Discussion Paper No. 818

BEHAVIORAL EFFICIENCY I: DEFINITION, METHODOLOGY AND DEMONSTRATION

Ronald M. Harstad

October 2011

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Behavioral Efficiency I: Definition, Methodology and Demonstration

Ronald M. Harstad*
Economics, University of Missouri
harstad@missouri.edu
ISER, Osaka University

September 4, 2011
Please consult before citing or circulating

Abstract

Economic experiments conducted in laboratories employing an induced-values methodology can report on allocative efficiencies observed. This methodology is limited by requiring the experimenter to know subjects' motivations, an impossibility in field experiments. Allocative efficiency implies a hypothetical costless aftermarket would be inactive. An outcome of an allocation mechanism is herein defined to be *behaviorally efficient* if an appropriate aftermarket is actually appended to the allocation mechanism and at most a negligible aggregate size of mutually beneficial gains is observed on the aftermarket. Methodological requirements for observation of behavioral efficiency or inefficiency are put forward. A simple field demonstration indicates when an increase in public good output can cover marginal cost in a mutually beneficial decentralization, without knowing valuations. Several empirical issues that arise with the methodology are noted.

C9; C93; D01; D61; D03; D46; Keywords: behavioral efficiency, field experiment methodology, allocative efficiency, revelation of valuations, aftermarkets

* I thank, without implicating, Marc Willinger for organizing experimental facilities and arrangements, Dimitri Dubois for assistance in conducting experiments, Dubois and Rafaele Preget for translating instructions, Willinger and Sophie Thoyer for organizing financial support, David Levine, Alvin Roth and Cathleen Johnson for useful suggestions, LAMETA at Universite de Montpellier for providing experimental facilities, and CNRS and INRA for financial support of the experiment, and ISER at Osaka University for support during the writing.

Behavioral Efficiency I: Definition, Methodology and Demonstration

Abstract

Economic experiments conducted in laboratories employing an induced-values methodology can report on allocative efficiencies observed. This methodology is limited by requiring the experimenter to know subjects' motivations, an impossibility in field experiments. Allocative efficiency implies a hypothetical costless aftermarket would be inactive. An outcome of an allocation mechanism is herein defined to be *behaviorally efficient* if an appropriate aftermarket is actually appended to the allocation mechanism and at most a negligible aggregate size of mutually beneficial gains is observed on the aftermarket. Methodological requirements for observation of behavioral efficiency or inefficiency are put forward. A simple field demonstration indicates when an increase in public good output can cover marginal cost in a mutually beneficial decentralization, without knowing valuations. Several empirical issues that arise with the methodology are noted.

C9; C93; D01; D61; D03; D46; Keywords: behavioral efficiency, field experiment methodology, allocative efficiency, revelation of valuations, aftermarkets

"My research agenda focuses on using field experiments to learn what social policies work, what do not, and why."[1]

--Dean Karlan, Yale University

1. Introduction

I begin by quoting Karlan, not as a straw horse, but as a definitive statement by a noted economist who has already contributed within this agenda. The verb "work" in this quote admits a broad interpretation, including possibly impacts of a policy on median income, income inequality, income security, intergenerational mobility and related opportunities, economic growth, educational attainment, as well as narrower impacts on family size, labor force participation, and many others. Whatever the list a particular researcher considers, to an economist an essential component of whether a policy "works" is how small or large a shortfall from allocative efficiency it attains.

Efficiency measures are commonplace in economic laboratory experimentation.[2] The restriction to laboratory experiments is critical, as the methodology allowing observations of the shortfall from allocative efficiency, *induced values*, can only be employed in a laboratory setting, and cannot help Karlan and other scholars observe "what social policies work … and why" in the field.

In many hundreds of laboratory experiments, financially motivated subjects transact abstract commodities (perhaps called X or "triangle" but not "tennis lessons" or "Big Macs") on markets observed by the experimenter. Absent any innate values, valuations of X

[1] Beginning text of his homepage, http://karlan.yale.edu/, visited 4/16/11.

[2] Cf. e.g., Smith [1962], [1976], Davis and Holt [1993], Kagel and Roth [1995], and original sources cited in the latter two works.

are *induced* by the experimenter, as "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." When transactions costs can be deemed negligible, inducing values allows the experimenter to calculate the set of Pareto efficient allocations, and attach a cardinal measurement to any observed shortfall from this set.

No corresponding measure is available in field experiments, where transactions might involve (usually subsidized) provision and/or allocation of: irrigation water, adult education, childcare, pollution permits, microfinance, insurance against background risk, or similar "naturally occurring" goods, services and contracts. In "framed" field experiments, these transactions occur on a market constructed and controlled by the experimenter; in "natural" field experiments, the experimenter observes but cannot control a naturally occurring market.[3] The experimenter can observe transacting behavior but not valuations or motivations, and so cannot calculate the Pareto set nor measure shortfalls from it.

This paper proposes a definition and a specific but broadly usable methodology to allow observations relevant to allocative efficiency in field experiments, which observe only behaviors that stem from unobserved motivations and preferences (and sometimes even incompletely observed feasibility constraints). Theoretical issues arising with this methodology are described. Following that, I provide a first field demonstration, observing whether an increase in output of a public good from an ad hoc starting point can be achieved as a mutually beneficial reallocation. Then several empirical issues that arise in the consideration of this methodology are discussed. A companion paper offers a concrete demonstration of the methodology in the simplest possible environment—an economy with a single, indivisible good—and illustrates how laboratory- and field- compatible efficiency measures might be distinct.

Among the finest examples of how far field experiments have been able to go in the direction of inferring efficiency conclusions from observations is Bohm [1984]. Peter Bohm convinced the Swedish federal government to let him control whether an indivisible public good would be produced or not, an office that would collect and provide certain statistical data to local governments. Randomly splitting the local governments into two groups, he announced rules to one group that gave them an incentive to understate willingness-to-pay (WTP), and to the other an incentive to overstate WTP. The observed sum of stated WTPs slightly exceeded cost, and mean WTP per capita was not markedly different between the two groups. Nothing Bohm could observe, however, would let him infer whether providing the statistical office would yield an efficiency gain.

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) Over a quarter century later, the methodology proposed herein seeks to move beyond (or back from?) science fiction.

[3] I am defining these terms slightly differently than Harrison and List [2004], as the issue of experimenter control seems more critical than whether subjects are aware they are in an experiment.

1.1. Pareto Efficiency Reinterpreted

In an economy with I individuals and C commodities, let x_i be the C-dimensional allocation to individual i, u_i his utility function, and $x = (x_1, \ldots, x_I)$. Then the usual definition of a Pareto-efficient allocation is that it satisfies

$$(P): \quad \text{Max}_x \; u_1(x), \text{ subject to: } u_j(x) \geq \underline{u}_j, \text{ all } j = 2,\ldots, I, \text{ and to feasibility conditions.}$$

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

This reinterpretation is informationally dissimilar: the maximization in (P) is clearly tied to knowledge of motivations and valuations (u_i, after all), while the counterfactual aftermarket is tied to hypothetical transactions (that is, to hypothetical behaviors). In principle, as transactions can be observed in field experiments, these experiments might avail themselves of aftermarkets.

2. A Definition

So I define *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 allocation mechanism[4] and at most a negligible aggregate size of mutually beneficial gains is observed on the aftermarket. Natural extensions of the definition 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 *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.

3. Aftermarket Methodology

The aftermarket referred to in this definition must be designed and implemented so as to support the intended normative interpretation. It likely aids first to set aside 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 (an illustration is in section 4).[5] [ii] Whatever its structure, a *resale* market (Zheng [2002]) does not suffice. A key terminological distinction: unlike resale markets (e.g., for US Treasury debt), an *aftermarket* necessarily involves *the same* economic *actors* as the original market (original allocation mechanism), none added and none absent.[6] Imagine $100K in 5-year T-notes is sold today, by their purchaser at a Treasury auction on the third Tuesday of the month before last, to a regional bank that did not compete in that particular Treasury auction. Today's resale in no way implies an inefficiency in the allocation that resulted from

[4] Any method, however informal, of reaching an allocation is herein labeled an allocation mechanism.

[5] Were repeating the same mechanism to suffice in some special circumstance, likely it would create needless confusion for subjects.

[6] This definition is clearly implied in the half-century-old reinterpretation of Pareto efficiency noted in section 1.1 above.

that auction. [iii] Later transactions involving an informationally distinct commodity cannot support interpretation of an earlier allocation as inefficient. 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 after the auction, the consigner of this particular car agrees to allow the winning bidder to return the car and be given 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-auction, leads to a realization that a quality issue unsuspected as of the original auction has since surfaced, thus to a different commodity being allocated at re-auction.

To shed light on behavioral efficiency, the aftermarket must be constructed so as to *identify* 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.[7] This requires *revelation*: that subjects' behavior in the aftermarket can be interpreted as revealing the border between potential transactions they prefer to make and prefer not to make, thus as revealing all relevant willingnesses-to-pay and willingnesses-to-accept.[8] 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. As formalized below, the aftermarket should be designed in such a way that a posited equilibrium behavior in the original allocation mechanism, together with truthful revelation in the aftermarket, constitute a perfect Bayesian equilibrium of the mechanism-cum-aftermarket game.

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 lab or field experiment, not to determine an allocation but merely to normatively categorize the allocation reached *before* the aftermarket is used. The simplification thus obtained is characterized in section 4.

This bears emphasis: for the purpose of behavioral efficiency interpretation, an aftermarket *does not* have to be an allocation mechanism.

That the aftermarket be approximately free of transactions costs, and that it be suitably transparent, necessarily have less exacting interpretations; these are arenas where employment of aftermarkets shifts from science, narrowly construed, in the direction of art.

---

7 I follow a century-old tradition in welfare economics that externalities to a transaction are either explicitly modeled or ignored. Thus, if allocations *A* and *B* differ only in that in *A*, a cocktail dress remains in company C's inventory, while in *B*, Ginger buys the dress from company C at a mutually beneficial price, then allocation *B* is considered Pareto-superior to *A*. This ignores the possibility that Ginger might later attend a party where seeing the new dress makes Rosemary less happy with her wardrobe. In contrast, where a field study is focused on externalities, the aftermarket must observe all possible disutilities resulting from recontracting, not just the transactors' evaluations.

8 I know of only two antecedent reports of 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 original and post-aftermarket efficiency.

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 all mutually beneficial transactions for which the perceived gains from trade exceed the perceived transactions costs.

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.[9]

Experimental psychology and laboratory experimental economic literatures yield insights into transparency that are extensive, although often anecdotal and always subjective.[10]

3.1. Required Nature of Potential Aftermarket Transactions

A useful aftermarket needs to observe behavior with respect to any potential transaction that may be mutually beneficial; this need will vary with the topics 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. 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.

Another important example arises when the original allocation mechanism determines a quantity of a public good to be produced, and an allocation of its production cost. 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, as seen in section 5.

Some field studies will require the aftermarket be designed so as to observe valuations of dynamically structured contracts. Allocations of common-pool resources are examples.

The usual characterization of efficiency via (P) above 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 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

---

[9] It is important that an aftermarket be constructed so that a subject engages observable behaviors whatever degree of satisfaction they have with the original allocation, so as to avoid "active participation hypothesis" concerns (Lei, Noussair and Plott [2001]).

[10] That price-clock-based ascending-price mechanisms are notably more transparent than sealed-bid mechanisms seems a reasonable inference to draw from the laboratory experiments reported in Harstad [2000]. For this reason, such ascending-price mechanisms are used both in section 5's demonstration and in the companion paper.

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.[11]

4. Theoretical Issues Raised

Suppose the allocation mechanism being studied is sufficiently formal to permit analysis as a game G.[12] Let $\mathbf{G}$ denote the game consisting of G followed by the aftermarket $G+$ that is being constructed to test behavioral efficiency of G. Then the aftermarket must be constructed so that $\mathbf{E} = \{E, E+\}$ is a perfect Bayesian equilibrium of $\mathbf{G}$, where E is the equilibrium usually focal in literature considering G, and $E+$ constitutes truthful revelation in $G+$. Otherwise, a behavioral efficiency (or inefficiency) conclusion is unwarranted.

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 two critical dimensions of flexibility generally unavailable in mechanism design: the field experiment [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.[13]

As a simple illustration, suppose a field experiment has observed a failure to reach an agreeable transaction in a bilateral bargaining situation. As motivations are unobserved, it is unknown whether a mutually beneficial bargain was possible. Myerson and Satterthwaite [1983] demonstrate that no mechanism can insure efficient outcomes when the potential seller's and potential buyer's valuations are private information. However, at least three distinct aftermarket constructions can observe whether the outcome was behaviorally efficient.[14] Each 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.[15] The experimenter has carefully explained in advance to the subjects what use will be made of their responses. Aftermarket version 1 will implement the transaction whenever B, her stated willingness-to-pay, exceeds A, his stated willingness-to-accept, with the pre-announced rule that she will pay A, and he

---

[11] Correspondingly, suppose an aftermarket appended to a mechanism allocating a given quantity of identical units of a private good were to observe that the 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 that 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 priced 2-unit (and perhaps 3-unit) trades as well as 1-unit trades, and allowed for multiunit trades to have multiple parties on the same side of the trade.

[12] This supposition is not trivial: some of the cultural incentive schemes discussed in Ostrom [1998] may be difficult to formalize as allocation mechanisms.

[13] 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.

[14] If the failure to reach a transaction occurred in a natural field experiment, the aftermarket would require a transition to a framed field experiment.

[15] 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.

will receive B. Version 1 is incentive-compatible, implements any transaction observed to be mutually beneficial, and requires the experimenter to cover the deficit $B - A$. Aftermarket version 2 draws a random variable R from a distribution exogenous to all information provided by this pair of subjects (perhaps uniform on [0.25 W, 1.75 W], where W is a publicly available average price from a prior survey of similar transactions in the economy), and transacts at random price R if $B \geq R \geq A$. Aftermarket version 2 is incentive-compatible, implements any transaction observed to be mutually beneficial *with positive probability*, and balances the budget. Aftermarket version 3 also draws a random variable R from an exogenous distribution, transacts if $B \geq R \geq A$, but she pays R and he receives 1.05 R, achieving incentive compatibility, implementing transactions observed to be mutually beneficial with positive probability, but requiring the experimenter cover a deficit of 0.05 R when transactions occur.

Note that mechanism design requirements can still impinge on experimental desiderata. In particular, consider an aftermarket construction which attempted to alter aftermarket 1 above by only transacting when the deficit $B - A$ did not exceed a maximum desired experimenter cash infusion M. This construction would no longer suffice for incentive compatibility: it is possible that the seller would attain the outcome of no transaction and no gain if he truthfully stated A, as $B - A$ might exceed M, while some overstatement $X \geq A$ 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 the buyer) would optimally trade off a lower gain in the event of transaction against a higher probability of a gain by some degree of overstatement. (Even if, instead of announcing M, the experimenter merely stated that the transaction would occur "unless the deficit were too large," the construction would still be insufficient to warrant conclusions as to behavioral efficiency.)

5. A Small Field Demonstration

5.1. FIELD CONTEXT: To focus the demonstration on the aftermarket, the initial allocation mechanism is submersed via an assumption that the outcome is production of one unit of public good, with the costs of the first unit's production covered from the experimenter's budget. The aftermarket then considers whether the cost of production of a second unit can be allocated to the perceived mutual benefit of all members of the economy. The aftermarket's construction does not balance the budget, allowing the experimenter to cover a deficit if it is observed that the sum of marginal benefits exceeds production cost of the second unit.

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. That is, [a] there was group exclusion but no individual exclusion in consumption, and [b] there was no rivalry in consumption. Either all subjects 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 studied consisted of six subjects. Eighteen subjects (in one session, twelve, due to no-shows) were in the room during a session, so that 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.

Subjects were seated at visually isolated computers. Instructions were passed out and read aloud, questions encouraged and answered.[16] 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 sum of the most each group member was willing to pay was at least €1.[17] 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).

5.2. METHODOLOGY IMPLEMENTATION: As mentioned, 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 "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 increase to two units.[18]

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 unit of candy, and each subject would pay €1 *minus* the sum of the other five Accept prices (or 0, whichever was larger).

This methodology implements the incremental version of the Vickrey-Clarke-Groves mechanism.[19] 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. Harstad [2000] has found the price-clock mechanism to be far more transparent than direct statement of willingness-to-pay in eliciting dominant strategy responses.

The incentive compatibility of this methodology warrants the conclusion that the group exhibits a behavioral inefficiency of the allocation—of one candy each—if the sum of Accept prices exceeds €1.

5.3. OBSERVATIONS: Thirteen of twenty-three groups exhibited a perceived mutual gain in increasing the distribution of candy from one unit to two units apiece. (This included six of the eleven groups that chose well before lunchtime, and seven of the twelve groups that chose shortly after 1:30 [or three after 3:30 pm], so there is no sign that chronobiology played a role.) The size of behaviorally revealed efficiency gains in these groups ranged from 4% to 72% of production cost (€1), averaging 33%.

---

[16] An English-language version of the instructions is available at http://harstad.missouri.edu/Instructs/.
[17] 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].
[18] The experiment was conducted in the Z-Tree programming environment (Fischbacher [2007]).
[19] Named for Vickrey [1961], Clarke [1971] and Groves [1973].

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

In none of the thirteen groups accomplishing 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 by the experimenter, but were revealed by behavior in the aftermarket (and the Vickrey-Clarke-Groves selection among those vectors actually used for payment).

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

Though there is evidence that it was transparent, 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 a private purchase of a second unit. Nor is it known whether any subjects were behaving altruistically, or the extent of any altruistic behavior. It is no more necessary to know 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 finds 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. Larger scale, more important field studies can exactly mimic the demonstration offered here, and relate behavioral efficiency observations to the methods used to determine levels of public good output.

5.4. LABORATORY OBSERVATION ON STRATEGIC TRANSPARENCY: Following the 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.[20]

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

[20] This assumes selfish preferences, a limitation the field experiment does not share.

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 the excess of production cost over the sum of the other five Accept prices if public good output were increased.[21]

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 = (Accept price – incremental value) was €0.00103, remarkably close to the €0.015 average it would have been had every subject exactly adopted the dominant strategy.[22] The standard error of Z was nonnegligible, €0.298; the 23 subjects for whom v_j = €0.06 had to click on Accept immediately (at €0.05) 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.

6. Remarks on Some of the Empirical Issues That May Arise

If there is a limit to the empirical questions that arise with behavioral efficiency, I haven't grasped it. Here's a sampling of those that are clear now.

Are Pareto-efficient allocations and behaviorally efficient outcomes necessarily distinct? Under the assumptions of the theory, the second paper in this series takes advantage of induced values to answer in the affirmative, finding a significant minority of observations that are either of these, but not the other. One side arises when the subject with the highest privately observed estimate of asset value wins the auction (as would occur in the symmetric, risk-neutral Bayesian equilibrium), thus attaining a Pareto efficient allocation, but nonetheless a mutual gain in the aftermarket is observed. The inverse also

[21] The sessions followed these observations with pilot experiments studying public-good allocation mechanisms that did not bear on the issues of this paper.

[22] 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 $Z = 0.02$.

arises when the auction winner is not the subject with the highest private information, but the aftermarket observes behavioral efficiency.

How large are the magnitudes of shortfalls from behavioral efficiency, and how might these be assessed? In each circumstance where I have been able to envision the outlines of an appropriate aftermarket design, any potential transaction perceived to be mutually beneficial that is observed provides an absolute magnitude of the perceived gain. To put this in percentage terms, as a shortfall from efficiency, requires being able to calculate the Pareto set. In field studies, the best hope is to observe behaviors in the original allocation that indicate the potential size of efficiency gains. The companion paper on auctions provides a treatment where such a comparison has a solid theoretical basis, and another with less foundation.

Does knowing that an aftermarket will follow an allocation mechanism affect subjects' behavior in the mechanism? A requirement for aftermarket design is that there is *in theory* no effect. This will be a potentially important question in every field context, and I expect to have to assess it *de novo*, at least until a large database of field aftermarkets has been compiled. It is possible to design demonstrations and adapt data analysis to shed some light on this. 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.

Will an aftermarket observe activity just because subjects assume they are supposed to do something? This question arose, for good reason, in reports of experiments observing financial bubbles in labs (Smith, Suchanek and Williams [1988]). 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, however. In aftermarkets, the 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 to do with a preference for the status quo that was reached in the original allocation mechanism.

Are aftermarket activities mistakes if the initial allocation reached should have been Pareto-efficient? In ordinary cases, it will be 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 assumptions that may well be unverifiable. For example, the Pareto efficiency characterization of the original allocation may be obtainable only under assumptions such as equilibrium behavior, risk neutrality, symmetric behavior, common knowledge of distributions of private information, selfish preferences, and others. In some situations, there may be serious questions about whether the aftermarket design was sufficiently transparent for the subject population, as for example when the subjects are illiterate. When transparency is adequate, I regard it as likely that aftermarket behaviors should be taken at face value.

Might behavioral efficiency determination depend on the structure of the aftermarket used for observation and identification? This will ever be an empirical possibility. When field budgets and subject population sizes permit, multiple aftermarket designs can be tested.

Might endowment effects lead to an inactive aftermarket even though the initial allocation was efficient? This could be imagined, although it is hard to say that inefficiencies exist, let alone identify and quantify them, when an appropriately constructed aftermarket observes inactivity. Plott and Zeiler [2005], [2007] demonstrate that the size of an endowment effect might almost be calibrable. Careful instructions (avoiding the term "winning bidder") in the companion paper on auction aftermarkets provide data strongly suggesting that the endowment effect is not a problem in that particular context.

Might other studied psychological biases and behavioral anomalies affect aftermarket observations? To a first approximation, an "anomaly" such as other-regarding preferences or hyperbolic discounting may equally impact both an original allocation mechanism and its aftermarket. For many such concerns, there is no reason to believe that they suddenly arise in aftermarkets following an allocation mechanism that went untouched by them. When there is evidence that some particular bias perceived to be relevant to a particular field study can reliably be redressed via education, it may well be best practice to educate first, then run the allocation mechanism followed by the aftermarket. There is considerable evidence suggesting persistence of some biases in the presence of education. Since it is only behaviors that can be observed in field settings, observing aftermarket activity in the presence of such biases may well be the most appropriate way to provide advice for policies that will be promulgated for the population being studied.

6. Concluding Remarks on the Meaning(s) of Behavioral Efficiency

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 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 policy *Y*. Indeed, it might even be widely accepted that advice to policymakers reporting and influenced by this finding 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 policy *Z*, this might also come to be accepted to play a role in policy advice despite divergent opinions as to its exact meaning.

References

Peter Bohm [1984], Revealing demand for an actual public good, *Journal of Public Economics* 24, 135-51.

Edward Clarke [1971], Multipart pricing of public goods, *Public Choice*, 11, 17-33.

Douglas D. Davis and Charles Holt [1993], *Experimental Economics*, Princeton University Press.

Urs Fischbacher [2007], z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics* 10, 171-78.

David M. Grether, R. Mark Isaac and Charles R. Plott [1979], Alternative methods of allocating airport slots: performance and evaluation, Civil Aeronautics Board Report, Pasadena, Calif.: Polynomics Research Laboratories, Inc.

David M. Grether, R. Mark Isaac and Charles R. Plott [1989], *The Allocation of Scarce Resources: Experimental Economics and the Problem of Allocating Airport Slots*, Westview Press, Boulder.

Theodore Groves [1973], Incentives in teams, *Econometrica* 41, 617-31.

Glenn W. Harrison, Ronald M. Harstad and E. Elisabet Rutstrom [2004], Experimental methods and elicitation of values, *Experimental Economics* 7, 124-40.

Glenn W. Harrison and John A. List [2004], Field experiments, *Journal of Economic Literature* 42, 1009-55.

Ronald M. Harstad [2000], Dominant strategy adoption and bidders' experience with pricing rules, *Experimental Economics*, 3 (2000), 261-80.

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

Vivian Lei, Charles N. Noussair, and Charles R. Plott [2001], Nonspeculative bubbles in experimental asset markets: lack of common knowledge of rationality vs. actual irrationality, *Econometrica* 69, 831-59.

Roger Myerson and Mark Satterthwaite [1983], Efficient mechanisms for bilateral trading, *Journal of Economic Theory* 29, 265-81.

Charles R. Plott and Kathryn Zeiler [2005], The willingness to pay-willingness to accept gap, the "endowment" effect," subject misconceptions, and experimental procedures for eliciting valuations, *American Economic Review* 95, 530-.

Charles R. Plott and Kathryn Zeiler [2007], Asymmetries in exchange behavior incorrectly interpreted as evidence of endowment effect theory and prospect theory? *American Economic Review* 97, 1449-.

Stephen J. Rassenti, Vernon L. Smith and Robert Bulfin [1982], A combinatorial auction mechanism for airport time slot allocation, *Bell Journal of Economics* 13, 402-.

Vernon Smith [1962], An experimental study of competitive market behavior, *Journal of Political Economy*, v. 52.

Vernon Smith [1976], Experimental economics: induced value theory, *American Economic Review*, v. 66.

Vernon Smith, Gerry Suchanek and Arlington Williams [1988], Bubbles, crashes and endogenous expectations in experimental spot asset markets, *Econometrica*, v. 58.

William S. Vickrey [1961], Counterspeculation, auctions and competitive sealed tenders, *Journal of Finance*, 16, 8–37.

Charles Z. Zheng [2002], Optimal auction with resale, *Econometrica* 70, 2197-2224.