Neurometrically Informed Mechanism Design *

Ian Krajbich[1,3]
John Ledyard[1]
Colin Camerer[1,2]
Antonio Rangel[1,2]

[1] HSS, California Institute of Technology, Pasadena, CA 91125
[2] CNS, California Institute of Technology, Pasadena, CA 91125
[3] Department of Economics, University of Zurich
* We thank Phil Reny for his helpful comments on earlier drafts.

Abstract. Several classic results have shown that it is impossible to design mechanisms that simultaneously satisfy efficiency, voluntary participation, and dominant strategy incentive compatibility. Modern neurometric technologies are making it possible to obtain noisy signals of subjects' preferences that could be used in the mechanism design problem. Here we relate the use of informative neurometric signals to existing theory in mechanism design and then show experimentally that such signals can be used to overcome the impossibility results. In particular, we show how to construct neurometrically informed mechanisms that use informative signals to incentivize agents to truthfully reveal their types. These mechanisms simultaneously satisfy efficiency, voluntary participation, and dominant strategy incentive compatibility, and are shown to elicit truth-telling behavior from subjects in two different public goods environments.

A. Introduction

In the classical mechanism design problem, a social planner wants to implement an allocation that maximizes some notion of social welfare given the preferences of the group. The planner's problem is difficult, however, because he does not have direct knowledge about the individuals' preferences. The best he can do is to implement a mechanism, which is characterized by a message space for each subject and a function mapping messages to outcomes. The mechanism solves the planner's problem if the outcomes induced in equilibrium, as a function of the underlying preferences, are the ones that the planner wants to implement.

A central preoccupation in the mechanism design literature has been to design mechanisms that satisfy three fundamental properties: efficiency, dominant strategy incentive compatibility, and voluntary participation. Efficiency requires that the outcome induced by the mechanism maximize the group's net expected utility (based on the underlying preferences) while also balancing the budget. This property is a basic requirement for a desirable mechanism. Voluntary participation requires that the expected utility from participating in the mechanism be positive for every subject regardless of her preferences and the actions of the other individuals. This property is desirable because it implies that all subjects benefit by participating. Finally, dominant strategy incentive compatibility requires that every subject choose the message that generates the planner's desired outcome, regardless of the messages and preferences of the other subjects. This property is highly desirable because it implies that subjects have a strong incentive to comply with the mechanism, even in the absence of common knowledge of rationality or beliefs.

Early on, a series of classic impossibility results showed that in most circumstances it is impossible to design a mechanism that satisfies efficiency, voluntary participation, and dominant strategy incentive compatibility (Hurwicz, 1972, Gibbard, 1973; Satterthwaite, 1975). The literature reacted to these impossibility results by investigating what could be

achieved if one or more of the properties were relaxed. For example, several articles have investigated what can be achieved if the requirement of voluntary participation is foregone and the dominant strategy requirement is relaxed to a Bayesian (d'Aspermont and Gerard-Varet, 1979) or a Nash Equilibrium (Groves and Ledyard, 1977; Maskin, 1999). Others have investigated relaxing the requirement of efficiency (Vickrey, 1961; Clarke, 1972; Groves, 1973). Although many of these results have generated profound insights into the nature of institutions and incentives, the solutions that they have generated fall short of the ideal criteria that originally motivated the literature.

More recently, work by Cremer and McLean, 1985 and 1988, and McAfee and Reny, 1992 showed using Bayesian incentive compatibility that "introducing arbitrarily small amounts of correlation into the joint distribution of private information among the players is enough to render private information valueless" (McAfee and Reny 1992, p. 395) and to allow the mechanism designer to fully extract expected rents. This is equivalent to achieving expected efficiency and voluntary participation. Their examples and applications are focused on situations with correlated information and Bayes incentive compatibility under common knowledge of beliefs and rationality. In a related paper, Riordan and Sappington (1988) derive conditions in which public ex-post signals of a seller's costs can be used to achieve a first-best outcome for the buyer.

This literature has been criticized for its reliance on the assumptions of common knowledge and 1-1 signals, which are extremely unlikely to hold in real situations, and because these mechanisms are highly susceptible to collusion among the agents. Furthermore, it has been shown that these assumptions are critical to the mechanisms' success and that without them it is generically impossible to achieve first-best outcomes (Neeman 2004, Heifetz & Neeman 2006, Robert 1991, Laffont & Martimort 2000, Barelli 2009). Indeed, McAfee & Reny themselves wrote that their results "cast doubt on the value of the current mechanism design paradigm as a model of institutional design" (p.400).

However, very recent research in neuroeconomics (see Final Remarks) has begun to shed light on the neural mechanisms responsible for subjective valuation, and with an understanding of these mechanisms it should be possible to measure neurally-based signals that correlate with value. Below, we will see that, with such "neurometric" signals, the common knowledge and collusion criticisms do not apply, and the 1-1 criticism is less likely to apply.

Because of major advances in neuroscience, it may soon be technologically feasible to implement neurometrically informed mechanisms (NIM), where an individual's tax depends both on their reported type and on an observed, stochastic signal of their type. It is straightforward to show that if these neurometric signals obey the basic conditions set forth by Cremer & McLean, McAfee & Reny, and Riordan & Sappington, we can combine their results, dispense with common knowledge assumptions and recover dominant strategy implementation for any mechanism design problem. For example, in Krajbich et al., 2009 we showed that experimentally induced valuations in a binary value public goods game could be predicted with 60% accuracy using a combination of fMRI and machine learning techniques. We then used that technology to implement a NIM where subjects truthfully revealed their types nearly 100% of the time.

Given the success of the "proof of concept" experiment and our rapidly advancing understanding of neuroscience, it is now important to investigate the practicality of implementing mechanisms that rely on noisy signals of individuals' types. Here, we use two distinct public goods experiments to investigate this class of mechanisms.

In the first experiment we test whether subjects adhere to the dominant strategy when there are many different types and actions available. This is an important test because the whole advantage of NIM depends on subjects understanding the details of the mechanism and following the dominant strategy of truth-telling, which is not always experimentally observed (need citations). It also serves as a useful illustration of how to construct these types of mechanisms using fairly simple proper scoring rules.

In the second experiment we test whether subjective voluntary participation is satisfied when players are risk- or loss-averse, by allowing players to opt-out of the mechanism. This is a critical issue for practical implementation since we know that most people are risk- and loss-averse, and these people need to be willing to adopt NIM in order for them to be useful. Simply assuring them of positive expected payoffs may not be enough to overcome high levels of risk- and loss-aversion in the face of stochastic payoffs.

Overall our results show that the mechanisms are robust to the large type and action space and to the degrees of loss and risk-aversion observed in most of our sample.

The paper is organized as follows. In Section B we review the basic impossibility results from classical mechanism design theory and introduce neurometrically informed mechanisms. In Section C we present the results of our two experiments. In Section D we discuss the scope and limitations of our results.

B. Theory

Consider environments with N individuals indexed by $i=1,\ldots,N$. Individuals have quasi-linear preferences denoted, $u_i(x,v_i)-t_i$, where $v_i \in V^i$ denotes player i's type, $x \in X \subset \Re^L$ is the allocation of resources for the group, and $t_i \in \Re$ denotes a payment from player i. We assume that the set of types for each individual, V^i, is a finite set. X denotes the set of feasible allocations. A mechanism is given by a message space $M = M_1 \times \ldots \times M_N$ and an outcome function g. Each individual i reports a message $m_i \in M_i$. By the Revelation Principle (A. Gibbard 1973, and R. Myerson, 1981) we can, without loss of generality, focus only on mechanisms in which $M^i = V^i$ for all i and for which $m^*(v)=v$ is a dominant strategy for all v.

We assume the planner is able to observe a neurometric signal $s \in S$ for each individual *after* they have announced their type. S is assumed to be a finite set.[1] The signals are distributed according to a density function conditional on the true types. A signal technology for an environment is thus given by a mapping $T : V \rightarrow \Delta(S)$ where $\Delta(S)$ is the set of probability densities on S. For the signaling technology to be useful, the signals have to be sufficiently informative. We assume that T is *1-1* from V to $T(V)$. Note that the signal likelihood function does not depend on the subject's messages. We assume that the subject does not know what signal the planner will observe, nor can the subject manipulate the signal in any way. All he knows is the density function $T(v_i)$.

The availability of the signals allows us to introduce an additional tax function that depends on both the signal and the subjects' reports. In particular, for any direct mechanism (V,g) we can define a *neurometrically informed mechanism* (NIM) (V,g,w), in which the outcome function is given by

$$g(s,m) = (x(m), t(m) + w(s,m)) \in X \times \Re^N$$

where $w(s,m)$ denotes the augmented tax.

A particularly useful class of neurometrically informed mechanisms are those constructed using person-by-person augmentation. In the first step we compute each individual's provisional tax $r(s_i,m_i)$ which only depends on her own message and signal. Second, to ensure efficiency, the surplus raised $\sum_{i=1}^{N} r(s_i, m_i)$ is redistributed back equally to all the individuals. This generates an augmented tax function

$$w_i(s,m) = \left(\frac{N-1}{N}\right)\left[r(s_i,m_i) - \left(\frac{1}{N-1}\right)\left(\sum_{j \neq i} r(s_j,m_j)\right)\right]$$

[1] This is done for ease of exposition. The results in this paper also hold for larger signal spaces.

The set of environments for the mechanism design problem is given by $\left\{X, u_1, ..., u_N, V^1, ..., V^N\right\}$. An allocation $x \in X$ is efficient in a quasi-linear environment, given the individuals' types $v = (v_1, ..., v_N)$, if and only if it maximizes $\sum_{i=1}^{N} u_i(x, v_i)$ and $\sum_i t_i + w_i = 0$. Note that efficient allocations satisfy the Pareto property, so there is no other feasible allocation that can make everyone better off.

The mechanism design literature has focused on direct revelation mechanisms satisfying three basic properties: efficiency, incentive compatibility[2] and voluntary participation. A direct mechanism (V,g,w) is <u>efficient</u> if and only if the allocation $(x(v),t(v),w(v))$ is efficient for all v.

The mechanism (V,g,w) is <u>dominant strategy incentive compatible</u> if and only if for all i, v, and m, we have that

$$u(x(v_i, m_{-i}), v_i) - t(v_i, m_{-i}) - E(w_i(s, v_i, m_{-i}) \mid v_i) \geq u(x(m_i, m_{-i}), v_i) - t(m_i, m_{-i}) - E(w_i(s, m_i, m_{-i}) \mid v_i)$$

where E denotes the expectation operator. This condition ensures that each individual's utility is highest by reporting $m_i = v_i$ regardless of the types and reports of the other players.

The mechanism (V,g,w) satisfies <u>voluntary participation</u> if and only if, for all i, v_i, and m_i,

$$u(x(v_i, m_{-i}), v_i) - t(v_i, m_{-i}) - E(w_i(s, v_i, m_{-i}) \mid v_i) \geq 0 .$$

This condition ensures that each individual receives a non-negative payoff from truthful reporting, regardless of the types and reports by the other players.

[2] From this point on we will use the phrases "dominant strategy incentive compatibility" and "incentive compatibility" interchangeably.

If the signals are uninformative (e.g. if there are no signals), then we are left with the following classic mechanism design result:

Fundamental Impossibility Theorem of Mechanism Design: *If V is rich enough, then there is no mechanism that is Efficient and Incentive Compatible and satisfies Voluntary Participation.*

A number of variations of this theorem have been provided. Hurwicz (1972) provides one of the first results for exchange environments. Gibbard (1973) and Satterthwaite (1975) show that the result holds when *V* includes all preferences over at least 3 alternatives. Green and Laffont (1979) consider quasi-linear preferences but include all possible preferences for the non-linear part. Walker (1980) shows that that the result still holds for quasi-linear preferences for which the non-linear part is concave. Hurwicz and Walker (1990) provide the most general version of the result for quasi-linear environments.

With this impossibility theorem, the original search for mechanisms that satisfy dominant strategy incentive compatibility, efficiency, and voluntary participation seemed doomed to failure. However, with the availability of a signaling technology, we can make use of results from Cremer & McLean, McAfee & Reny, and Riordan and Sappington to overcome the impossibility result.

In brief, with an informative signaling technology, having an agent reveal their type, *v,* is equivalent to having them reveal the probability distribution *T(v).* There is a well-known way to induce revelation of a probability density using proper scoring rules (G.W. Brier, 1950, T. Gneiting and A.E. Raftery, 2007). To create our augmented mechanism, we simply let $r(s_i, m_i)$ be a proper scoring rule. To use a scoring rule to induce agents to reveal their true types, we must make the incentives to reveal at least as great as the gains from non-revelation. This can be done with any proper scoring rule by simply increasing

the magnitude of the payoffs until the expected penalty for any misreport is greater than the corresponding benefit from misreporting in the signal-less mechanism.[3]

While any proper scoring rule can lead to an incentive compatible augmented mechanism, it is not necessarily true that the mechanism will then also satisfy voluntary participation. A sufficient condition that ensures voluntary participation, is for $E(r(s,v)) =$ C for all v. This can be guaranteed if the signal satisfies the convex independence condition of Cramer and McLean (1985):

CM Condition: The signal technology satisfies $\forall v' \in V, T(v) \notin co\{T(v)\}_{v \in V, v \neq v'}$ where coA is the convex hull of A.

One simple example arises when the supports of $T(v)$, $\bar{S}(v) = \{s \in S \mid T_s(v) > 0\}$, are different for all v. When $v \neq v' \Rightarrow \bar{S}(v) \neq \bar{S}(v')$, we can use the simplest possible scoring rule, which assigns a score of 0 to any signal in the support of $T(v)$, and a score of -1 to any other signal. A second case occurs if the signal distribution maintains its shape and simply shifts around as v changes. If $T_s(v) = T_{s+\mu}(v+\mu)$ for all $\mu = v' - v, \forall v, v' \in V$, then for scoring rules such that $r(T(v),s)=r(T(v+\mu),s+\mu)$ it will be true that $E(r(T(v),T(v))) = E(r(T(v+\mu),T(v+\mu)))$. One such scoring rule is the logarithmic rule, $r(T(v),s)=ln\ T_s(v)$.

We summarize the possibility results for neurometrically informed mechanisms in

A Possibility theorem for neurometrically informed mechanisms. *Given a 1-1 signal technology $T:V \rightarrow \Delta(S)$ and quasi-linear environments with finite type spaces:*

a) *given any direct revelation mechanism (V,g), there is a person-by-person augmented mechanism (V,g,w) that is incentive compatible and yields the same expected outcome,*

[3] The precise details can be found in an appendix.

b) *given any direct revelation mechanism* (V,g) satisfying voluntary participation, if the *signal technology* $T:V \rightarrow S$ satisfies the CM condition, *there is a person-by-person augmented mechanism* (V,g,w) *that is incentive compatible, satisfies voluntary participation, and yields the same expected outcome.*

These results show that with access to a sufficiently informative signal technology any efficient allocation rule can be implemented in dominant strategies using a tax function that satisfies voluntary participation. This stands in sharp contrast with one of the most fundamental theorems of mechanism design which shows that, in the absence of such signals, there is no standard mechanism that is efficient, incentive compatible and satisfies voluntary participation.

Note that for neurometrically informed mechanisms, incentive compatibility relies on the timing of the signal technology. At the time i is choosing her message, she does not know what signal s will be observed by the planner and thus must base her choice on the likelihood function of the signaling technology. Incentive compatibility states that all individuals are willing to report their true types regardless of the reports by the other subjects, given the expected payoffs induced by the augmentations. Under our incentive compatibility condition, truth telling is the best response even if others misreport. We do not require common knowledge of either rationality or beliefs. We get dominant strategy behavior instead of ex-post incentive compatibility because the signals are measured directly from the individual, rather than from other individuals in the group. This means that unlike Cremer & Mclean and McAfee & Reny, these results hold both with independent types and with interdependent types.

Voluntary participation relies on quasi-linearity but also on the timing of the signal technology. At the time i is choosing her participation decision, she does not know what signals s will be observed by the planner and thus must base her choice on the likelihood function of the signaling technology. Voluntary participation states that all truth-telling individuals are willing to participate regardless of the reports by the other subjects, given

the expected payoffs induced by the augmentations. We are normalizing u so that the value of the individual's outside option is 0.

Notice several useful properties of the class of neurometrically informed mechanisms. First, all of the possibility results can be obtained from person-by-person augmentation, which makes the computation and description of the mechanisms relatively easy. Second, in the dominant strategy equilibrium where $m = v$, the expected utility to any individual is exactly the utility they get under honest reporting in the un-augmented mechanism. Thus, in the dominant strategy equilibrium the augmented taxes do not on average cause any wealth redistribution. Third, the mechanisms are balanced even *after* the signal. That is, they do not take any resources from the system, they only redistribute. However, although they satisfy voluntary participation *before* the signals are observed, they may not *after* the signals. We confront this experimentally in Section C.

We have shown that NIMs do not suffer from previous common knowledge and collusion criticisms of this literature since incentive compatibility and voluntary participation are dominant strategies. Indeed, the only knowledge requirement is that the planner knows the distribution of signals for each type, and that he/she communicates those distributions to the agents, who believe that distribution and believe that they cannot manipulate the signals. However, we have not fully addressed the additional criticism that the signaling technology may not be 1-1. This critical assumption requires that no two distinct types have identical signal distributions. What distinguishes two different types is that there exist two resource allocations x_A and x_B for which one type chooses x_A and for which the other type chooses x_B. In order to make such choices, both agents must evaluate options x_A and x_B, compare them, and make a choice. If we can assume that agents differ only in the values that they assign to options and *not* in the way that they compare the options (at least not in a way that systematically correlates with type) then we must conclude that the two agents' brains would respond differently while evaluating the two options. This difference in brain activity must be measurable, since it produces two different choices, and so in principle there must exist some signaling technology that is 1-1 with type.

It is important to note that in order to achieve any first-best allocation, the mechanism requires a signal space that is at least as large as the type space. What this means is that a simple lie-detector technology is only sufficient for an environment with two types. While it may be possible to achieve first-best allocations in situations where this requirement is not met, Riordan & Sappington (1988) tell us that we will be quite limited in what allocations we will be able to achieve. Other non-neurometric signals that one might naturally think of suffer from the same problem. In addition, non-neurometric signals are likely to either be manipulable or known to the agent. Either of these cases would render the signals useless to the planner, since the agent could then condition their response on the outcome of the signal.

C. Experiments: Neurometrically informed public goods.

In this section we present the results of two experiments designed to test the results described above. Testing the results with real subjects is important for two reasons.

First, the theory assumes that subjects are risk neutral, whereas experimental tests show that most subjects exhibit risk and/or loss aversion (Kahneman and Tversky, 1979). This is a potential problem for the theory because, since the signal technology is stochastic, subjects face a distribution of taxes, even if they report truthfully. As a result, voluntary participation and/or incentive compatibility in terms of expected utility might be violated for some individuals.

Second, although the neurometrically informed mechanisms that we have proposed are relatively simple, there is a concern about whether subjects will understand that it is in their best interest to report truthfully. This concern is justified, for example, by previous results showing that some subjects do not bid truthfully in second price auctions (Coppinger, Smith, and Titus, 1980).

Because the goal of these experiments is to investigate subjects' behavior in response to the incentives and the presence of an informative signal of value, we chose to conduct these experiments in a normal experimental economics laboratory, rather than in an fMRI scanner. To do so, we used the laboratory computers to simulate noisy signals correlated with each subject's type, *as if* it was being generated by an informative neurometric technology. Consequently, we had to make assumptions about the signal distributions, balancing simplicity with realism. In the first two-type experiment we assumed a signal distribution analogous to the distribution we actually measured from the fMRI scanner in Krajbich et al. 2009, but with a moderate improvement in overall accuracy. In the second multi-type experiment we assumed a censored uniform distribution of signals centered on the subjects' true type. While a normal distribution would have been more realistic than a censored uniform distribution, the simplicity of the uniform distribution allowed us to most clearly explain the signaling technology and the mechanism to the subjects and the reader. Future work will need to investigate the best way to construct NIMs based on different signal distributions.

C.1. Experiment 1

The first experiment is designed to test whether the voluntary participation and incentive compatibility properties of a variant of the neurometrically informed mechanism proposed in Krajbich et al. (2009) are robust to the introduction of typical levels of risk and loss aversion.

Environment. Groups of five subjects need to decide whether or not to produce a public good which has a cost of \$25. The preferences of subject i are given by $v^i g - y^i$, where g = 0,1 is the level of the public good, v^i={\$1,\$9} is the value for the public good, and y^i is the net tax paid by the individual.

We assume that the social planner has access to a signal technology with the following properties: $p(s = 1 \mid v = 1) = p(s = 9 \mid v = 9) = 0.8$ and $p(s = 1 \mid v = 9) = p(s = 9 \mid v = 1) = 0.2$.

In other words, the signal equals the true value with 80% probability and signals the other value with 20% probability.

Neurometrically informed mechanism. We consider a simple direct revelation mechanism in which subjects simultaneously report their values m^i = {\$1,\$9} and then the planner receives a signal for each of them s^i = {\$1,\$9}. The public good is ruled to induce efficiency over the reported values. Therefore, the public good is built whenever three or more subjects report a \$9 value. Each subject pays a gross tax $t^i(m^i) + r^i(s^i, m^i)$ as given by the following table:

		message	
		\$1	\$9
signal	\$1	receive \$3.67	pay \$9
	\$9	pay \$14.67	pay \$9

If the amount raised from the gross taxes does not equal the cost of the public good, then the difference is evenly collected from (deficit) or returned to (surplus) the five subjects. In Krajbich et al. (2009) we showed that the resulting mechanism satisfies efficiency, voluntary participation and dominant strategy incentive compatibility for risk neutral subjects.

Subjects. *N*=50 subjects participated in the study. They were recruited from Caltech's student population. In addition to their experimental earnings, subjects were paid \$20 for their participation.

Experimental task. In each experimental session, subjects played two rounds of the experimental task described below, once with a \$1 value for the public good, and once with a \$9 value (both presented simultaneously). In each round subjects had to make two sequential decisions after observing their private value. First, they had to cast a binary Yes-No vote for whether they wanted the neurometrically informed mechanism to be played. These votes mattered because with probability 0.2 * $I_{\#\text{No votes}}$ the mechanism was

not played. Second, they had to report a value m^i which was binding only if the mechanism was played.

Note the logic of the experiment. In the voting decision, it is a dominant strategy for the subject to vote Yes if and only if her expected utility from the neurometrically informed mechanism is positive. As a result, the vote decision allows us to determine for each subject whether the mechanism satisfies voluntary participation. Furthermore, since the vote does not affect the rules of the mechanism conditional on it being played, we can elicit this measure without affecting the subsequent incentives.

There was no feedback between rounds. The experiment was carried out with pen and paper. After the decisions were collected from all the subjects in the experimental session, each subject's decision was randomly paired with four other subjects from the group, and they were paid for one randomly chosen group outcome. This ensured that the distribution of values were independent across rounds and subjects.

The experimental instructions (included in Appendix I, available on-line) provide an in-depth description of the rules of the mechanism and the payments that they induced under different contingencies.

Measuring risk aversion. Immediately after the main experimental task, subjects were asked to play a simple gambling task designed to estimate their risk and loss aversion parameters (Wang, Camerer, Filliba, 2010). In particular, subjects made 40 different hypothetical binary choices between lotteries (Sokol-Hessner et al., 2009). Their responses were used to estimate the parameters of the following prospect theoretic model (Kahneman & Tversky, 1979) using Bayesian estimation:

$$U(x) = \begin{cases} x^{\rho} & \text{if } x \geq 0 \\ -\lambda \mid x \mid^{\rho} & \text{if } x < 0 \end{cases}$$

where ρ is a measure of the degree of risk aversion, and λ is a measure of the degree of loss aversion. Note that risk aversion is decreasing in ρ, but loss aversion is increasing in λ.

In a departure from Wang, Camerer and Filliba 2010, we paid subjects using a novel mechanism. We told subjects that while their 40 choices were hypothetical and would not count, we would use their hypothetical choices to determine their risk preferences. At the end of the experiment a new set of options would be drawn and we the experimenters would make a choice for the subjects based on what we think they would have chosen, given their choices on the 40 hypothetical trials. Therefore, it is in a subject's best interest to respond truthfully in every trial, since every choice could influence the decision we make for the subject in the payment round. This mechanism is incentive compatible while also addressing a possible concern with paying subjects for one random round with the adaptive algorithm from Wang, Camerer and Filliba, which is that subjects may realize that they can influence the options that they see and try to earn higher payoffs by choosing strategically.

Results. The aggregate results suggest that for most subjects the mechanism was incentive compatible: 88% of subjects reported truthfully when they had a value of $9, and 98% reported truthfully when they had a value of $1.

We next tested if there was a systematic relationship between the coefficients of risk and loss aversion and subjects' choices to misrepresent their values. A logit regression of an indicator function for truth-telling on the coefficients of risk and loss aversion for each subject estimated no effect for loss aversion (p=0.27) or for risk aversion (p=0.42).

The aggregate results also suggest that most of the subjects believed that the mechanism satisfies voluntary participation: 100% of the subjects voted Yes when they had a value of $9 and 78% voted Yes when they had a value of $1.

We next tested if there was a systematic relationship between the coefficients of risk and loss aversion and subjects' beliefs that the mechanism did not satisfy voluntary participation. A logit regression of an indicator function for No-votes on the coefficients of risk and loss aversion for each subject estimated no effect for the coefficient of loss aversion (p=0.5) or for the coefficient of risk aversion (p=0.55).

Figure 1 provides an additional test of this relationship. The dashed line in the upper graph depicts the voluntary participation constraint, as a function of the coefficients of risk and loss aversion, for individuals with v_i= 1. Note that the constraint is satisfied in the lower left and it is violated in the upper right. Each point in the graph represents a subject's parameters and her vote. As can be seen in the figure,, a vast majority of the subjects fall to the left of or very near the indifference curve. Only 14% of subjects have sufficiently high loss aversion where voluntary participation becomes an issue. Furthermore, only 2 out of those 7 subjects actually vote No in the public goods game. Also, note that the voluntary participation calculation changes depending on the gross taxes paid by the other players. Therefore, here we have considered the 'worst case scenario' for the subject, where the expected sum of the other players' gross taxes is minimized, and therefore so is her budget-balancing refund. If voluntary participation is satisfied in this scenario, then it will be satisfied in all scenarios.

For the case where v_i=9, a truth-telling subject faces no uncertainty in his gross tax and earns a positive payoff in the worst case scenario, and so voluntary participation will always be satisfied. This is consistent with the experimental data. Together, these results allow us to conclude that violations of voluntary participation are associated with large degrees of loss-aversion, which are observed in a small fraction of the population.

C.2. Experiment 2

The second experiment is designed to test the performance of the neurometrically informed mechanism in more complex domains in which there are a large number of types.

Environment. Groups of five subjects need to decide how much of a public good to produce. The preferences of subject i are given by $v_i \log(z) - y_i$ where v_i denotes the public good type, z is the total amount invested in the public good, and y_i is the net tax paid by the subject. We assume that every subject can have one of 20 possible types, so that $V = \{1,2,...,20\}$.

There is a signal technology with the following properties: if the subject's true type is v_i then the signal is uniformly distributed on $[v_i - 10, v_i + 10]$.

It is straightforward to show that efficiency requires producing a level of public goods given by $z(v) = \sum_{i=1}^{n} v_i$ and raising an equivalent amount of taxes given by $\sum_{i=1}^{n} y_i(v) = z(v)$.

An efficient allocation of particular interest is the one characterized by a Lindahl Equilibrium (Lindahl, 1958). In this allocation, i's contribution is given by her marginal benefit for the public good times the level of the public good (both calculated at the efficient level $z(v)$), which in this environment equals v_i.

Standard Lindahl Mechanism **(SLM)**. A Lindahl mechanism is a direct revelation mechanism (V,g) that implements the optimal level of the public good given the reports (i.e., $z(m) = \sum_{i=1}^{n} m_i$) and funds it using the Lindhal taxes (i.e., $y_i(m) = m_i$). For these environments, the SLM is identical to the standard Voluntary Contributions Mechanism. (See Ledyard 1995.) It is straightforward to show that the Standard Lindahl Mechanism is efficient and satisfies voluntary participation, but is not dominant strategy incentive compatible.

Augmented Lindahl Mechanism **(ALM).** We construct the Augmented Lindahl Mechanisms by applying the person-by-person augmentation methods described in

Section B. In particular, we use the gross augmented tax function $r_i(s_i,m_i) = \frac{N}{N-1}(s_i - m_i)^2$ and so the net augmented tax function is given by

$$w_i(s,m) = (s_i - m_i)^2 - \frac{1}{N-1}\sum_{j \neq i}(s_j - m_j)^2 .$$

The net tax paid by the subjects is given by

$$t_i(s,m) = m_i + w_i(s,m).$$

For this signal technology, $E(r) = -(\frac{1}{20})(\frac{N}{N-1})\int_{v-10}^{v+10}(s-m)^2 ds = -\frac{N}{N-1}((v-m)^2 + \frac{100}{3})$. Thus $E(r(s,m)) < E(r(s,v))$ for all m ≠ v and $E(r(s,v)) = -(\frac{N}{N-1})(\frac{100}{3})$. Therefore, we know that for risk and loss-aversion neutral subjects it satisfies incentive compatibility and voluntary participation.

Subjects. N = 30 Caltech undergraduates participated in the experiment. In addition to their earnings during the task, they paid a $15 fee.

Experimental task. In each experimental session 10 subjects participated in 40 rounds of decision-making, twenty with the SLM, and twenty with the ALM. Twenty subjects played the rounds with the SLM before the ALM, the other ten subjects played the opposite order.

In each round the private value was randomly and independently drawn for each subject from a uniform distribution on V. For each mechanism subjects remained in the same group of five, but between mechanisms subjects were randomly re-matched. To allow for learning, at the end of each round subjects were told their earnings for that round as well as the group's total contribution to the public good (given by $\sum_{i=1}^{5} m_i$). In the ALM,

subjects were also told the value of their signal for that round and the components of their total tax t_i.

The experimental instructions (included in Appendix II, available online) provide an in-depth description of the rules of both mechanisms as well as the payments that they induce under different contingencies. An important feature of the instructions is that we explicitly explain to the subjects that truth-telling maximizes their expected payoffs regardless of the decisions made by the other subjects. It is important to emphasize that while this aspect of the instructions is not an explicit requirement of the theory (which assumes that subjects know this), making sure that subjects fully understand key aspects of the distribution of payoffs is an integral part of applied mechanism design.

Results. In Figure 2 we report the deviations between truthful and actual reporting (given by $v_i - m_i$), as a function of treatment and experimental round. Whereas there was no under-reporting in the ALM condition (mean=0.02, se=0.05, p=0.73), there was significant under-reporting in the SLM (mean=5.50, se=0.73, two-sided t-test p=0.0007). Furthermore, a mixed effects regression of the size of the deviations on round number for the ALM case reveals no learning effect (p=0.59), but a similar regression for the SLM shows that under-reporting increases with round (beta=0.10, p=0.0005). These results clearly illustrate the power of the neurometrically informed mechanisms: whereas the ALM elicits near-perfect truth-telling by the subjects, there is substantial free-riding in the SLM and it gets worse over time. The SLM results are consistent with standard linear Voluntary Contribution Mechanism results (See Ledyard 1995.)

In Figure 3 we display the efficiency of the allocations induced by both mechanisms. The figure plots efficiency as a function of round for each case, where efficiency is defined as

$$Efficiency = \frac{\text{actual total group payoff}}{\text{optimal total group payoff}}.$$

Average efficiency was 99.95% (se=0.02%) on the ALM and 88% on the SLM (SE=3%). Thus, the ALM generated significantly larger efficiencies (p=0.0097 two-sided *t*-test). Furthermore, a mixed effects regression of efficiency on round number for the ALM case revealed no time trends (p=0.32), whereas a similar regression for the SLM showed that efficiency decreased with time (beta= -0.0032 , p<0.0007), which translates to a drop of efficiency of 6.4% over 20 trials. These results show that whereas the ALM institution generated nearly perfect efficiency in every round, the SLM generates substantial inefficiencies that worsen over time.

Because we draw new values each round, it may not be clear from Figure 3 whether the efficiencies for the SLM are due to high contributions or high values. In Figure 4, we compare the efficiencies from the actual reports and to those that would occur in the Nash Equilibrium for the SLM and ALM. That is, we plot

$$\frac{\textit{actual group payoff} - \textit{Nash equilibrium payoff}}{\textit{optimal group payoff} - \textit{Nash equilibrium payoff}}.$$

It is important to realize that since we redraw values each round and since information about values is private, there is no game-theoretic reason why subjects should play the Nash Equilibrium which usually has the highest valued type reporting their true value and all others reporting $m = 1$. Nevertheless, from the data displayed in Figure 2 it is not possible to reject the hypothesis that, on average, the outcomes are the Nash Equilibrium outcomes (p<0.9). Further, unlike standard linear Voluntary Contribution Mechanism results, we do not see initial contributions that are higher than the Nash Equilibrium that then decline over time (p<0.7). Here, contributions are near to the Nash equilibrium from beginning to end. Consistent with the observations from Figure 2, the efficiencies of the ALM are at the optimum. The Augmented Lindahl Mechanism provides a significant improvement over the Standard Lindahl Mechanism.

D. Final Remarks

A fundamental assumption behind the classic impossibility results in mechanism design is that the only way the planner can gain information about individual preferences is by eliciting them behaviorally through a cleverly constructed mechanism. Although this has been a valid assumption for the last 30 years, modern neurometric technologies are now making it possible to obtain direct noisy signals of subjects' preferences.

Specifically, functional magnetic resonance imaging (fMRI) can now be used to measure levels of activity throughout the brain. Many of these studies seek to establish which brain regions' activity consistently correlates with behavior or stimuli. In the realm of neuroeconomics, much work has already been done to identify brain structures involved in the computation of value. The medial orbital frontal cortex (mOFC) has emerged as the leading candidate for computing decision value based on work in humans (Plassmann, O'Doherty & Rangel, 2007, Plassmann et al., 2008, Litt et al., 2010, Hare et al., 2008, Valentin, Dickinson & O'Doherty 2007, Hare, Camerer & Rangel 2009, Tom et al., 2007, Levy et al., 2010, Kable & Glimcher, 2007, Chib et al., 2009) and in primates (Padoa-Schioppa & Assad, 2006, Padoa-Schioppa, 2009, Padoa-Schioppa & Assad, 2008, Wallis & Miller, 2003). Furthermore, components of value have been associated with other notable brain structures, such as reward prediction and surprise ("prediction error") in the striatum (Schultz, Dayan & Montague, 1997, Hare et al., 2008, McClure, Berns & Montague, 2003, O'Doherty et al., 2004, Pagnoni et al., 2002, Knutson et al. 2001, Hsu et al., 2009), risk in the insula (Preuschoff, Quartz and Bossaerts, 2008, Singer, Critchley & Preuschoff, 2009), and other reward components such as framing effects in the amygdala (Gottfried, O'Doherty & Dolan 2003, Murray, 2007, Baxter & Murray, 2002, De Martino et al., 2006, De Martino, Camerer & Adolphs, 2010).

To look at one particularly relevant example, Chib et al. (2009) asked subjects to make purchasing decisions for foods, non-food consumables and monetary gambles while in the MRI scanner. They found that for all three classes of goods, activity in the mOFC correlated with the subjects' pre-scanner BDM bids for the items. Together with the

other studies mentioned above, these results argue strongly for the involvement of mOFC in value computation.

A related literature has instead focused on using pattern classification techniques (also called machine learning) on biological signals such as fMRI measures of brain activity, in order to classify/predict mental states or behavior. Neuroeconomics papers in this literature have successfully predicted subjects' choices outside the fMRI scanner, based on brain activity in response to the individual goods in the scanner (Lebreton et al., 2009, Levy et al., 2010, Smith, Bernheim, Camerer & Rangel, forthcoming) as well as differentiated between probabilistic and intertemporal valuation (Clithero, Carter & Huettel, 2009). This field of "mind reading" (also called neural decoding) is rapidly advancing and the accuracy of the measurements is steadily increasing (Haxby et al, 2001; Cox and Savoy, 2003; Kamitani and Yong, 2005; Polyn et al, 2005; Norman et al, 2006; Haynes et al, 2007; O'Toole et al, 2007; Pessoa and Padmala, 2007; Serences and Boynton, 2007; Kay et al, 2008, Clithero, Carter & Huettel, 2009, Levy et al., 2011).

Also, notice that unlike conventional lie detectors, which detect the physiological arousal generated by telling a lie (which can potentially be reduced through extensive training), fMRI technology can produce continuous measures of brain activity that correlate with subjects' gut reactions to stimuli. This added flexibility is necessary for implementing different resource allocations.

Here we have argued that using these technological advances in neuroscience can have a profound impact on the mechanism design problem: they make it possible to design mechanisms that implement any desired allocation while satisfying voluntary participation and dominant strategy incentive compatibility. This stands in sharps contrast with the classic impossibility results showing that without such signals it is generally impossible to design such mechanisms.

The use of neurometric or biologically-based signals also helps us avoid the major criticisms of Cremer-McLean and McAfee-Reny, since our mechanisms do not depend

on or involve player's beliefs about others. As long as subjects believe in the technology, it is a dominant strategy for them to report their type honestly.

In most of the mechanism design literature we assume that individuals are risk- and loss-neutral and concern ourselves only with expected payoffs. However, if these mechanisms are to be adopted for actual use, they must be robust to risk preferences. To address this issue we investigated these mechanisms in two public goods experiments. The results show that subjects follow the dominant strategy of truth-telling, and that voluntary participation and incentive compatibility are easily accommodated for the range of loss- and risk-aversion parameters observed in most of our subject population.

Several aspects of our results deserve more discussion. First, how informative does the signal technology have to be for our general possibility result to hold? The key assumption, namely that the signal technology satisfy the CM condition, is quite weak. For example, in the two-type case it is satisfied whenever the likelihood function over signals is different for different values of the true preferences, which can be satisfied even if the probability of the high signal given that the type is high is barely above chance. One caveat, though, is that the neurometrically informed mechanisms generate payoff variability that depends on the realization of the signal. Furthermore, the lower the quality of the signal, the larger the amount of variability that will be present in the payoffs, which could be a problem if subjects exhibit sufficient amounts of loss-aversion. Fortunately, current trends in neurometric technologies suggest that high-accuracy signals might be available in the near future (Haxby et al, 2001; Cox and Savoy, 2003; Kamitani and Yong, 2005; Polyn et al, 2005; Norman et al, 2006; Haynes et al, 2007; O'Toole et al, 2007; Pessoa and Padmala, 2007; Serences and Boynton, 2007; Kay et al, 2008, Clithero, Carter & Huettel, 2009, Levy et al., 2011). Additionally, taking repeated measures or combining different technologies promises to improve signal quality, even in the absence of technological progress.

Second, what are likely sources of signals in future applications? The findings in Krajbich et al. (2009), as well as those in the references in the previous paragraph,

suggest that fMRI could be a good source of high-quality signals. However, such measurement technologies remain expensive, and thus it will be important to explore other less expensive sources of signals, such as electroencephalography (EEG), pupil-dilation, skin-conductance, and facial electromyography (EMG) (Tassinary and Cacioppo, 1992; Lang et al, 1993; Aboyoun and Dabbs, 1998; Bradley, 2000; Dimberg et al, 2000; Partala et al, 2000; Bradley et al, 2008). Preliminary work in our laboratory suggests that these signals might be able to provide sufficiently informative signals, at least in simple contexts. It is also important to emphasize that our results also apply to non-neurometric signals. For example, the results also apply to a situation in which the planner has sufficiently informative priors about individual subjects' preferences based on previous behavioral or demographic data. However, the caveat with such signals is that individuals cannot know the values of the signals or be able to manipulate them, so non-neurometric signals would have to be collected covertly.

Third, what if it is only possible (perhaps due to costs) to obtain signals from a subset of the members in the group? This problem can be addressed using 'random augmentation'. This works as follows. Subjects make their choices for the neurometrically informed mechanism without knowing if a signal will be available for them. Afterwards, the planner randomly selects a subset of the group and obtains signals only for them. The augmented taxes can be redefined so that everybody's incentives at the initial phase are the same as in the case of full signal monitoring. For example, suppose that the planner selects each individual for signal extraction with a constant probability q. It is easy to see that an augmented tax given by

$$\widehat{w}^i(s,m) = \frac{1}{q} w^i(s,m)$$

induces the same incentives as in the case of full monitoring (where $w^i(s,m)$ are the augmented taxes for the q=1 case). This idea is reminiscent of the literature on auditing, where a principal induces an agent to truthfully report by probabilistically auditing them

and charging a high fee if misrepresentation is found. See for example, Border and Sobel (1987) and Baron and Besanko (1984).

We conclude on an optimistic note. The rapid and likely rise of neurometric technologies has the potential to make it feasible to design much better mechanisms and institutions in a large number of applications. Importantly, as demonstrated here and in Krajbich et al. (2009), the development of such institutions will require a careful combination of methods from neuroscience and computer science and ideas and models from economics. In particular, the insights of mechanism design theory developed over the last 30 years are likely to be critical to make progress in designing new classes of neurometrically informed mechanisms.

FIGURES

Figure 1. Individual voluntary participation constraint as a function of the individuals risk and loss-aversion parameters for the case of $v_I = 1$. The dashed line depicts the voluntary participation constraint, as a function of the coefficients of risk and loss aversion. Note that the constraint is satisfied in the lower left and it is violated in the upper right. Each point in the graph represents a subject's parameters and her vote. Note that the constraint is calculated in the 'worst case scenario' for the subject, where the sum of the gross taxes paid by the other players is minimized. So if voluntary participation is satisfied in this case, it will be satisfied in every other scenario as well.

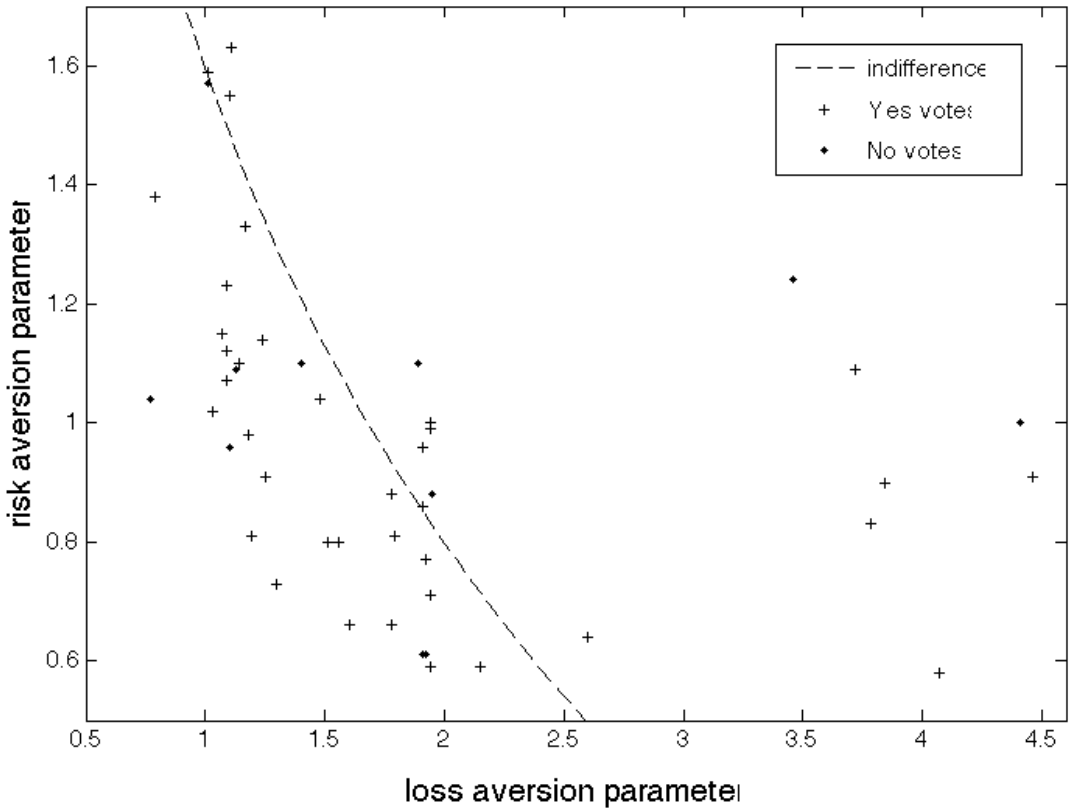


Figure 2. Under-reporting in Experiment 2 by mechanism and trial.

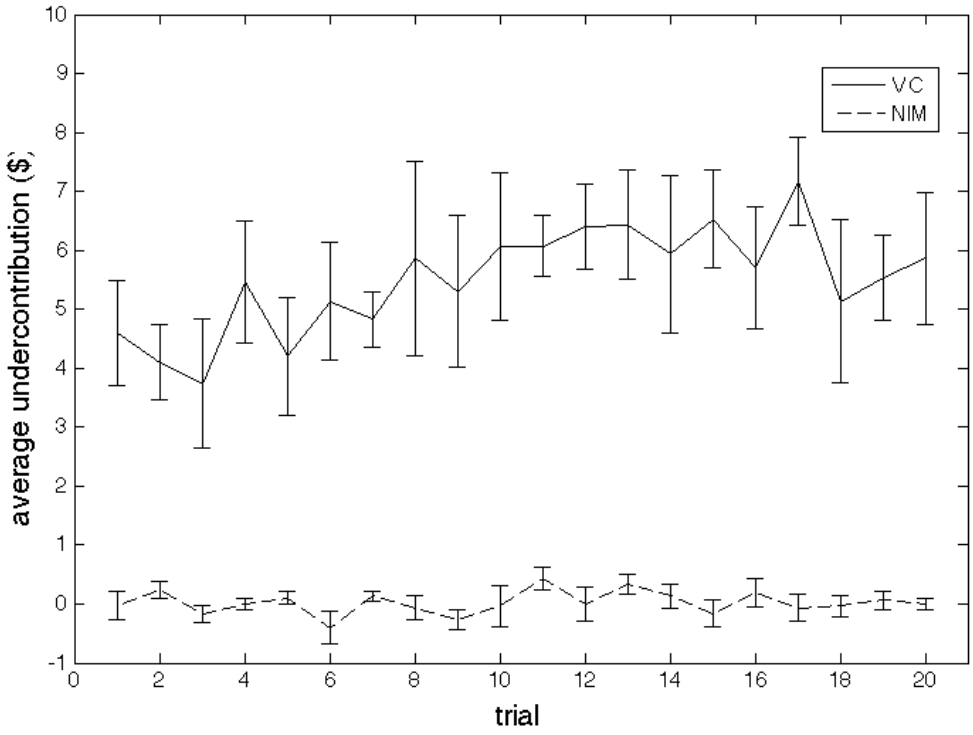


Figure 3. Average efficiency in Experiment 2 by mechanism and trial.

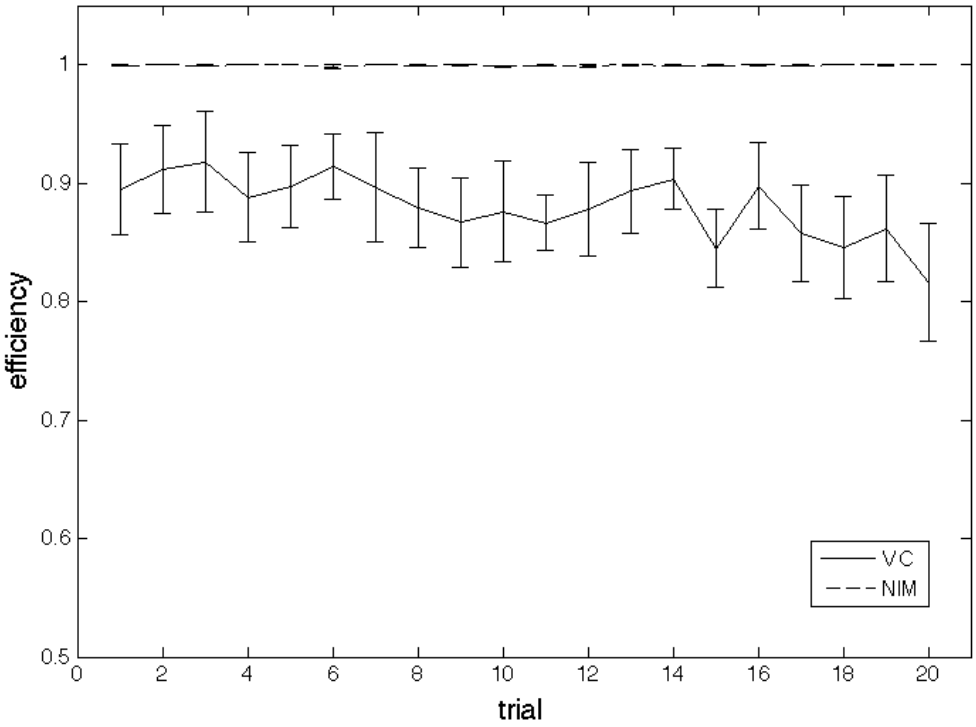
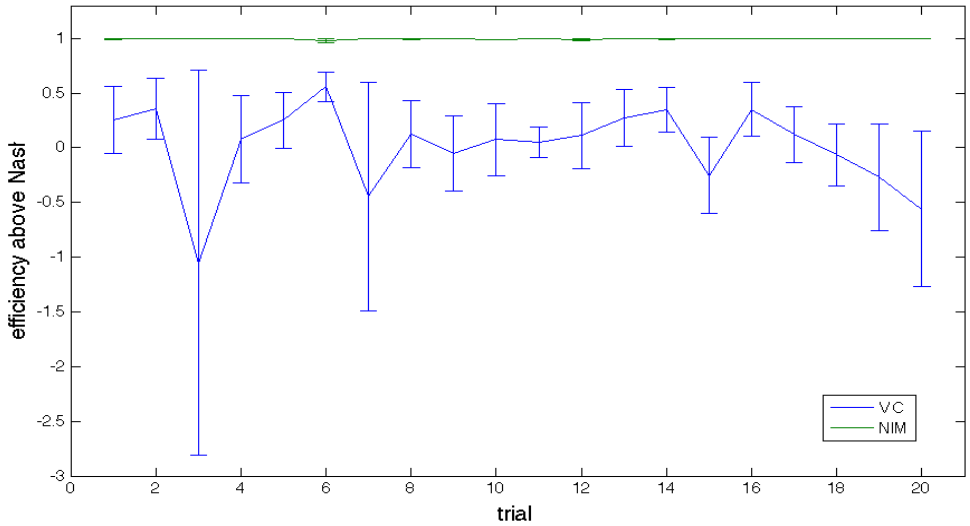


Figure 4: Comparing Nash Equilibrium to actual choices in Experiment 2 by mechanism and trial.



References

Aboyoun, D.C. & Dabbs, J.M. (1998) "The Hess pupil dilation findings: Sex or novelty?" Social Behavior and Personality: an international journal 26, 415-419

Baron, D.P. & Besanko, D. (1984) "Regulation, asymmetric information, and auditing" Rand Journal of Economics 15, 447-470.

Bergstrom, T., Blume, L., & Varian, H. (1986) "On the Private Provision of Public Goods" Journal of Public Economics 29, 25-49

Border, K.C. & Sobel, J. (1987) "Samurai Accountant: A Theory of Auditing and Plunder" The Review of Economic Studies 54, 525-540

Bradley, M.M. (2000) "Emotion and Motivation" in Handbook of Psychophysiology (ed. J.T. Cacioppo, L.G. Tassinary & G.G. Berntson) (Cambridge University Press)

Bradley, M.M., Miccoli, L., Escrig, M.A. & Lang, P.J. (2008) "The pupil as a measure of emotional arousal and autonomic activation" Psychophysiology 45, 602-607

Brier, G.W. (1950) "Probabilistic Forecasts of Precipitation in terms of Quantiles using NWP Model Output", Monthly Weather Review 132, 338-347

Clarke, E. (1972) "Multi-part Pricing of Public Goods" Public Choice 11, 17-33

Coppinger, V., Smith, V.L. & Titus, J. (1980) "Incentives and Behavior in English, Dutch, and Sealed-Bid Auctions" Economic Inquiry 18, 1-22

Cox, D.D. & Savoy, R.L. (2003) "Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex", Neuroimage 19, 261-270

Cremer, J. & McLean, R.P. (1985) “Optimal Selling Strategies Under Uncertainty for a Discriminating Monopolist When Demands are Interdependent” Econometrica 53, 345-361

Cremer, J. & McLean, R.P. (1988) “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions” Econometrica 56, 1247-1257

D'Aspremont, C., & Gerard-Varet, L. (1979) “Incentives and Incomplete Information" Journal of Public Economics, 11, 25-45.

Dimberg, U., Thunberg, M. & Elmehed, K. (2000) “Unconscious facial reactions to emotional facial expressions” Psychological Science 11, 86-89

Fehr, E. & Gachter, S. (2002) “Altruistic punishment in humans” Nature 415, 137-140

Gibbard, A. (1973) “Manipulation of voting schemes: A general result” Econometrica 41, 587-601

Green, J. & Laffont, J.J. (1977) “Characterization of satisfactory mechanisms for the revelation of preferences for public goods” Ecnometrica 45, 427-438

Gneithing, T. & Raftery, A.E. (2007) “Strictly Proper Scoring Rules, Prediction, and Estimation” Journal of the American Statistical Association 102, 359-378

Groves, T. (1973) "Incentives in Teams" Econometrica 41, 617-631

Groves, T. & Ledyard, J. (1977) “Optimal Allocation of Public Goods: A Solution to the 'Free-Rider' Problem” Econometrica 45, 783-809

Harsanyi, J.C. (1967) “Games with incomplete information played by ‘Bayesian’ players” Management Science 14, 159-182

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L. & Pietrini, P. (2001) "Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex" Science 293, 2425-2430

Haynes, J., Sakai, K., Rees, G., Gilbert, S., Frith, C. & Passingham, R. (2007) "Reading Hidden Intention in the Human Brain" Current Biology 17, 323-328

Herrmann, B., Thoni, C. & Gachter, S. (2008) "Antisocial punishment across societies" Science 319, 1362-1367

Hurwicz, L. (1972) "On Informationally Decentralized Systems" in Decision and Organization, eds. B. McGuire and R. Radner, Amsterdam: North Holland 297-336

Hurwicz, L. & Walker, M (1990) "On the Generic Nonoptimality of Dominant-Strategy Allocation Mechanisms: A General Theorem That Includes Pure Exchange Economies" Econometrica 58, 683-704

Kahneman D. & Tversky, A. (1979) "Prospect Theory: An Analysis of Decision Under Risk" Econometrica 47, 263-291

Kamitani, Y. & Tong, F. (2005) "Decoding the visual and subjective contents of the human brain" Nature Neuroscience 8, 679-685

Kay, K.N., Naselaris, T., Prenger, R.J. & Gallant, J.L. (2008) "Identifying natural images from human brain activity" Nature 452, 352-355

Krajbich, I., Camerer, C.F., Ledyard, J. & Rangel, A. (2009) "Using neural measures of economic value to solve the public goods free-rider problem" Science 326, 596-599

Lang, P.J., Greenwald, M.K., Bradley, M.M. & Hamm, A.O. (1993) “Looking at Pictures: Affective, Facial, Visceral, and Behavioral Reactions” Psychophysiology 30, 261-273

Ledyard, J.O. (1995) “Public Goods: A Survey of Experimental Research” Handbook of Experimental Economics, Princeton University Press

Lindahl, E. (1958) “Just taxation-a positive solution” in Richard Musgrave and Alan Peacock, editors, Classics in the Theory of Public Finance, Macmillan, London 98–123

Maskin, E.S. (1999) “Nash Equilibrium and Welfare Optimality” The Review of Economic Studies, 66, 23-38

McAfee, R.P., & Reny, P.J. (1992) “Correlated Information and Mechanism Design” Econometrica 60, 395-421

Myerson, R. (1981) ``Optimal Auction Design." Mathematics of Operations Research, 6, 58--73

Nikiforakis, N. (2008) “Punishment and counter-punishment in public good games: Can we really govern ourselves?” Journal of Public Economics 92, 91-112

Norman, K.A., Polyn, S.M., Detre, G.J. & Haxby, J.V. (2006) “Beyond mind-reading: multi-voxel pattern analysis of fMRI data” Trends in Cognitive Sciences 10, 424-430

O’Toole, A.J., Jiang, F., Abdi, H., Penard, N., Dunlop, J.P. & Parent, M.A. (2007) “Theoretical, Statistical, and Practical Perspectives on Pattern-based Classification Approaches to the Analysis of Functional Neuroimaging Data” Journal of Cognitive Neuroscience 19, 1735-1752

Partala, T., Jokiniemi, M. & Surakka, V. (2000) "Pupillary responses to emotionally provocative stimuli" in The 2000 symposium on eye tracking research & applications 123-129 (ACM Palm Beach Gardens, Florida, United States)

Pessoa, L., Padmala, S. (2007) "Decoding Near-Threshold Perception of Fear from Distributed Single-Trial Brain Activation" Cerebral Cortex 17, 691-701

Polyn, S.M., Natu, V.S., Cohen, J.D. & Norman, K.A. (2005) "Category-Specific Cortical Activity Precedes Retrieval During Memory Search" Science 310, 1963-1966

Satterthwaite, M. (1975) "Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions" Journal of Economic Theory 10, 187-217

Serences, J.T. & Boynton, G.M. (2007) "Feature-Based Attentional Modulations in the Absence of Direct Visual Stimulation" Neuron 55, 301-312

Sokol-Hessner, P., Hsu, M., Curley, N.G., Delgado, M.R., Camerer, C.F., Phelps, E.A. (2009) "Thinking like a trader selectively reduces individuals' loss aversion" Proceedings of the National Academy of Sciences of the United States of America 106, 5035-5040

Tassinary, L.G. & Cacioppo, J.T. (1992) "Unobservable Facial Actions and Emotions" Psychological Science 3, 28-33

Vickrey, W. (1961) "Counterspeculation, Auctions, and Competitive Sealed Tenders" Journal of Finance 16, 8-37

Walker, M. (1980) "On the Nonexistence of a Dominant-Strategy Mechanism for Making Optimal Public Decisions" Econometrica 48, 1521-15

Wang, S., Camerer, C. & Filliba, M. (2010). “Bayesian Adaptive Design for Optimal Elictation of Risk Preferences”, Caltech working paper.