

The Hurwicz Program, Past and Suggestions for the Future*

Andrew Postlewaite[†] and David Schmeidler[‡]

March 23, 2018

Abstract

There are two parts in our note. The first highlights Leo's contribution that earned him the Nobel prize in economics, from a personal point of view. The second discusses further questions that can naturally be addressed using Leo's conceptual framework.

*Forthcoming, *Social Design. Essays in Memory of Leonid Hurwicz*, Walter Trockel, ed., fall 2018.

[†]University of Pennsylvania, apostlew@econ.upenn.edu

[‡]Tel-Aviv University, davidschmeidler@gmail.com.

Mechanism Design

The modern neoclassical consumer model was formulated in Scandinavia between the world wars, but modern theory started essentially with the publication by Arrow and Debreu (1954) of the proof of existence of competitive (Walrasian or price) equilibrium. The first conceptual contribution by Leo (1962) was to separate the economic variables into two groups: The *environment*, as he termed it, includes the characteristics of economic agents, initial endowments, preferences, and production sets, and the allocation mechanism, i.e., the methods or institutions by which the society organizes the exchange of commodities and makes production and consumption decisions. Next (in the same paper) he introduced and formally defined concepts like: *performance correspondence*, *implementation*, *incentive compatibility*, *informational decentralization*, *equilibrium of a mechanism*, etc.(and proved theorems relating these terms). Until about the end of the sixties the theory, still named mathematical economics, dealt mainly with the properties of competitive equilibria including stability and convergence. See for example Arrow and Hahn. (1971).

In Leo's framework, the competitive mechanism is only one of many possible. In the theoretical literature, alternatives to price equilibrium were mostly considered in cases of market failure such as public goods, nonconvexities, etc.¹ Mechanism design has been also applied to allocations within firms and organization. However the main goal of the mechanism design research was not just to design mechanisms, but to investigate which combinations of desired properties of the performance correspondence can be implemented by a mechanism with desired properties. We should recall here that a mechanism is defined for a whole class of environments, such as neoclassical environments with some fixed number of commodities. The mapping that assigns equilibria outcomes (allocations) to environments is termed the performance correspondence (of this mechanism). Any correspondence from environments to outcomes is termed social choice correspondence (SCC). Thus a performance correspondence is a SCC implemented by the above mechanism.

The noncooperative game theory developed during the fifties and the sixties, and in the seventies it entered the economic theory and partially

¹Historically, alternatives to the market mostly originated from socialist utopias starting with utopian socialism, continuing with Marx and Engels' socialism and communism, and the attempts to implement them, from the USSR in 1917 to North Korea today (Jan 2018). In mainstream economics these utopias were relegated to economic history and the history of economic thought.

or mostly replaced competitive price equilibria as its principal tool of research. In the seventies Leo, and others who joined the mechanism design research program, replaced the abstract concept of mechanism with game forms, and the abstract equilibria with equilibria of strategic games like Nash equilibria, strong equilibria, dominant strategies equilibria Bayesian equilibria, etc.² Already by the late sixties Shapley and Shubik constructed a descriptive model of the market: a game form whose strategic equilibria coincided asymptotically with competitive equilibria, but suggested more realistic equilibria for oligopolistic markets. Many variants of this game form were suggested and investigated.

A central feature of Leo's contribution was to reverse the search for desirable mechanisms. Instead of inquiring the properties of equilibria of a specific game form, he started from the desired properties of a SCC: Does there exist a mechanism whose performance correspondence satisfies these desiderata? In addition, some desired properties were sometimes prescribed, for example, the domain of the environments, the informational requirements of the mechanisms including the type of the equilibrium, etc.

An important related area where mechanism design extended and re-defined the scope of research is the voting\social choice theory. It started with Arrow's cardinal impossibility result (1952) and continued with the by now classical Gibbard (1973) and Satterthwaite (1975) results showing the impossibility of a straightforward (Farquharson 1969, (1963)) voting rule. Concurrently, majority voting rules modelled along the lines of those used in parliament were investigated. The mechanism design approach asks whether there are mechanisms (voting rules) whose strategic equilibria (i.e., the performance correspondences) satisfy certain desiderata. Maskin (1977, 1991) showed that 'monotonicity' of a SCC is a necessary and almost sufficient condition for it to be possible to implement the SCC via Nash equilibria.

While these results about social choice mechanisms are of first order importance, the impetus for Leo's first work on mechanism design was motivated by the Lange-Lerner debates about the viability of socialism (Lange (1942), Lerner (1944) cite Hayek?). The debate was, to a large extent, whether a centralized system could uncover information dispersed among many agents, and use that information to achieve Pareto efficient outcomes. Formalizing this, Leo considered a set of pure exchange economies consisting of a finite number of agents, each of whom was characterized by a nonnegative initial endowment of the goods in an economy and a utility function

²Leo was not very keen of mixed strategies, vNM utility or Bayesian equilibrium.

over possible consumption bundles. A performance function on this set of pure exchange economies is then a function that maps each economy (a finite collection of agents) into an equilibrium redistribution of the agents' endowments. In Hurwicz (1972) Leo showed the impossibility of a performance function that lead to individually rational Pareto efficient allocations when the equilibrium notion was dominant strategy. The literature turned the question of whether there were performance functions with desirable properties when Nash equilibria was the solution concept, where quite general characterizations were obtained.

Much of the literature motivated by Leo's early work focused on the possibility of implementing Pareto efficient performance functions, motivated by the debates about planned economies. *What* performance functions can be implemented is obviously important, but *how* a performance function is implemented is no less important. Leo's conceptual framework separates the performance correspondence one might want to arise from the game form that governs individuals' behavior, allowing one to investigate the properties of the game form (the institutions that provide incentives for behavior) separately from the properties of the performance (the equilibria arising from the institutions).

There are two broad reasons to care about the properties of a game form that implements a given performance function: The analyst may care about the game form, and the agents participating in the game form may care.

The analyst's concern. As noted above, much of the mechanism design literature asks whether a performance function can be implemented in Nash equilibrium, that is, is there a game form for which the Nash equilibrium outcomes coincide with the outcomes specified by the performance function. The reliance on Nash equilibrium as the solution concept did not arise because it was particularly compelling, but rather, because it seemed the "least flawed" solution concept that gave interesting insights.³

Maskin's (1977) seminal paper mentioned above illustrates potential conceptual problems with Nash equilibrium as a solution concept. Roughly, the paper gives sufficient and (nearly) necessary conditions on a performance function for it to be implementable in Nash equilibria. Sufficiency is shown by constructing a game form whose equilibrium outcome is the outcome proposed by the performance function satisfying the sufficient conditions. An agent's strategy in the constructed game form includes a precise description of all agents' preferences, both other agents' preferences and her own. If all agents' agree on the profile of preferences, the outcome is that prescribed by

³See Jackson (1992) for an early paper along this line.

the performance function. The game form is cleverly designed so that the only possible equilibrium is that the agents agree.

This works fine when (as implicitly assumed) the profile of preferences is common knowledge among the agents. But the central idea is applied to pure exchange economies, the game form is highly discontinuous, and the slightest deviation by a single agent can lead to very bad outcomes.⁴ So, while the Nash equilibrium solution captures the incentives among agents, it is unreasonable to think of it as being plausible for this game form implementing, say, the (constrained)⁵ Walrasian performance function.^{6,7}

An analyst might well think that while Nash equilibrium is appropriate for some game forms, but prefer a game form that was continuous as this would avoid disastrous outcomes when small deviations from equilibrium play. Postlewaite and Schmeidler (1978) analyze a *continuous* game form for pure exchange economies in which there are Nash equilibria arbitrarily close to Walrasian equilibrium allocations when the number of agents is sufficiently large.

One might rate this game form as preferable to a Maskin-type game form on this basis, but less desirable on two counts. First, Nash outcomes are not fully Pareto efficient.⁸ Second, while when there are many agents there is a Nash outcome that is close to the Walrasian outcome, there are other equilibria that lead to no trade. This is in contrast to the performance of a Maskin-type game form which does not have such less desirable equilibrium outcomes.

There are other important characteristics of game forms besides continuity (or lack thereof) and multiplicity of equilibria. As mentioned above, in Maskin-type game forms an agent's strategy includes announcing the vector of preferences for the participating agents. As the number of agents gets large, the size of her messages grows proportionately. In addition to

⁴See Hurwicz, Maskin and Postlewaite (1995) for details.

⁵**Constrained Walrasian equilibria are essentially price and allocations for which all agents are maximizing subject to their budget sets and the feasibility of their demands. For simplicity, we will drop the "constrained" and refer simply to the Walrasian correspondence.**

⁶One might argue that if the problem is that agents may not, in fact, know precisely the preferences of all agents in the economy, one should then include in the basic model agents' *beliefs* about the preferences. This, however, doesn't really solve the underlying problem. Postlewaite and Schmeidler (1986) show that when following this path, one can accomplish the analogous Bayes Nash implementation for exchange economies with asymmetric information using a similarly discontinuous game form.

⁷See Eliaz (2002) for a discussion of this and related issues.

⁸Postlewaite and Wettstein (1989) demonstrate a continuous game form somewhat resembling Maskin-type game forms whose outcomes are constrained Walrasian.

the implausibility that she would have this information, there is the sheer difficulty of acting upon it. In contrast, in the Postlewaite and Schmeidler (1978) paper mentioned above, an agent's strategy is a vector of amounts of goods she wishes to put up for sale and the amount she is willing to spend for each good she wishes to buy. This is a vector of dimension twice the number of goods independent of the number of agents.⁹

The discussion so far has outlined features of a game form under consideration that an analyst might look at in evaluating the plausibility of Nash equilibrium as a solution concept for the game form. Another feature is the information an agent needs to determine her best response to other agents' strategies. In Maskin-type game forms an agent needs to know precisely the strategies of each and every agent to determine her best response, while in the Postlewaite and Schmeidler (1978) game form, agents need only predict the *sum* of other agents strategies. That, along with the continuity, might give greater plausibility to the Nash outcome of the game form.

Before turning to participants' possible concerns about properties of the game forms employed in the design of mechanisms, it is useful to mention how analysts often derive optimal mechanisms. A common technique is to invoke the revelation principle¹⁰

A participant's concern. The discussion above dealt with the analyst's concerns, driven primarily on the suitability of Nash equilibrium as the solution concept. In addition to those concerns, the participants of the game form might have concerns that are separate from questions Nash equilibrium. If I were an agent in a pure exchange economy who had a choice of what game form I would like to govern reallocation, I would care about many of the properties that the analyst cares about. I would like the game form to be continuous so I didn't need to worry about small trembles on my part or by other agents; I would like a game form that wouldn't entail my needing to predict all other agents' individual strategies in detail; I would prefer a game form where. I would also care about how complex my strategies in the game form, for example, am I to choose a finite dimensional vector? Do I need to choose from more complicated sets when there are more agents involved?

In addition to the properties of interest to the analyst, I would like to know how "risky" the game form is that is, how badly off could I be in a worst-case event? Suppose there is a given game form that implements

⁹See Hurwicz and Reiter (2002) and Mount and Reiter (1974) for a discussion of related issues.

¹⁰See Wikipedia https://en.wikipedia.org/wiki/Revelation_principle .

the Walrasian outcome for pure exchange economies. Suppose that I am an agent in a pure exchange economy and I play my part of a Nash equilibrium for this economy. For the game form in Hurwicz, Maskin and Postlewaite, the outcome may be the worst possible - my endowment is confiscated and I consume nothing. I would prefer an alternative game form that implemented Walrasian outcomes, but in which I could guarantee an outcome that was at least as good as my initial endowment even if other agents didn't play their Nash strategies (if such a mechanism existed). (Leo coined the term *non-confiscatory* for game forms which guaranteed that agents were guaranteed not to be worse off than their initial endowment, but this constraint was on equilibrium outcomes not, as suggested here, that this constraint hold should other agents choose non-equilibrium strategies.)

There are (at least) three notions of how I might guarantee that I will not be worse off than at my initial endowment: (i) for any strategy vector of other agents, I have a strategy that guarantees an outcome at least as good as my initial endowment; (ii) a stronger notion, that I have a strategy that for any strategy choice of other agents I will not be worse off than at my initial endowment; and (iii) an even stronger notion, that my equilibrium strategy guarantees I will not be worse off than at my initial endowment.

The game form in Postlewaite and Schmeidler (1978) satisfies the stronger notion, that an agent has a strategy that uniformly across all possible strategies other agents may choose leaves me as well off as with my initial endowment. Unfortunately though, that strategy leaves me with my initial endowment regardless of others' strategies. One would like a game form that has the desirable property (a strategy that leads to an outcome at least as good as my initial endowment) but also (at least sometimes) leads to gains relative to my endowment. One can imagine a game form analogous to that in Postlewaite and Schmeidler (1978), but one in which an agent chooses a demand function, and the outcome of the game form is a Walrasian equilibrium for the vector of demand functions chosen.¹¹ This has many of the desirable properties of the game form in Postlewaite and Schmeidler (1978) – the outcome function is (upper hemi) continuous, agents need only predict the sum of other agents' strategies to compute a best response, and the agent has a (natural) strategy that assures an outcome at least as good as her initial endowment (choose her honest demand function). But unlike the strategy that guarantees an outcome as good as the initial endowment in Postlewaite and Schmeidler (1978), announcing my true demand function typically gives a gain relative to my initial endowment. In fact, if all

¹¹Roberts and Postlewaite (1976) analyze such a game.

agents announce their true demand functions the outcome is the Walrasian outcome for the given exchange economy, and consequently Pareto efficient. The game form in which agents choose demand functions has a serious defect relative to the game form in Postlewaite and Schmeidler (1978), however. While for large economies agents (usually) gain little by deviating, whatever the other agents do, the Nash equilibrium outcome can be far from the Walrasian outcome. To our knowledge it is not known whether the Walrasian correspondence can be implemented with a game form for which an agent who plays her equilibrium strategy can be guaranteed an individually rational outcome.

While both the analyst and the participant might care about the aspects above of a game form that implements a particular SCC, there are other aspects that the analyst might be less interested in than a participant. I would like institutions that lead to efficient outcomes, but in addition I care about the process by which outcomes arise. For example, I prefer to share as little information about myself as possible, given the goal of implementing the Walrasian correspondence. An agent's equilibrium strategy in the game form in Hurwicz, Maskin and Postlewaite includes the agent's true preferences, while an agent's equilibrium strategy in the game form that also implements the Walrasian correspondence in Postlewaite and Wettstein has the agent revealing only his net trade at that Walrasian price. There might be instrumental reasons for wanting to reveal as little as possible in an implementing game form, such as a fear that information I reveal might be used to my detriment in the future. But separately from instrumental concerns, an agent might have a *direct* preference to maintain as much privacy as possible.

A participant may also care about the range of outcomes that he can effectuate when other agents play their part of a Nash equilibrium. By definition, at a Nash equilibrium, the outcome I get is as good or better than any of the others available to me given the play of others. But for two different game forms that give rise to the same equilibria, when others play their part of a Nash equilibrium the range of outcomes achievable as I choose different strategies in one may be larger than the range in another. For example, in the game form implementing the Walrasian correspondence in Hurwicz, Maskin and Postlewaite, at a Nash equilibrium I can achieve all feasible allocations that are no better than my Nash equilibrium outcome. In the game form in Postlewaite and Wettstein that implements the Walrasian correspondence, at a Nash equilibrium I can achieve only the feasible outcomes that give me a bundle that cost less than my Walrasian equilibrium bundle

(at the Walrasian equilibrium price), a (typically) strictly smaller set of outcomes. Hence, I would prefer the latter if I would like a smaller choice set available at equilibrium and the former if I like a larger set.

Dealing with the Walrasian correspondence as we have above is relatively easy for (at least) two reasons. First, an agent cares only about the bundle of goods that he consumes, and is indifferent about other agents' consumption. Second, agents' choices and the outcomes that result from those choices are precisely defined. When we move to interesting real-world mechanism design problems we see limitations of this framework.

The creation of the constitution of the United States is a leading example of a real-world problem in which a set of agents met to design institutions for a new country. The actors in the venture were very intelligent and knowledgeable, and engaged in prolonged heated discussion about the institutions they were creating. The power of the to-be-formed central government to levy taxes was one of the most contested issues. It was imperative that taxes to support an army be included if the system was to survive. Previous central authorities relied on voluntary contributions of the independent states. Predictably, free riding crippled the central authority. While this was universally recognized, many of the delegates charged with designing the constitution were very apprehensive of granting the central government too much power given the recent experience under British rule.

The conflicts among the delegates writing the new constitution illustrate two problems in mechanism design that typically do not show up in the standard academic mechanism design literature. First, while the constitutional delegates had different preferences over the outcomes that would result from the new constitution, much of the debate centered not so much on which outcomes were preferable, but instead on which outcomes *were likely* to arise from different sets of rules. It wasn't possible to completely describe the actions available to various players, or even what outcome would result from a given set of actions agents might take. In our language, there was no general agreement about what would be the Nash equilibrium outcomes from any proposed constitution.

This is not an issue in the academic mechanism design literature, as the "rules of the game" for writing academic papers in this area more or less require that the game form be specified precisely.¹² A necessary step in transferring the mechanism design methodology to many real-world prob-

¹²We are not arguing that there are *no* problems for which the game form can be precisely specified; there are, for example, problems in designing computer protocols where this issue may not arise. Rather, we suggest that there are important real-world problems where this issue is of first order importance.

lems is to formalize participating agents' difficulty in predicting equilibrium outcomes for the proposed mechanism.¹³

References

Arrow, K. and F. Hahn, (1971), *General Competitive Analysis*. San Francisco: Holden-Day.

Arrow, K. J. and Debreu, G. (1954), "Existence of an equilibrium for a competitive economy," *Econometrica*. 22 (3): 265–290.

Eliasz, K. (2002), "Fault tolerant implementation," *Review of Economic Studies* 69, 589-610.

Farquharson, R. (1969), "Theory of Voting," Yale University Press, New Haven.

Gibbard, A. (1973), "Manipulation of voting schemes: a general result," *Econometrica* 41, 587-602.

Hurwicz, L. (1951), "Theory of Economic Organization," *Econometrica*, 19, pp. 54.

Hurwicz, L. (1955), "Decentralized Resource Allocation," Cowles Commission Discussion Paper: Economics No. 2112.

Hurwicz, L. (1960a), "Optimality and Informational Efficiency in Resource Allocation Processes," Chapter 3 in K. J. Arrow, S. Karlin, and P. Suppes (eds.), *Mathematical Methods in the Social Sciences*, Stanford University Press, Stanford, pp. 27-46.

Hurwicz, L. (1960b), "Conditions for Economic Efficiency of Centralized and Decentralized Structures," in G. Grossman (ed.), *Value and Plan*, University of California Press, Berkeley, pp. 162-183.

Hurwicz, L. (1966), "On Decentralizability in the Presence of Externalities," paper presented at the San Francisco meeting of the Econometric Society (unpublished).

Hurwicz, L. (1969), "On the Concept and Possibility of Informational Decentralization," *American Economic Review*, 59, pp. 513-534.

Hurwicz, L. (1970), "Organizational Structures for Joint Decision Making: a Designer's Point of View," in M. Tuite, R. Chisholm and M. Radnor (eds.), *Interorganizational Decision-Making*, Aldine Press, Chicago.

Hurwicz, L. (1971), "Centralization and Decentralization in Economic Processes," Chapter 3 in A. Eckstein (ed.), *Comparison of Economic Systems: Theoretical and Methodological Approaches*, University of California Press, Berkeley.

¹³One's first reaction might be to model agents' uncertainty by putting a prior over agents' uncertainty and utilizing Bayes equilibrium as the solution concept. This doesn't seem to be a realistic solution to problems like writing constitutions.

Hurwicz, L. (1972), "On informationally decentralized systems", in Radner and McGuire, *Decision and Organization*. North-Holland, Amsterdam.

Hurwicz, L., E. Maskin and A. Postlewaite (1995), "Feasible Implementation of Social Choice Correspondences by Nash Equilibria," in J. Ledyard (ed.) *Essays in Honor of Stanley Reiter*, Kluwer Academic Publishers, pp. 367-433.

Hurwicz, L., R. Radner, and S. Reiter (1970), "A Stochastic Decentralized Resource Allocation Process," (unpublished).

Hurwicz, L., and S. Reiter (2006), *Designing Economic Mechanisms*, Cambridge University Press, New York.

Jackson, M. (1992), "Implementation in undominated strategies: A look at Bounded Mechanisms," *Review of Economic Studies* 59(4) pp. 757-775.

Lange, O. (1942), "The Foundations of Welfare Economics," *Econometrica*, 10, 215-228.

Lerner, A. P. (1944), *The Economics of Control*, New York.

Maskin, E. (1977), "Nash equilibrium and welfare optimality". Paper presented at the summer workshop of the Econometric Society in Paris, June 1977. Published 1999 in the *Review of Economic Studies* 66, 23-38.

Mount, K. and S. Reiter (1974), "The informational size of message spaces," *Journal of Economic Theory* 8, pp. 161-192.

A. Postlewaite and D. Schmeidler (1978), "Approximate Efficiency of Non-Walrasian Nash Equilibria," *Econometrica*, Jan., Vol. 46, No. 1, pp. 127-137.

Postlewaite, A. and D. Schmeidler (1986), "Implementation in Differential Information Economies," *Journal of Economic Theory* 39, pp. 14-33.

Postlewaite, A. and D. Wettstein (1989), "Implementing Constrained Walrasian Equilibria Continuously," *Review of Economic Studies* 56, pp. 603-612.

Roberts, J. and A. Postlewaite (1976), "The incentives for price-taking behavior in large exchange economies," *Econometrica* 44, pp. 115-129.

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