Mechanism Design Theory and Economic Institutions: Some Reflections

Dilip Mookherjee

Abstract We highlight the shortcomings of mechanism design theory to understand key institutional design issues such as the allocation of authority, organizational structure or safeguards against abuse of power. This requires incorporation of imperfect commitment, collusion, and costs of communication and information processing.

Mechanism design theory has been remarkably successful over the past half century in many areas of economic theory, such as auctions, voting and market design, to name just a few. Yet how relevant has it been to understanding institutions such as firms, organizations or governments? They have advanced our ability to model two out of the three key components of institutions identified by Williamson (1985): asymmetric information and opportunism. But not much progress has been made on the third component: bounded rationality.

To illustrate, any theory of institutions will have to address the question of allocation of authority: who is delegated what authority, who reports to whom, what are the safeguards against 'abuse of power'. In the absence of bounded rationality, broadly interpreted, a version of the Revelation Principle applies (as exemplified in the work of Myerson 1982): the performance of any specific allocation of authority can be replicated by a comprehensive contract (administered by a court or algorithm) in which all members report their information to the administrator, await instructions on what to do, and have incentives to be honest and obedient. The central administrator would process all the information received, compute the solution to an optimization problem, and then issue instructions to all members. Underlying assumptions include ability of the Principal (organization owners) to commit to a mechanism, absence of costs of communicating or processing information, and absence of collusion among agents. In such a world, the notion of 'allocation of authority' makes no sense. Analogous to the Modigliani–Miller Theorem in corporate finance, this asserts the 'irrelevance of institutions' in classical environments of the sort usually studied in mechanism design theory *sans* bounded rationality.

D. Mookherjee (✉)
Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA
e-mail: dilipm@bu.edu

To escape this quandary, the theory has to venture forth into difficult territory, where the assumptions underlying the Revelation Principle do not apply. The most significant of these pertain to bounded rationality, which Herbert Simon emphasized as key to understanding organizational structures. What exactly does bounded rationality mean? Two important dimensions are (i) limited capacity to store and process information, and (ii) costs of communication. The former implies limits on cognitive capacity of central administrators, which could rationalize distribution of information processing tasks to members. The latter implies the difficulty of members communicating all they know to the central administrator, which provides an alternative rationale for distributing decision-making authority to members so as to take advantage of their specialized non-communicable information. Similar arguments were made by critics of socialism in the celebrated debates of the 1930s such as Friedrich von Hayek.

However, Hayek did not draw attention to the opportunistic ('abuse of power') hazards inherent in systems with decentralization of authority. This would seem to be a key tradeoff underlying the design of allocation of authority. Studying this trade-off necessitates a theory of costs of communicating, storing and processing information. Progress was made in this respect by the literature on the 'message space size' led by Hurwicz (1960, 1977), Mount and Reiter (1974), and the literature on delay and information processing led by Bolton and Dewatripont (1994), Radner (1992, 1993) and van Zandt (1996). However, these theories abstracted from the modeling of incentives. Hence they could not be used to study the tradeoff between communication and information processing advantages of decentralization with resulting incentive problems. A few recent studies have attempted to study this tradeoff in specialized settings, but there is a long way to go. Examples are models of mechanism design that incorporate communication costs (Blumrosen et al. 2007; Fadel and Segal 2009; Kos 2012, 2013; Green and Laffont 1986; Melumad et al. 1992, 1997; Mookherjee and Tsumagari 2014), or the control of collusion (Baliga and Sjostrom 1998; Celik 2009; Dessein 2002; Faure-Grimaud et al. 2003). In my view, this is an important direction mechanism design theory needs to confront in order to make progress in studying institutions.

Another aspect of bounded rationality pertains to the inability of humans to foresee all possible future contingencies, and plan for them at some initial point of time. In such settings, a preference for 'flexibility' naturally emerges, whereby the Principal leaves open the possibility of adapting later decisions to the arrival of information as the future unfolds, rather than committing to all decisions upfront. Once the door is left open for such adaptation, the Principal would have the scope for possibly opportunistic ex post renegotiations that generate adverse ex ante incentives for agents. Many authors have argued this can give rise to a preference for delegation of authority by the Principal, or imposition of constraints on the Principal's own authority. Yet it is hard to formalize the source of the problem (i.e., unforeseen contingencies) that give rise to the commitment problem. It is equally difficult to develop a theory of how the Principal or agents can evaluate the outcomes of different allocations of authority in contingencies they cannot foresee.

There has been a profusion of recent literature in the field of organizational economics, relying mainly on theories of 'incomplete' contracts that rule out Myersonian revelation mechanisms by assumption (Aghion and Tirole 1997; Dessein 2002). The complexity of modeling bounded rationality motivate such an approach, raising questions regarding its foundations (as in the 1999 *Review of Economic Studies* symposium). Incomplete contract models rely on ad hoc restrictions on mechanisms, thereby raising concerns regarding the robustness of results to these assumptions. Very few authors succeed in modeling the underlying 'frictions' from first principles, as is the norm in mechanism design theory. This has resulted in an unfortunate divergence between the latter and contemporary models in organizational economics. The challenge is to develop models that are both derived from first principles, and tractable enough to generate interesting insights.

References

Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy, 105*(1), 1–29.

Baliga, S., & Sjostrom, T. (1998). Decentralization and collusion. *Journal of Economic Theory, 83,* 196–232.

Bolton, P., & Dewatripoint, M. (1994). The firm as a communication network. *Quarterly Journal of Economics, 109*(4), 809–839.

Blumrosen, L., Nisan, N., & Segal, I. (2007). Auctions with severely bounded communication. *Journal of Artificial Intelligence Research, 28,* 233–266.

Celik, G. (2009). Mechanism design with collusive supervision. *Journal of Economic Theory, 144*(1), 69–95.

Dessein, W. (2002). Authority and communication in organizations. *Review of Economic Studies, 69,* 811–839.

Faure-Grimaud, A., Laffont, J.-J., & Martimort, D. (2003). Collusion, delegation and supervision with soft information. *Review of Economic Studies, 70,* 253–280.

Fadel, R., & Segal, I. (2009). The communication cost of selfishness. *Journal of Economic Theory, 144*(5), 1895–1920.

Green, J., & Laffont, J. (1986). Incentive theory with data compression. In W. Heller, R. Starr & D. Starrett (Eds.), *Essays in honor of Kenneth Arrow* (Vol. 3). Cambridge: Cambridge University Press.

Hurwicz, L. (1960). Optimality and informational efficiency in resource allocation processes. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences* (pp. 27–46). Stanford: Stanford University Press.

Hurwicz, L. (1977). On the dimensional requirements of informationally decentralized pareto-satisfactory processes. In K. J. Arrow & L. Hurwicz (Eds.), *Studies in resource allocation processes* (pp. 413–424). Cambridge: Cambridge University Press.

Kos, N. (2012). Communication and efficiency in auctions. *Games and Economic Behavior, 75,* 233–249.

Kos, N. (2013). Asking questions. *Games and Economic Behavior, 87,* 642–650.

Melumad, N., Mookherjee, D., & Reichelstein, S. (1992). A theory of responsibility centers. *Journal of Accounting and Economics, 15*(4), 445–484.

Melumad, N., Mookherjee, D., & Reichelstein, S. (1997). Contract complexity, incentives and the value of delegation. *Journal of Economic and Management Strategy, 6*(2), 257–289.

Mookherjee, Dilip, & Tsumagari, M. (2014). Mechanism design with communication constraints. *Journal of Political Economy, 122*(5), 1094–1129.

Mount, K., & Reiter, S. (1974). The informational size of message spaces. *Journal of Economic Theory, 8*(2), 161–191.

Myerson, R. (1982). Optimal coordination mechanisms in generalized principal agent problems. *Journal of Mathematical Economics, 10,* 67–81.

Radner, R. (1992). Hierarchy: The economics of managing. *Journal of Economic Literature, 30,* 1382–1415.

Radner, R. (1993). The organization of decentralized information processing. *Econometrica, 61*(5), 1109–1146.

van Zandt, T. (1996). Decentralized Information Processing in the Theory of Organizations. In M. Sertel (Ed.), *Contemporary economic development reviewed, volume 4: The enterprise and its environment*. London: Macmillan Press.

Williamson, O. (1985). *The economic institutions of capitalism*. New York: Free Press.