Overcoming Incentive Constraints by Linking Decisions

by Matthew O. Jackson and Hugo F. Sonnenschein *

Extended Abstract for the P2P conference

1 Introduction

Over the past fifty years we have learned that the possibility of efficient social interaction depends not only on resources and technology, but equally and most critically on incentive constraints, including participation constraints, and the ability of social institutions to mediate those constraints. Thus, voting systems, markets, financial contracts, P2P systems auction forms, public goods decisions, and a host of other practical arrangements are now often formulated as Bayesian games, and judged in terms of their ability to mediate incentive constraints.

We demonstrate that the limitations that incentive constraints impose on the attainment of socially efficient outcomes can be made to disappear when several problems are linked. For instance, it is well known that fully efficient trade is impossible between a buyer and a seller if their values are private information. Sellers have an incentive to inflate the cost and buyers have an incentive to understate their valuation, and these incentives necessarily lead to trade failure with a non-negligible probability (e.g., see Myerson and Satterthwaite 1983). We exploit the idea that when several independent social decision problems are linked, or when there are several independent aspects of a given problem, then it makes sense to speak of "rationing" or "budgeting" an agent's representations. For instance, when a buyer and seller bargain over several items, we do not allow the buyer to claim to have a low valuation for each item, nor do we allow the seller to claim to have a high cost for all of the items. The critical insight, is that by linking the bargaining over the items, we can discover which items are relatively more costly for the seller and which items are relatively more valued by the buyer. The rationing or budgeting of announcements leads to a tendency, which we make precise, for agents to be as truthful in their representation as is possible. This helps us to overcome incentive constraints and yields more efficient decisions.

In more formal language, we consider an abstract Bayesian collective decision problem and an ex ante Pareto efficient social choice function f that indicates the collective decision we would like to make as a function of the realized preferences of the n agents. Let $(u_1, u_2, \ldots, u_n)$ denote the ex ante expected utilities that are achieved under f. Such an ideal f will generally not be implementable because of incentive constraints. Now, consider K copies of the decision problem, where agents' preferences are additively separable and independently distributed across the problems. We show that as K becomes large it is possible to essentially implement f on each problem and thus achieve the target utilities $(u_1, u_2, \ldots, u_n)$ on each problem. Even when K is small there are typically substantial utility gains from considering the K problems together.

We establish this result by constructing a general mechanism that has each agent present a K-vector of preferences, and then the decision on each of the K problems is made according to f. The key is that we require that agents present vectors whose distribution of types across problems mirrors the underlying distribution of their preferences. The agents are not, for example, allowed to represent themselves as having a "bad draw" on more than the expected number of problems on which they "should" have a bad draw. We show that in the limit there is no gain from lying. In fact, for *every* K there is an equilibrium in which all agents are telling the truth as fully as the constraint on their representations permits. Moreover, we show that *all* equilibria of the linking mechanisms

*Jackson is at the Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, California 91125, USA, http://www.hss.caltech.edu/∼jacksonm/Jackson.html, email: jacksonm@hss.caltech.edu , and Sonnenschein is at the Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, USA, email: h-sonnenschein@uchicago.edu. Financial support under NSF grants SES-9986190 and SES-0316493 is gratefully acknowledged, as is support from the University of Chicago where Jackson was visiting when this research began.

converge to the target utility levels.

We should emphasize that while we make use of the law of large numbers, the heart of the matter is to figure out who is of what type on which problems. A large number of independent problems makes it likely that the realized frequency of agents' preferences mirrors the underlying distribution of preferences, but this cannot guarantee that agents will find it in their interest to truthfully announce those preferences. It is the interplay between the efficiency of the underlying social choice function and the fact that agents' announcements are budgeted to match the underlying distribution that makes it in each agent's best interest to be as truthful as she can.

2 A Theorem on Linking Decisions

Consider n agents who are involved in making decisions.

A decision problem is a triple $\mathcal{D} = (D, U, P)$.

Here D is a finite set of possible alternative decisions; $U = U_1 \times \cdots \times U_n$ is a finite set of possible profiles of utility functions $(u_1, \ldots, u_n)$, where $u_i : D \to I\!R$; and $P = (P_1, \ldots, P_n)$ is a profile of probability distributions, where P_i is a distribution over U_i. We abuse notation and write $P(u)$ for the probability of u.

Note that the decision problem may involve many alternatives and the utility functions may have any possible structure.[1]

The u_i's are drawn independently across agents.

A *social choice function* on a social decision problem $\mathcal{D} = (D, U, P)$ is a function $f : U \to \Delta(D)$, where $\Delta(D)$ denotes the set of probability distributions on D.

This is interpreted as the target outcome function.

Let $f_d(u)$ denote the probability of choosing $d \in D$, given the profile of utility functions $u \in U$.

A social choice function f on a decision problem $\mathcal{D} = (D, U, P)$ is *ex ante Pareto efficient* if there does not exist any social choice function f' on $\mathcal{D} = (D, U, P)$ such that

$$\sum_u P(u) \sum_d f'_d(u) u_i(d) \geq \sum_u P(u) \sum_d f_d(u) u_i(d)$$

for all i with strict inequality for some i.

[1] The results extend easily to infinite settings through finite approximations.

Linking Mechanisms

Given a base decision problem $\mathcal{D} = (D, U, P)$ and a number K of linkings, a *linking mechanism* (M, g) is a message space $M = M_1 \times \cdots \times M_n$ and an outcome function $g : M \to \Delta(D^K)$.

A linking mechanism is a mechanism that works on a set of decision problems all at once, making the decisions contingent on the preferences over all the decisions rather than handling each decision in isolation. Here M_i is a message space for agent i, and can be an arbitrary set.

We let $g_k(m)$ denote the marginal distribution under g onto the k-th decision, where $m \in M$ is the profile of messages selected by the agents.

When we link K versions of a decision problem $\mathcal{D} = (D, U, P)$, an agent's utility over a set of decisions is simply the sum of utilities. So, the utility that agent i gets from decisions $(d^1, \ldots, d^K) \in D^K$ given preferences $(u_i^1, \ldots, u_i^K) \in U_i^K$ is given by $\sum_k u_i^k(d^k)$.

We assume that the randomness is independent across decision problems. Given independence and additive separability, there are absolutely no complementarities across the decision problems. The complete lack of interaction between problems guaranteed that any improvements in efficiency that we obtain by linking the are not due to any correlation or complementarities.

Strategies and Equilibrium

A *strategy* for agent i in a linking mechanism (M, g) on K copies of a decision problem $\mathcal{D} = (D, U, P)$ is a mapping $\sigma_i^K : U_i^K \to \Delta(M_i)$.

We consider Bayesian equilibria of such mechanisms.

Approximating Efficient Decisions through Linking

Given a decision problem $\mathcal{D} = (D, U, P)$ and a social choice function f defined on $\mathcal{D}$, we say that a sequence of linking mechanisms defined on increasing numbers of linked problems, $\{(M^1, g^1); (M^2, g^2), \ldots, (M^K, g^K), \ldots\}$and a corresponding sequence of Bayesian equilibria $\{\sigma^K\}$ *approximate* f if

$$\lim_K \left[\max_{k \leq K} Prob\left\{g_k^K(\sigma^K(u)) \neq f(u^k)\right\}\right] = 0.$$

Thus, a sequence of equilibria and linking mechanisms approximates a social choice function if for large enough linkings of the problems, on every problem the probability that the equilibrium outcome of the linking mechanism results in the same decision as the target social choice function is arbitrarily close to one.

Strategies that Secure a Utility Level

To show that *all* equilibria of our linking mechanisms converge to being efficient, we show that there exists a strategy that guarantees that an agent's payoff will be above a certain level, regardless of what strategy other players employ.[2]

More formally, consider an arbitrary mechanism (M, g) on K linked decision problems. A strategy $\sigma_i : U_i^K \to M_i$ *secures a utility level* $\overline{u}_i$ if

$$E\left[\sum_{k \leq K} u_i(g^k(\sigma_i, \sigma_{-i})\right] \geq K\overline{u}_i$$

for *all* strategies of the other agents σ_{-i}.

The Linking Mechanisms

Consider K linked problems. Each agent announces utility functions for the K problems. So this is similar to a direct revelation mechanism. However, the agent's announcements across the K problems must match the expected frequency distribution. That is, the number of times that i can (and must) announce a given utility function u_i is $K \times P_i(u_i)$. The choice is then made according to f based on the announcements.

More formally, find any approximation P_i^K to P_i such that $P_i^K(v_i)$ is a multiple of $\frac{1}{K}$ for each $v_i \in U_i$, and the Euclidean distance between P_i^K and P_i (viewed as vectors) is minimized.

Agent i's strategy set is $M_i^K = \{\widehat{u}_i \in (U_i)^K$ s.t. $\#\{k : \widehat{u}_i^k = v_i\} = P_i^K(u_i)K$ for each $v_i \in U_i\}$.

Thus, agents must announce a vector of types across problems that matches the true underlying frequency distribution of their types.

The decision of g^K for the problem k is simply the target f operated over the announced types. That is, it is $g^K(m) = f(\widehat{u}^k)$, where $\widehat{u}_i^k$ is i's announced utility function for problem k under the realized announcement $m = \widehat{u}$.[3]

[2] This differs from the idea of a dominant strategy, and is closer in spirit to the concept of minimax strategy, although in a different setting.

[3] This is not quite the complete description of the mechanism. We need to realign the probability of choosing decisions to be as if announcements were exactly P_i rather than P_i^K, which is described in the proof. We also need an adjustment when there are at least three agents, that eliminates certain collusive equilibria and ensures that all equilibria converge to the desired targets.

Approximate Truth

The constraint of announcing a distribution of utility functions that approximates the true underlying distribution of types will sometimes force an agent to lie about their utility functions on some problems, since their realizations of utility functions across problems may not have a frequency that is precisely P_i. Nevertheless, strategies that are as truthful as possible subject to the constraints, turn out to be useful strategies for the agents to employ, and so we give such strategies a name.

A strategy is *approximately truthful* if the agent's announcements always involve as few lies as possible. Formally, $\sigma_i : U_i^K \to M_i^K$ is *approximately truthful* if

$$\#\{k \mid \sigma_i^K(u_i^1, \ldots, u_i^K) \neq u_i^k\} \leq \#\{k \mid m_i^k \neq u_i^k\}$$

for all $m_i \in M_i^K$ and all $(u_i^1, \ldots, u_i^K) \in U_i^K$.

A Theorem on Approximating Efficient Decisions through Linking

Let $\overline{u}_i = E\left[u_i\left(f(u)\right)\right]$, and let $\overline{u} = (\overline{u}_1, \ldots, \overline{u}_n)$ denote the ex ante expected utility levels under the target social choice function. These are the targets for the utility level that we would like to implement.

Theorem 1 *Consider a decision problem $\mathcal{D}$ and an ex ante Pareto efficient social choice function f defined on it. There exists a sequence of linking mechanisms (M^K, g^K) on linked versions of the decision problem such that:*

(1) There exist a corresponding sequence of Bayesian equilibria that are approximately truthful.

(2) The sequence of linking mechanisms together with these corresponding equilibria approximate f.

(3) Any *sequence of approximately truthful strategies for an agent i secures a sequence of utility levels that converge to the ex ante target level $\overline{u}_i$.*

(4) All *sequences of Bayesian equilibria of the linking mechanisms result in expected utilities that converge to the ex ante efficient profile of target utilities of $\overline{u}$ per problem.*

The proof of the theorem appears in the full paper, Jackson and Sonnenschein (2003).

Theorem 1 holds for any ex ante efficient social choice functions that we target. As such, f can satisfy any number of auxiliary properties, such as participation constraints (also known as individual rationality constraints), fairness, etc.

COROLLARY **1** *Consider any ex ante efficient f that satisfies a strict participation constraint of any sort: ex ante, interim or ex post. Consider the two-stage linking mechanisms where in a first stage agents are allowed to decide whether or not to participate (after having learned their types). For every K, there exists an approximately truthful equilibrium*[4] *of the modified linking mechanism such that the resulting social choice function satisfies an ex post (and thus interim and ex ante) participation constraint, and the sequence of these equilibria approximate f.*

3 Related Mechanisms and Literature

One of the main conclusions from our results is that one should expect to see a linking of decisions problems in practice, as it can lead to substantial efficiency gains. It is thus not surprising that one can find, scattered about in the literature, examples of mechanisms that link problems. Some of these turn out to be cousins of the general mechanisms we have described here, but in the context of some particular problem.

Indeed, our initial discussions on this topic were spurred by trying to understand how the creative and innovative storable votes mechanism of Casella (2002) works. Casella's setting is one where a society makes binary decisions repeatedly over time, and in each period operates by a vote, choosing the alternative garnering the majority of votes. Her storable votes mechanism is one where an agent may store votes over time. So, an agent may choose not to vote in period 1, and then would have two votes at his or her disposal in period 2. Votes that are not cast by some agent in any period are stored for that agent's future use. Casella shows that in some cases, an equilibrium of the storable votes mechanism offers a Pareto improvement over separate votes. While there are many equilibria to this mechanism, and it is not clear that Pareto improvements are always present,[5] the idea behind the storable votes mechanism as a tool to get some gauge of the intensity of preferences, is what inspired our investigation into the linking of decisions.[6] Of course, the value added here is in showing how linking can be done to get full efficiency in the limit, how it can be done so that all equilibria converge, and that it applies to essentially any decision problem, not just a binary voting problem. In the context of the binary voting setting, our linking mechanisms also suggests some potential improvements relative to a storable votes mechanism. This comes from the way we force agents to ration their announcements. This might be thought of as giving voters a series of votes of different powers (corresponding to the relative intensities of their preferences). They must spend exactly one vote on each problem. The key is that this additional rationing (eliminating the discretion of agents to freely pile up votes) leads to efficiency as an outcome and as the only viable outcome in the limit. More generally, depending on how variable preferences are, one could admit additional votes of various point levels, to match the distribution of intensities of preference.

Another setting where one sees mechanisms linking decisions across problems is one where a group of agents is splitting up a set of objects. In such a context, under some symmetry assumptions, McAfee (1992) shows the limiting efficiency of a mechanism where agents take turns selecting objects.[7] With large numbers of objects, and symmetric type distributions across agents, this would lead to approximately the same outcomes as a linking mechanism that sought to give objects to agents with the highest valuation.

Let us close with some final remarks on the relation to some other literature that the linking of decisions

[4]The equilibrium notion is now Perfect Bayesian Equilibrium, as we are dealing with a two stage mechanism.

[5]Experimental studies by Casella and Palfrey (2003) indicate that agents spend storable votes at least roughly in the right ways and realize some Pareto gains relative to standard voting mechanisms.

[6]Hortala-Vallve (2003) (independently) studies what he calls "qualitative voting," which is a variation on Casella's storable votes that allows the transfer of votes freely across problems, whereas storable votes can only be stored for future use rather than borrowed from the future. As a result, in a two-intensity world some equilibria of his mechanism will Pareto dominate those of the storable votes mechanism.

[7]A related mechanism is discussed by Pesendorfer (2000) in the context of a set of bidders colluding in a sequence of auctions trying to decide who should win which auctions. Se also, Blume (19), Campbell (1998), and Chakraborty, Gupta, and Harbaugh (2002), as well as earlier work on multi-market collusion by Bernheim and Whinston (1990).

might have brought to mind.

When thinking about voting problems and linking decisions, it is natural to think of log-rolling.[8] Indeed there is some flavor of trading across decisions that is inherent in the linking mechanisms. However, logrolling generally has to do with some coalition (often a minimal majority) making trades in order to control votes, and usually at the expense of other agents. Logs are rolled in the context of majority voting mechanisms across different problems, which points out the important distinction that the mechanism itself is not designed with the linking in mind. This leads to a contrast between the benefits of linking mechanisms and the dark side of logrolling.

Finally, another place where some linking of decisions occurs is in the bundling of goods by a monopolist. The idea that a monopolist may gain is selling goods in bundles rather than in isolation is was pointed out in the classic paper by Adams and Yellen (1976). Moreover, this gain can be realized when preferences over the goods are independent (see McAfee, McMillan and Whinston (1979)), can be enhanced by allowing for cheap talk where information about rankings of objects is communicated (see Chakraborty and Harbaugh (2003)), and in fact in some cases the monopolist can almost extract full surplus by bundling many goods (see Armstrong (1999)[9]). Indeed, applying the linking decisions to the case of a bundling monopolist we can obtain (a strengthening of) Armstrong's result as a corollary to Theorem 1 by having the monopolist be agent 1 and the buyer be agent 2 and letting f be that the monopolist sells the good to the buyer at the buyer's reservation price whenever the reservation value is less than the cost of the good.

[8] For some of the classics on this subject, see Tullock (1970) and Wilson (1969), as well as the discussion in Miller (1977).

[9] See also Fang and Norman (2003) who examine the efficiency gains from the bundling of excludable public goods.

References

Adams, W.J. and J.L. Yellen (1976) "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 90:3, 475-498.

Armstrong, M. (1999) "Price Discrimination by a Many-Product Firm," *Review of Economic Studies*, 66:1, 151 -168.

Benoit, J.-P. and V. Krishna (1985) "Finitely Repeated Games," *Econometrica*, 53, 905–922.

Bernheim, B.D. and M.D. Whinston (1990) "Multimarket Contact and Collusive Behavior," *Rand Journal of Economics*, 21, 1–26.

Billingsley, P. (1968) *Convergence of Probability Measures*, Wiley, New York, New York.

Campbell, C.M. (1998) "Coordination in Auctions with Entry," *Journal of Economic Theory*, 82, 425-450.

Casella, A. (2002) "Storable Votes," mimeo: Columbia University.

Casella, A. and T.R. Palfrey (2003) "Storable Votes," mimeo: Columbia University and Caltech.

Chakraborty, A., N. Gupta, and R. Harbaugh (2002) "Ordinal Cheap Talk in Common Value Auctions," mimeo: Baruch College CUNY, University of Michigan, and Claremont McKenna College.

Chakraborty, A. and R. Harbaugh (2003) "Cheap-Talk Comparisons in Multi-Issue Bargaining," *Economics Letters*, 78, 357-363.

Crémer, J. and R. McLean (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica,* 56, 1247–1257.

d'Aspremont, C. and L.-A. Gerard-Varet (1979) "Incentives and Incomplete Information," *Journal of Public Economics,* 11, 25–45.

Fang, H. and P. Norman (2003) "An Efficiency Rational for Bundling of Public Goods," Cowles Foundation Discussion Paper no 1441.

Hortala-Vallve, R. (2003) “Qualitative Voting,” mimeo: LSE.

Jackson, M.O. (2003) “Mechanism Theory,” in *Optimization and Operations Research*, edited by Ulrich Derigs, in the *Encyclopedia of Life Support Systems*, EOLSS Publishers: Oxford UK,[http://www.eolss.net].

Jackson, M.O. and H.F. Sonnenschein (2003) “Overcoming Incentive Constraints by Linking Decisions,” mimeo: Caltech, http://www.hss.caltech.edu/∼jacksonm/linking.pdf.

May, K. (1952) “A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision,” *Econometrica*, 20, 680-684.

McAfee, R.P., (1992) “Amicable Divorce: Dissolving a Partnership with Simple Mechanisms,” *Journal of Economic Theory*, 56:2, 266-293.

McAfee, R.P., J. McMillan, and M.D. Whinston (1979) “Multiproduct Monopoly, Commodity Bundling and the Correlation of Values,” *Quarterly Journal of Economics*, 100:3, 371-383.

Miller, N.R. (1977) “Logrolling, Vote Trading, and the Paradox of Voting: A Game-Theoretical Overview,” *Public Choice* , 30, 51-75.

Pesendorfer, M. (2000) “A Study of Collusion in First-Price Auctions,” *Review of Economic Studies*, 67, 381-412.

Tullock, G. (1970) “A Simple Algebraic Logrolling Model,” *American Economic Review*, 60:3, 419-426.

Wilson, R. (1969) “An Axiomatic Model of Logrolling,” *American Economic Review*, 59, 331-341.