some allocation mechanism, with the experiment observing an aftermarket. A price clock ticked up on subjects’ computer screens, increasing by 2 euro cents, initially every 4 seconds, after 8 euro cents, increasing every 2 seconds. Subjects were asked simply to watch the clock so long as the price was one which they were willing to pay in order to have the group increase public good output from one unit to two, and then to click on the “Accept” button on the screen as soon as the next tick of the clock would yield a price that they were not willing to pay in return for the second unit.

Before the clock was run, the outcome function was carefully explained to subjects. If the sum of the six “Accept” prices was at least €1, each subject in the group would be given a second packet of candy, and each subject would pay $\text{max}\{€1 \text{ minus the sum of the other five Accept prices}, 0\}$.

This methodology implements the incremental version of the Vickrey-Clarke-Groves (VCG) mechanism. If a subject is certain of the amount of euros which he would be willing to pay to have the public good output increased from one to two, then it is a dominant strategy to click on the Accept button at the multiple of 2 euro cents nearest his willingness-to-pay.

The incentive compatibility of this methodology warrants the conclusion that the economy exhibits a behavioral inefficiency of the initial allocation—of one unit of public good—if the sum of Accept prices exceeds €1. This aftermarket structure runs a deficit (of at most €1); a balanced-budget aftermarket, likely less transparent, can be designed (cf. Appendix A.3).

---

6.3. Field Observations

Thirteen of twenty-three groups exhibited a perceived mutual gain in increasing the distribution of candy from one unit to two units apiece.\textsuperscript{44} Sizes of behaviorally revealed efficiency gains in these groups ranged from 4\% to 72\% of production cost (€1), averaging 33\%.

The other ten groups found one unit of public good to be behaviorally efficient relative to the sole alternative of two units. Sums of Accept prices in these groups ranged from 56\% to 98\% of cost, averaging 86\%.

In none of the thirteen groups attaining the behavioral efficiency improvement could public good production cost have been covered by a mutually acceptable uniform tax. Rather, only through person-specific pricing could the public good increment be mutually beneficial. Of course, acceptable person-specific price vectors (not uniquely determined in any of the thirteen groups) could not have been known to the experimenter, but were revealed by behavior in the aftermarket (and the VVCG selection among those vectors actually used for payment).

The highest stated willingness-to-pay was €0.54; 18 of 138 subjects chose an Accept price of €0.04 or less, another 45 €0.16 or less.

Though there is evidence that it was transparent (cf. Appendix D), it is of course not known whether subjects adopted the dominant strategy of revealing their willingnesses-to-pay. Instructions made it clear that subjects were to evaluate not a second unit of Haribo candy for their own consumption, but a second unit of public good production. Nonetheless, it is unknown whether any subject selfishly placed the same value on second units for all group members as on his private purchase of a second unit for himself only. Nor is it known whether any subjects were behaving altruistically, or the extent of any altruistic behavior. It is no more necessary to know

\textsuperscript{44} This included 6 of the 11 groups that chose well before lunchtime, and 7 of the 12 groups that chose shortly after 1:30 [or 3 after 3:30 pm], yielding no sign that chronobiology played a role.
their motivations than it would be necessary to know why a consumer purchased a shirt in order to evaluate the allocative efficiency of a shirt market. This aspect justifies treating a laboratory setting as a simple field experiment.

Although the setting was simple almost to the point of contrivance, and the stakes miniscule, this demonstration indicates that, at least in the case of public good allocation, the concept can be taken to the field. Whether an adjustment in public good output can be accomplished—via a mutually beneficial decentralization of adjustment costs—can be inferred from observations solely of behavior, provided the aftermarket used to observe those behaviors is appropriately designed.

7. Empirical Issues

Necessary conditions (section 2.5) are nearly the only issue relating to aftermarkets that can be settled theoretically. One more: Might a disparity in which willingness-to-accept dramatically exceeds willingness-to-pay prevent aftermarkets from informing us about efficiency? No, definitively; for any particular valuation sought in an aftermarket, only one of WTA and WTP is relevant.\(^{45}\)

Empirical issues are unlimited. For example, whether different populations (farmers in Peru vs. Arizona undergraduates) might show different relative behavioral efficiencies in first- and second-price auctions, or whether different aftermarket structures both satisfying necessary conditions will necessarily rank two

---

\(^{45}\)This confusion is not rare. If, e.g., a public good increase from \(Q\) to \(Q+1\) is considered, the valuations observed in the aftermarket will be WTP’s, with the increase recommended if their sum exceeds marginal cost. The purely hypothetical WTA’s for a reduction from \(Q+1\) to \(Q\) are of no import. In considering a decrease from \(Q\) to \(Q-1\), valuations observed are WTA’s, with the decrease recommended if marginal cost exceeds their sum; the WTP’s for an increase from \(Q-1\) to \(Q\) are irrelevant.

For another example, a losing bidder in a multiunit auction is only considering buying a unit in the aftermarket, only his WTP is relevant. A winner of \(Q_{\text{max}}\), the maximal allowed number of units, is only considering selling a unit in the aftermarket, only his WTA is relevant. A winner of \(Q < Q_{\text{max}}\) units may be considering both buying a \(Q+1^{\text{st}}\) unit, for which only his WTP for \(Q\) to \(Q+1\) is relevant, and also selling a unit, which only involves his WTA for \(Q\) to \(Q-1\) (not his WTP nor his WTA for \(Q+1\) to \(Q\)).
economic policies correspondingly. \textsuperscript{46} I limit myself to three more that have come up in seminars and discussions.

Will an aftermarket observe activity just because subjects assume they are supposed to do something?\textsuperscript{47} Might other studied psychological biases and behavioral anomalies affect aftermarket observations?\textsuperscript{48}

Are aftermarket activities mistakes if the initial allocation reached should have been Pareto-efficient? It will be usually difficult if not impossible to label particular aftermarket behaviors mistakes. Only when conducted in induced-values demonstrations will it be possible to determine whether the original allocation reached prior to the aftermarket was in fact Pareto efficient, and even then usually only under stringent assumptions that may well be unverifiable. In some studies, there could be serious questions about whether the aftermarket design was sufficiently transparent for the subject population, as for example when the subjects

\textsuperscript{46} A couple seminar commenters have suggested a broad way in which aftermarket observations might be of limited value in assessing efficiency of outcomes from allocation mechanisms or policy choices: that preferences of subjects may not be stable across an allocation mechanism and the appended aftermarket. If engaging in behaviors that might potentially have transactional impacts \textit{per se} renders preferences unstable, it is unclear that any transactional outcomes can normatively compared (even to autarky), and indeed that any normative economic research—whether theoretical, experimental, or empirical analysis of historical data—can be meaningful.

\textsuperscript{47} This question arose, for good reason, in reports of experiments observing financial bubbles in labs (Smith, Suchanek and Williams, 1988, and many later papers). The particular “active participation hypothesis” raised by Lei, Noussair and Plott (2001) (that subjects engage in irrational activity because the experimental setup limited rational behavior to inactivity) need not be a concern here. In aftermarkets, anticipated designs will always have subjects do something, in essence engage in valuation activities. Even if the initial allocation might have been Pareto efficient, there will be a financial incentive to engage in the valuation activities. It should always be possible to structure them so that a zero valuation behavior has nothing special to do with a preference for the status quo that was reached in the original allocation mechanism.

\textsuperscript{48} To a first approximation, an “anomaly” such as attention to sunk costs, other-regarding preferences, or hyperbolic discounting, may similarly impact both an original allocation mechanism and its aftermarket. For many, there is no reason to believe that they suddenly arise in aftermarkets following an allocation mechanism that went untouched by them. Since it is only behaviors that can be observed in field settings, observing aftermarket activity in the presence of such biases may well be a most appropriate way to provide advice for policies that will be promulgated for the population being studied.
are illiterate. When transparency is adequate, I regard it as sensible that aftermarket behaviors be taken at face value.

8. Field Readiness

A field study of a single-asset auction could exactly mimic the procedures of section 4 to obtain evidence on whether the efficient acquirer won the auction and if not, the size of the inefficiency that arose, even in cases where existence of equilibrium is in doubt (cf. Jackson [2009]) or is incalculable (subjects’ beliefs about rivals’ valuations are unknown) or unobservable (whenever motivations and valuations are unknown to the experimenter).49

Fairly straightforward complications of the aftermarket design used here can accommodate observing efficiency shortfalls for mechanisms seeking to allocate multiple homogeneous assets. Consider the example of the two-unit uniform-price and Vickrey auctions of List and Lucking-Reiley (2000). Potential buyers in that aftermarket could be asked for [i] a pair of Accept Bids if seeking to buy, [ii] a single Accept Bid for a second unit and also an Accept Ask for possible sale if one asset had been acquired, or [iii] a pair of Accept Asks if two assets had been acquired. Aftermarket rules could specify that neither party to a trade determined the price: whenever an Accept Bid by a rival fell between an Accept Bid and a lower Accept Ask, it set the price for that transaction, and a random price was consulted when necessary. (Details are in Appendix A.1.) In larger, semi-competitive markets for homogeneous assets, a variant on a call market could serve as an aftermarket (so long as no trader were seeking both to buy more and to sell some of what he had obtained), with the highest quantity where the demand price exceeded the supply price transacted, buyers paying the price of the last accepted supply unit, sellers

49 It may be good practice first to familiarize bidders in that field setting with the aftermarket procedures, perhaps by auctioning off an unrelated, less expensive “demonstration” commodity and then running the aftermarket with the demonstration commodity, prior to conducting the aftermarket to be used to gauge behavioral efficiency.
receiving the (higher) price of the last accepted demand unit, and the experimenter covering the deficit.

Similarly, the aftermarket design used in section 6.2 is ready for more substantial field usage in any pure-public-good study in which that design is deemed sufficiently transparent for the studied population. It is straightforward to observe whether a decrease in public good output from an allocation mechanism outcome $Q$ to $Q - 1$ can remit production cost savings in person-specific, mutually beneficial rebates. For example, as the price clock increases, each subject would be asked to click on the smallest rebate that would compensate for the reduction in output, with VCG rebates implemented if the sum of accepted rebates were at most the production cost savings. Should an increase [decrease] in public-good production be found mutually advantageous, it could be possible to repeat the aftermarket to see if $(Q + 2) [(Q - 2)]$ would represent a behavioral efficiency gain relative to $(Q + 1) [(Q - 1)]$.

9. Final Remarks

Observation of a properly designed aftermarket provides an efficiency measure of the preceding outcome with few assumptions about underlying motivations of behavior. In addition to field experiments, this feature can be valuable in laboratory experiments.50

It is a luxury of a parsimonious theory that economists who might disagree about the role of Pareto efficiency—how important is, or perhaps even whether it is desirable—nonetheless agree on the definition of the term and its meaning. I do not see how the terminology of empirical, behavioral studies can have the same luxury. Thus, even if the definition of behavioral efficiency offered here becomes widely

---

50 It would allow inferences even if completeness of induced valuations were subject to question, as when subjects may have altruistic, egalitarian, spiteful or other motivations, or native aggressiveness or passivity exceeding that of other subjects, or unobserved beliefs about rivals' motivations or behavior, any to a degree unknown to the experimenter.
accepted, it seems naïve to hope that its meaning will achieve any universal interpretation.

A perhaps less naïve hope is the following. Suppose, starkly, that policy Y is observed in a particular context (including a particular subject population and their characteristics) to robustly yield behaviorally efficient outcomes, while alternative policy Z in the same context robustly yields outcomes with significantly large shortfalls from behavioral efficiency. Then it might be widely agreed that, at whatever levels of sophistication underlie their perceptions and whatever level of transparency the aftermarket offers, subjects perceive mutual gains from trade that policy Z does not capture while perceiving no uncaptured mutual gains from trade following implementation of Y. Indeed, it might even be widely accepted that advice to policymakers reporting and influenced by this finding (among others) could be an improvement over advice reporting and influenced by field experiments that obtain no observations about allocative efficiency. Less starkly, when the observed size of shortfalls from behaviorally efficient outcomes are in context robustly smaller for policy Y than for Z, this might also come to be accepted for playing a role in policy advice despite divergent opinions as to its exact meaning.

Appending behavioral efficiency observations can dimensionally enrich field studies.

ACKNOWLEDGMENTS

I thank, without implicating, the Economic Science Laboratory at the University of Arizona and LAMETA at Universite de Montpellier for providing facilities, Cathleen Johnson and Marc Willinger for organizing experimental facilities and arrangements, Eric Cardella, Raymond Chiu and Dimitri Dubois for assistance in conducting experiments, Katsunori Yamada for assistance in data analysis, Dubois and Rafaele Preget for translating instructions, Willinger and Sophie Thoyer for organizing financial support, Glenn Harrison, John Ledyard, David Levine, David Reiley, Alvin Roth, Larry Samuelson and over 800 seminar attendees for useful suggestions, Philippe Jehiel for posing a challenge, the
Rhoads Foster Professorship Endowment of the University of Missouri, the Freshwater Group Research Fund of the University of Arizona, CNRS and INRA for financial support of the experiments, and ISER at Osaka University for support during initial writing.

APPENDIX A. AFTERMARKET EXAMPLES

A.1. TWO-UNIT AUCTION

Two identical assets are simultaneously auctioned to N bidders, each of whom values a second unit less than the first unit. After the auction:

[A] When one bidder acquired both units, in the aftermarket he would be asked to click on a price that was the lowest price at which he would be willing to sell his second acquired unit (his “Accept 2nd Ask”), and then click again at the (no smaller) lowest price at which he would be willing to sell his first unit under the assumption that the second unit will have been sold in the aftermarket (his “Accept 1st Ask”). Each losing bidder would be asked to click on the highest price at which he would be willing to buy the acquirer’s first unit as his second unit to acquire (his “Accept 2nd Bid”) under the assumption that he would already have successfully acquired a unit in the aftermarket, and then click again at the (no smaller) highest price at which he would be willing to buy a unit as his first unit to acquire (his “Accept 1st Bid”).

Aftermarket outcome rules would specify the following: [i] If the Accept 2nd Ask exceeded all Accept 1st Bids, there would be no transaction (the only behaviorally efficient aftermarket outcome). [ii] Or if any combination of at least five Accept Bids (combining 1st’s and 2nd’s) exceeded the Accept 1st Ask, the two highest Accept Bids would each acquire a unit, both priced at the highest still-unfilled Accept 1st Bid. [iii] Or if the second-highest Accept 1st Bid exceeds the Accept 2nd Ask, the highest Accept 1st Bid acquires a unit, priced at a randomly chosen Accept 1st Bid that falls between the highest Accept 1st Bid and the Accept 2nd Ask; following this, [iii.a] if the second-highest among \( C = \{ \text{the acquiring bidder’s Accept 2nd Bid and the still-unfilled Accept 1st Bids} \} \) exceeds the Accept 1st Ask, the highest in \( C \) buys the auction winner’s remaining asset, priced at the second-highest element of \( C \); [iii.b] if exactly one element of \( C \) exceeds the Accept 1st Ask, that bidder buys the winning bidder’s remaining unit at an exogenous random price iff the price would be acceptable to both given the prices they accept; [iii.c] otherwise the second unit is not transferred in the aftermarket. [iv] The remaining possibility is that only the highest of the
remaining Accept 1st Bids exceeds the Accept 2nd Ask; in this case one unit is transferred at an exogenous random price iff the price would be acceptable to both given the prices they accept.

\(A.2. \text{Values-externalities Auction}\)

Consider the following situation, simplified while still addressing the major complication. Bidders \(A, B, C\) and \(D\) compete for an indivisible asset. \(A\) and \(B\) are competitors in the same industry; if \(C\) or \(D\) wins, the asset will be used in a way that does not change competitive conditions in that industry. As in section 3, bidders \(j = a, b, c, d\) have both private-value \(X_j\) and common-value components \(Y_j\) to their valuation of the asset.\(^{51}\) In addition, bidders \(j = a, b\) each value the outcome where either bidder \(C\) or \(D\) wins higher than where the in-industry competitor wins by an amount \(Z_j > 0; Z_a\) and \(Z_b\) may not be equal (these are the values externalities). A losing \(C\) or \(D\) is indifferent over which rival bidder wins. The aftermarket designer is aware of \(C\)'s and \(D\)'s indifference and aware that there are positive amounts \(Z_a, Z_b\), but does not know the amounts. Assume the aftermarket designer knows some \(L < \infty\) that bounds the \(Z_j\)'s and absolute values of all possible values of all four bidders for all possible outcomes of the aftermarket.

To handle the most complicated case, let \(A\) win the auction.\(^{52}\) In the aftermarket, a price clock ticks up on all bidders' screens. Bidder \(A\) has two buttons: he is asked to click the yellow button at the lowest price at which he would sell to bidder \(C\) or \(D\), and the red button at the lowest price he would sell to bidder \(B\). Each of \(B, C\) and \(D\) is asked to click on the highest price he is willing to pay to buy the asset from \(A\). On a separate price clock, starting from 0, \(B\) is asked to click on the highest amount that he is willing to pay in order to facilitate transferring the asset from \(A\) to \(C\). Let \(S_B\) and \(S_{CD}\) be the minimum prices \(A\) states to be acceptable to sell to \(B\), and to \(C\) or \(D\); let \(G_B, G_C\) and \(G_D\) be the maximum prices \(B, C\) and \(D\) state to be acceptable to buy from \(A\), and \(F_A\) be the maximum \(B\) indicates he is willing to pay in order to have the asset transferred from \(A\) to \(C\) or \(D\).

Aftermarket rules can then be established as follows:

\(Case I:\) There is only one transfer that would be mutually beneficial. For a transfer from \(A\) to \(C\) or \(D\), the aftermarket incorporates the feature that \(B\) has indicated willingness to pay up to \(F_A\) to see this transfer happen, in determining whether it is mutually beneficial or not.

\(^{51}\) The private- and common-value components may be based on separate, independent estimates or knowledge.

\(^{52}\) The aftermarket if \(B\) wins corresponds exactly, that if \(C\) or \(D\) wins is similar.
Case I.1: A non-transacting losing bidder stated a willingness-to-pay to buy from \( A \) that is acceptable to both parties of the mutually beneficial transfer. I can use that stated price for the transfer price. (If the aftermarket recipient is \( C \) or \( D \), that \( B \) would have been willing to pay to make this transfer happen is in this case ignored.)

Case 1.2: There is no such price. Then draw random numbers from exogenous distributions set in advance: \( P \) from a superset of the range of mutually acceptable prices, and \( Q \) from \([0, K]\) where \( K \) is sure to exceed \( F_A \). However, if the potential aftermarket buyer is \( B \), then set \( Q = 0 \). Now the aftermarket transfer occurs if and only if: \{ \[i\] \( P \) is acceptable to the recipient of the asset, \[ii\] \( Q \leq F_A \), and \[iii\] \( P + Q \) is an acceptable price for \( A \) to sell \}.\(^{53}\) Otherwise, despite observing that a mutually gainful potential transfer exists, it is not implemented. Because \( Q \) and not \( F_A \) affected the transfer price, \( F_A \) only affected whether the transfer took place (so possibly not implementing is crucial to giving \( B \) the correct incentives).

Case II: both \( C \) and \( D \) announced accept prices that, given \( F_A \), imply a mutually beneficial transfer (i.e., either could willingly pay a price that is within \( F_A \) of \( A \)'s willingness-to-accept, \( S_{CD} \)), while there is no price at which \( B \) can buy from \( A \) for mutual benefit. While that implies \( G_B < S_B \), it does not imply \( G_B \leq S_{CD} \). Without loss of generality, assume \( C \) outbid \( D \) in the aftermarket. The aftermarket transfers the good from \( A \) to \( C \), or makes no transfer.

Case II.1: \( G_B \leq S_{CD} \). If \( G_D > S_{CD} \), the transaction is simply that \( C \) pays \( A \) the price \( G_D \) and the asset is transferred. (Again, ignore the feature that \( B \) gains from the transaction.) If \( G_D < S_{CD} \), then if \( G_D + Q < S_{CD} \) or \( Q > F_A \), there is no transaction; else \( C \) pays \( G_D \) to \( A \) and \( B \) pays \( Q \) to \( A \).

Case II.2: \( G_B > S_{CD} \). \[a.1\] If \( G_B > G_C \), then \( a.1 \) if \( G_D > S_{CD} \), \( C \) buys from \( A \) at price \( G_D \); \[a.2\] if \( G_D \leq S_{CD} \), if \( G_D + Q < S_{CD} \) or \( Q > F_A \), there is no transaction; else \( C \) pays \( G_D \) to \( A \) and \( B \) pays \( Q \) to \( A \). \[b.1\] If \( G_B \leq G_C \), then \( b.1 \) if \( G_C > G_D + Q > G_B \) and \( Q < F_A \), then \( C \) pays \( G_D \) to \( A \) and \( B \) pays \( Q \) to \( A \); \[b.2\] otherwise, \( C \) pays \( G_B \) to \( A \) and \( B \) pays nothing.

Case III: Announced accept prices imply a mutually beneficial transaction from \( A \) to either of \( B \) and one of \( C, D \)—without loss of generality, let it be \( C \). Denote the revealed (i.e., from accept prices) gains from trade \( H_B = G_B - S_B \) and \( H_C = G_C + F_A - S_{CD} \), let \( H = H_B + H_C \).

Consider first possibilities for transacting from \( A \) to \( B \). If \( S_B \leq G_C \leq G_B \), then label it “possible for \( B \) to buy at price \( G_C \)” let \( E_{BG} \) denote this event. If not, draw a random number \( P_B \) from an exogenous distribution designed in advance to be a superset of the range

\(^{53}\) Though not necessary, the transfer happens to be budget-balancing: \( B \) pays \( Q \), the recipient pays \( P \), \( A \) receives \( P + Q \).
of mutually acceptable prices. If $S_B \leq P_B \leq G_B$, then label it “possible for $B$ to buy at price $P_B$” let $E_{BP}$ denote this event. Let $B_{pos}$ denote that it is possible to transfer to $B$, via $E_{BG}$ if possible, and otherwise via $E_{BP}$. The remaining event, “not $E_{BG}$” together with $P_B$ falling outside $[S_B, G_B]$, is denoted $B_{imposs}$.

Consider next possibilities for transacting from $A$ to $C$. If $S_{CD} \leq G_B \leq G_C$, then label it “possible for $C$ to buy at price $G_B$;” let $E_{CG}$ denote this event. If not, but $S_{CD} - Q \leq P_B \leq G_C$ and $Q \leq F_A$, then label it “possible for $C$ to buy at price $P_B$, and $B$ pay $Q$;” let $E_{CPQ}$ denote this event. If neither of these is possible, draw a random number $P_C$ from an exogenous distribution designed in advance to be a superset of the range of mutually acceptable prices. If $S_{CD} \leq P_C \leq G_C$, then label it “possible for $C$ to buy at price $P_C$;” let $E_{CP}$ denote this event.

Then, in event $B_{imposs} \cap C_{imposs}$, there is no aftermarket transaction. In event $B_{pos} \cap C_{pos}$, the asset is transferred from $A$ to $B$ via the first of the possibilities contained in this event (i.e., via $G_C$ if possible, otherwise $P_C$). In event $B_{imposs} \cap C_{pos}$, the asset is transferred from $A$ to $C$ via the first of the possibilities in the order defined. That leaves event $B_{pos} \cap C_{imposs}$. In this event, a random number $J$ is drawn uniformly on $[0,1]$.

If $J < H_B/H$, the aftermarket sells to $B$, via $G_C$ if possible, otherwise $P_C$. If $J > H_B/H$, the aftermarket sells to $C$ via the earliest possibility in the above order.

Remark: The critical element for a proper aftermarket to be designable is that in event $B_{pos} \cap C_{pos}$, with probability up to $\frac{1}{2}$, the aftermarket transacts a mutually beneficial trade despite the fact that another trade was identified with a greater gain from trade, and the greater-gain transaction was not implemented.

Case IV: all three losing bidders indicate mutually beneficial trades. Without loss of generality, assume $G_D < G_C$. Now transacting to $D$ is ignored, and the procedure basically follows case III, with the minor added complication that each time $P_B$ [P_C] is considered as a price for $C$ [B] to pay, $G_D$ will instead be used if it is mutually acceptable. Case IV also may conduct a transaction when another shows a larger gain from trade, and may indeed transact to $B$ with positive probability when both transactions to $C$ and to $D$ show larger gains from trade.
Case V: if no transaction shows a gain from trade, none is made. (As in the paper, this is the only case found to be behaviorally efficient.)

Note that no player’s behavior in the aftermarket determines the price at which he transacts, nor does B's behavior determine the payment he might make to attain a transfer to C or D. For any player to deviate from truthful revelation either has no impact or works to his detriment.\textsuperscript{54}

A.3. No-deficit Public-good Aftermarket

An alternative aftermarket that does not run a deficit could be designed as follows, for \(N\) subjects and a public-good unit production cost \(C\). Each period, draw \(N - 1\) random numbers from a uniform distribution on \([0,1]\). Let \(w = (w_0, w_1, w_2, \ldots, w_{N-1}, w_N)\) be defined by \(w_0 = 0, w_N = 1\), with the \(w_i\)'s in between being the \(N - 1\) random draws, ordered ascendingly. Then set \(f_j = w_j - w_{j-1}\), \(j = 1, \ldots, N\), cost shares summing to 1.

An Accept price \(P_j\) is obtained from each subject \(j\) just as in section 6.2, let \(V = \text{the sum of the N P's.}\) This aftermarket increases public-good output by one unit if and only if the following requirements are met: [i] \(V \geq C\), and [ii] \(P_j \geq f_j C\), all \(j = 1, \ldots, N\). If so, each subject \(j\) pays \(f_j C\) for the unit increase, exactly covering production cost, and indicated to be acceptable to each subject. If requirement [i] is met, but [ii] is not, the aftermarket uncovers and measures \((V - C)\) a mutually beneficial reallocation, but does not implement it.

Appendix B: Session Protocol

Each experimental session ran 150 minutes and followed a multi-phase protocol, to build the desired treatment step-by-step from simpler games. After instructions regarding the whole session and the first phase, that first phase exposed subjects to the software of the aftermarket, without introducing the word “aftermarket.” In phase 1 (4-5 periods), each subject was informed of a list of all five private values of the abstract asset (told which was his value), which were drawn i.i.d. uniform on \([\$5, \$10]\). Per instructions, one subject was chosen at random to be the acquirer, the others bidders.\textsuperscript{55} As just described, the acquirer was asked to click on an Accept Ask, the four bidders to click on Accept Bids. Then the aftermarket rules above were used to determine payoffs for the period, which were simply

\textsuperscript{54} This construction responds to an assertion by Philippe Jehiel (whom I thank for the challenge) that it is impossible to construct an aftermarket in such cases.

\textsuperscript{55} To generate possible gains from trade frequently, the program chose the acquirer from the highest, second-highest, \ldots, lowest private values with probabilities \(\{1/8, 1/8, 1/4, 1/4, 1/4\}\).
asset value minus transaction price for the buyer, and transaction price minus asset value for the seller, if there was a transaction, and zero for all non-transacting subjects.

Further instructions were distributed and read before each following phase. Phase 2 (6-7 periods) introduced private information, with subjects’ private values not revealed to all group members until end-of-period feedback (anonymously, as always). Phase 3 (6-7 periods) introduced two changes: [i] all five subjects were now bidders asked to select Accept Bids (that is, in a “closed-clock” variant of an English auction56), and [ii] the private values were now affiliated (as in section 3, except that \( V_j = X_j \)). Phase 4 (6-7 periods) set aside the price clock, introducing bidding in a sealed-bid auction (first- or second-price, depending on the session). Phase 5 (8-11 periods) introduced affiliated values, via \( V_j = (3/4)X_j + (1/4)C \) as in section 3.

All this led to the phase of principal interest, phase 6, which re-introduced the software from the first two phases, but with the acquirer being the bidder who acquired the asset in the sealed-bid auction, and the following price-clock activity now called an aftermarket. Phase 6 was generally limited by the time constraint, 6-11 periods. The session ran faster when there were fewer groups (the software always waits for the last subject in the session to bid, to peruse feedback, etc.); in four of the sessions, we were able to run a final phase 7. Phase 7 had aftermarkets only in even-numbered periods, with the sealed-bid auction the final determination of period profits in odd-numbered periods. In the other nine sessions, phase 6 was the final phase.

**APPENDIX C: DISTINCT MEASURES?**

What can be said about how well behavioral efficiency concurs with the conventional theoretical definition of allocative efficiency? As the auction experiments used induced values, they can yield insights into the differences between these measures. That is, critically assume subjects are all risk-neutral and induced motivations are complete (in particular, assume away interdependent preferences, nonpecuniary preferences, any innate relevant personality differences such as aggressiveness, and satiation in cash). Then in a Pareto-efficient allocation, the asset is acquired by the subject with the highest asset-value estimate.

A first-price [second-price] auction reached an induced-values Pareto-efficient allocation in 90% [92%] of observations. In most of these observations, behavioral efficiency was also

---

56 The closed-clock variant provides no feedback about rival bidders’ behavior while the clock ticks.
observed in the aftermarket, and the inverse: Pareto-inefficient auctions usually exhibited behaviorally inefficient outcomes, mutual gains observed in the aftermarkets.

To track this degree of concurrence, define the induced-values shortfall to be \{maximum less realized induced gains from trade\}, that is, the highest induced asset value minus the auction winner’s induced asset value. In first-price [second-price] auctions, this averages $0.408 [0.302]. This might be considered the theoretical measure of the size of inefficiencies.

Treating a shortfall of at most $0.50 (one price clock tick in the aftermarket) as small, the following degree of concurrence between induced-values shortfalls and behavioral efficiency shortfalls is seen. In first-price [second-price] auctions, in 58% [50%] of 203 [142] observations, both measures show efficiency (a $0 shortfall), in 64.5% [59.9%] of observations, both measures show at most a small shortfall; 11.3% [11.3%] of observations show more-than-small induced-value shortfalls but at-most-small behavioral shortfalls; 15.3% [20.4%] of observations show at-most-small induced-value shortfalls but more-than-small behavioral shortfalls; and in 8.9% [8.5%] of observations, both measures show more-than-small shortfalls.\(^{57}\) This degree of concurrence I consider strong enough to suggest usefulness of behavioral efficiency as a normative criterion for evaluation of the first-price and second-price auction formats.

The two distinctions from concurrence were both observed in minorities of the observations. [i] In 15% of first-price auctions and 24% of second-price auctions, the induced-values efficient acquirer was the high bidder, so the outcome could be assumed Pareto-efficient, yet the aftermarket found a mutual gain could arise from a transfer to an outbid rival with a lower induced estimate of asset value (this includes over a dozen observations where the second-price shortfall was small but positive). [ii] For both first-price and second-price auctions, 16% of observations found an induced-values inefficient acquirer submitting the high bid and then clicking on an Accept Ask that exceeded all Accept Bids, including that selected by the (outbid) induced-values efficient acquirer whose estimate exceeded his (again, this includes dozens of observations where the induced-values shortfall was small but positive, some a few pennies).

While some variant of an endowment effect might lead to the second way in which behavioral efficiency has been found distinct from Pareto efficiency, it bears notice that

\(^{57}\) Nonetheless, the simple correlation between these two measures is significantly positive but not high: 0.158 [0.129] for first-price [second-price] auctions. Presumably, the induced-values measure being a multiple of $0.01 while the behavioral measure is a multiple of $0.50 has a lot to do with the low correlation.
distinction [i], as prevalent or more so, is essentially the reverse of an endowment effect. More importantly, being able to observe which auction outcomes are Pareto efficient and thus observe these distinctions depends on having induced values and assumed motivational completeness and risk neutrality. Using aftermarkets to observe the size and frequency of shortfalls from behavioral efficiency requires none of these.

**APPENDIX D: STRATEGIC TRANSPARENCY**

Following the public-good demonstration, since the subjects were in an experimental laboratory, a simple induced-value laboratory phase was added. Subjects were given instructions about the allocation of an abstract public good. The only value of this public good was monetary utility to each individual subject that had been specified by the experimenter.58

To mimic the field demonstration, an ad hoc mechanism that was suppressed set initial public good output to 7 units, and each group was asked whether to increase output to 8 units, at an incremental production cost of €3. Each subject was privately told the incremental value $v_i$ to her or him of the increase from 7 to 8 units; the distribution of these incremental values was not announced, although it was announced that the incremental values were not all the same. These six values summed to less than the incremental cost (a random decision, as was the 7-unit starting point); one randomly chosen subject had an incremental valuation equal to €0.7, one equal to €0.06, four equal to €0.42 (that four had the same incremental value was not known to the subjects until results were reported). Division of subjects into groups was via a new random draw, independent of the draw in the field experiment; this feature was announced.

In all other respects, the induced-value procedure was identical to that of the field demonstration: a clock ticked up on all screens (by a multiple of €0.05), subjects were asked to click “Accept” at the highest price they were willing to pay to increase public good output from 7 to 8 units (their incremental value was shown on the screen as the price ticked up), and were told beforehand that an individual group member’s personal cost of this increase, which would happen if and only if the sum of Accept prices were at least the production cost, would be the €3 production cost less the sum of the Accept prices of the other five group members. It was carefully explained that their payoff for this decision would be zero if the amount of public good were not changed, and would be their incremental value less

---

58 This assumes selfish preferences, a limitation the field demonstration does not share.
the excess of production cost over the sum of the other five Accept prices if public good output were increased.\footnote{The sessions followed these observations with pilot experiments studying public-good allocation mechanisms that did not bear on the issues of this paper.}

As before, it is a dominant strategy to set one’s Accept price equal to the multiple of €0.05 closest to one’s incremental value. Only with incremental values induced (thus in the lab, not the field), is it possible to see whether subjects did this. Most did not exactly hit this dominant strategy, although on average the statistic $Z = (\text{Accept price} - \text{incremental value})$ was €0.00103, remarkably close to the −€0.015 average it would have been had every subject exactly adopted the dominant strategy.\footnote{Because the Accept bid had to be a multiple of €0.05, in each group the one subject with incremental valuation €0.06 had a dominant strategy $Z = -€0.01$, and the four subjects with incremental valuation €0.42 had a dominant strategy $Z = -€0.02$.} The standard error of $Z$ was not large, €0.298; the dominant strategy for 23 subjects for whom $v_j = €0.06$ was to click on Accept immediately (at €0.05); they were clicking later than this, and averaged $Z = €0.38$. This was compensated for by the 23 subjects for whom $v_j = €0.7$, perhaps not waiting for the price clock to reach that high; they averaged $Z = -€0.32$. Still, over 80% of 138 subjects were within €0.3 of $Z = 0$, nearly half of those within €0.1. For all twenty-three groups, the six Accept prices summed to less than the €3 production cost, so for all groups, seven units of public good was behaviorally efficient relative to the alternative of eight units. Evidence in the induced-value setting for inexperienced subjects to be unable to understand the incentives they faced in the field demonstration is unpersuasive, indeed quite limited.

APPENDIX E. CONTRASTING PREDICTIONS

Auction theory predicts aftermarket activity with positive probability following first-price auctions, but with zero probability following second-price auctions. It is straightforward to show that the unique risk-neutral, symmetric Bayesian equilibrium of auction-cum-aftermarket (either first- or second-price) is to submit one’s equilibrium bid in the auction and to select an Accept Bid or Ask in the aftermarket most nearly equal to one’s rational Bayesian-updated willingness-to-pay or -accept.
In this equilibrium, publicly announcing the price attained in a first-price auction informs each losing bidder (but not the winner) of the amount by which his bid lost. Whenever a bidder lost by a sufficiently small margin, rational updating leads to his willingness-to-pay exceeding the winning bidder’s willingness-to-accept (as he knows of a second estimate nearly as high as the winning bidder’s estimate, which can be inferred from the price set by the winner’s monotonic equilibrium bid function).  

No similar occurrence is possible following announcement of the price in second-price auctions. Here the price reveals the private information of the second-highest bidder, who Bayesian updates on the basis of learning that one rival estimate was higher and three lower, and this leads to a willingness-to-pay that exceeds his equilibrium bid, while the winning bidder’s updating leads to a willingness-to-accept that is less than his equilibrium bid. However, the second-price equilibrium is envy-free: these two adjustments of willingness-to-pay and -accept sum to less than the difference between the two highest bids, not changing the ordinal ranks.

To my knowledge, prior auction experiments have either induced private values (independent, as in [Appendix B] phases 1 and 2, or affiliated, as in phases 3 and 4) or a common value (modifying section 3 so that $V_i = C$, hence there is no inefficiency generated should the bidder with the highest estimate be outbid). Nonetheless, in both settings, bidders have bid significantly above the risk-neutral symmetric Bayesian equilibrium (Kagel, 1995), and (more pertinent here) have exhibited more heterogeneity in this overbidding in second-price than in first-price.

---

61 It would be awkward, but probably feasible in the lab, to end the auction with a private announcement to each bidder whether he won or lost, but withhold the auction price until the aftermarket finishes. To do so would remove this theoretical aspect of first-price auctions, but divert the experimental auction from any real-world auction or any experience with which subjects might have some familiarity.

62 Though stated slightly differently, the results in this and the previous paragraph are not new, and can be pieced together from Milgrom (1981), Milgrom and Weber (1982) and Harstad and Bordley (1996).
auctions. Thus, prior laboratory experiment results predict more aftermarket activity following second-price than first-price auctions.

REFERENCES


---

63 Kagel and Levin (1986) report on 199 first-price, common-value auctions, and Kagel, Levin and Harstad (1995) on 154 second-price, common-value auctions. To adjust for varying number of bidders, I calculated a statistic for each session that takes the frequency with which the high signal holder was the high bidder and subtracts 1/n. The weighted (by number of auctions) average of these statistics was 50.93 for first-price auctions and 38.73 for second-price auctions.


