Rev. Econ. Design (2009) 13:77–100
DOI 10.1007/s10058-009-0075-x

ORIGINAL PAPER

The human side of mechanism design: a tribute to Leo Hurwicz and Jean-Jacque Laffont

Daniel McFadden

Received: 26 June 2007 / Accepted: 30 January 2009 / Published online: 26 February 2009

Abstract This paper considers the human side of mechanism design, the behavior of economic agents in gathering and processing information and responding to incentives. I first give an overview of the subject of mechanism design, and then examine a pervasive premise in this field that economic agents are rational in their information processing and decisions. Examples from applied mechanism design identify the roles of perceptions and inference in agent behavior, and the influence of systematic irrationalities and sociality on agent responses. These examples suggest that tolerance of behavioral faults be added to the criteria for good mechanism design. In principle-agent problems for example, designers should consider using experimental treatments in contracts, and statistical post-processing of agent responses, to identify and mitigate the effects of agent non-compliance with contract incentives.

Keywords Mechanism_design · Principal-agent_problem · Incentives · Agent_compliance

JEL Classification D000 · D600 · D610 · D710 · D800 · C420 · C700

This paper was first presented in the Jean-Jacques Laffont Lecture, Toulouse, October 2006, and a condensed version was presented on the occasion of Leo Hurwicz's 90th birthday, April 2007. I am indebted to Eric Maskin for useful comments, to Florian Heiss, Charles Manski, Rosa Matzkin, and Joachim Winter for collaborative research on control of response errors in economic surveys, and to the National Institute on Aging for research support.

D. McFadden (✉)
Department of Economics, University of California, Berkeley, CA 94720-3880, USA
e-mail: mcfadden@econ.berkeley.edu
URL: http://www.econ.berkeley.edu/~mcfadden/index.shtml

1 Introduction

The study of mechanism design, the systematic analysis of resource allocation institutions and processes, has been the most fundamental development in economics in the last half-century, revealing the roles of information, communication, control, incentives, and agent processing capacity in decentralized resource allocation, and allowing identification of sources of market failure. This paper is a tribute to Leo Hurwicz, who first recognized the core issue of mechanism design in resource allocation problems and formalized its theoretical foundations, and to Jean-Jacques Laffont, who was at the center of the translation of the foundational economic theory into the language and tools that today appear in game theory, in studies of the organization of firms and markets, and in the applied economics of regulation, taxation, and public good provision. Thinking about transactions among economic agents in terms of information and incentives now threads through and connects pure and applied research across economics. The discipline itself has been transformed, from observers and commentators on economic systems to architects who design incentives and engineer, implement, and test institutions.

Section 2 of this paper gives a nutshell review of mechanism design theory and applications, and its central place in economics.[1] Section 3 examines a pervasive premise in mechanism design theory that economic actors respond rationally to the incentives embedded in a mechanism. In reality, mistakes that agents make in processing and drawing inferences from communications and information, and in exercising control and responding to incentives, can undermine the ideal efficiency of mechanisms, making it important to consider the robustness of mechanisms involving human agents. There are shifting attitudes regarding the relative merits of innovative but sometimes over-exuberant unregulated markets versus steady but sometimes plodding and inconsistent government regulation, but the principles of mechanism design make a few standards clear. While there are many markets best left to unregulated competition among the participants, there are some that are crippled by inconsistencies in information, control, incentives, and behavior, and require social management. Sections 4 through 6 of the paper give applied mechanism design examples in which the human side of mechanism design matters, and econometric and behavioral tools may help to improve mechanism robustness.

2 Mechanism design in a nutshell

Figure 1 is a schematic of major economic topics that have developed from or been enriched by mechanism design theory. The seminal contributions of Leo Hurwicz and Jacob Marschak at the end of the 1950's, the contemporaneous contributions of Ken Arrow and Gerard Debreu, and parallel works by Bill Vickery and Herb Simon, identified three critical elements in decentralized resource allocation—information, incentives, and the computational and control limits of agents. Today, most econo-

[1] Excellent general surveys of current issues in the theory of mechanism design can be found in Klemperer (2004), Krishna (2002), Maskin (2004), and Milgrom (2004).

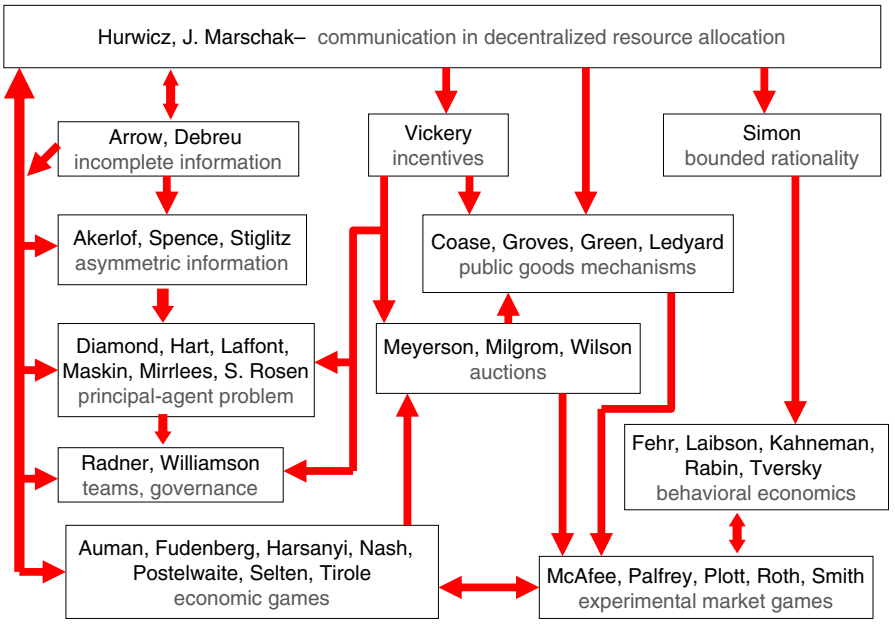

Fig. 1 Mechanism design in a nutshell

mists think of Leo Hurwicz, Ken Arrow, Bill Vickery, and Herb Simon as the founders respectively of mechanism design theory, information economics, incentive theory, and behavioral economics.

From the contributions of Ken Arrow to the economics of information, and the integration of the theories of information and resource allocation by Gerard Debreu, the role of information and communication developed further in the studies by George Akerlof, Michael Spence, and Joe Stiglitz of asymmetric information, by Peter Diamond, Oliver Hart, Jean-Jacques Laffont, Eric Maskin, Jim Mirrlees, and Sherwin Rosen of principal/agent problems and the design of efficient contracts, and by Roy Radner and Oliver Williamson of teams and governance. The theory of games also contributed fundamentally to these topics, as well as benefitting from them. Bob Auman, Drew Fudenberg, John Harsanyi, John Nash, Andy Postelwaite, Reinhart Selten, and Jean Tirole are some of the important contributors to this subject.

The second major branch is the theory of incentives, emphasized in the work of Vickery, and applied to the problem of social decisions on public goods by Ron Coase, Jerry Green, Ted Groves, Jean-Jacques Laffont, and John Ledyard, and the design of auctions by Roger Meyerson, Paul Milgrom, and Robert Wilson. These topics can also be viewed as applications of economic games.

The third major branch is the analysis of bounded rationality, the limited ability of economic agents to process information and consistently advance their self-interest. From the early study of this subject by Herb Simon, the fields of behavioral economics and experimental game theory have developed, with notable contributions to the former by Ernst Fehr, Danny Kahneman, David Laibson, Matt Rabin, and Amos

Tversky, and to the latter by Vernon Smith, Charles Plott, Tom Palfrey, Al Roth, and Preston McAfee.

There are important interconnections between the three major branches of mechanism design, with incentive theory playing a major role in principal-agent problems and in governance, and the reliability of information playing a major role in public goods decisions and economic games. Studies of auctions and of public good provision fueled the development of behavioral economics and experimental game theory. Particularly valuable integrative contributions to the field of mechanism design were made by Ken Arrow, Peter Diamond, Jean-Jacques Laffont, Eric Maskin, Roy Radner, and Jean Tirole, who recognized the span of mechanism design theory across the three major branches and ranged far beyond my assignments to boxes. Indicators of the impact on economics of mechanism design and the associated field of game theory are the eighteen Nobel Memorial prizes awarded through 2008 to names listed in Fig. 1. I think it is safe to predict that this subject will remain at the center of economics over this century, many new names and topics will be added, and more Nobel prizes will come.

The focus of the first formalization of mechanism design theory by Hurwicz (1960) was on the communication required to provide enough reliable information to each agent to achieve efficient one-time resource allocation. An elegant extension of the formal theory to incorporate stochastic and dynamic elements is accomplished by introducing information sets and state-dependent preferences; see Arrow (1953), Debreu (1959), Diamond (1967), Hurwicz et al. (1975), and Radner (1972). This extension produced important insights, but it left implicit the *process* that economic agents adopt to collect and draw inferences from statistical information. An alternative analysis starts with agents who face the real-time *econometric* problem of collecting data from communications, distinguishing signal from noise, and learning about their environment. Modern communications technology illustrates the usefulness of this approach—the bandwidth required to stream music or images is greatly reduced because one need transmit only enough information to reconstruct *changes* with sufficient resolution. Mean information requirements determined by sequential error-correction are typically far less than worst case requirements. Analogously, the communication needed to support nearly efficient trade from a reference or status quo allocation may be substantially less than that required to determine an efficient allocation from scratch.

3 The premise of rational action

A premise that pervades the theory of mechanism design is that economic agents act in their individualistic self-interest. The binding constraint on efficient resource allocation is then the amount and reliability of the information agents receive on the nature of goods and the interests of other agents. This premise has two important implications. First, if mechanism design provides suitable institutions, communication channels, and incentives to ensure that information is reliable, then the self-interest of rational agents will ensure efficient resource allocation. Second, while planners lack the bandwidth, computational capacity, and incentives to manage efficient centralized resource allocation, individual agents do have sufficient capacity to deal optimally with the more limited information relevant to them and decisions they face.

In reality, human agents operate as untrained statisticians in a stochastic environment, displaying systematic behavioral flaws in perceptions and choices, and may fail to recognize or act in their self-interest. Then, characteristics of mechanisms that should make no difference to rational agents, such as complexity, framing of information, transparency of consequences, and choice defaults, can in reality have a substantial impact on the stability and efficiency of resource allocation. These considerations are not new. The processing limits of economic agents have been studied since the works of Simon, Kahneman, and Tversky, and figure regularly in behavioral economics. Bayesian learning models are common in experimental game theory. Nevertheless, there continue to be new findings on the nature of human decision-making, and opportunities for further research on designing mechanisms that are tolerant of human error.

Among human limitations are bounded attention, memory limits, reasoning limits, and sociality. Simon (1971) noted that people have an attention budget, saying that ". . .a wealth of information creates a poverty of attention". We see in the market the effects of attention failures—heavy advertising and teaser pricing for new customers, and low switching rates for quasi-durables like mobile phones and insurance. Memory is systematically distorted. We recall first and last events more easily than intermediate ones. We remember remarkable coincidences, but not remarkable non-coincidences, and this leads us to infer patterns and correlations even when events are independent. Reasoning is also systematically distorted. We strongly discount the future and the risk of improbable events, and often regret it in hindsight. We use exemplars and analogies to guide decisions rather than reasoning through the consequences. We discount ambiguous information, focusing on aspects of alternatives where comparisons are easy, and using lexicographic screening to reduce data collection and computation tasks. Finally, we are guided by *sociality*—the proclivity of humans to imitate others and to be guided by social norms for reciprocity and altruism. Sociality is not necessarily antagonistic to rationality, but it confounds individualistic self-interest with social norms and perceptions of group preferences, blurs the lines from individual decisions to consequences, and invites moral hazard in which consumers do not experience the full marginal benefits and costs of "hidden" actions that are masked by group behavior.

Consider for example, the behavior of riders in a bicycle race. They often affiliate voluntarily with the *pellaton*, a mass of riders that provides an energy-saving, choice-limiting environment. In many circumstances, the pellaton promotes efficient resource allocation by reducing requirements for information collection and processing, providing error-correcting choice algorithms for its members that make it highly stable. However, by reducing sensitivity to small shocks, pellaton behavior can allow bubbles and Ponzi schemes to develop, which can induce chaotic instability in response to large shocks. A major challenge to understanding economic behavior is to explain the economic, social, and stochastic factors that cause pellatons of economic agents to form and break apart.

Mechanism design theory is one of the more abstruse subjects in economics, and even now it may not strike many applied economists as a useful day-to-day guide to practical problems of economic policy. I think that in fact its organizing concepts are of fundamental practical importance, and policy economics is well served by

recognizing its connections to this unifying theory. I turn in the following sections of this paper to three practical problems in applied mechanism design that illustrate the value of melding design theory and information on human behavior to develop robust mechanisms that function well in reality.

4 Optimal tax schedules with noisy signals on agent characteristics

Optimal tax policy seeks tax schedules that minimize welfare-decreasing distortions of economic activity. A leading example is the classical problem of designing income tax schedules that minimize distortions in labor supply; see Mirrlees (1971), Diamond and Mirrlees (1971, 1972), and Saez (2001). Green and Laffont (1986) place this problem in a broader context of imperfect control of agents. Optimal schedule design often requires information on consumer preferences or activities that are not perfectly observable to the taxing authority, so that schedules must be based on noisy signals. Noise may come from the mechanics of monitoring and measurement, and from consumer mistakes and misrepresentation in self-reports. Erroneous signals have a direct social cost, and in addition may distort the incentives for consumers to give accurate self-reports. For example, systems that depend on self-reports that are subject to audits and penalties may perform poorly if audits themselves are subject to error, or if the threat of audit and penalty is too weak an incentive to induce truthful reports.

Papers that concentrate on noise in the mechanics of measurement include Varian (1974) and Chakraborty and McAfee (2008). Varian considers the costs of errors in statistical property assessment for property taxes, and devises an estimation procedure that mitigates these costs. Chakraborty and McAfee consider the second-best optimal Pigovian tax on an externality-generating activity by a consumer when the activity is measured with noise. An interesting question for mechanism design is how to minimize welfare losses when both the principle and agents can introduce measurements in a costly appeals process for error mitigation.

Feinstein (1990, 1991, 1999) studies self-reported tax submissions and compliance using auditor effects, which are essentially experimental treatments on the probability that errors will be detected, to estimate levels of compliance; see also Dubin et al. (1992) and Kim (2005). The interesting problem in mechanism design is to ask how imperfect enforcement alters the incentives for economic activity as well as the reporting process, and what tax schedules are optimal when the effects of enforcement are taken into account. Suppose a consumer underpays her taxes by an amount t, is audited with probability p, and if audited is charged $(1 + \gamma)t$, where γ is a penalty rate. If $\gamma > (1 - p)/p$, then a rational risk-averse consumer will not underpay. If audits are costly, then the taxing authority gains by making p very small, with γ large as a consequence. However, a systematic irrationality in consumer behavior is misperception of very low probability events, and refusal of fair insurance against large, very low probability losses. In the face of this behavior, audit probabilities in a successful enforcement design cannot become too small. An experimental design that varies audit probabilities and penalties could quantify this behavior and be used to determine an efficient enforcement design.

5 Economic juries for public projects

Consider the problem facing a social planner who must select among various public projects. To estimate the social values of these projects, the planner may select an *economic jury* of consumers, and elicit stated values from the jury members, using incentive mechanisms to minimize welfare losses from selection bias, reporting error due to strategic misrepresentation or carelessness, and statistical variation. The use of a jury for public projects decisions was first suggested by Green and Laffont (1979) in their analysis of the Groves–Clarke mechanism. There are several reasons to prefer juries rather than a population census to estimate social values. One is that incentive mechanisms to induce truthful value reports from the whole population may require income transfers that are inconsistent with general equilibrium balance. This difficulty is eliminated if non-jurors are assigned residual income. Second, a population-wide elicitation will as a result of attrition lead to a *de facto* self-selected jury. It is statistically sounder to control jury selection through random sampling and fees for participation. Third, if juries are small and members are elicited independently, then there is less opportunity for formation of coalitions that can upset the incentive-compatibility of mechanisms. However, stated values of jurors may fail to reflect the public value even with an incentive-compatible elicitation mechanism if the jury is not representative, or if jurors fail to receive, recognize, and respond rationally to the incentives they face. Humans are inconsistent in their response to low probability events, and the possibility of being pivotal may be too remote in large juries or a census to induce rational response. Other factors that may influence juror behavior, and may be as important as economic incentives, are the degree to which interpreting questions and forming responses requires cognitive effort, including the recall of facts and experiences from memory and construction of previously unexpressed preferences, and non-economic incentives for strategic misrepresentation, including the influence of norms for "socially responsible" behavior.

As background for the problem of eliciting juror values, consider the history of direct elicitation of preferences. This subject dates to a 1932 paper written by the psychologist Leon Thurstone at the instigation of his University of Chicago colleague Henry Schultz. Thurstone's proposal to elicit indifference values was rejected by leading economists of that day, including Frisch, Hotelling, and Friedman. The concerns of these critics were that subjects freed of the discipline of completing market transactions would fail to take measured account of prices and budget, and would use their responses to posture, or to express attitudes and opinions, making the stated preferences unreliable for predicting market behavior. The possibility of using stated preferences disappeared from economics, and there was little further development of these methods until the mid-1960's, when the approach, renamed *conjoint analysis*, began to be explored as an applied tool in psychometrics, market research, and transportation research. These developments emphasized construction of preference maps through presentation of multiple choices set by experimental design. For private goods that are familiar, or given sufficiently rich description, conjoint analysis with embedded incentives has proven to be a reliable tool for predicting market demand, and it is widely used in the design of new products.

A largely independent development of stated preference methods, called *contingent valuation* (CV) and focused on eliciting preferences for public goods, occurred in resource economics (Davis 1963; Randall et al. 1974). The method has been promoted and used somewhat uncritically as a tool for valuing resource damage, and there is a large and contentious literature on its validity, but methodologically it is simply a form of conjoint analysis with a truncated design for the experimental presentation of alternatives. Hence, the concerns of its critics are those for stated preference methods in general, with added concerns about consumers' ability to generate preferences for unfamiliar public goods, separate self-interested preferences from broader preferences incorporating social norms, respond consistently in hypothetical versus real choice settings, and respond predictably to hypothetically incentive-compatible framing of survey tasks.

Three distinct aspects of direct elicitation of preferences in conjoint analysis are (1) the *elicitation frame*, or context and format of the question and requested response, (2) the *implementation frame*, or link between jury responses and the (subjective) probability that a policy will be implemented, and (3) the *payment vehicle*, specifying the tax a juror would bear for implemented projects. Aspects of the elicitation frame are whether the juror is trained or experienced in making jury judgments, whether open-ended or referendum (yes/no) responses are called for, and whether the context encourages or discourages altruistic behavior. The implementation frame is *consequential* if there is a positive (subjective) probability that a juror's stated value will be *pivotal* in determining whether a project is supplied, and *hypothetical* if there is no direct link between jury response and the eventual decision on project supply. The payment vehicle imposes taxes earmarked to cover the cost of implemented projects, and the tax imposed on jurors may be *coupled* to their stated values, or *decoupled*.

The selection and motivation of juries is a principal-agent problem. *Ex ante*, jurors may be given incentives for participation, effort, and truthful responses. Experimental treatments can be embedded in the incentive mechanisms to facilitate *ex post* statistical analysis to identify and mitigate response errors. Considering incentive mechanisms and statistical mitigation in tandem can improve the reliability of information collected from juries. I first consider briefly the question of agent participation and selection, and then turn in more detail the use of incentive-compatible mechanisms in a consequential preference elicitation.

Agent participation in the case of a single agent has been studied by Grossman and Hart (1983), Jewitt (1988), and Laffont and Martimort (2000, 2002, Ch. 3,5). Philipson (1997), Philipson and Malani (1999), Philipson (2001) and Ryu et al. (2005) show that sample recruitment is a similar problem, except that rather than elicit the participation of a single agent, the principal now wants to control selection bias by recruiting as representative a jury of agents as possible. Factors entering this problem are the costs of contacting prospects and eliciting information from jurors, the effect of fees on participation, and the costs of a non-representative sample. McFadden (2009) analyzes juror participation and response using a variant of the bivariate selection model originally introduced by Heckman (1979), and analyzed by Imbens and Newey (2002) and Chesher (2005). This model extends the univariate selection analysis of Philipson (1997), drawing upon Imbens and Manski (2004) and Manski (2005) for results on set identification, and on Matzkin (1992, 2006) for non-parametric estimators. The major

findings of this research are that relatively high participation fees, of an order of magnitude greater than is typically encountered in survey research, are needed for optimal control of selection bias under common "worst case" conditions, and that juries of modest size are sufficient for public project decisions, essentially because when the social choice is clear-cut, this can be determined easily even with a small sample, and when it is close, the social regret from a mistake is small. An important feature of this problem is that participation fees, presentation of public projects, and incentives are treatments under the control of the planner that can be designed to identify and mitigate response errors.

Next consider elicitation mechanisms. Consider a public project decision problem with projects x in a finite feasible set X. Let p denote the vector of private good prices. Assume that aggregate consumer income $F(p)$ is independent of x and convex and conical in p, and the per capita cost of the project is a function $r(x, p, \zeta)$ that is concave and conical in p and depends on a variable ζ that is private information to the planner. In general, X and the cost $r(x, p, \zeta)$ can be configured so that x indexes combinations of sub-projects that can be complements or substitutes.[2] Assume a population of consumers $n = 1, \ldots, N$ who have heterogeneous Gorman polar indirect utility functions with parallel Engle curves,

$$V_n(p, x) = [F_n(p) - t_n - B_n(p) + w_n(x, p)]/A(p),$$

where $F_n(p)$ is this consumer's income, t_n is a lump-sum net tax, $B_n(p)$ is a heterogeneous concave conical non-decreasing function of p, $w_n(x, p)$ is the heterogeneous value consumer n attaches to project x, a function that is convex and conical in p, and $A(p)$ is a common concave conical non-decreasing price index. Then, (F_n, B_n, w_n) define the consumer type. At least under some conditions, the characteristics (F_n, B_n) are revealed through market good demands, but w_n is private information to consumer n; see Green and Laffont (1977). This particular preference field has the property of linear transferrable utility that is required for basic incentive-compatible mechanisms, and the property that individual preferences aggregate to a single representative consumer with the social indirect utility function

$$V(p, x) = [F(p) - T - B(p) + N \cdot w(x, p)]/A(p), \tag{1}$$

where $T = \sum_{n=1}^{N} t_n$, $B(p) = \sum_{n=1}^{N} B_n(p)$, and $w(x, p) = N^{-1} \sum_{n=1}^{N} w_n(x, p)$; see Chipman and Moore (1980, 1990), McFadden (2004). In equilibrium, taxes must cover the cost of the supplied public projects, the balance condition $T = N \cdot r(x, p, \zeta)$. Competitive equilibrium with socially optimal supply of public projects is characterized by $(p^*(\zeta), x^*(\zeta))$ satisfying the saddle point

$$\min_{p \in P} \max_{x \in X} [F(p) - B(p) + Nw(x, p) - Nr(x, p, \zeta)]/A(p). \tag{2}$$

[2] For example, $X = \{(0, 0), (1, 0), (0, 1)\}$ corresponds to two mutually exclusive sub-projects, while $X = \{(0,), (1, 0), (1, 1)\}$ corresponds to two sub-projects such that the second can be done only if the first is done.

This formulation of the problem allows cost and willingness-to-pay for a public project to depend on private good prices. Then, the optimal provision of public projects depends in general on p, and public project choice will enter the determination of equilibrium prices. However, note that if $r(x, p, \zeta) = A(p)r_0(x, \zeta)$ and $w(x, p) = A(p)w_0(x)$, then the saddle point problem (2) separates into a minimization in p and a maximization in x, and each is determined independently of the other.

Suppose the planner draws a jury of size J at random from the population, so that selection due to non-participation is not an issue, and suppose the consumers are numbered so that the jurors are consumers $j = 1, \ldots, J$. The elicitation of stated values is a game of incomplete information played by the planner and these jurors, with each juror having private information on her true value, and the planner having private information on the costs of projects. I consider two implementation frames and payment vehicles that will induce truthful valuations, provided jurors understand and respond to the mechanism incentives; this analysis is drawn in part from Green et al. (1998).

Groves–Clarke–Green–Laffont (GCGL) Mechanism: Originating in the works of Groves and Loeb (1975) and Clarke (1971), and stated for juries by Green and Laffont (1978), this mechanism requires that the planner announce a tax schedule $t_j = r_j(x, p, \zeta)$ for juror j that depends on the project x, on p, and on the factor ζ that when realized determines the net tax payment $r_j(x, p, \zeta)$. Each juror then reports a stated value schedule $w'_j(x, p, \zeta)$ to the planner; in principle, this schedule can depend on ζ. The project implemented maximizes the inner term in (2) with the unknown true social value per capita $w(x, p)$ replaced by its jury average, so that

$$x'(p, \zeta) = \text{argmax}_{x \in X} \left(J^{-1} \sum_{j=1}^{J} w'_j(x, p, \zeta) - r(x, p, \zeta) \right). \tag{3}$$

The net benefit of x to juror n, taking the tax into account, is $w_n(x, p) - r_n(x, p, \zeta)$. This juror then maximizes individualistic self-interest by reporting the strategic value

$$w'_n(x, p, \zeta) = w_n(x, p) - r_n(x, p, \zeta) + Jr(x, p, \zeta) - \sum_{j \le J \& j \ne n} w'_j(x, p, \zeta), \tag{4}$$

so that (3) reduces to $x'(p, \zeta) = \text{argmax}_{x \in X}(w_n(x, p) - r_n(x, p, \zeta))$ and yields a social choice criterion that is congruent with this juror's individualistic utility. Now, if the tax the planner imposes on n is

$$r_n(x, p, \zeta) = Jr(x, p, \zeta) - \sum_{j \le J \& j \ne n} w'_j(x, p, \zeta), \tag{5}$$

the result in (4) is that $w'_n(x, p, \zeta) \equiv w_n(x, p)$, so that it is a dominant strategy, independent of the behavior of other jurors and of ζ, for juror n to report her true value. In this setup, non-jurors will be taxed for the residual necessary to achieve the equilibrium balance condition that the cost of implemented projects is covered,

$$\sum_{n=J+1}^{N} r_n(x, p, \zeta) = (N - J^2) r(x, p, \zeta) + (J - 1) \sum_{j=1}^{J} w'_j(x, p, \zeta),$$

It may be necessary in this mechanism to use additional lump-sum transfers, which can depend on ζ and p, but not on x, to ensure that jurors and non-jurors all have net incomes sufficient to cover committed expenditures. For juries of modest size in large populations, sufficiency for jurors will be the primary concern, as the average impact on non-jurors will be close to the per capita real cost of the project $r(x, p, \zeta)$, which can be taken to be generally affordable by definition of X. Lump-sum transfers to jurors then may have the dual purpose of assuring juror participation and ensuring that the incentive-compatible mechanism is feasible.

The GCGL mechanism is a *provision point mechanism* that ties implementation to an average value that exceeds a specific cost threshold. If each juror believes there is a positive probability that an implementation decision will be made, that if it is made it will maximize the average jury net payoff, and there is a positive probability of a configuration of reports of others and costs that would make her response pivotal, then the argument above verifies that it is a dominant strategy for each jury member to report her true value, so the mechanism is strongly individually incentive-compatible; see Palfrey and Srivastava (1989, 1991). Note that if subjects believe and understand the implementation frame and tax function in the GCGL jury mechanism, then features of the elicitation frame, such as whether values are reported as functions of x, p, and ζ, as open-ended responses to elicitations at specific x, p, ζ values, or as yes/no responses to threshold questions, should not matter. The mechanism will lead to efficient provision of public projects, up to what in general will be a modest loss of accuracy from jury sampling noise and selection that can be controlled by the planner's choice of jury size and participation fees.

Becker–DeGroot–Marschak–Palfrey–Rosenthal (BDMPR) Mechanism: A second incentive-compatible mechanism that is natural for referendum elicitations is an adaptation of the Becker–DeGroot–Marschak auction mechanism that has been used in public goods games by Palfrey and Rosenthal (1990, 1994), Ledyard and Palfry (1994), and Green et al. (1998), and tacitly by Hoehn and Randall (1987). Suppose there is a single project, so $X = \{0, 1\}$. Suppose each juror understands that her tax if project $x = 1$ is implemented is the per capita real cost $r(1, p, \zeta)$, and believes an implementation frame stating that her valuation can alter the probability of implementation, making her pivotal. This belief may be induced by language such as "when the cost per person of providing $x = 1$ is finally determined, then the probability of implementation increases with the plurality in this jury who favor the project at this cost". Let $w'_n(1, p)$ denote the cost threshold at which juror n would vote to support $x = 1$, so that approval of the project at realized cost $r(1, p, \zeta)$ is indicated by $\mathbf{1}(w'_n(1, p) - r(1, p, \zeta) > 0)$. Juror n's subjective probability of implementation is

$$\mathbf{E}_{-n} \Psi_n \left(\sum_{j=1}^{J} [\mathbf{1}(w'_j(1, p) - r(1, p, \zeta) > 0)]/J \right), \tag{6}$$

where Ψ_n is a non-decreasing function determined by jury instruction and juror n's beliefs, and $\mathbf{E}_{-n}$ is this juror's subjective expectation regarding the thresholds of other jurors. Juror n is pivotal if either Ψ_n is a strictly increasing function, or if Ψ_n is non-decreasing and non-constant, and its expectation with respect to the reports of others is strictly increasing. The last possibility includes conventional voting rules such as majority rule, provided each consumer's subjective beliefs about others is sufficiently diffuse so that she believes her vote *might* be pivotal. Juror n's expected utility is then

$$\mathbf{E}_\zeta(w_n(1,p) - r(1,p,\zeta)) \cdot \mathbf{E}_{-\mathrm{n}}\Psi_n\left(\frac{\mathbf{1}(w'_n(1,p) - r(1,p,\zeta)) + \sum_{j\neq n \& j\leq J}\mathbf{1}(w'_j(1,p) - r(1,p,\zeta))}{J}\right). \tag{7}$$

Make the critical assumption that $w_n(1, p)$ is in the interior of the support of the random per capita cost $r(1, p, \zeta)$ induced by the unknown cost factor ζ. Then, any report $w'_n(1, p) < w_n(1, p)$ lowers the probability of implementation for some events that are desirable for n, and any report $w'_n(1, p) > w_n(1, p)$ raises the probability of implementation for some events that are undesirable for n. Then, truth-telling in the referendum vote is a dominant strategy. The difference between this setup and the GCGL one is that in this case, the effect of being pivotal operates through the probability of provision rather than through the payoff conditioned on achieving a provision point. Again, elicitation format does not matter—asking directly for the subject's threshold $w'_n(1, p)$ or obtaining it indirectly in various referendum setups should lead to the same answer. It is important that a juror's required payment given implementation is independent of the stated threshold. If, alternately, the payment is coupled to $w'_n(1, p)$ through a payment function that is increasing in $w'_n(1, p)$, then the subject has an incentive to "free ride" by under-reporting $w_n(1, p)$.

The BDMPR mechanism can be extended to multiple alternatives if the planner conducts a series of independent elicitations that compare each possible project portfolio with the baseline alternative $x = 0$, and if the juror responds to each elicitation myopically, overlooking the strategic possibility that understating values on less preferred alternatives may increase the probability that more preferred alternatives are implemented. However, if jurors are not myopic, then this mechanism encounters the usual difficulties of strategic manipulation in sequential voting.

While the BDMPR mechanism is incentive-compatible, it is not efficient, as there is no guarantee that project x will be implemented if and only if $w(x, p)-r(x, p, \zeta) > 0$. Specifically, the referendum vote used in the mechanism cannot recognize when a few jurors with thresholds far above cost should in the utilitarian calculus outweigh a larger number of jurors with thresholds just below cost.

Other mechanisms: A number of alternative mechanisms are available for making public project decisions that vary in the conditions under which they are (approximately) incentive-compatible, and in the (approximate) efficiency of their implementation rules. A benchmark that is not incentive-compatible and suffers from free-riding is the voluntary contribution mechanism. Morgan (2000) proposes a variant in which tickets in a large-payoff lottery are tied to voluntary contributions, and provide an

incentive that mitigates free-riding; see Pecorino and Temimi (2007). More generally, interweaving portfolios of private goods and public projects, and designing elicitations of stated values with stochastic implementation of some (private good) components, may present jurors with a problem that is easy to solve consistently only by being truthful. It is also possible to vary the strength of incentives by rewarding consistency across jurors, as in the powerful provision point mechanism of Groves and Ledyard (1977, 1980), which can be adapted to jury-based valuation under more general preferences than Gorman preferences.

Two critical requirements in the public projects provision mechanisms just described are that each juror recognize and act upon her ability to directly influence her net income and the supply of public projects through her self-reported values, and that she *not* recognize and act upon the strategic opportunities her report offers for the indirect determination of her income function and private goods prices. Gibbard (1973) and Satterthwaite (1975) show that in general no non-dictatorial balanced mechanism in an economy with a finite number of consumers can be strategy-proof, so the restrictions on juror beliefs and behavior necessary to assure that a mechanism is strongly individually incentive compatible require something less than total rationality and understanding. In particular, the mechanisms will perform poorly if jurors act strategically to influence private goods prices, or alternately fail to recognize they may be pivotal, fail to recognize the consequences of their actions when they are pivotal, or allow non-economic incentives to override economic ones. There is a large empirical literature on consumer behavior in various economic environments in the laboratory and in the field, including studies of public good provision mechanisms. I will give a very selective review of findings that shed some light on the ability of consumers to recognize and exploit choice opportunities in their own self-interest, in the presence of the incentives that naturally appear in markets, and in laboratory settings where incentives can be designed that should lead to specific behaviors if consumers can process information and choose rationally.

Behavior in public good games: Mechanisms of the GCGL or BDMPR type are effective in obtaining truthful information if consumers recognize the opportunities provided by the choice alternatives they are offered, and seek to maximize individualistic utility that satisfies the restrictions the elicitation mechanism requires (e.g., risk-neutral, Gorman polar preferences with parallel Engle curve), uncomplicated by sociality. The behavioral question is whether consumers meet these standards. Some of the most striking evidence comes from voluntary contribution systems for public goods, the ultimatum and trust games, and auctions. An early paper of Bohm (1972) found that "free riding" was uncommon even in circumstances where the incentive structure invited it. Shafir and Tversky (1992) found that the dominated strategy of cooperation is often played in the prisoner's dilemma game, apparently induced by superstitious beliefs. Fehr and Schmidt (1999), Fehr and Falk (2002), Fehr et al. (2002), Fehr and Fischbacher (2002, 2004), Fehr and Gachter (2004), and others have found that in the ultimatum and trust games, many participants are motivated by social norms to play dominated strategies. These results suggest broadly that in circumstances where there is a perceived mutual benefit from cooperation, consumers have altruistic motives, superstitious beliefs, and social norms for reciprocity and fairness that may override pure self-interest. On the other hand, there is considerable evidence that in

the purely competitive circumstances of second-price auctions, where the compatibility of the incentives in the auction with truth-telling is transparent and there are no strong social norms against winning, consumers tend to bid their true values; see Smith (1980), Harstad (1990), Friedman and Rust (1993), List and Lucking-Reiley (2002), and Garratt et al. (2004).

In both the GCGL and BDMPR mechanisms, the probability that an individual juror is pivotal falls with jury size. If jurors display the common behavioral pattern of sometimes ignoring low-probability events, then compliance with these mechanisms will fall as jury size increases. Studies of behavioral response to the GCGL mechanism find that it does not induce wide-spread truth-telling in small untrained juries, but compliance increases sharply when subjects are trained and given detailed information on the payoff structure. There is also an indication that compliance falls in larger juries where the pivotal income adjustment does not loom as large and the advantages of the dominant strategy are obscured; see Attiyeh et al. (2000), Cason et al. (2003), and Kawagoe and Mori (2001). Chen and Plott (1996) find that compliance in the related Groves–Ledyard mechanism depends significantly on the penalty parameter in that mechanism, indicating that the magnitude of the incentive matters. Palfrey and Rosenthal (1990) find that with training, small juries show good compliance when the public good game is played with the BDMPR mechanism in referendum voting form.

Summarizing these results, there appear to be three main factors that determine whether consumers will comply with individual incentives: (1) whether the game is purely competitive, versus one in which benefits of cooperation are recognized and lead to responses influenced by social norms; (2) whether the mechanism is substantially individualistic and transparent, or is obscured by institutions or the actions of other players; and (3) whether or not the penalties for deviating from a compliant response are strong and obvious. Thus, second-price auctions are generally sufficiently competitive and the incentives for truth-telling are sufficiently individualistic and transparent, to induce compliance. By contrast, public goods games require considerable training and clear information on payoffs to avoid erratic, non-compliant responses. In this respect, the BDMPR mechanism, or the Groves–Ledyard mechanism with a substantial penalty, appear to have some transparency advantage over the GCGL mechanism. These factors imply for survey research applications where it is difficult to provide strong incentives and training for direct preference elicitation, compliance with the incentives of strategy-proof mechanisms is problematic, and except for purely individualistic decisions such as private good choices, responses are likely to be influenced by social norms. Consequently, it is unclear that one can obtain more reliable information in surveys using weakly incentive-compatible mechanisms than using a purely hypothetical but scientifically worthy framework that evokes social norms for honesty and reciprocity.

Evidence on the reliability of Contingent Valuation (CV) responses: Elicitation of stated preferences, and particularly the CV method, have been the focus of most of the concentrated attention in economic survey research on the reliability of responses and the effect of hypothetical versus real incentives. The primary concerns have been the incentive compatibility properties of alternative elicitation formats, and the issue of "hypothetical bias" and survey methods that minimize this bias.

Incentive compatibility of CV elicitations of value has been a continuing concern of environmental economists; see Randall et al. (1974), Randall et al. (1983), Hoehn and Randall (1987), and Carson and Groves (2007). Careless treatment of incentive issues, particularly failure to distinguish clearly between circumstances where incentives are hypothetical or real, and to distinguish between the theoretical incentive compatibility of mechanisms and behavioral compliance, have led to confusion in the resource economics literature regarding the influence of elicitation formats, and the relevance of private good choice behavior to public good choice behavior; e.g., the claim by Hoehn and Randall (1987), Carson and Groves (2007), and Loomis et al. (1996) that only a referendum format can potentially elicit incentive-compatible responses. The discussion of incentive compatibility given in Green et al. (1998) and in this paper shows that when a CV elicitation is presented within a consequential implementation frame that has a credible possibility that the respondent is pivotal, then both the GCGL or BDMPR jury mechanisms are incentive-compatible, and there are no first-order differences in incentive-compatibility between elicitation formats that employ the same payment vehicle. However, the transparency of the mechanisms may interact with elicitation format and with the training needed for subjects to be aware of their payoffs.

The reliability of stated preferences and their predictive power has been studied in market research, and applied areas such as transportation research; see McFadden (1980), Ben-Akiva and Morikawa (1990), Louviere et al. (1999, 2000), Shen (2005), and Train and Wilson (2005). In most cases, preferences for private goods such as new consumer products are examined. Questions have centered on the format of the elicitations, particularly the "richness" of the description of choice alternatives, the form of response (e.g., choice, ranking, rating, referendum WTP, open-ended WTP), the design of multiple elicitations, and cross-analysis of revealed preferences. Methods for studying these questions include study of the internal consistency of multiple stated preferences (e.g., transitivity, monotonicity, diminishing returns), consistency between stated and revealed preferences, and predictability of real choices from stated preferences, either to subsequent offerings within the survey or to subsequent market experience.

A very broad summary of the findings are that stated preferences for private goods in a well-designed conjoint analysis are generally consistent with revealed preferences, or can be made so by calibration. The incentives provided by a positive probability of a follow-up transaction may increase compliance, but compliance without incentives is not bad, and compliance with incentives is not perfect; see Camerer and Hogarth (1999). Stated preferences can be influenced by the framing and presentation of attributes. For example, Tversky et al. (1988) show that the decision format can change the *prominence* given to different attributes of alternatives. In choice among products, price is given more weight in a direct choice task than it is when consumers are asked to specify an attribute level that makes two alternatives indifferent. Further, price is often given more prominence in stated preferences than it is in revealed preferences, probably because it provides a common and familiar quantitative low-effort standard for comparison. There is a strong *status quo* or *endowment* effect in stated preferences, sometimes termed the WTP/WTA gap, and while this also appears in revealed preferences, its importance varies. When goods in a stated choice experiment are unfamiliar or sparsely described, the expressed preferences are more erratic. An overall conclusion is that stated preferences for private goods collected within an experimen-

tal design that provides a good sense of verisimilitude are generally consistent with and predictive for revealed preferences, even without positive incentives for truth-telling. However, stated preferences for unfamiliar goods are erratic, partly because of the difficulty of providing sufficiently cogent descriptions of these products to make the choice problem realistic and induce the effort needed to approximate real market behavior, and partly because consumer preferences among unfamiliar objects are a construction project, poorly formed and unstable until contextual cues, experience, and perceptions come together to fix their form.

For the public goods that are commonly the target of CV surveys, such as recreational facilities, uncontaminated groundwater, and seabirds, most studies suggest that hypothetical bias is significant. The methods used for this assessment include internal consistency of WTP elicitations that vary by extent, adding up, and context, but most importantly the relationship between stated willingness to contribute and actual contributions. Elicitation format influences responses, and it is possible that subjects are influenced by the nominal incentive compatibility of some hypothetical formats. However, altruism, social norms, and perceptual anomalies are more likely explanations for the observed patterns; see List and Gallet (2001), Venkatachalam (2004), and Kahneman et al. (1999). Champ et al. (2002) find that payment vehicle (e.g., referendum on mandatory tax, unspecified voluntary donation, and voluntary contribution with provision-point mechanism for implementation) matters in a hypothetical, but perhaps taken as realistic, elicitation of WTP for acquisition of park land in Boulder, Colorado. Cummings et al. (1995) and Cummings et al. (1997) find in a laboratory CV experiment conducted under hypothetical and real conditions that subjects are not usually truthful in referendum responses. Lusk and Schroeder (2004) find significant hypothetical bias in WTP for beef steaks. Loomis et al. (1996, 1997) find strong hypothetical bias in experiments comparing hypothetical CV and real second-price auctions. Rondeau et al. (1999) and Poe et al. (2002) compare hypothetical referendum WTP with that obtained from a provision point mechanism, and find a smaller gap than in experimental comparisons with a voluntary contribution mechanism. Other authors finding significant hypothetical bias include Azevedo et al. (2003), Bennet et al. (2004), Champ and Bishop (2001), Cummings et al. (1997), Diamond and Hausman (1994), Johannesson et al. (1998), Loomis et al. (1996, 1997), and McFadden (1994). Authors finding limited hypothetical bias include Carlsson and Martinsson (2001), Carson et al. (2001), Frykblom (2000), Frykblom and Shogren (2000), Haab et al. (1999), Whitehead (2002), and Willis and Powe (1998). An overall assessment is that studies finding the least bias focus on private goods, and that proponents of CV find fewer problems with hypothetical bias than do critics.

A significant issue in CV elicitation for public goods, and a possible factor in hypothetical bias, is that respondents construe hypothetical tasks as asking for "socially responsible" values that reflect an altruistic attribution of the benefit of a public project to others in addition to the individual's personal value, whereas consequential tasks calling for payment focus attention on individualistic value. Such behavior is consistent with statements from consumers that voluntary contributions and other altruistic acts provide a "warm glow". Put another way, altruistic motives may be overwhelmed when private incentives are strong, but may reassert themselves when private incentives are weak or context encourages attention to the advantages of cooperation and reciprocity.

A number of authors have suggested variations on the CV method that appear to have less hypothetical bias, or provide a basis for calibration to remove this bias. List (2002) investigates choice experiments for a private good and a public good contribution. This approach is consistent with the general methods of conjoint analysis used in market research, so that the finding that private good choices conform to truth-telling is not surprising, but the carry-over to the voluntary contribution task is, and the details of List's mechanism may prove instructive to designers of WTP elicitations. However, calibration is an imperfect method for overcoming hypothetical bias, because it must rely on comparison commodities that may not be good proxies for the target good. For example, Fox et al. (2003) find that calibration factors are commodity-specific.

The primary lesson for mechanism design theorists from the studies of contingent valuation is that it matters whether a preference elicitation is consequential, and hypothetical bias can be substantial, but the differences seem to be more complex than just differences in rational response between a strongly incentive compatible elicitation and one that is incentive neutral. In particular, individualistic versus social framing of the choice task, and consumer experience with the relevant preference judgments, seems to matter.

Consumer response to large incentives: At a basic level, the fact that humans can function and survive in market economies indicates that they recognize and act upon the economic incentives they face. However, there is a long-standing question in economics as to whether this comes from conscious, relentless preference maximization, or from less coherent and organized use of heuristics that give satisfactory results in most circumstances. In familiar settings, these alternatives models of behavior may be largely indistinguishable, but in an unfamiliar setting such as play of a public goods game or making a choice among new products and services, heuristics may be incompatible with rational response to the incentives in the situation. Then, it is useful to look for designed or natural experiments where consumers are confronted with novel decisions and their responses can be assessed against rational standards. The answers can help to guide mechanism design—can it rely on economic incentives alone, or is a degree of paternalism needed to inform, train, and coax consumers to act in their self-interest?

There is considerable evidence that in familiar decision-making circumstances where self-interest really matters, consumers are approximately rational. Studies of choice among lotteries with large payoffs by Binswanger (1980) and Attanasio et al. (2006) have been found to conform closely to postulates of rational decision-making under uncertainty. List (2003) and Garratt et al. (2004) find that experienced market decision-makers show few behavioral anomalies. In summary, these findings suggest that to assure that a mechanism will induce compliant, rational responses from human agents, it is important to keep the design *simple*, the goals *selfish*, and the payoff incentives *substantial*.

6 Consumer behavior in health insurance markets

In the ongoing public policy debate on the merits of providing health insurance through a market with competition between insurers rather than a government-run "single

payer" system, the advantages claimed for the competitive market are that it frees consumers to take responsibility for their own risks and well-being by making their own health care choices, with the resulting consumer demands and supplier competition producing prices that induce efficient health resource allocation. However, conflicts between competitive underwriting[3] and social sentiments for fairness in ex post outcomes, moral hazard,[4] and adverse selection[5] often lead in practice to private insurance markets that do a poor job of pooling risk and promoting efficient resource allocation, with limited consumer choice and incomplete or unaffordable coverage. The public often becomes the "insurer of last resort" for health care, with much to lose if the market breaks down.[6] The public response is to regulate, and in some cases subsidize, these markets to guarantee access and affordability. The most comprehensive attempt in the United States at such a public/private partnership is the new Medicare Part D market for prescription drug insurance for seniors, a bellwether for market-based health care reforms in this country. A study of this market by Florian Heiss, Daniel McFadden, and Joachim Winter (hereafter, HMW) asks whether publically managed competitive health insurance markets on the Part D model are sufficiently successful in offering choice, efficiency, and equity to make them an attractive alternative to a "single payer" system; see Winter et al. (2006), Heiss et al. (2006, 2007, 2009), and McFadden et al. (2008).

The Medicare Part D program works through voluntary enrollment in one of a menu of private plans. Immediately prior to the start of the program in 2006, HMW surveyed consumers and asked their enrollment intentions. That survey also collected data on prescription drug use, which determines whether the program would be immediately beneficial to a risk-neutral consumer. Immediately after the open enrollment period

[3] Competitive underwriting sets premiums equal to the actuarially fair expected value of benefits to be paid under a policy, conditioned on all available information on the consumer, plus a loading for administrative costs and profit that is determined competitively.

[4] Moral hazard occurs when a consumer insured against risk does not face the full marginal cost of losses, and as a consequence devotes less effort to avoiding or minimizing loss. When this level of effort is hidden from the insurer, the consumer does not carry the full burden of or fully pay for the consequences of her actions, and is also unable to certify or be rewarded fully for effort that reduces expected loss. Insurers sometimes attribute moral hazard to a consumer's "carelessness, incompetence, recklessness, indifference to loss, dishonesty, or fraudulent nature". Adjusting effort to its marginal reward is rational for a consumer, and unless the consumer breaches contract terms, is not dishonest. However, hidden effort impedes market efficiency by precluding underwriting based on this effort; see Spence and Zeckhauser (1971).

[5] Belli (2001) characterizes adverse selection as "strategic behavior by the more informed partner in a contract, against the interest of the less informed partner. In the health insurance market it is relevant because each individual chooses among the set of contracts offered by insurance companies according to his/her expected probability of using health services. In brief, those who foresee an intense use of health services will tend to choose more generous plans than those who expect a more limited use of them. In the extreme, for each premium and degree of coverage, those who will decide to purchase that particular health insurance contract are those who expect to have health expenditure greater or equal to the premium paid. Then, whatever the premium, the insurance company may end up with a loss on each customer".

[6] In many insurance markets, the government effectively becomes an insurer of last resort through reinsurance and guarantees, and moves to minimize counter-party risk. A current case in point is pellaton-like behavior by major financial institutions in the face of a Ponzi scheme of laundering risk using unregulated derivatives. Stabilizing regulation is needed in such markets to avoid chaotic unraveling in response to large shocks.

Table 1 Medicare part D prescription drug insurance enrollment behavior

2005 Drug bill	Enrollment choice	Total (%)	2006 Net benefit		Expected present value		Irrational	
			Negative (%)	Positive (%)	Negative (%)	Positive (%)	Min (%)	Max (%)
$0	No	36.3	32.3	**4.0**	5.8	**30.5**		
	Yes	63.7	56.9	6.8	**10.9**	52.9		
	Total	14.1	89.1	10.9	16.6	83.4	**4.0**	**41.3**
(0,$1250]	No	19.4	3.9	**15.5**	0.0	**19.4**		
	Yes	80.6	12.8	67.9	**0.1**	80.5		
	Total	22.2	16.7	83.3	0.1	99.9	**15.5**	**19.5**
($1250,∞)	No	5.7	0.0	**5.7**	0.0	**5.7**		
	Yes	94.3	0.2	94.2	**0.0**	94.3		
	Total	63.7	0.2	99.8	0.0	100.0	**5.7**	**5.7**
All	No	13.0	5.4	**7.6**	0.8	**12.2**		
	Yes	87.0	11.0	76.0	**1.6**	85.4		
	Total	N = 721	16.4	83.6	2.4	97.6	**7.6**	**13.8**

ended, HMJ surveyed these consumers again and asked their enrollment choices. The program was new and complex, and the consequences of choices ambiguous, so that consumers were at risk of procrastinating past the enrollment period, or of making poor decisions. Table 1 summarizes the findings from this study on enrollment behavior among those who had to make an active enrollment decision. The table is weighted to correct for attrition; see McFadden et al. (2006).

In this table, active deciders are classified by their annual pharmacy bills in 2005. Within each pharmacy bill category, the percentages enrolling are given in the third column, with the "Total" rows giving the percentage of the sample in each category. Columns 4 and 5 break the sample down by whether, given their 2005 age, health status, and pharmacy bills, enrollment is expected to have an immediate positive expected net benefit in 2006. Consumers who fail to enroll in the face of an immediate expected net benefit are fairly clearly irrational. Columns 6 and 7 break the sample down by whether a dynamic stochastic program gives a net positive expected present value for immediate enrollment, taking into account the consumer's expected health and mortality, and the penalties for delayed enrollment. This program uses health status and prescription drug use transitions estimated from the Medicare Current Beneficiary Study, a rotating panel of 40,000 consumers enrolled in Medicare. Consumers who fail to enroll when this expected present value is positive, or enroll when this expected present value is negative, are probably irrational, although it is possible that some of these classifications are due to subjective beliefs, discount rates, or private information that are not reflected in the dynamic stochastic program. In principle, risk aversion would induce higher enrollment rates than the maximization of expected present value would predict. However, the predicted enrollment rates are sufficiently high to make the possible impact of risk aversion very small. Boldface entries in the table correspond to choices that are arguably irrational. Columns 8 and 9 give lower and upper bounds

on the percentage of consumers making irrational choices. The overall conclusion of the study is that only a small minority of consumers, between 7.6 and 13.8 percent, made enrollment decisions that were clearly contrary to their self-interest. Further, many of these were consumers with low or zero prescription drug use in 2005 for whom the consequences of a non-optimal choice were small in expected present value terms. There is however, a hard core of about 7.6 percent of consumers who failed to enroll in the face of substantial immediate incentives to do so. These results are consistent with the proposition that most, but not all, consumers faced with substantial incentives respond rationally, but there is a fringe who without assistance will make choices that are clearly not in their self-interest.

7 Conclusions

In overview, I conclude that when incentives are large, consumer behavior shows little deviation from rationality, not only in familiar choice settings, but surprisingly even in complex, unfamiliar ones. There are exceptions. The quality of decision-making is heterogeneous, and there will usually be a fringe of consumers who are unable to get it right. When choices involve remote future consequences, uncertainty, or affect, this fringe grows. However, when incentives are small or unclear, less effort goes into determining best choices, and irrelevant factors play a larger role. Consumers are surprisingly truthful in circumstances where they don't need to be, but they may not supply the concentration and effort required to be accurate. Unfortunately, most economic surveys fit the case of small or unclear incentives, with little built-in control of effort and accuracy. The use of incentive theory, for example the Philipson and Malani (1999) suggestion to reward responses that are validated, is a promising avenue for bringing economic consumers up to the task of providing the information needed to implement the broad program of mechanism design set out by Leo Hurwicz, Jean-Jacques Laffont, and others for organization of resource allocation for public projects and private goods in a world of imperfect information. However, inconsistency in consumer response to incentives, particularly when their consequences are perceived as small or ambiguous, appears to be a problem that needs to be taken into account in drawing policy conclusions from principal-agent theory.

References

Arrow K (1953) Le role des valeurs boursieris pour la repartition la meilleure des risques. Econometrie 11:41–47

Attanasio O, Barr A, Cardenas J (2006) Playing games in surveys. UCL working paper

Attiyeh G, Franchosi R, Issac R (2000) Experiments with the pivot process for providing public goods. Public Choice 102:95–114

Azevedo C, Herriges J, Kling C (2003) Combining revealed and stated preferences: consistency tests and their interpretation. Am J Agric Econ 85:525–537

Belli P (2001) How adverse selection affects the health insurance market. Policy research working paper 2574
Ben-Akiva M, Morikawa T (1990) Estimation of switching models from revealed preferences and stated intentions. Trans Res A 24:485–495
Bennet M, Provencher B, Bishop R (2004) Experience, expectations and hindsight: evidence of a cognitive wedge in stated preference retrospectives. University of Wisconsin Staff Paper 468
Binswinger H (1980) Attitudes toward risk: experimental measurement in rural India. Am J Agric Econ 62:395–407
Bohm P (1972) Estimating demand for public goods: an experiment. Eur Econ Rev 2:111–130
Camerer C, Hogarth R (1999) The effects of financial incentives in experiments: a review and capital-labor-production framework. J Risk Uncertain 19:7–42
Carlsson F, Martinsson P (2001) Do hypothetical and actual marginal willingness to pay differ in choice experiments. J Environ Econ Manag 41:179–192
Carson R, Flores N, Meade N (2001) Contingent valuation: controversies and evidence. Environ Resour Econ 19:173–210
Carson R, Groves T (2007) Incentive and informational properties of preference questions. Environ Res Econ 37:181–210
Cason T, Saijo T, Sjostrom T, Yamato T (2003) Secure implementation experiments: do strategy-proof mechanisms really work? California institute of technology working paper
Cason T, Saijo T, Yamato T, Yolotani K (2003) Non-excludable public good experiments. Purdue University discussion paper
Chakraborty I, McAfee P (2008) Let the punishment fit the crime: enforcement with error. California institute of technology working paper
Champ P, Bishop R (2001) Donation payment mechanisms and contingent valuation: an experimental study of hypothetical bias. Environ Res Econ 19:383–402
Champ P, Flores N, Brown T, Chivers J (2002) Contingent valuation and incentives. Land Econ 78:591–604
Chen Y, Plott C (1996) The Groves-Ledyard mechanism: an experimental study of institutional design. J Public Econ 59:335–364
Chesher A (2005) Nonparametric identification under discrete variation. Econometrica 73:1525–1550
Chipman J, Moore J (1990) Acceptable indicators of welfare change: consumer's surplus analysis, and the gorman polar form. In: Chipman J, McFadden D, Richter M (eds) Preferences, uncertainty, and optimality: essays in honor of Leonid Hurwicz. Westview Press, Boulder
Chipman J, Moore J (1980) Compensating variation, consumer's surplus, and welfare. Am Econ Rev 70:933–949
Clarke E (1971) Multipart pricing of public goods. Public Choice 11:19–33
Cummings R, Elliott S, Harrison G, Murphy J (1997) Are hypothetical referenda incentive compatible. J Polit Econ 105:609–621
Cummings R, Harrison G, Rutstrom E (1995) Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive-compatible. Am Econ Rev 85:260–266
Davis R (1963) Recreation-Planning as an economic problem. Nat Resour J 3:239–249
Debreu G (1959) The theory of value. Cowles Commission Monograph
Diamond P (1967) The role of a stock market in a general equilibrium model with technological uncertainty. Am Econ Rev 62:759–776
Diamond P, Mirrlees J (1971) Optimal taxation and public production. Am Econ Rev 61:8–27 and 261–278
Diamond P, Mirrlees J (1972) Optimal taxation and public production. Am Econ Rev 62:238
Diamond P, Hausman J (1994) Contingent valuation: is some number better than no number. J Econ Perspectives 8:45–64
Dubin J, Graetz M, Wilde L (1992) State income tax amnesties: causes. Quart J Econ 107:1057–1070
Fox J, Shogren J, Hayes D, Kliebenstein J (1998) CVM-X: calibrating contingent values with experimental auction markets. Am J Agric Econ 80:455–465
Friedman D, Rust J (1993) The double auction market: institutions, theories, and evidence. Addison-Wesley, Reading
Fehr E, Gachter S (2004) Cooperation and punishment in public goods experiments. Am Econ Rev 90:980–994
Fehr E, Fischbacher U (2004) Social norms and human cooperation. Trends Cogn Sci 8:185–190
Fehr E, Fischbacher U, von Rosenbladt B, Schupp J, Wagner GG (2002) A nation-wide laboratory: examining trust and trustworthiness by integrating behavioural experiments into representative surveys. Schmollers Jahrbuch 122:519–542

Fehr E, Fischbacher U (2002) Why social preferences matter: the impact of non-selfish motives on competition, cooperation, and incentives. Econ J 112:1–33
Fehr E, Falk A (2002) Psychological foundations of incentives. Eur Econ Rev 46:687–724
Fehr E, Schmidt K (1999) A theory of fairness, competition, and cooperation. Quart J Econ 69:817–868
Feinstein J (1990) Detection controlled estimation. J Law Econ 33:233–276
Feinstein J (1991) An econometric analysis of income tax evasion and its detection. RAND J Econ 22:14–35
Feinstein J (1999) Approaches for estimating noncompliance: examples from federal taxation in the United States. Econ J 109:F360–F369
Fox J, Shogren J, Hayes D, Kliebenstein J (2003) CVM-X: calibrating contingent values with experimental auction markets. In: Shogren J (ed) Experiments in environmental economics, vol 1. International library of environmental economics and policy, Ashgate pp 445–455
Frykblom P (2000) Willingness to pay and the choice of question format: experimental results. Appl Econ Lett 7:665–667
Frykblom P, Shogren J (2000) An experimental testing of anchoring effects in discrete choice questions. Environ Resour Econ 16:329–341
Garratt R, Walker M, Wooders J (2004) Behavior in second-price auctions by highly experienced eBay buyers and sellers. U.C. Santa Barbara working paper
Gibbard A (1973) Manipulation of voting schemes: a general result. Econometrica 41:587–602
Green D, Jacowitz K, Kahneman D, McFadden D (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods. Resour Energy Econ 20:85–116
Green J, Laffont J (1986) Partially verifiable information and mechanism design. Rev Econ Stud 53:447–456
Green J, Laffont J (1979) Incentives in public decision making. Amsterdam, North Holland
Green J, Laffont J (1978) A sampling approach to the free rider problem. In: Sandmo A (ed) Essays in public economics. Lexington Books, Lexington
Green J, Laffont J (1977) Characterization of satisfactory mechanisms for the revelation of preferences for public goods. Econometrica 45:427–438
Grossman S, Hart O (1983) An analysis of the principal-agent problem. Econometrica 51:7–45
Groves T, Ledyard J (1980) The existence of efficient and incentive compatible equilibria with public goods. Econometrica 48:1487–1506
Groves T, Ledyard J (1977) Optimal allocation of public goods: a solution to the 'Free Rider' problem. Econometrica 45:783–809
Groves T, Loeb M (1975) Incentives and public inputs. J Publ Econ 4:211–226
Haab T, Huang J, Whitehead J (1999) Are hypothetical referenda incentive compatible? A Comment. J Polit Econ 107:186–196
Harstad R (1990) Dominant strategy adoption, efficiency and bidders' experience with pricing rules. Experimental Economics, pp 261–280
Heckman J (1979) Sample selection bias as a specification error. Econometrica 47:153–161
Heiss F, McFadden D, Winter J (2006) Who failed to enroll in Medicare, and why. Health Aff 25:w344–w354
Heiss F, McFadden D, Winter J (2007) 'Mind the gap! consumer perceptions and choices of medicare part D prescription drug plans'. NBER working paper 13627
Heiss F, McFadden D, Winter J (2009) 'The workings of a private health insurance market: enrollment decisions, plan choice, and adverse selection in medicare part D'. University of Munich working paper
Hoehn J, Randall A (1987) A satisfactory benefit-cost indicator for contingent valuation. J Environ Econ Manag 3:1441–1482
Hurwicz L (1960) Optimality and informational efficiency in resource allocation. In: Arrow K, Karlin S, Suppes P (eds) Mathematical methods in social sciences. Stanford University Press, Stanford
Hurwicz L, Radner R, Reiter S (1975) A stochastic decentralized resource allocation process. Econometrica 43:187–221 and 363–93
Imbens G, Manski C (2004) Confidence intervals for partially identified parameters. Econometrica 72:1845–1857
Imbens G, Newey W (2002) Identification and estimation of triangular simultaneous equations models without additivity. NBER working paper 285
Jewitt I (1988) Justifying the first-order approach to principal-agent problems. Econometrica 56:1177–1190
Johannesson M, Liljas B, Johansson P (1998) An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions. Appl Econ 30:643–647

Kahneman D, Ritov I, Schkade D (1999) Economic preferences or attitude expressions? an analysis of dollar responses to public issues. J Risk Uncertain 19:203–235
Kawagoe T, Mori T (2001) Can the pivotal mechanism induce truth-tessing? an experimental study. Publ Choice 108:331–354
Klemperer P (2004) Auctions: theory and practice. Princeton University Press, Princeton
Krishna V (2002) Auction theory. Academic Press, San Diego
Kim Y (2005) Audit misperception, tax compliance, and optimal uncertainty. J Publ Econ 7:521–541
Laffont J, Martimort D (2002) The theory of incentives. Princeton University Press, Princeton
Laffont J, Martimort D (2000) Mechanism design with collusion and correlation. Econometrica 68:309–342
Ledyard J, Palfry T (1994) Voting and lottery drafts as efficient public goods mechanisms. Rev Econ Stud 61:327–355
List J (2003) Does market experience eliminate market anomalies. Quart J Econ 118:41–71
List J (2002) Using choice experiments to value non-market goods and services. University of Maryland working paper
List J, Lucking-Reiley D (2002) Bidding behavior and decision costs in field experiments. Econ Inquiry 40:611–619
List J, Gallet C (2001) What experimental protocol influence disparities between actual and hypothetical stated values? Environ Resour Econ 20:241–254
Loomis J, Brown T, Lucero B, Peterson G (1997) Evaluating the validity of the dichotomous choice question format in contingent valuation. Environ Resour Econ 10:109–123
Loomis J, Brown T, Lucero B, Peterson G (1996) Improving validity experiments of contingent valuation methods: results of efforts to reduce the disparity of hypothetical and actual willingness to pay. Land Econ 72:450–461
Louviere J, Hensher D, Swait J (2000) Stated choice methods: analysis and application. Cambridge University Press, Cambridge
Louviere J, Hensher D, Swait J (1999) Conjoint analysis methods in the broader context of preference elicitation methods. In: Gustafson A, Herman A, Huber F (eds) Conjoint measurement: methods and applications. Springer, Berlin, pp 279–318
Lusk J, Schroeder T (2004) Are choice experiments incentive compatible? a test with quality differentiated beef steaks. Am J Agric Econ 86:467–482
Manski C (2005) Social choice with partial knowledge of treatment response. Princeton University Press, Princeton
Maskin E (2004) The unity of auction theory. J Econ Lit 42(4):1102–1115
Matzkin R (2006) Nonparametric identification. In: Heckman J (ed) Handbook of Econometrics VI (forthcoming)
Matzkin R (1992) Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. Econometrica 60:239–270
McFadden D (2009) Economic jury recruitment and management. U.C. Berkeley working paper
McFadden D, Heiss F, Winter J (2008) Consumer-Directed health care: can consumers look after themselves. Swiss J Econ Stat 144:287–309
McFadden D, Heiss F, Jun B, Winter J (2006) On testing for independence in weighted contingency tables. Medium Econom Appl 14:11–19
McFadden D (2004) Welfare economics at the extensive margin: giving gorman polar consumers some latitude. U.C. Berkeley working paper, June 2004
McFadden D (1994) Contingent valuation and social choice. Am J Agric Econ 76:689–708
McFadden D (1980) Econometric models for probabilistic choice among products. J Bus 53:S13–S29
Milgrom (2004) Putting auction theory to work. Cambridge University Press, Cambridge
Mirrlees J (1971) An exploration in the theory of optimum income taxation. Rev Econ Stud 38:175–208
Morgan J (2000) Financing public goods by means of lotteries. Rev Econ Stud 67:761–784
Palfrey T, Srivastava S (1991) Nash implementation using undominated strategies. Econometrica 59:479–501
Palfrey T, Srivastava S (1989) Mechanism design with incomplete information: a solution to the implementation problem. J Polit Econ 97:668–691
Palfrey T, Rosenthal H (1990) Testing game-theoretic models of free-riding: new evidence on probability bias and learning. California institute of technology working paper
Palfrey T, Rosenthal H (1994) Repeated play, cooperation, and coordination: an experimental study. Rev Econ Stud 61:545–565

Pecorino P, Temimi A (2007) Lotteries, group size, and public good provision. J Public Econ Theory 9:451–466
Philipson T (2001) Data markets, missing data, and incentive pay. Econometrica 69:1099–1111
Philipson T (1997) Data markets and the production of surveys. Rev Econ Stud 64:47–72
Philipson T, Malani A (1999) Measurement errors: a principal investigator-agent approach. J Econom 91:273–298
Poe G, Clark J, Rondeau D, Schulze W (2002) Provision point mechanisms and field validity tests of contingent valuation. Environ Res Econ 23:105–131
Radner R (1972) Existence of equilibrium of plans, prices, and price expectations in a sequence of markets. Econometrica 40:289–303
Randall A, Hoehn J, Brookshire D (1983) Contingent valuation surveys for evaluating environmental assets. Nat Resour J 23:635–648
Randall A, Ives B, Eastman C (1974) Bidding games for valuation of aesthetic environmental improvements. J Environ Econ Manag 1:132–149
Rondeau D, Schulze W, Poe G (1999) Voluntary revelation of the demand for public goods using a provision point mechanism. J Publ Econ 72:455–470
Ryu E, Couper M, Marans R (2005) Survey incentives: cash vs In-Kind; Face-to-Face vs. Mail; Response Rate vs Non-Response error. Int J Publ Opinion Res 18:90–106
Saez E (2001) Using elasticities to derive optimal income tax rates. Rev Econ Stud 68:205–229
Satterthwaite M (1975) Strategy-Proofness and arrow's conditions: existence and corresponence theorems for voting procedures and social welfare functions. J Econ Theory 10:187–217
Shafir E, Tversky A (1992) Thinking through uncertainty: nonconsequential reasoning and choice. Cogn Psychol 24:449–474
Shen J (2005) A review of stated choice method. Osaka University working paper
Simon H (1971) Computers, communications and the public interest. In: Martin Greenberger (ed) Computers, communications, and the public interest. Johns Hopkins Press, Baltimore, pp 40–41
Smith V (1980) Experiments with a decentralized mechanism for public good decisions. Am Econ Rev 70:584–599
Spence M, Zeckhauser R (1971) Insurance, information, and individual action. Am Econ Rev 61:380–387
Train K, Wilson W (2005) Econometric analysis of stated-preference experiments constructed from revealed-preference choices. University Of California, berkeley working paper
Tversky A, Sattath S, Slovic P (1988) Contingent weighting in judgment and choice. Psychol Rev 95:371–384
Varian H (1974) A Bayesian approach to real estate assessment. In: Fienberg S, Zellner A (eds) Studies in bayesian econometrics and statistics. North-Holland, Amsterdam, pp 195–208
Venkatachalam L (2004) The contingent valuation method: a review. Environ Impact Assess Rev 24:89–124
Whitehead J (2002) Incentive incompatibility and starting-point bias in iterative valuation decisions. Land Econ 78:285–297
Willis K, Powe N (1998) Contingent valuation and real economic commitments: a private good experiment. J Environ Plan Manag 41:611–619
Winter J, Balza R, Caro F, Jun B, Matzkin R, McFadden D (2006) Medicare prescription drug coverage: consumer information and preferences. Procee Natl Acad Sci 103:7929–7934