Organizations in the Aggregate Economy

Esteban Rossi-Hansberg*

Economists generally agree that an important feature of any modern macroeconomic theory is an explicit aggregation of the microeconomic behavior of all agents in the economy. In the last century, the profession has gone from the formulation of some general aggregate relationships governing the evolution of the economy to detailed theories that explicitly incorporate the observed heterogeneity in many characteristics of agents and firms. Adding these microeconomic details has resulted in new insights on policy, as well as better and more detailed descriptions of modern economies. Although it is obviously important to recognize that an aggregate economy is formed by individual agents making explicit decisions, the standard aggregate models still abstract from a precise description of how these agents interact in small (or not so small) groups to produce, live, and consume. Most economic activity occurs in intermediate levels of aggregation: organizations. Firms or plants, but also cities, families, international production chains, political parties, and religious organizations, among many others, are examples of such organizations. Most of my recent research has concentrated on incorporating these organizations into general equilibrium theories in order to understand their implications for aggregate outcomes.

A starting point of this agenda is an understanding of how organizations affect economic growth. There is a set of fairly consistent facts for developed economies that suggest that the long-term growth rate of these economies is fairly stable over time. This suggests the need for theories that exhibit constant returns to scale in the factors that can be accumulated over time—a feature that most endogenous growth theories share. How can intermediate levels of organization affect the required linearity in aggregate production? Mark Wright and I argue that the organization of agents in cities is closely related to aggregate technologies with constant returns to scale.1 Agents organize production and their lives in cities because they obtain benefits from agglomeration: there are increasing returns at the local level. We claim that to reconcile the increasing returns at the local level with constant returns at the aggregate level, one must understand the role of cities. We thus propose a theory in which the number and sizes of cities react to industry productivity shocks in a way that exhausts the increasing returns at the local level and yields constant returns in the aggregate. According to our findings, cities are the reason to obtain aggregate balanced growth, but our mechanism also yields a size distribution of cities that very closely resembles the one observed in the data. This is a stark example in which considering intermediate organizations (in this case, spatial agglomeration in cities) is fundamental to understanding aggregate outcomes.

A similar argument can be made for considering the formation and characteristics of firms. In the presence of any form of fixed costs of production it is hard to rationalize constant returns to scale at the plant level. Given this, how can we reconcile the organization of economic activity in establishments of particular sizes with the behavior of the aggregate economy? Again we need to show that aggregation yields the necessary linearity in production. The key, however, is that we can do so while still being consistent with the rich data on establishment sizes and dynamics. Wright and I argue that the size distribution and growth patterns of establishments can be aggregated in a way that implies balanced growth but that also explains important cross-sectoral differences in the observed establishment growth rates.2 The key, we argue, is to recognize that fluctuations in prices of industry-specific factors lead to changes in the use and accumulation patterns of these factors, which in turn affect the employment size and growth of establishments.

All these theories incorporate a rich pattern of economic organization in industries, cities, and production plants, but they model individuals, and their human capital, as an industry-specific but uniform factor of production. For the theory, hiring ten workers with some knowledge is the same as hiring one expert and nine people with no knowledge, as long as total units of human capital are kept fixed. This distinction is important, because modeling the production process as simply a function of total inputs implies that the organization of agents in specific teams is irrelevant. In contrast, we may believe that sometimes having knowledge concentrated in one individual is more efficient than having it evenly distributed, because we can organize the tasks of production in ways that will exploit knowledge more intensively. The important distinction, I believe, is that knowledge is embedded in individuals and these individuals have limited time. Organizations maximize their access to the knowledge embedded in individuals by not making them do standard tasks that other, less able agents can also per-

---

*Rossi-Hansberg is a Research Associate in the NBER’s Program on Economic Fluctuations and Growth and International Trade and Investment and a Professor of Economics and International Affairs at Princeton University. His profile appears later in this issue.
form. Heterogeneous agents form production teams and the characteristics of this process, and the ability and knowledge of their co-workers, determines their productivity and compensation. Of course, once incorporated into an aggregate framework, these considerations have key implications for the distribution of wages in the economy, the size and characteristics of production teams, and aggregate productivity. In a series of papers, Luis Garicano and I have studied the implications of hierarchical organization for wage inequality, the cross-sectional implications of changes in information and communication technology (ICT), and, more recently, the effect of ICT on innovation and growth.

Incorporating the explicit decision of whom, and not only how many, workers to hire implies that the distribution of abilities in the population has important consequences for total output and efficiency. Naturally, it also has important consequences for trade and the international organization of production. An economy with many talented workers can organize more efficiently by using some less talented individuals to perform the most common and straightforward tasks. This implies that there are gains from trading production tasks internationally. Pol Antràs, Garicano, and I show that in such a framework the organization of international teams leads to more wage inequality in the south, gains from trade, and merchandise trade deficits in the north. We also study the role of intermediate managerial skills in generating trade in tasks across countries. We find that having intermediate skilled agents is essential to being a good target for offshoring if communication technology in the target country is not particularly good. The logic is simple: intermediate managers are necessary to save on international communication costs, but their opportunity cost is to setup a firm on their own (which increases with the quality of communication technology).

The papers discussed above take the stand that the tasks required for production are heterogeneous in terms of their difficulty or the ability or knowledge needed to solve them. Therefore, the location of their performance is determined by the distribution of skills in the different countries. But many tasks required for production are heterogeneous in other dimensions not related to ability or knowledge. A particularly important dimension for trade is heterogeneity in offshoring costs, namely, the cost of performing a task away from a firm’s headquarters. Some tasks can require very basic skills but need to be performed close to headquarters (like janitorial or transportation services), while others require substantial knowledge but can easily be performed far away from headquarters (like tax accounting, or many business services). Heterogeneity in offshoring costs implies that, as tasks performed by workers of all talents can be offshored, any worker could win or lose from offshoring. Gene Grossman and I show that if the motive for trade is cost differences then, under some conditions, reductions in the costs of offshoring will lead to Pareto gains in the source country. All workers may gain from offshoring, because firms that intensively use low-skill tasks obtain a relative cost advantage, even though tasks performed by, say, low-skilled workers are being offshored. This leads to a higher demand for low-skill labor and a higher low-skill wage. Of course, there may be other, more standard effects on factor prices if there are labor mobility frictions or if task trade leads to changes in relative prices.

It is clear from this research that modeling the production problem and the resulting organization in more detail has allowed us to better understand the aggregate implications of globalization. The emergence of global production chains is an organizational phenomenon that can have important effects on factor prices, income levels, trade patterns, and growth. Antràs and I recently reviewed the progress made by this literature in exploring the role of organizations in trade. As we argue in our survey, most of the available research on the international organization of production in aggregate theories studies the case of trade in tasks between a developed and an underdeveloped country (north-south offshoring). Trade in tasks is, however, by no means restricted to north-south relationships. It is probably more important between developed economies. Of course, the motive for trade is less transparent in this case. One needs to incorporate some form of increasing returns. One avenue is to use internal increasing returns as in the “New Trade Theory.” The difficulty with this approach is that it implies full specialization at the task level—an unappealing implication. Another avenue, which Grossman and I have favored in our work, is to use external increasing returns. Of course, the potential problem with using externalities in a competitive model is the existence of multiple equilibria attributable to standard coordination problems. We show that having a continuum of tasks, all of which are required for production, and letting firms outsource tasks and compete in prices, are enough to eliminate the potential multiplicity of equilibria. The reason is that an outsourcing firm can break an equilibrium in which the performance of the task is concentrated in a sub-optimal location by moving on its own, competing in prices, and obtaining the whole world market for the task. Importantly, since tasks are small relative to total production, the firm does not gain monopoly power. We use this framework to study the specialization pattern of tasks across developed economies. Our main result is that if two countries trade tasks, and are identical except for their size, the smaller one will specialize in the tasks with the low offshoring costs and will have lower wages.

To conclude, my work emphasizes the need to incorporate a rich set of organizations in our theories in order to understand aggregate phenomena. As is evident from this discussion, many
aspects of intermediate organizations have not yet been explored. Key among them are the dynamic implications of organization for factor investment and knowledge acquisition. Organizations are a key determinant of the ability to innovate and exploit efficiently available technologies. How can we study growth without paying full attention to the way in which these organizations develop over time?