

# THE IMPACT OF TRADE ON ORGANIZATION AND PRODUCTIVITY\*

LORENZO CALIENDO  
ESTEBAN ROSSI-HANSBERG

A firm's productivity depends on how production is organized. To understand this relationship we develop a theory of an economy where firms with heterogeneous demands use labor and knowledge to produce. Entrepreneurs decide the number of layers of management and the knowledge and span of control of each agent. As a result, in the theory, heterogeneity in demand leads to heterogeneity in productivity and other firms' outcomes. We use the theory to analyze the impact of international trade on organization and calibrate the model to the U.S. economy. Our results indicate that, as a result of a bilateral trade liberalization, firms that export will increase the number of layers of management. The new organization of the average exporter results in higher productivity, although the responses of productivity are heterogeneous across these firms. Liberalizing trade from autarky to the level of openness in 2002 results in a 1% increase in productivity for the marginal exporter and a 1.8% increase in its revenue productivity. Endogenous organization increases the gains from trade by 41% relative to standard models. *JEL* Codes: D21, D24, F12, F13.

## I. INTRODUCTION

Production requires organization. To produce, firms need to hire managers who organize inputs, solve the problems that arise in production, and plan for the future, among other tasks. As a result, the productivity of a firm depends on the way its production is organized. Of course, how much a firm is willing to invest in organizing its production by hiring a hierarchy of managers, with different amounts of knowledge and particular tasks within the firm, depends on the scale of its output. It would be clearly inefficient for a small family business to have the same organizational structure as a large multinational firm that sells all around the world. These differences in organization are obvious for anyone that has seen the complicated organizational hierarchy of a firm like IBM and, at the same time, has bought anything in a mom-and-pop shop around the corner. The implications of

\*We thank Pol Antràs, Kyle Bagwell, Robert Barro, Paco Buera, Arnaud Costinot, Luis Garicano, Elhanan Helpman, Wolfgang Keller, Alex Monge, Steve Redding, Steve Yeaple, three anonymous referees, and many seminar participants for useful conversations and comments. Caliendo thanks the International Economics Section at Princeton University.

© The Author(s) 2012. Published by Oxford University Press, on behalf of President and Fellows of Harvard College. All rights reserved. For Permissions, please email: journals.permissions@oup.com

*The Quarterly Journal of Economics* (2012), 1393–1467. doi:10.1093/qje/qjs016.

Advance Access publication on April 13, 2012.

organizational decisions for the gains from international trade and other changes in the economic environment are less obvious. In fact, organizational structure is absent in virtually all international trade theories.<sup>1</sup> How do these organizational differences affect productivity? How are they reflected in the within-firm distribution of wages and skills? How are organization and productivity impacted by firm decisions that change its scale, like the decision to export?

We study the internal organization of firms in a product market. The market is formed by a continuum of differentiated products. The demand for these products varies depending on their individual characteristics. An entrepreneur can enter the market after designing her product, which conveys a fixed cost. A particular design is associated with a random level of demand for the product. Once the entrepreneur observes the resulting demand for the product she envisioned, she can organize production to serve local and international markets. The decision of how to organize production in turn determines the productivity of the firm as well as its input use, price, and profits.

We model an economy in which production requires labor and knowledge. We follow Garicano (2000) and Garicano and Rossi-Hansberg (2004, 2006, 2012) and model firms as knowledge-based hierarchies where managers deal with exceptions. In this production structure workers in the firm provide work time. For each unit of work time they can *potentially* produce  $A$  goods. Problems have to be solved for output to be realized. To solve problems agents acquire knowledge. Acquiring knowledge is costly, so in general, it is not efficient for the firm to make workers learn how to solve infrequent problems. Instead, hierarchies are created where managers learn how to solve the less common problems—the “exceptions”—and workers deal with the routine problems. When workers are confronted with a problem they do not know how to solve, they ask a manager. The organizational structure determines the knowledge acquired by agents, the order in which different agents confront problems,

1. This is clearly the case for the two main models in international trade today, namely, Melitz (2003) and Eaton and Kortum (2002), as well as virtually all other available theories. Two exceptions are the work of Marin and Verdier (2008, 2010) and Conconi, Legros, and Newmen (2008). There is also a variety of papers that emphasize contracting frictions inside the firm to derive conclusions on the boundaries of the firm (see Antràs and Rossi-Hansberg 2009 for a review of the literature on organizations and trade).

and how they communicate with each other. Firms economize on the use of knowledge by leveraging the knowledge of managers. A larger firm can have more than one layer of managers, where managers learn less common problems the higher they are in the hierarchy. All firms are required to hire a top manager, the CEO or entrepreneur. Of course, adding a layer of management is costly, because one has to pay managers that do not generate production possibilities but only solve problems. As a result, adding layers involves an extra cost that is only worth paying if the firm produces a high enough quantity. Adding an extra layer can be thought of as reducing the marginal cost of the firm, in exchange for increasing the fixed cost. Nevertheless, given that we provide a fully microfounded theory of the organization of production, the menu of fixed costs and marginal costs faced by the firm is endogenous and depends on the cost of acquiring and communicating knowledge as well as other characteristics of the economy.

We incorporate this production technology into an economy in which agents have constant elasticity of substitution (CES) preferences and where firms compete with each other monopolistically. Hence, firms determine the price that maximizes their profits. As is common with this structure, they set prices using a constant markup rule. However, given that they are making simultaneous decisions about their organizational structure, their marginal cost is now an endogenous variable. In a closed economy, the only exogenous and heterogeneous characteristic of a firm is the level of demand for its product, which we parameterize with a demand shifter  $\alpha$ . In an open economy, the firm can decide to export to other markets as well. We fully characterize the cost minimization problem of a firm given a demand level. Crucially, the level of demand determines the productivity of a firm by determining its optimal organizational structure.

Our theory exploits the basic idea that the productivity of a firm is the result of the way production is organized given the level of demand for its product. This link between demand and productivity is present in any production technology in which a firm needs to make some costly investment decision to expand production. A classic example is the decision to expand by increasing the number of plants. Another one is the decision to invest in new technologies (as in Yeaple 2005; Costantini and Melitz 2008; Atkeson and Burstein 2010; Lileeva and Trefler 2010; Bustos 2011; Rubini 2011). We model the decision to form a management hierarchy by adding new layers of management,

since it allows us to study a rich set of internal characteristics of the firm, as well as the effect of learning costs and communication technology on firms' outcomes.

There is another link between the level of demand and productivity that is not present in other papers in the literature, namely, the explicit modeling of the discreteness of many of these investment decisions. Take, for example, the decision of how many plants to use for production. If each plant has an identical minimum efficient scale (MES),<sup>2</sup> the first effect is not present. However, because the number of plants the firm chooses to build is discrete, the productivity of the firm depends on its production level per plant relative to the MES of a plant (if average cost at the plant are convex). In this sense, the fact that the decision of how many plants to build is discrete will result in heterogeneous responses in productivity to positive demand shocks (or in differences in productivity across firms).<sup>3</sup>

In our theory, both of these mechanisms are present. The second effect arises naturally because the number of layers in the firm is discrete. However, because more layers allow the firm to manage the acquisition of knowledge of its labor force more efficiently, more layers also imply a reduction in the minimum average cost of the firm.

The article is related to the literature on knowledge-based hierarchies. The main distinction between this work and this literature is that we model production hierarchies in a product market where firms produce differentiated products. This allows us to study the internal organization and productivity of heterogeneous firms in general equilibrium. Furthermore, it allows us to study the effect of a bilateral trade liberalization on organization and firm productivity. The result is a multicountry model that can be readily contrasted with the data and delivers clear testable implications.<sup>4</sup>

2. The minimum efficient scale is defined as the scale of production at which a firm minimizes average costs given a number of layers.

3. The literature on multiproduct firms has considered the decision to add or drop products and the within-firm revenue productivity changes that result from this type of selection. See for example Bernard, Redding, and Schott (2010) and Mayer, Melitz and Ottaviano (2011).

4. Antràs, Garicano, and Rossi-Hansberg (2006, 2008) study the implications of offshoring on wage inequality using a similar model of organizations. In contrast to this article they model an economy in which firm heterogeneity is the result of worker heterogeneity and study how the ability to form international teams

The vast literature on trade and heterogeneous firms has been successful at providing a framework to analyze individual firm experiences in an international trade context. However, in this literature, a firm is just a technology to produce goods at a given marginal cost. The basic equilibrium framework of our model is similar to the one used in this literature but incorporates endogenous productivity heterogeneity across firms and wage heterogeneity across and within firms. In fact, we show that Melitz (2003) is a special case of our model when knowledge becomes costless.

The article aims to speak to the empirical literature on the response of firm organization and productivity to trade liberalization. The papers by Clerides, Lach, and Tybout (1998) and Bernard and Jensen (1999) show that larger and more productive firms export. They also provide evidence that firms that decide to export are larger and have a larger probability of survival, but they do not find evidence that exporting leads to higher productivity.<sup>5</sup> In our framework firms that export are larger in terms of employment, but their productivity does not necessarily rise. Other papers, such as Baldwin and Gu (2003), Van Biesebroeck (2004), and De Loecker (2007, 2011), have found more significant effects of trade liberalization on productivity.<sup>6</sup> These apparently contradictory findings can be rationalized with our theory. We find that the effect of trade on a firm's productivity is heterogeneous and depends on how large firms are before trade liberalization, the size of the new market, export costs, as well as information and communication technology. For example, nonexporters reduce their scale of production compared to autarky. Nonexporters that were producing in autarky above their MES gain in productivity, whereas nonexporters that were producing at the MES or below reduce their productivity unless they reduce the number of management layers in the organization.

---

changes the composition of two-layer production teams, all of which produce a single good. Here we abstract from exogenous worker heterogeneity and let knowledge acquisition and the number of layers be endogenously determined. Perhaps more important, we model an economy with firm heterogeneity in demand in which trade is the result of products being differentiated as in Melitz (2003).

5. Other papers that document similar findings are Delgado, Fariñas, and Ruano (2002) and Bernard and Jensen (2004).

6. Again, other papers are Aw, Chung, and Roberts (2000), Hallward-Driemeier, Iarossi, and Sokoloff (2002), Fernandes and Isgut (2006), Park et al. (2009), Aw Roberts, and Winston (2007), and Aw, Roberts, and Xu (2008).

Heterogeneous responses in productivity as a result of a bilateral trade liberalization have been recently documented by Lileeva and Trefler (2010). They study the responses of productivity and technology upgrading for Canadian plants after the Canada–United States Free Trade Agreement went into effect. They show that plants with low pre-agreement productivity invested more in information and communication technologies, as well as product innovation. Any of these investments could be attributed to changes in their organizational structure, but the first two are more closely related. They also document that new exporting firms that experienced productivity gains increased their domestic sales relative to nonexporters, as implied by our theory. A recent paper by Bustos (2011) also finds a heterogeneous response in technology upgrading. She finds that reductions in tariffs induce process innovation mostly in the third quartile of the firm size distribution and not only in new but also in old exporters.

Regarding the internal reorganization of firms, Guadalupe and Wulf (2010) present evidence that increasing competition as a result of a bilateral liberalization of international trade leads to flatter hierarchies, something our model can explain for the nonexporting firms. They also show that the opportunity to sell in more markets can lead to firms with more management layers, although the results here are more noisy. Given the heterogeneity in the responses to trade liberalization in our theory, looking at just average responses necessarily leads to less accurate results. So again, our theory can rationalize these findings. More important, by providing a structural framework of trade with organization we can relate the different dimensions of firms' actions and so can guide future empirical research on the determinants of within-firm responses to trade policy and technological changes.

The rest of the article is structured as follows. The next section presents the production technology and solves the cost minimization problem of firms. We also study the profit maximization problem of firms and the equilibrium of the model in a closed economy. Section III presents our results for the open economy. In Section IV we calibrate the model to the U.S. economy and perform several counterfactuals. Finally, Section V concludes. We include an analysis of alternative measures of productivity in Appendix A and relegate all proofs to Appendix B.

## II. THE MODEL

The economy is formed by  $\tilde{N}$  agents that supply one unit of time. All agents are ex ante identical and order consumption baskets according to constant elasticity of substitution preferences given by

$$(1) \quad U(x(\cdot)) = \left( \int_{\Omega} \alpha^{\frac{1}{\sigma}} x(\alpha)^{\frac{\sigma-1}{\sigma}} M \mu(\alpha) d\alpha \right)^{\frac{\sigma}{\sigma-1}}$$

where  $x(\alpha)$  denotes the consumption of variety  $\alpha$ ,  $M$  is the mass of products available to the consumer,  $\mu(\cdot)$  the probability distribution over the available varieties in  $\Omega$ , and  $\sigma > 1$  the elasticity of substitution. Note that  $\alpha$  is a demand shifter, so agents like varieties with higher  $\alpha$  better.

Agents can work either in the production sector or in the schooling sector. Since agents are ex-ante identical in equilibrium, all of them obtain an equilibrium wage  $w$  as compensation for the unit of time they supply in the market (they are also compensated for the knowledge they acquire, as we will discuss later). We call agents that want to start a new firm entrepreneurs. An entrepreneur pays a fixed cost  $f^E$  to design its product. After the product is developed, the demand for the product,  $\alpha$ , is realized from a cumulative distribution function given by  $G(\alpha)$ . This draw determines the demand schedule faced by the firm. The entrepreneur then creates an organization if her draw is high enough to obtain profits given the fixed production cost  $f$ .

Production requires labor and knowledge. Within an organization agents divide into two main classes: workers and problem solvers or managers. Workers supply their unit of time to generate a production possibility. For this production possibility to become output, the organization needs to solve a problem. Solving problems requires knowledge. In particular, each production possibility is linked to a problem drawn from the cumulative distribution  $F(\cdot)$ , with decreasing density. A realization  $z$  implies that to solve the problem the worker needs to have acquired a set of knowledge that includes  $z$  as an element. If the worker's knowledge includes  $z$ , the problem is solved and the production possibility becomes  $A$  units of output. If the worker cannot solve the problem, she asks a manager one layer above for a solution, if the firm has hired such a manager. Suppose it has. Then the manager spends  $h$  units of time listening to the worker's problem (so  $h$  is the cost of communication) and solves the problem if her



knowledge includes  $z$ . If the lowest layer of managers cannot solve the problem, the worker has the option to send the problem one layer higher, and so on, until the problem reaches the top layer, formed by a single agent, namely, the entrepreneur.<sup>7</sup>

The organizational design of the firm involves determining the number of layers as well as the knowledge and number of agents in each layer. The cost of learning is linear in the Euclidean length of the knowledge set. So if an agent learns how to solve problems in the interval  $[0, z]$ , she has to pay  $wcz$ . That is, the cost of knowledge is  $wc$  per unit of knowledge. The cost of knowledge is proportional to the wage, since learning requires teachers in the schooling sector that earn  $w$ . So learning one unit of knowledge requires  $c$  units of time of a teacher at wage  $w$ .<sup>8</sup> Of course, since  $F(\cdot)$  has a decreasing density, agents at the bottom of the hierarchy, which attempt to solve the problem first, learn the most common problems, and agents in higher layers learn more rare and exceptional problems progressively. We assume that knowledge is not cumulative so to solve a problem  $z$  the agent needs to know an interval that includes  $z$ , but we do not force the agent to learn how to solve all problems between  $z$  and 0.<sup>9</sup>

7. The assumption that all firms employ a unit of time at the top layer is important. It restricts small firms to have simple organizational structures. Otherwise, a firm that employs only a few units of time in total could have the same organizational structure as the largest firm, by hiring tiny fractions of the time of managers at all layers. In practice there are some constraints to the divisibility of the units of manager's time that a firm can hire. These constraints are represented in our theory by the assumption that the entrepreneur uses her full endowment of time in her firm.

8. Note that this interpretation implicitly assumes that every period workers need to be retrained. Since we consider a general equilibrium model in steady state that is calibrated to the share of education workers, one can reinterpret  $c$  as the cost of training workers times the share of workers that die every period. In that case, the cost of knowledge is  $c = \delta_M \tilde{c}$  where  $\delta_M$  is the population mortality rate and  $\tilde{c}$  is the cost of training workers when only new workers have to be trained. We return to this point in the calibration of the model, where our calibration under this interpretation implies that the stock of knowledge in units of workers is equivalent to about four times the population.

9. The cumulative case, in which all agents have to learn intervals that include the most common problems yields qualitatively very similar results. It is not clear to us, that one assumption dominates the other. Do CEOs need to know how to do the tasks of production workers even if in equilibrium they never do them? Of course, assuming cumulative knowledge would imply larger and increasing learning costs and, therefore, wages that are necessarily increasing in the layer of the employee.



Consider first the case of a self-employed entrepreneur, a one-worker firm. The entrepreneur draws problems from the known exponential distribution of problems,  $F(\cdot)$ , and acquires knowledge  $[0, z]$  to solve them. For simplicity, throughout the article we let  $F(z) = 1 - e^{-\lambda z}$  and we discuss the implications of this choice as we describe our results.<sup>10</sup> The parameter  $\lambda > 0$  regulates how common the problems faced in production are. Higher  $\lambda$  implies that problems are less costly to solve because the distribution has more mass close to zero.<sup>11</sup> Since the entrepreneur has only one unit of time and she is the sole worker in the firm, she has to spend her time generating production possibilities. This implies that even if she knows how to solve all problems, her output is bounded by the amount of output generated by a single production possibility, namely,  $A$ . The parameter  $A$  reflects the overall technology of the economy. So if the knowledge acquired by the entrepreneur is  $z_0^0$ , the fraction of problems that she solves is  $F(z_0^0) = 1 - e^{-\lambda z_0^0}$ , and the expected output she produces is  $A[1 - e^{-\lambda z_0^0}]$ . Throughout, we let  $z_L^\ell$  denote the knowledge of an agent in layer  $\ell$  of a firm with highest layer  $L$  (and therefore  $L + 1$  layers, since we denote the layer of workers by 0). We follow the same notation convention for all variables.

Suppose now the entrepreneur decides to hire some workers and form a hierarchical organization with two layers where she is at the top of the organization. The entrepreneur hires workers trained to solve the more routine problems,  $z_1^0$ . Denote by  $n_1^0$  the number of workers (or units of time) she hires. Because she is the only top manager,  $n_1^1 = 1$ . Each worker solves a fraction of problems  $[1 - e^{-\lambda z_1^0}]$  and the problems that they cannot solve are directed to the entrepreneur. Since the entrepreneur spends  $h$  units of time communicating with each worker that sends her a problem to solve, it follows that she can deal with at most  $\frac{1}{h}$  problems. Her time constraint is then given by  $hn_1^0 e^{-\lambda z_1^0} = n_1^1 = 1$  and so she can hire at most  $n_1^0 = \frac{e^{\lambda z_1^0}}{h}$  workers.

We can easily generalize this structure to a firm with  $L + 1$  layers in which workers ask managers to solve their problems sequentially. This flow of questions is optimal, as proven by

10. The decreasing density of problems is just a normalization, since agents always choose to learn how to solve the most common problems first. The exponential specification is not essential but adds some tractability to the model.

11. In the limit when  $\lambda \rightarrow \infty$ , knowledge is not an input in production since all problems can be solved by agents with an infinitesimal amount of knowledge.

Garicano (2000) as long as agents that cannot solve a problem do not know who can solve it. We assume throughout that this is the case. The span of control of each manager depends on the knowledge of the managers/workers below them, namely,  $\frac{n_L^\ell}{n_{L+1}^\ell} = e^{\lambda z_L^\ell}$ , for all  $0 < \ell < L$ , and  $\frac{n_L^0}{n_L^1} = \frac{e^{\lambda z_L^0}}{h}$ . We can then apply the law of large numbers to this problem and use the expectation of output as the realized level of output.<sup>12</sup>

We solve the organizational problem of the firm in two stages. We determine first the cost-efficient way of producing a given quantity. The solution to the cost minimization problem characterizes the cost function. Then, given a particular draw of  $\alpha$ , the entrepreneur acts like a monopolist and decides how to price the goods and how much output to supply.

### II.A. Cost Minimization

Consider a firm that produces a quantity  $q$ . The variable cost function of such a firm is given by

$$(2) \quad C(q; w) = \min_{L \geq 0} \{C_L(q; w)\}$$

where  $C(q; w)$  denotes the minimum variable cost of producing  $q$  and  $C_L(q; w)$  is the minimum cost of producing  $q$  with an organization with  $L + 1$  layers. Since the fixed costs of production do not change the organizational problem of the firm, in this section we consider only the variable cost  $C(q; w)$ , to which we would need to add fixed costs to obtain total costs. We do so in the profit maximization section. Throughout, to simplify the exposition, we refer to this function as the “cost function.”

Consider the minimization problem of an organization with  $L \geq 0$  layers:

$$(3) \quad C_L(q; w) = \min_{\{n_L^l, z_L^l\}_{l=0}^L \geq 0} \sum_{l=0}^L n_L^l w [cz_L^l + 1]$$

12. To apply the law of large numbers (LLN), suppose that workers generate a production possibility in  $dt$  units of time. Each of these production possibilities can yield  $Adt$  units of output and managers spend  $hdt$  units of their time with each problem that reaches them. Then, as  $dt \rightarrow 0$ , the number of problems generated by the worker in a unit of time,  $\frac{1}{dt}$ , converges to  $\infty$ . The LLN then implies that the worker can solve a fraction  $F(z_L^0)$  of them, producing  $AF(z_L^0)$  units of output. The rest of the problems are sent to the manager above, who spends a fraction  $h(1 - F(z_L^0))$  of her unit of time on them.

subject to

$$(4) \quad A[1 - e^{-\lambda Z_L^l}] n_L^0 \geq q,$$

$$(5) \quad n_L^l = n_L^0 h e^{-\lambda Z_L^{l-1}} \text{ for } L \geq l > 0,$$

$$(6) \quad \text{and } n_L^L = 1,$$

where  $Z_L^l = \sum_{\ell=0}^l z_L^\ell$  is the cumulative knowledge of the firm up to layer  $l$ . Because the knowledge of agents in two distinct layers never overlaps, if an agent learns an interval of size  $z_L^l$ , it knows how to solve problems in the interval  $[Z_L^{l-1}, Z_L^{l-1} + z_L^l]$  or  $[Z_L^{l-1}, Z_L^l]$ . Agents at lower layers learn more common problems to economize on the time of their managers, since the distribution of problems can always be normalized to have a decreasing density, like the exponential. So, for a given number of layers,  $L$ , the entrepreneur decides the units of time to hire in each layer ( $n_L^l$ ) and the knowledge of the employees at each layer ( $z_L^l$ ) subject to the amount of output she wants to produce and the time constraints. The entrepreneur supplies one unit of time, so  $n_L^L = 1$ .<sup>13</sup>

Figure I presents an example of the average cost and the marginal cost of an organization with  $L = 0$  and 1. With  $L = 0$ , as the firm approaches its maximum capacity of production,  $A$ , the average cost converges to infinity (only one-worker firms have a maximum capacity). If there was someone else in the organization who knew how to solve some of the more exceptional problems, the organization would be able to produce more output. This creates an incentive to create an organization with more layers. This is evident in Figure I where eventually the lowest minimum cost is achieved by adding a layer of managers.

Note that since  $n_L^L = 1$ ,  $n_L^0 = \frac{e^{\lambda Z_L^{L-1}}}{h}$ . The first-order condition with respect to  $z_L^0$  is

$$(7) \quad c + \lambda(cz_L^0 + 1) - \frac{\phi}{w} \lambda A + \frac{\theta_L^0}{wn_L^0} \leq 0.$$

13. This assumption is not just a normalization. It determines the boundaries of the hierarchy, and it implies that organizations cannot be just replicated at the minimum efficient scale by letting  $n_L^L < 1$ . If we only impose the restriction that the unit of CEO-time is greater than one,  $n_L^L \geq 1$ , the cost function would have flat parts at the level of the minimum efficient scale where  $n_L^L > 1$ . Since each agent is endowed with one unit of time, our restriction that  $n_L^L = 1$  implies that CEOs are dedicated to only one firm. The case of  $n_L^L > 1$  can be interpreted as partnerships.



not add anything to the production process. Under a restriction on parameters, summarized in Assumption 1 in Appendix B, we can guarantee that entrepreneurs and workers will also have positive knowledge. If the parameter restriction does not hold, it could be the case that either workers have zero knowledge or workers and entrepreneurs have zero knowledge. The intuition for the result is as follows. There are two possible ways the firm can reduce production, either by decreasing the knowledge of workers and problem solvers at each layer of the organization, or by reducing the number of layers of the organization. Assumption 1 guarantees that the cost of communication is large enough relative to the cost of learning, so the firm always prefers to reduce the number of layers before setting the knowledge of its employees to zero. This does not change any of the results that follow, but it does imply that we would need to carry the non-negativity constraints for the knowledge of workers and entrepreneurs throughout. We prove this in Proposition B.1, which we relegate to Appendix B.

With these results in hand, let us set the Lagrange multiplier for the non-negative knowledge constraint to zero,  $\theta_L^l = 0$  for all  $l$ .<sup>14</sup> We first characterize the marginal cost curve given  $L$ . Note that the marginal cost of a firm is equal to the value of  $\phi$  (the Lagrange multiplier of the output constraint). Hence, the first order condition with respect to  $z_L^l$  yields

$$(10) \quad MC_L(q; w) = \frac{\partial C_L(q; w)}{\partial q} = \phi = \frac{wch}{\lambda A} e^{\lambda z_L^l}.$$

As we can see, the marginal cost is an increasing function of the knowledge of the entrepreneur,  $z_L^l$ , the cost of acquiring knowledge relative to the knowledge requirements in production,  $\frac{wc}{\lambda}$ , and communication technology  $h$ . Recall that the higher  $h$  is, the worse are employees at communicating problems and so the higher the marginal cost. Naturally, the marginal cost is a decreasing function of the level of technology,  $A$ . We prove later that, given the number of layers  $L$ , the marginal cost of a firm is increasing in output,  $q$ . The result implies that the knowledge and number of agents at all layers is increasing in  $q$ . In turn, the increase in knowledge implies that the span of control of

14. Strictly speaking we still could have zero knowledge for  $q$  values and layers that are never optimally chosen by the firm. However, given that these ranges are irrelevant for our purposes, we proceed as if the constraint never binds.

managers,  $\frac{n_l^l}{n_{l+1}^l}$ , increases, although the knowledge of the entrepreneur relative to the total knowledge in the firm—their autonomy measured as  $\frac{z_L^L}{Z_L^L}$ —decreases with  $q$ . Intuitively, given the number of layers, increasing the quantity produced implies that the chosen organizational structure becomes more restrictive and so the cost of producing an extra unit increases. Therefore, investing in a new layer is a way of expanding the production capacity of the firm.

The cost function of a firm with  $L$  layers is given by

$$(11) \quad C_L(q; w) = \frac{wc}{\lambda} \left( \frac{h}{A} e^{\lambda z_L^L} q + \left( 1 - e^{\lambda z_L^{L-1}} \right) + \lambda z_L^L + \frac{\lambda}{c} \right),$$

where  $z_L^L$  and  $z_L^{L-1}$  are the solutions to the first-order conditions (8) and (9).<sup>15</sup> The average cost function of a firm with  $L$  layers is then given by  $AC_L(q; w) = \frac{C_L(q; w)}{q}$ .

The average cost functions for different  $L$ 's all share a number of characteristics. They are all homogeneous of degree 1 with respect to factor prices. The two factor prices are  $cw$  for knowledge and  $w$  for a worker's time. So this implies that the cost functions are homogeneous of degree 1 in  $w$ . The choice of knowledge levels is invariant to  $w$ , because it affects all costs proportionally. Finally, the average cost curves are convex in  $q$ . When production is small, an increase in production reduces average costs up to the MES. Further increases in output increase average cost: The firm needs to provide too much knowledge to agents in low layers of the organization to achieve such a large scale. The following proposition characterizes the cost function given a number of layers  $L$ .

**PROPOSITION 1.** Given the number of layers  $L$ :

- (1) The marginal cost,  $MC_L(q; w)$ , the knowledge of all employees,  $z_L^l$  for all  $l$ , the number of employees at all  $l < L$ ,  $n_l^l$ , and the the span of control of managers at all  $l$ ,  $\frac{n_l^l}{n_{l+1}^l}$ , all increase with  $q$ ; while the entrepreneur's autonomy,  $\frac{z_L^L}{Z_L^L}$ , decreases with  $q$ .

15. The cost function for the case of  $L = 1$  is slightly different. In particular,

$$C_1(q; w) = \frac{wc}{\lambda} \left( \frac{h}{A} e^{\lambda z_1^1} q + \left( 1 - \frac{e^{\lambda z_1^0}}{h} \right) + \lambda z_1^1 + \frac{\lambda}{c} \right).$$

- (2) The cost function and the corresponding factor demands have the following properties: (i).  $MC_L(q; w) > 0$ , for all  $L$  and  $q$ , and so the cost function is strictly increasing. (ii). The cost function is homogeneous of degree one in  $w$ . (iii). Conditional factor demands are homogeneous of degree 0 in  $w$ . (iv). The average cost curve is convex in  $q$ , has a minimum  $q_L^*$  where it crosses the marginal cost curve, and converges to infinity when  $q \rightarrow 0$  or  $q \rightarrow \infty$ .

Figure I illustrates part 2 of Proposition 1. It shows the convex average cost functions for organizations with one and two layers, together with the corresponding marginal cost curves. Now consider the changes in the cost function as the firm decides to increase the number of layers. We show later that when a firm increases the number of layers, it also increases discontinuously the quantity produced. Consider  $q^*$ 's at the point of a switch in layers (when the argmin of (2) increases). The marginal cost at this point always jumps down discontinuously and so does the knowledge and span of control of all employees. The reason is simple: when the firm increases the number of layers, it invests in the ability to produce more output. Adding a layer of management allows the firm to economize on the knowledge of the many agents in the lower layers. When firms add a layer, to increase output accordingly, they also expand the number of workers in each layer and reduce the overall autonomy of the entrepreneur.

Figure II presents an example of the average cost curves for an organization with up to three layers ( $L = 2$ ). As we can see, the minimum average cost is decreasing in the number of layers. Furthermore, the level of output that minimizes average cost is increasing with the number of layers. The intuition is that firms with more layers are able to economize on the knowledge acquisition of their employees by leveraging on the knowledge of its managers at higher scales of output. The firm is paying more "fixed" costs to have a large organization and this allows it to produce at lower average costs, as long as its scale is large enough. This results imply that the average cost curves cross between the MES of the firms with adjacent numbers of layers. This is relevant for our purposes, since it implies that the average cost function at the MES, and the region around it, will be part of the lower envelope that determines the average cost curve when the



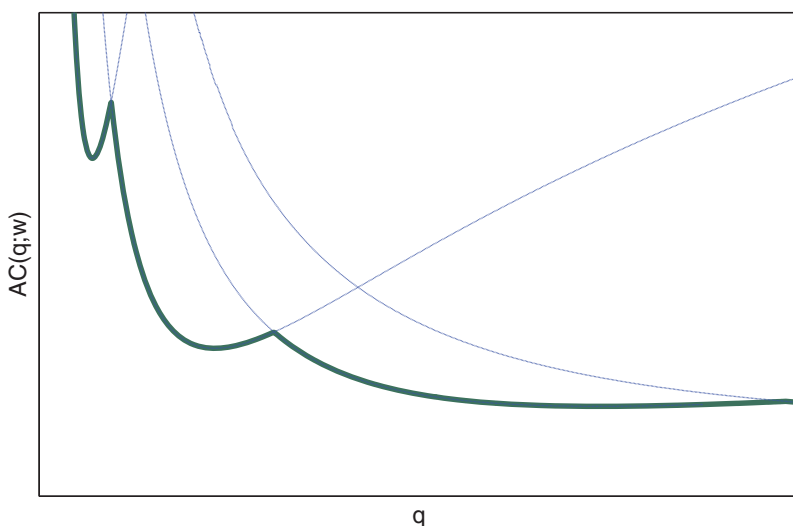


FIGURE II  
The Cost Function

number of layers is optimized as well, as in equation (2). We characterize the cost function across layers in the next proposition.

**PROPOSITION 2.** The minimum average cost for a given number of layers decreases with the number of layers and the level of output that achieves this minimum,  $q_L^*$ , increases with the number of layers. The average cost curves of an organization with  $L$  and  $L + 1$  layers cross in the interval  $(q_L^*, q_{L+1}^*)$ . At the crossing, firms add a layer of management, and the marginal cost,  $MC_L(q; w)$ , the knowledge of all employees,  $z_L^l$  for all  $l$ , the span of control of managers,  $\frac{n_L^l}{n_{L+1}^l}$ , and the entrepreneur's autonomy,  $\frac{z_L^L}{z_{L+1}^L}$ , decrease discontinuously with  $q$ ; while the number of employees at all  $l$ ,  $n_L^l$ , increase discontinuously with  $q$ .

The results we have proven in this section imply that the average cost function that solves the problem in (2) has the shape illustrated in Figure II. This curve is the lower envelope of the average cost curves given the number of layers. It has an overall decreasing shape, but it is not monotone. Increases in  $q$  can increase average costs, since they move the firm away from the MES but not by enough to trigger a change in the number of

layers. The shape of this function is critical in what follows and for our main results on the impact of trade on organization and productivity.<sup>16</sup>

Measuring productivity at the firm level involves more than just looking at marginal costs. The firm is using labor and knowledge as inputs in production. Hence, to measure productivity we need to aggregate variable factors. Of course, the cost function gives us an appropriate way to do so. Throughout, we measure the productivity of a firm producing a quantity  $q$ ,  $a(q)$ , as the inverse of variable average costs, that is, average costs without considering the fixed costs of production. This is done by taking the inverse of the  $AC$  function we defined and characterized already, but setting the wage to one (since the input bundle is measured in units of workers so as to keep factor prices constant). Hence, our measure of firm-level productivity is given by  $a(q) = \frac{q}{C(q;1)} = \frac{1}{AC(q;1)}$ , where the average cost is net of any fixed costs of production. We discuss other measures of productivity in Appendix A.

Communication technology,  $h$ , and the cost of acquiring knowledge about a fraction of problems,  $\frac{c}{\lambda}$ , are key to determining the average cost function. Figures III and IV show an example of the effect of changes in  $h$  and  $\frac{c}{\lambda}$ , respectively. First note that in both cases higher  $h$  or  $\frac{c}{\lambda}$  increases the average cost for all  $q$ 's beyond  $L=0$ . When the entrepreneur decides to work on her own ( $L=0$ ),  $h$  plays no role in determining average costs, since there is no communication inside the firm. In contrast, the average cost curve increases with  $\frac{c}{\lambda}$  for all  $q$ , including the ones where the entrepreneur decides to be self-employed, because even then she has to acquire some knowledge. Note also that the higher  $h$  or the lower  $\frac{c}{\lambda}$ , the flatter are the average cost curves beyond layer zero. As we increase the number of layers to more than one, moving beyond the MES is more costly the higher any of these costs.

16. To examine our results further, consider the case in which firms can choose continuously their number of layers. Then, the average cost function would be the lower envelope of average cost functions that change continuously. In fact, the average cost function would be formed only by the minimum of the average costs. Since we know that the minimum of the average cost is decreasing in the number of layers, the resulting average and marginal cost functions would be continuous and strictly decreasing in quantities. Such a model is a straightforward extension of our setup, but the interpretation of a continuum of layers is challenging and hard to reconcile with reality, since we observe firms choosing a discrete and small number of layers or classes of employees (as well as plants, products, etc.). Hence, this extension eliminates an essential part of a firm's organizational problem: namely, the many discrete decisions of firms.

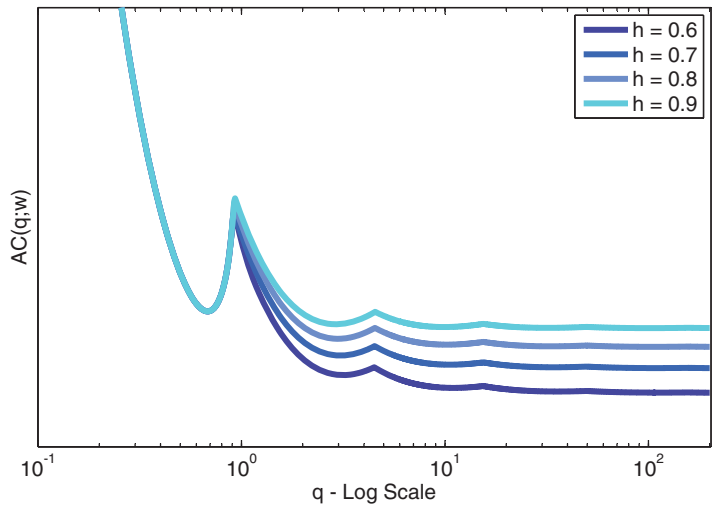


FIGURE III  
Changing  $h$  ( $\frac{c}{\lambda} = 1$ )

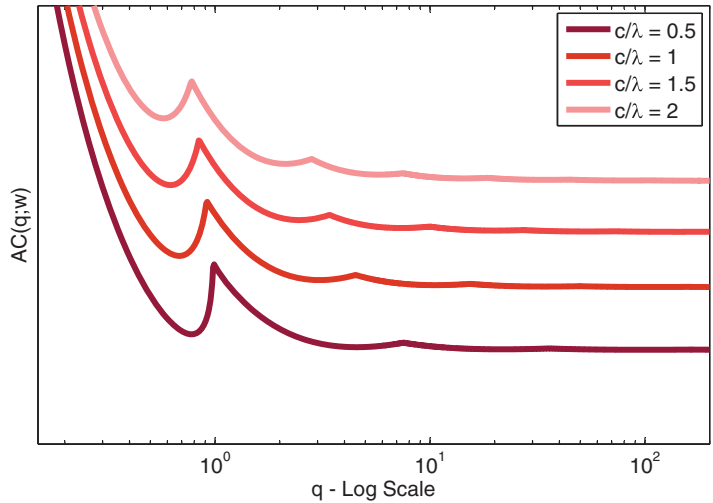


FIGURE IV  
Changing  $\frac{c}{\lambda}$  ( $h = 0.7$ )

We now turn to the characterization of the cost function when  $\frac{c}{\lambda} \rightarrow 0$ , that is, what happens in this economy when knowledge becomes either unnecessary to solve problems or extremely cheap. Remember that the role of organization in our economy is to economize on the acquisition and communication of knowledge. In this case organization has no purpose. This implies that there would be no firm with more than two layers. Firms would still have two layers instead of one because otherwise their production is limited to  $A$ . In fact, if we eliminate the possibility of self-employment, one can show that in this limit the cost function is as in Krugman (1980) or Melitz (2003) with an additional fixed cost. Namely, the marginal cost is constant, the average cost is decreasing due to fixed costs, and the cost function involves a fixed cost plus a variable cost equal to output times the marginal cost. Hence, in the limit without self-employment, our technology includes the standard technology used in trade models as a special case. Overall, since the model specification that follows is identical to that in Melitz (2003), with heterogeneity in demand rather than in marginal costs, the special case of our model when  $\frac{c}{\lambda} \rightarrow 0$  and when self-employment is not an option is isomorphic to Melitz (2003).<sup>17</sup> The result is proven in Proposition 3.

**PROPOSITION 3.** In the limit when  $\frac{c}{\lambda} \rightarrow 0$  and  $L \geq 1$ , the variable cost function is given by  $C(q; w) = \frac{qw}{A}$  and so  $AC(q; w) = MC(q; w) = \frac{w}{A}$ . Furthermore, in the limit productivity is given by  $a(q) = A$ .

In this limit the model generates another fixed cost, so total fixed costs are given by  $w(1+f)$ . Since the cost function, as we defined it, includes  $w$  but not  $wf$ , we just need to subtract  $w$ . That is, because the variable average cost is given by  $\frac{w}{A}$ , our measure of productivity results in exactly  $A$ . Hence, for this case our

17. The need to eliminate the possibility of self-employment to obtain the standard technology arises from the fact that when  $L=0$  the entrepreneur can act as manager and worker at the same time. Allowing for self-employment would amount to having an extra technology in Melitz (2003) where a firm can produce up to  $A$  without paying the fixed cost of production. In this case the variable cost function at the limit when  $\frac{c}{\lambda} \rightarrow 0$  would be given by

$$C(q; w) = \begin{cases} w & \text{if } q \leq A \\ q \frac{w}{A} & \text{if } q \geq A. \end{cases}$$

measure of productivity coincides exactly with the one in Melitz (2003) and others.<sup>18</sup> Clearly when  $\frac{c}{\lambda} > 0$ , productivity is not solely determined by  $A$  but is a function of how much the firm produces (and the corresponding organizational structure) given its demand draw  $\alpha$ , as well as communication and information technology. That is, in our framework, differences in demand are mapped into differences in technology via the organizational decisions of the firm.

### II.B. Profit Maximization

Consider the profit maximization problem of an entrepreneur with demand draw  $\alpha$  and the cost function,  $C(q; w)$ , discussed in the previous section. The entrepreneur chooses the quantity supplied in the market and its price subject to her demand draw. Households have the iso-elastic preferences in equation (1), and so the demand schedule faced by an entrepreneur with draw  $\alpha$  is given by  $p(\alpha) = q(\alpha)^{-\frac{1}{\sigma}}(\alpha R)^{\frac{1}{\sigma}}P^{\frac{\sigma-1}{\sigma}}$ , where  $R$  denotes total revenues in the economy,  $q(\alpha) = \tilde{N}x(\alpha)$ , since quantity produced is equal to the quantity demanded by each agent times the number of agents, and  $P$  denotes the ideal price index, which we normalize to one throughout. The problem of an entrepreneur with draw  $\alpha$  is

$$\begin{aligned} \pi(\alpha) &= \max_{q \geq 0} \tilde{\pi}(q; w) = \max_{q \geq 0} p(\alpha)q - C(q; w) - wf \\ (12) \quad &= \max_{q \geq 0} q^{1-\frac{1}{\sigma}}(\alpha R)^{\frac{1}{\sigma}} - C(q; w) - wf, \end{aligned}$$

where in the last line we substituted the demand schedule and where  $f$  denotes the fixed costs of production. The first-order condition of this problem implies that the price is given by a constant markup over marginal cost,

$$(13) \quad p(\alpha) = \frac{\sigma}{\sigma - 1} MC(q(\alpha); w).$$

In contrast to standard models, here the marginal cost, and therefore the price, depend on the quantity produced. The quantity

18. Of course, in Melitz (2003), as in the limit  $\frac{c}{\lambda} \rightarrow 0$  of our model, productivity is only constant as long as the inputs used to pay the fixed costs are not taken into account.

supplied is then implicitly determined by

$$(14) \quad q(\alpha) = \alpha R \left( \frac{\sigma}{\sigma - 1} MC(q(\alpha); w) \right)^{-\sigma}.$$

Firms with different  $\alpha$ 's charge different prices and supply different quantities of goods to the market. It is easy to show using equation (14) and Propositions 1 and 2 that  $q(\alpha)$  is an increasing function of  $\alpha$ . Furthermore, when the firm decides to increase the number of layers, the marginal cost falls discontinuously and so  $q(\alpha)$  jumps up. The elasticity of quantity with respect to the decline in marginal cost as we change the number of layers is given by  $-\sigma$ , and the elasticity of the price with respect to the jump in marginal cost is 1. In fact, since the price is proportional to marginal costs, it is increasing in  $\alpha$  given a number of layers  $L + 1$  and jumps down when the firm increases the number of layers. Hence, prices are not monotone in  $\alpha$ , which implies that larger firms (firms with higher  $\alpha$  and therefore  $q(\alpha)$ ) do not necessarily charge higher or lower prices.

**PROPOSITION 4.** The quantity produced,  $q(\alpha)$ , and revenue,  $r(\alpha) = p(\alpha)q(\alpha)$ , increase continuously with  $\alpha$  for a given number of layers and increase discontinuously when the firm increases the number of layers. The price,  $p(\alpha)$ , increases continuously with  $\alpha$  for a given number of layers and decreases discontinuously when the firm increases the number of layers.

The previous result is intuitive. Given  $\alpha$ , the entrepreneur decides the number of layers of the organization. The pricing rule is to set prices such that marginal revenue is equal to marginal cost. As we increase  $\alpha$ , the marginal revenue schedule changes and the entrepreneur reoptimizes, finding in some cases profit maximizing to restructure the organization and increase the number of layers. Consider an  $\alpha$ , say  $\hat{\alpha}$ , such that the entrepreneur is just indifferent between producing with  $L$  or with  $L + 1$  layers. Note that by Proposition 2, the organization with  $L + 1$  layers has a lower marginal cost than the organization with  $L$  layers at  $\hat{\alpha}$ . This implies that the quantity supplied under organization  $L$  cannot be the optimal quantity supplied under organization  $L + 1$ . In particular, since the  $L + 1$  organization has a lower marginal cost, the entrepreneur produces a higher quantity. A parallel logic then implies the corresponding changes in prices and quantities proven in Proposition 4.

The profits of a firm with demand realization  $\alpha$  (so not including the cost of entry) are given by

$$(15) \quad \pi(\alpha) = r(\alpha) - C(q(\alpha); w) - wf = \frac{r(\alpha)}{\sigma} - w \left( \frac{c}{\lambda} \left( 1 - e^{\lambda z_L^{L-1}(\alpha)} \right) + cz_L^L(\alpha) + 1 + f \right),$$

where  $\pi(\alpha) = \tilde{\pi}(q(\alpha); w)$  for the function  $q(\alpha)$  solves the fixed point in equation (14). Note that because  $C(\alpha; w)$  is not strictly convex due to changes in the number of layers, to maximize profits we need to maximize the profits given a number of layers and then compare these local maxima to determine the global maxima of the overall profit function. It is easy to prove that the profit function for each individual layer is concave and that  $\tilde{\pi}(0; w) = -w(1 + f)$  and  $\lim_{q \rightarrow \infty} \tilde{\pi}(q; w) = -\infty$ , so the local maxima of the problem in equation (12) are well defined and determined by the first-order conditions of the problem. Note also that since we are comparing local optima that change continuously with  $\alpha$ , revenues and profits are continuous and increasing in  $\alpha$ . We prove these results in Proposition 5.

**PROPOSITION 5.** Given  $L$ , the profit function  $\tilde{\pi}(q(\alpha); w)$  is strictly concave in  $q$ . Furthermore,  $\pi(\alpha)$  is increasing and continuous in  $\alpha$ .

### *II.C. Equilibrium in the Closed Economy*

Entrepreneurs pay a fixed cost of entry  $f^E$ , in units of labor, that buys them a draw from the known distributions of  $\alpha$ 's,  $G(\alpha)$ . We let  $G(\alpha)$  be a pareto distribution with coefficient  $\gamma$ , so  $G(\alpha) = 1 - \alpha^{-\gamma}$ . Let  $M^E$  denote the total mass of firms that enter in a given period. Firms that enter decide whether to produce or exit. Let  $M$  denote the mass of firms operating in equilibrium. Each period a fraction  $\delta$  of operating firms exits and is replaced by the same number of new entrants.

The value of operating a firm with draw  $\alpha$  is then given by

$$v(\alpha) = \max \left\{ 0, \sum_{t=0}^{\infty} (1 - \delta)^t \pi(\alpha) \right\} = \max \left\{ 0, \frac{\pi(\alpha)}{\delta} \right\}$$

where  $\pi(\alpha)$  is given by equation (15). Note that the fixed cost of production determines a cutoff level  $\bar{\alpha}$  such that firms with  $\alpha > \bar{\alpha}$  produce. In particular, the marginal firm,  $\bar{\alpha}$ , makes zero profits.



Thus, the zero cutoff profit condition for the marginal firm defines the threshold  $\bar{\alpha}$ ,

$$(16) \quad \pi(\bar{\alpha}) = 0.$$

Of the mass of new entrants, the fraction that operates is  $[1 - G(\bar{\alpha})]M^E$ . In a stationary equilibrium, the new entrants replace the number of firms that exist, and so  $[1 - G(\bar{\alpha})]M^E = \delta M$ , where  $M$  denotes the total number of active firms. Conditional on entry, the distribution of  $\alpha$  is given by  $\mu(\alpha) = \frac{g(\alpha)}{[1 - G(\bar{\alpha})]}$  if  $\alpha > \bar{\alpha}$  and  $\mu(\alpha) = 0$  otherwise, where  $g(\cdot)$  is the density of  $G(\cdot)$ .

Free entry implies that the expected profits are zero accounting for the entry costs, that is

$$(17) \quad \int_0^\infty v(\alpha)g(\alpha)d\alpha = \int_{\bar{\alpha}}^\infty \frac{\pi(\alpha)}{\delta}g(\alpha)d\alpha = wf^E.$$

In equilibrium, labor markets clear. Labor is used for several purposes, as workers and managers in a firm, as teachers to train workers and managers, and to design new products and cover the fixed costs of production. Because the total mass of agents in the economy is given by  $\tilde{N}$ , the labor market clearing condition is given by

$$(18) \quad \tilde{N} = \frac{M}{1 - G(\bar{\alpha})} \left( \delta f^E + \int_{\bar{\alpha}}^\infty (C(q(\alpha); 1) + f)g(\alpha)d\alpha \right).$$

Finally, equilibrium in the goods market implies that

$$(19) \quad R = w\tilde{N}.$$

The general equilibrium of this economy is characterized by the demand threshold of the firm that obtains zero profits,  $\bar{\alpha}$ , real wages,  $w$ , revenue  $R$ , and the mass of firms that operate  $M$ . Remember that we have normalized the price index  $P$  to 1. These four equilibrium variables are obtained by solving the four equations (16), (17), (18), and (19). We show the following.

**PROPOSITION 6.** There exists a unique stationary equilibrium.

An equilibrium of this economy exhibits a useful property due to the assumed CES demand system. In equilibrium an increase in total population from  $\tilde{N}$  to  $\tilde{N}'$  leads to higher wages given by  $w' = (\frac{\tilde{N}'}{\tilde{N}})^{\frac{1}{\sigma-1}}w$ , but does not change the entry threshold  $\bar{\alpha}$  or the quantities produced. Hence, we can easily show that in equilibrium higher population leads to higher prices, revenues,

and operating profits of active firms. In fact, prices, revenues, and profits all change proportionally by the factor  $(\frac{\tilde{N}}{N})^{\frac{1}{\sigma-1}}$ . Of course, as is standard in this literature, these changes come from the gain in variety associated with the higher population and the increase in the number of firms that the increase in population triggers. The organizational structure enters the result only in that it leads to a cost function that is homogeneous of degree 1 in wages. However, because quantities produced are not affected by increases in  $\tilde{N}$ , the organizational structure of firms is unaffected.

**PROPOSITION 7.** In equilibrium a larger population size does not affect the entry threshold or the quantities produced (and therefore the organizational structure), but increases the wages, prices, revenues and operating profits of all active firms.

### III. OPEN ECONOMY

Consider now an economy with a domestic (D) and a foreign (F) country, which can trade with each other. Throughout, we assume that the demand draw of a firm applies to both markets. That is, tastes are identical across countries. Entrepreneurs in each country decide how much to supply to the local market and how much to sell abroad. Entrepreneurs in the domestic country pay a fixed cost  $f_{DD}$  to produce. If they want to supply the foreign market they also need to pay a fixed cost  $f_{DF}$ . Similarly, firms in the foreign country need to pay a fixed cost  $f_{FF}$  to produce and  $f_{FD}$  to export. Trading goods is costly. Let  $\tau_{DF} = \tau_{FD} > 1$  be the “iceberg” trade cost incurred by firms. These costs entail a physical loss of resources. Finally,  $\tilde{N}_i$  for  $i = D, F$  denotes the fixed supply of labor in country  $i$ . As a result of the fixed cost of production and exporting, there exist thresholds  $\bar{\alpha}_{ij}$  such that if  $\alpha < \bar{\alpha}_{ij}$ , the entrepreneur  $\alpha$  from country  $i$  does not find it profitable to supply to market  $j$  and vice versa.

Let  $x_{ij}(\alpha)$  be the demand of an agent in country  $j$  for goods  $\alpha$  produced in country  $i$  for  $i, j \in \{F, D\}$ , and  $p_{ij}(\alpha)$  denotes its price. Since agents now have access to goods from both countries, the preferences in equation (1) become

$$U_i(x_{ii}(\cdot), x_{ij}(\cdot)) = \left( M_i \int_{\bar{\alpha}_{ii}}^{\infty} \alpha^{\frac{1}{\sigma}} x_{ii}(\alpha)^{\frac{\sigma-1}{\sigma}} \mu_{ii}(\alpha) d\alpha + M_j \int_{\bar{\alpha}_{ji}}^{\infty} \alpha^{\frac{1}{\sigma}} x_{ji}(\alpha)^{\frac{\sigma-1}{\sigma}} \mu_{ji}(\alpha) d\alpha \right)^{\frac{\sigma}{\sigma-1}}$$

for  $i, j \in \{F, D\}$ , where  $\tilde{\alpha}_{ij}$  is the demand draw of the marginal firm that supplies goods to market  $j$  from market  $i$ .  $M_i$  now denotes the number of operating firms in market  $i$  and  $\mu_{ji}(\alpha) = \frac{g(\alpha)}{(1-G(\tilde{\alpha}_{ji}))}$  is the conditional distribution of varieties available in market  $i$  and produced in country  $j$ .

The budget constraint of agents in  $i$  is given by

$$M_i \int_{\tilde{\alpha}_{ii}}^{\infty} p_{ii}(\alpha) x_{ii}(\alpha) \mu_{ii}(\alpha) d\alpha + M_j \int_{\tilde{\alpha}_{ji}}^{\infty} p_{ji}(\alpha) x_{ji}(\alpha) \mu_{ji}(\alpha) d\alpha = w_i.$$

Then, total demand for country  $j$ 's products in country  $i$  is given by

$$(20) \quad \tilde{N}_i x_{ji}(\alpha) = \alpha \left( \frac{p_{ji}(\alpha)}{P_i} \right)^{-\sigma} \frac{R_i}{P_i},$$

for all  $i$ , where the price index is given by

$$(21) \quad P_i = \left( M_i \int_{\tilde{\alpha}_{ii}}^{\infty} \alpha p_{ii}(\alpha)^{1-\sigma} \mu_{ii}(\alpha) d\alpha + M_j \int_{\tilde{\alpha}_{ji}}^{\infty} \alpha p_{ji}(\alpha)^{1-\sigma} \mu_{ji}(\alpha) d\alpha \right)^{\frac{1}{1-\sigma}}.$$

As before, we normalize the price index of the domestic country to 1, so  $P_D = 1$ .

Consider first the problem of a firm with demand draw  $\alpha$  in country  $i$  that does not export. It solves,

$$\pi_{ii}(\alpha) \equiv \max_{x_{ii} \geq 0} \left\{ p_{ii}(\alpha) \tilde{N}_i x_{ii} - C(q_i; w_i) - w_i f_{ii} \right\}$$

subject to equation (20). For firms that supply only to the domestic market,  $q_i(\alpha) = \tilde{N}_i x_{ii}(\alpha)$ , and so the first-order condition is given by

$$(22) \quad q_i(\alpha) = \alpha R_i P_i^{\sigma-1} \left( \frac{\sigma}{\sigma-1} MC(q_i(\alpha); w_i) \right)^{-\sigma},$$

which implies

$$p_{ii}(\alpha) = \frac{\sigma}{\sigma-1} MC(q_i(\alpha); w_i).$$

Denote by  $\pi_{ij}(\alpha)$  the profits of a firm with demand draw  $\alpha$  in country  $i$  that sells in the domestic and foreign markets. The profit maximization problem becomes

$$\pi_{ij}(\alpha) \equiv \max_{\{x_{ii}, x_{ij}\} \geq 0} \left\{ p_{ii}(\alpha) \tilde{N}_i x_{ii} + p_{ij}(\alpha) \tilde{N}_j x_{ij} - C(q_i; w_i) - w_i (f_{ii} + f_{ij}) \right\}$$

subject to equation (20) and where total production is given by

$$(23) \quad q_i(\alpha) = q_{ii}(\alpha) + q_{ij}(\alpha) = \tilde{N}_i x_{ii}(\alpha) + \tau_{ij} \tilde{N}_j x_{ij}(\alpha),$$

where  $q_{ij}(\alpha)$  denotes the quantity produced by  $i$  for market  $j$ . The cost function  $C(\cdot; w_i)$  solves the cost minimization problem described in Section II. The first-order conditions of this problem implicitly define the quantities sold in each market,

$$(24) \quad \tilde{N}_i x_{ii}(\alpha) = \alpha R_i P_i^{\sigma-1} \left( \frac{\sigma}{\sigma-1} MC(q_i(\alpha); w_i) \right)^{-\sigma},$$

and

$$(25) \quad \tilde{N}_j x_{ij}(\alpha) = \alpha R_j P_j^{\sigma-1} \left( \frac{\sigma}{\sigma-1} \tau_{ij} MC(q_i(\alpha); w_i) \right)^{-\sigma}.$$

In contrast with the standard model,  $x_{ii}(\alpha)$  and  $x_{ij}(\alpha)$  enter the marginal cost function through  $q_i(\alpha)$  as well. That is, a firm's level of total production affects its marginal cost and therefore how much it sells in each market. Importantly, the decision to export affects the cost of production of the goods sold in the local market.<sup>19</sup> Hence, the price in each market is given by a constant markup over marginal cost, namely,

$$p_{ij}(\alpha) = \frac{\sigma}{\sigma-1} \tau_{ij} MC(q_i(\alpha); w_i) = p_{ii}(\alpha) \tau_{ij}.$$

### III.A. Equilibrium in the Open Economy

As in the closed economy, the zero profit condition for the marginal firm defines the threshold of entering firms in country  $i$ ,  $\bar{\alpha}_{ii}$ . So,  $\pi_{ii}(\bar{\alpha}_{ii}) \geq 0$ . Similarly, indifference between exporting or selling domestically defines the threshold to export  $\bar{\alpha}_{ij}$ , so  $\pi_{ij}(\bar{\alpha}_{ij}) = \max\{0, \pi_{ii}(\bar{\alpha}_{ij})\}$ . In parallel to the closed economy, the free entry condition is then given by

$$\int_{\bar{\alpha}_{ii}}^{\bar{\alpha}_{ij}} \frac{\pi_{ii}(\alpha)}{\delta} g(\alpha) d\alpha + \int_{\bar{\alpha}_{ij}}^{\infty} \frac{\pi_{ij}(\alpha)}{\delta} g(\alpha) d\alpha = w_i f_i^E,$$

19. This implies that even when  $f_{ij} > f_{ii}$  all firms in the economy could enter the exporting market. Of course, if  $f_{ij}$  is large enough, then only the most productive firms will export. This is a key distinction with Melitz (2003) where, for the case of two symmetric countries, all firms will export if and only if  $f_{ij} \leq f_{ii}$ .

where  $f_i^E$  denotes the fixed cost of entry in market  $i$ . Labor market clearing is guaranteed by

$$\tilde{N}_i = \frac{M_i}{1 - G(\tilde{\alpha}_{ii})} \left( \delta f_i^E + \int_{\tilde{\alpha}_{ii}}^{\infty} (C(q_i(\alpha); 1) + f_{ii})g(\alpha)d\alpha + \int_{\tilde{\alpha}_{ij}}^{\infty} f_{ij}g(\alpha)d\alpha \right). \quad (26)$$

All these conditions hold for  $i \neq j$  and  $i, j \in \{D, F\}$ . Finally, goods market clearing is guaranteed by  $R_i = w_i N_i$  for  $i \in \{D, F\}$ . It is easy to show that the labor market clearing condition, the free entry condition, and the price index normalization together imply that trade is balanced. An equilibrium is then given by a vector  $(\bar{\alpha}_{DD}, \bar{\alpha}_{DF}, \bar{\alpha}_{FF}, \bar{\alpha}_{FD}, \bar{M}_D, \bar{M}_F, \bar{P}_D, \bar{P}_F, \bar{w}_D, \bar{w}_F, \bar{R}_D, \bar{R}_F)$ , which solves the eight conditions above together with the two goods market clearing conditions,  $\bar{P}_D = 1$ , and  $\bar{P}_F$  determined by equation (21).

The next proposition guarantees gains from trade as well as an increase in the quantity produced of all exporters and a decline in the quantity produced of all nonexporters. Given that we are normalizing  $\bar{P}_D = 1$ ,  $\bar{w}_D$  measures welfare in the domestic country. Now consider the effect of an increase in  $w_i$  on the quantity produced of nonexporters. We can use equation (22) after substituting  $R_i = w_i N_i$  to study the effect of  $w_i$  on  $q_i(\alpha)$ . Note first that as shown in Proposition 1,  $MC(\cdot; w_i) = w_i MC(\cdot; 1)$ . Hence, an increase in wages increases the right-hand side of equation (22) because  $\sigma > 1$ . If the marginal cost did not depend on  $q_i(\alpha)$ , this would immediately imply that the quantity produced by nonexporters would decrease. In our model there are two other effects. Consider first the case of a firm that does not change the number of layers as a result of the change in wages. In this case the direct effect of  $w_i$  is dampened by the fact that the decrease in  $q_i(\alpha)$  decreases marginal costs as proven in Proposition 1, which makes firms produce more. However, this new effect can never dominate the first, since it operates through the original effect on  $q_i(\alpha)$ . If the original effect is strong enough, or the firm is close to switching layers, the original effect may trigger firms to decrease their number of layers. In this second case, the original decline in  $q_i(\alpha)$  as a result of the rise in  $w_i$  increases discretely the marginal cost of the firm, which reinforces the direct effect and makes the firm sell even less. So the response of quantities produced as a response of changes in wages relative to the standard

model is dampened for firms that do not change the number of layers and amplified for firms that do. The same logic applies to the domestic sales of exporting firms, although exporters now have access to a new market. Using Proposition 7, we can prove that access to this new market always increases their sales.

**PROPOSITION 8.** In equilibrium a bilateral trade liberalization increases welfare in both countries. Furthermore, the quantity produced of all nonexporters decreases and the quantity produced of all exporters increases.

Proposition 8 is useful to understand the effect of trade on firm productivity and the internal organization, since in Section II we characterized the cost minimization of firms as a function of their quantity produced. The productivity of exporters can increase or decrease depending on a firm's production relative to the MES of its number of layers. Opening trade to a larger country, which increases sales further, tends to raise productivity as it increases layers of management and moves producers to average cost curves with lower minima. Hence, on average, productivity tends to increase for exporters. Nonexporters could also gain or lose in terms of productivity. The smallest firms always become less productive. Others might become more productive, because they get closer to the MES of their chosen  $L$ . The average cost curve has a general convex shape, which implies that the productivity gains that result from opening to trade are larger for small exporters than for large exporters (although there are firms that increase and others that decrease their productivity).<sup>20</sup>

**COROLLARY 9.** A bilateral trade liberalization increases weakly the number of management layers of all exporters and decreases weakly the number of layers of all nonexporters.

Corollary 9 is the direct result of Propositions 2 and 8. All exporters produce more after a trade liberalization and firms that produce more minimize their cost when they have weakly more management layers. The converse is true for nonexporters. With

20. The heterogeneous response of firm productivity in our model imply that average productivity can increase or decrease depending on the distribution of  $\alpha$ 's. We show in Appendix A that for the calibrated economy with a Pareto distribution of  $\alpha$ 's, the general shape of the average cost function leads to an average productivity gain for exporters and average productivity loss for nonexporters.

these results in hand, Propositions 1 and 2 can be used to understand the effect of bilateral trade liberalization on knowledge, income, number of workers, autonomy, and other layer-specific and firm characteristics. We turn now to a calibrated version of the model where we study a richer set of firm and economy-wide responses to a bilateral trade liberalization.

#### IV. CALIBRATION AND COMPARATIVE STATICS

For the quantitative exercises we require values of the following set of parameters:  $f_i^E$ ,  $f_{ii}$ ,  $f_{ij}$ ,  $h$ ,  $\frac{c}{\lambda}$ ,  $\gamma$ ,  $\sigma$ ,  $A$ ,  $\tilde{N}_i$ ,  $\delta$ ,  $\tau_{ij}$ . We specialize the model to a world with two symmetric countries and calibrate the parameters such that the model matches several moments in the U.S. data for 2002. We use the reference values in the literature for the parameters  $\delta$ ,  $\sigma$ , and  $t_{ij}$ . We set the elasticity of substitution to 3.8 following Bernard et al. (2003) and Ghironi and Melitz (2005). For the level of trade costs,  $\tau$ , we use 1.3 and for the exogenous firm exit rate we use 10%, both values obtained from Ghironi and Melitz (2005) and used for the United States for 2002. We set  $\tilde{N}_i$  equal to the total number of employees in the manufacturing sector and the educational sector as described below.

We normalize  $f_{ii} = 1.1$  and calibrate the values of  $f_i^E$ ,  $f_{ij}$ ,  $h$ ,  $\frac{c}{\lambda}$ ,  $A$  and  $\gamma$  to match five moments in the U.S. data and the Pareto coefficient of the size distribution of firms producing manufactured goods in the United States. The moments that we match are the share of expenditure on domestic goods, total expenditure, the fraction of firms that export, the average size of firms, and the share of education employees in the economy.

The fraction of firms that export is not directly reported in the U.S. Census Bureau. Bernard et al. (2007) estimate that 18% of the U.S. manufacturing firms exported in 2002. We use this figure since we restrict our attention to the manufacturing sector.

In the model some of the labor used in production is used to train workers. To take this into account we create a sector that is a bundle of the manufacturing sector and the share of the educational services sector used in the manufacturing sector. To do so, we first calculate the share of education employees using the figures reported by the Career Guide to Industries (CGI) from the Bureau of Labor Statistics (BLS) Current Population Survey for 2008. The CGI reports the number of employees per occupations



in different industries.<sup>21</sup> The total fraction amounts to 11.8% of total private employment in 2008. We calculate the share of total gross production from the education sector using the STAN - OECD database for Structural Analysis. We apply these shares to the manufacturing sector measures of employees and gross production. In this way we obtain that  $\tilde{N} = 16.48$  million.<sup>22</sup>

The average firm size is calculated as the total number of employees over the total number of firms from the manufacturing sector.<sup>23</sup> The total number of workers from the manufacturing sector reported by the 2002 Statistics of U.S. Businesses (SUSB) from the U.S. Census Bureau is 14,393,609 and the total number of manufacturing firms is 318,126. Therefore, the average size is 45.24. Finally, we use the TRAINS database to obtain data on imports from the manufacturing sector and the STAN database for gross production from the bundled sector. We find that the share of expenditure on domestic goods is 78.9% and total expenditure is equal to \$5.1 billion.

We search over the parameter space for parameter values that match the discussed moments, using as a loss function the norm of the percentage deviation difference between the model and the data. Table I presents the moment conditions that we match and the values they take in our preferred calibration.<sup>24</sup>

The calibrated parameter values are presented in Table II.<sup>25</sup> The cost of knowledge,  $\frac{c}{\lambda}$ , and the communication technology,  $h$ ,

21. We calculate the share of education employees by the number of workers in the Educational Services sector over total private employment in 2008.

22. A similar calibration but with multiple asymmetric countries requires two extra variables for each country relative to a standard framework, namely, education employment shares and average firm sizes for each country. These two moments can be used to calibrate the two extra parameters of our model, namely,  $c$  and  $h$ . Here, we apply the share of education employees from 2008 to the total number of employees in the manufacturing sector in 2002 since the figure for 2002 is not available on the CGI website.

23. We do not include the education workers as part of the manufacturing sector to calculate the average size of firms. We also do not include them in the moment calculated in the model.

24. In the model, as well as for the calibration, the maintained assumption is that knowledge is fully flexible, and so firms can decide freely the level of knowledge of workers at each layer. Of course, in reality there might be restrictions that limit the flexibility of firms to change the knowledge and wages of their employees in the short run. In this sense, the results in this section should be understood as long-run effects of changes in the economic environment.

25. The Pareto coefficient that matched the size distribution of firms has a value lower than 1; therefore, the mean and the variance of the distribution are

TABLE I  
CALIBRATED MOMENTS

Moments	Data	Model
Share of firms that export	18.0	17.53
Average size of firms	45.2	45.44
Total expenditure	5.1	5.10
Share of expenditure on domestic goods	78.9	74.94
Share of education employees	11.8	11.85
Size distribution of firms, pareto coefficient	-1.095	-1.094

TABLE II  
PARAMETER VALUES

Parameters	$A$	$f^E$	$f_{ij}$	$\gamma$	$\frac{c}{\lambda}$	$h$
Values	0.26	35.1	5.4	0.9	0.225	0.26

satisfy Assumption 1. Hence, in the calibrated exercises there are no firm with zero knowledge at any layer.<sup>26</sup> In their paper on the empirical effect of trade on the number of management layers Guadalupe and Wulf (2010, hereafter GW) consider a sample of 230 large manufacturing firms in the United States. for the period of 1986 to 1999. We use data on the entire population of manufacturing firms for the year 2002, so the period and samples are not directly comparable. Nevertheless it is interesting to compare the characteristics of firms in their sample relative to what our model yields. GW show that the average span of control was 7 in 1999 and growing over time. In the model we obtain 7.97 for 2002. Similarly, GW document that the average depth, measured as the number of positions between the division manager and the CEO, was 3 in 1999, which is similar to the 2.71 average layers in the model. So without targeting these moments, the model does

not well defined. Nevertheless, since the support of the distribution used in the computations is finite, all the moments we compute are well defined.

26. In the previous section we suggested an alternative interpretation for the per period cost of knowledge  $c$ ; it can be reinterpreted as  $\tilde{c}\delta_M$  where  $\tilde{c}$  is the cost of training workers in an economy where only new workers have to be trained every period and  $\delta_M$  is the the annual death rate. The U.S. death rate is 0.008% and so the cost of training new workers takes a value of 28, implying a stock of human capital equivalent to 4 times the population of the economy.

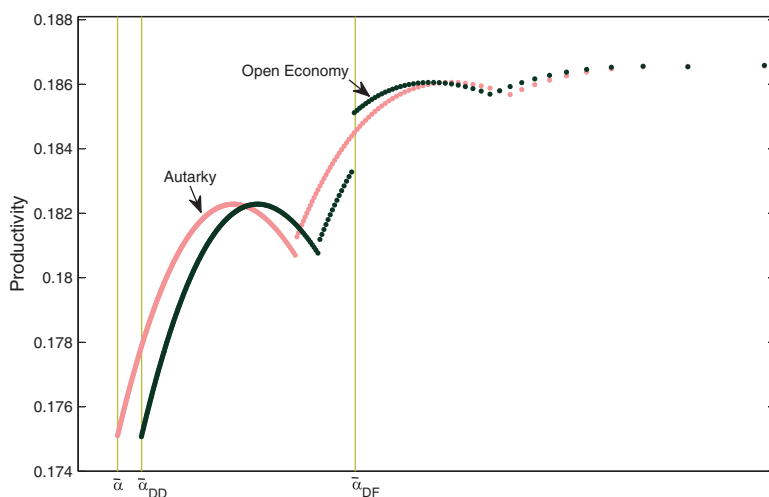


FIGURE V

Productivity for all  $\alpha$  in Autarky and Open Economy

well in generating hierarchies that match the data in these two important dimensions.

Throughout we compare our results with a version of our model where we let  $\frac{c}{\lambda} \rightarrow 0$  and so our model is isomorphic to Melitz (see Proposition 3).<sup>27</sup> Figure V presents productivity as a function of the idiosyncratic demand parameter  $\alpha$  for the case of the calibrated open economy and an economy in autarky with identical fundamental parameter values. All light curves refer to the implications of the model in autarky and dark curves refer to the calibrated open economy. Thus, the difference between the curves quantifies the effect of a bilateral trade liberalization. The vertical lines in the graph mark the thresholds  $\bar{\alpha} < \bar{\alpha}_{DD} = \bar{\alpha}_{FF} < \bar{\alpha}_{DF} = \bar{\alpha}_{DF}$ . Namely, they mark the entry threshold in autarky, the entry threshold in the two trading economies, and the export threshold in both economies. The figure shows productivity, measured as  $\frac{q(\alpha)}{C(q(\alpha); 1)}$ , and illustrates how the productivity of firms changes as we move to the open economy. We discuss alternative measures of productivity in Appendix A. In the parallel Melitz (2003) limit, productivity is constant and

27. We calibrate the parameters of the limiting model to match the first four moments of Table I. This yields  $A = 0.17$ ,  $f^E = 2.8$ ,  $f_{ij} = 3.2$ , and  $\gamma = 1.5$ .

equal to  $A$ , which in the calibration yields a productivity of 0.17. Hence the productivity of firms in the model with organization is larger for all firms.

Note that the model generates the fact that exporters are more productive and that more productive firms are larger. As in any model with fixed exporting costs, only the firms with high enough  $\alpha$ 's export, and some firms exit as we open the economy. Among the firms that enter in both scenarios, the ones with the lowest demand realizations experience a reduction in productivity as a result of the trade liberalization. They sell less, at higher prices, and obtain less profit. Note the difference with the standard model in which trade liberalization has no effect on productivity and productivity is identical for all firms (given our setup with exogenous heterogeneity in demand, not productivity). As  $\alpha$  increases from its lowest values, productivity increases in both scenarios, since firms are still producing below the MES of a firm with two layers (in this calibration all firms have at least two layers). Firms in autarky reach the MES of a two-layer organization for a lower  $\alpha$  than in the open economy (as implied by Corollary 9). Hence, as  $\alpha$  increases further, some nonexporting firms experience a gain in productivity.

The relationship is reversed once the firms change to an organization with three layers. At this point the discontinuous increase in  $q(\alpha)$  leads to a discontinuous change in the cost, which results in a discontinuity in productivity (recall Proposition 4). As we move beyond the export threshold  $\bar{\alpha}_{DF}$ , the productivity of exporting firms jumps up discretely in the open economy. The jump implies an increase in productivity for the marginal firm of 1%. Firms with  $\alpha$ 's just before the threshold experience a loss in productivity as a result of trade and firms with  $\alpha$ 's just after the threshold experience a gain. Since exporting firms switch layers for smaller  $\alpha$ 's than in autarky, there are some large exporting firms that lose productivity as a result of trade. Our parameterization of the model implies that, as firms get larger and have more layers, the gains or losses in productivity are progressively smaller. Productivity in both scenarios converges to a common constant.<sup>28</sup>

Figure V illustrates the heterogeneous responses in productivity as a result of a trade liberalization from autarky to the

28. These features are a consequence of agents learning the most routine problems first, namely, a distribution of problems with decreasing density.

calibrated economy. In general, the effects on productivity are below 1%. However, note that these changes in productivity would not exist if we were to move from autarky to an economy without any fixed trade costs (see Figure X later). In that case, opening to trade would amount to an increase in population that raises wages, and the number of firms, but it leaves the quantities produced, and therefore the organizational structure and productivity, unchanged (see Proposition 7). Hence, the productivity implications of our model are the result of the reorganization of production given the change in  $q(\alpha)$  generated by the presence of fixed and variable trade costs.

Figure VI presents four plots with the main firm outcomes. The first plot presents the average cost function in both scenarios. For the range of  $\alpha$ 's presented, firms have between two and four layers and the marginal firms that export have three layers. Note how the average cost function becomes flatter as we increase  $\alpha$ , which implies that the gains from organizational change beyond four layers are minimal. The main difference between the average cost functions in both scenarios is their level. This change is the result of an increase of 8.1% in wages in the open economy relative to autarky, which increases average costs. The panel also presents the flat horizontal average cost curves in the limit economy. Clearly, costs in the Melitz (2003) economy are larger for all firms.

The second panel presents firm's profits, which, as expected, are monotonically increasing in  $\alpha$ . Profits decrease for all nonexporters and some of the marginal exporters, and increase for the large exporters. These effects are standard in the literature and remain in our setup. The third figure presents production. Nonexporters produce more in autarky than in the open economy and the difference is larger for firms that change layers. Exporters produce more overall, again with the difference between trade and autarky larger for firms in the open economy that change layers.

The third panel of Figure VI also presents production for the export and domestic markets. Exporters produce less for the domestic market than in autarky. Note the increase in production as a result of trade. This is the main effect that leads to the productivity gains that we have underscored here. The fourth figure presents the domestic prices charged by all firms (foreign prices are just multiplied by  $\tau_{DF}$ ). Propositions 1 and 2 describe the behavior of the marginal cost as a function of  $q(\alpha)$  and, as we

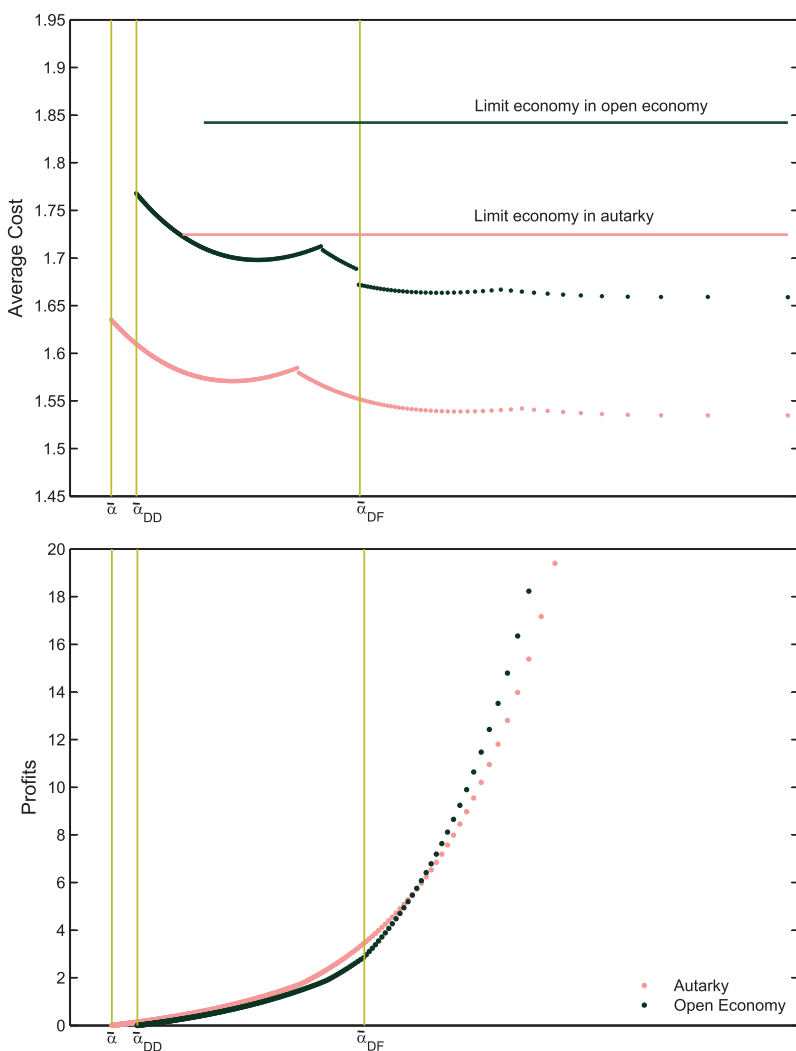


FIGURE VI

Average Cost, Production, Profits, and Prices

discussed in Section III, our CES demand system leads to prices that are proportional to marginal costs. Since  $q(\alpha)$  is increasing in  $\alpha$  as we saw in the third panel, we know that prices increase with  $\alpha$ . As firms change layers, prices fall discontinuously as described in Proposition 4. The price function also becomes flatter

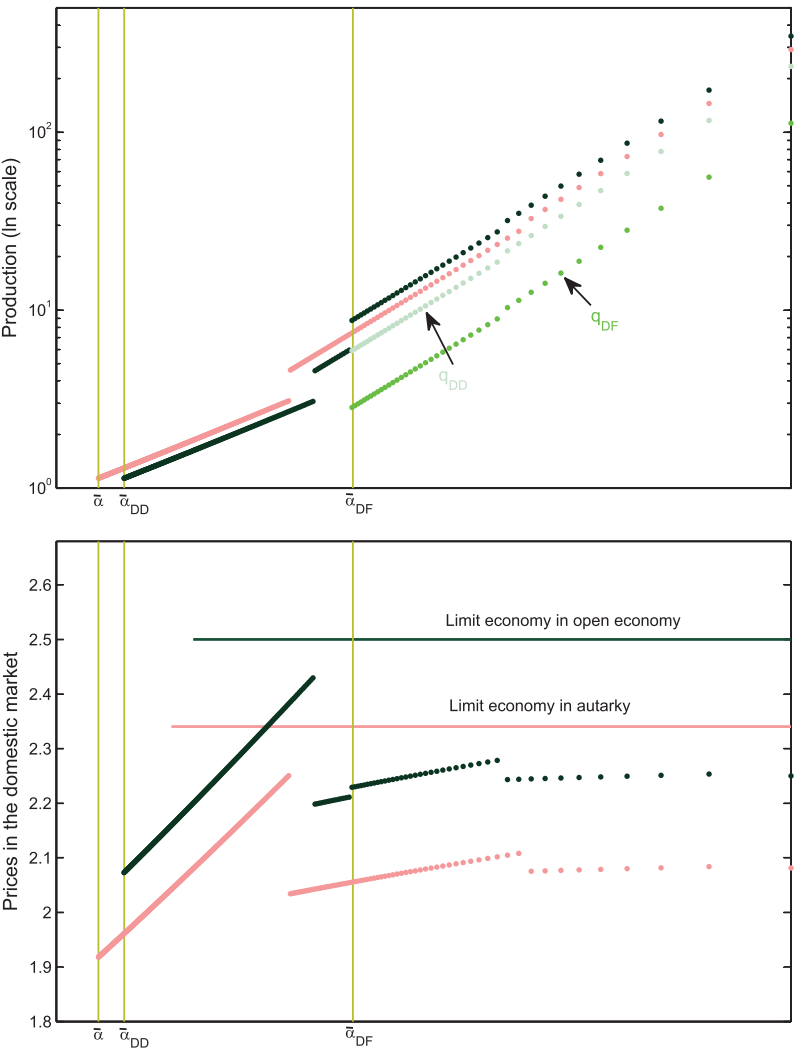


FIGURE VI  
Continued

as we increase the number of layers and eventually converges to a constant as we let  $\alpha \rightarrow \infty$ . Note also the jump in prices for the marginal exporting firm in the open economy. This is the result of the jump in quantity, which increases its marginal cost and therefore its price. We can also calculate revenue productivity by



multiplying our productivity measure by the relevant prices. In the standard model domestic prices are constant in  $\alpha$  and higher due to the higher marginal costs, as also illustrated in the figure.

Throughout this section we have compared the gains from trade in our theory to the case when we let the cost of knowledge ( $\frac{\epsilon}{\lambda}$ ) converge to zero. In that limit economy, which is isomorphic to Melitz (2003), knowledge is not a factor of production and so trade does not trigger changes in organization. As derived in Arkolakis, Costinot, and Rodríguez-Clare (2012), the gains from trade in the limit economy are given by  $(D_{ii})^{-\frac{1}{\gamma}(\sigma-1)}$  where the parameter  $\gamma$  is calibrated to 1.5.<sup>29</sup> With this calibrated value of  $\gamma$  the elasticity of imports with respect to trade costs in the limit economy is 4.2. In our model this elasticity is not constant, but is equal to 2.9 for the calibrated economy. The resulting gains from trade in the limit economy are 5.7%. Compared to the 8.1% welfare gains from trade that we obtain in our model, we conclude that changes in organization lead to gains from trade that are 2.4 percentage points larger or 41% larger than in the Melitz (2003) model. This significant increase can be interpreted as the effect of within-firm organizational change on the gains from trade.

Figure VII shows the impact of moving from autarky to the calibrated economy for the distribution of firm sizes, knowledge, income, and productivities. The first panel presents the distribution of firm's employment size and how it fits the data. We plot the size distribution using a standard log-log plot of the cumulative distribution in which a Pareto distribution would appear as a straight line with slope of minus its shape coefficient. As is evident from the figure, the model does a nice job in fitting the distribution of firm sizes. This is, of course, partly the result of our assumption that the idiosyncratic demand shocks are distributed Pareto and the fact that we are calibrating the shape parameter of this distribution to fit the data. However, the fit is by no means guaranteed in a model as rich as this. It is clear from the figure that in autarky firm sizes are smaller and the tail of the distribution is thinner, indicating fewer large firms.

The third panel of Figure VII presents a histogram of the distribution of knowledge in the economy in our two scenarios.

29. Given  $\sigma$ ,  $\tau_{ij}$ , the share of firms that export,  $\frac{M_{ij}}{M_i}$ , and the share on domestic expenditure  $D_{ii}$ , we calibrate  $\gamma$  to match  $D_{ii}$  using the following equilibrium

$$\text{condition } D_{ii} = \frac{(\tau_{ij})^{\sigma-1} \left( \frac{M_{ij}}{M_i} \right)^{\frac{1}{1/\gamma}}}{\left( (\tau_{ij})^{\sigma-1} \left( \frac{M_{ij}}{M_i} \right)^{\frac{1}{1/\gamma}} + \frac{M_{ij}}{M_i} \right)}.$$

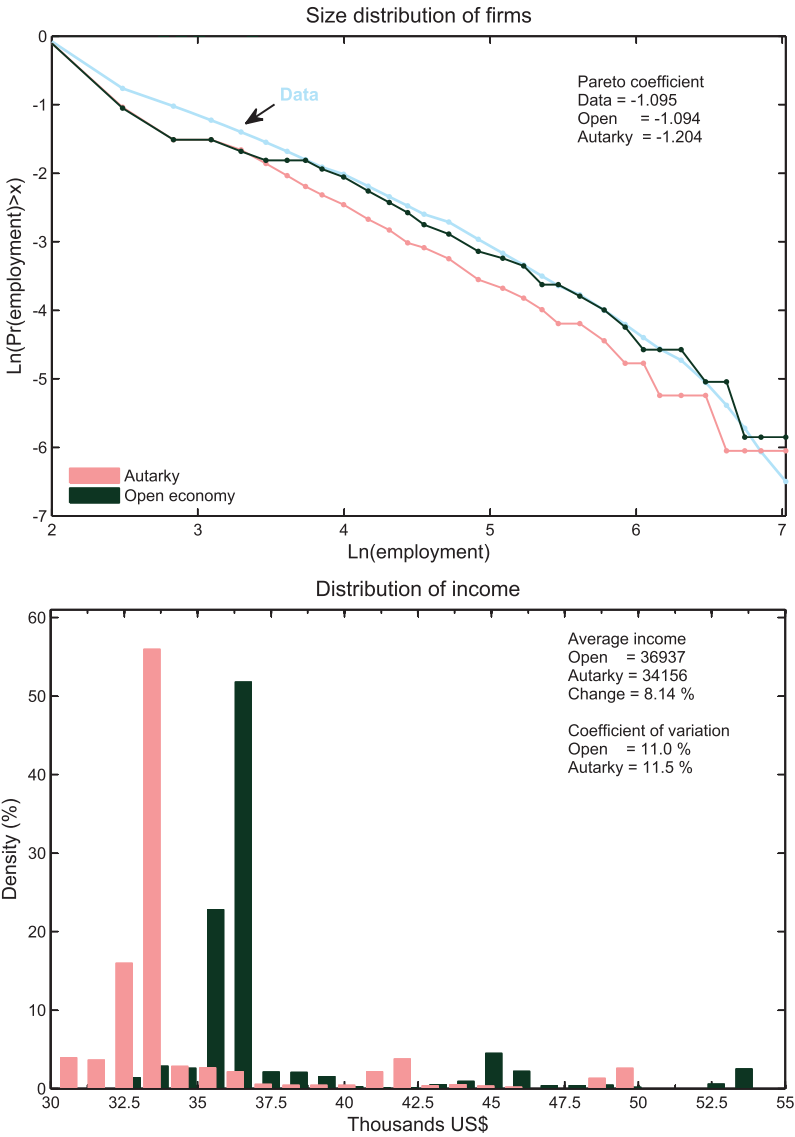


FIGURE VII

Firm Size, Knowledge, Income, and Productivity Distributions

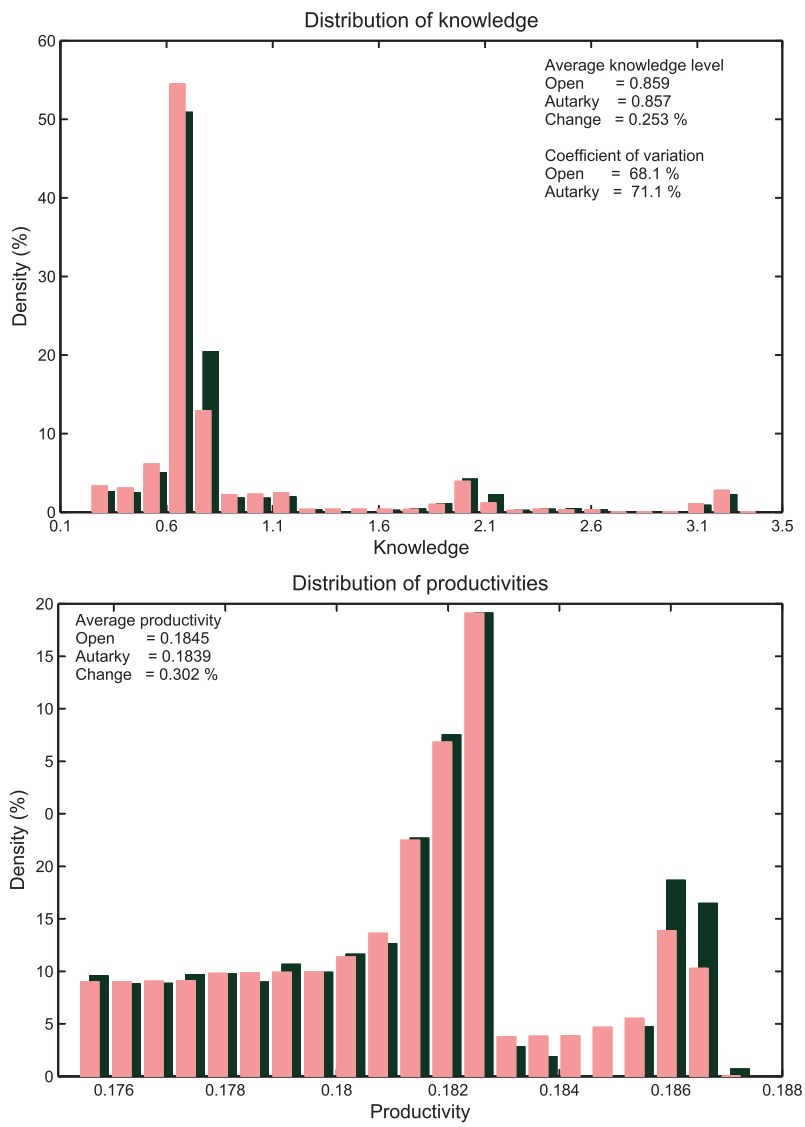


FIGURE VII  
Continued

The figure indicates that there are substantially more workers with low levels of knowledge in the open economy, although there are more agents with intermediate levels of knowledge. The effect is the result of the increased number of layers of exporting firms in the open economy. These firms introduce more layers and, as a result, require their workers to know less. However, since they hire more intermediate managers, there is more mass of agents with intermediate levels of knowledge than in autarky. The agents with the highest levels of knowledge are all entrepreneurs and are present in both economies. The second panel of Figure VII presents the distribution of income. The distribution of income follows closely the distribution of knowledge in the economy. That is, income is given by  $w_i(cz_L^\ell + 1)$ . We know that  $w_i$  increases as we open the economy, which shifts the distribution of income in the open economy to the right relative to the closed economy.

Finally, the fourth panel presents the distribution of productivities in the economy. The figure shows that the fraction of small productivity firms in both economies remains largely unchanged. Recall from Figure V that the productivity of the smallest firms decreases, but the entry threshold also shifts to the right. Hence, since the shape of the Pareto distribution of  $\alpha$ 's is invariant to truncation, we obtain similar shares of small productivity firms in both economies. The main difference between this conclusion and the one we reached discussing Figure V is that the latter conditions on  $\alpha$  whereas the former does not. The figure also illustrates the increase in the mass of the most productive firms as we open to trade. Essentially, firms in the middle of the distribution increase their productivity, thereby raising the mass at the tail. This is the main effect of a trade liberalization: it makes firms with intermediate levels of demand more productive than in autarky because it incentivizes a reorganization of workers and their knowledge within the firm.

#### IV.A. *Effects of Trade on Internal Organization*

So far our discussion has been limited to firms' and economy-wide changes in outcomes as we open to trade. We now turn to describe the internal reorganization of firms as determined by Propositions 1 and 2. Propositions 1 and 8 tell us that new exporters that do not change layers increase the knowledge, wages or income, and number of employees at all layers. In contrast, non-exporters with the same layers decrease all of them. Figure VIII

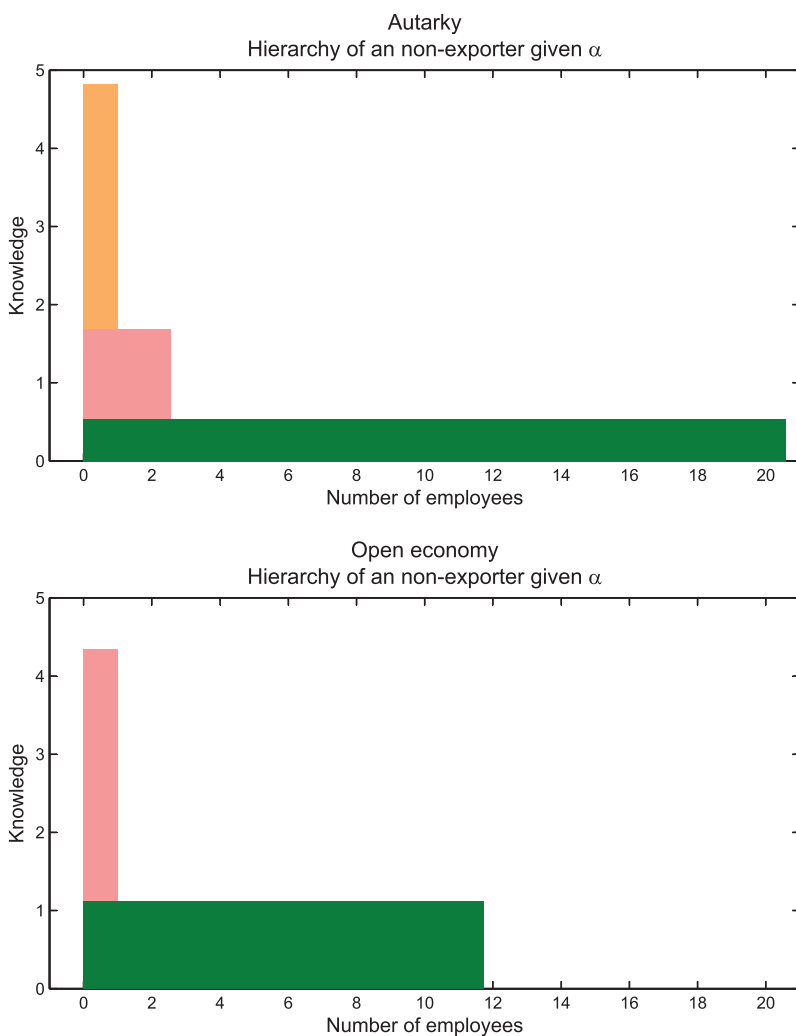


FIGURE VIII

Changes in the Organizational Design after a Bilateral Trade Liberalization

presents examples of the reorganization of a nonexporter and an exporter that do change their number of layers as the economy opens to trade. Each layer is represented by a rectangle where the height represents the knowledge of workers in that layer, and the length represents the number of agents working in that layer.

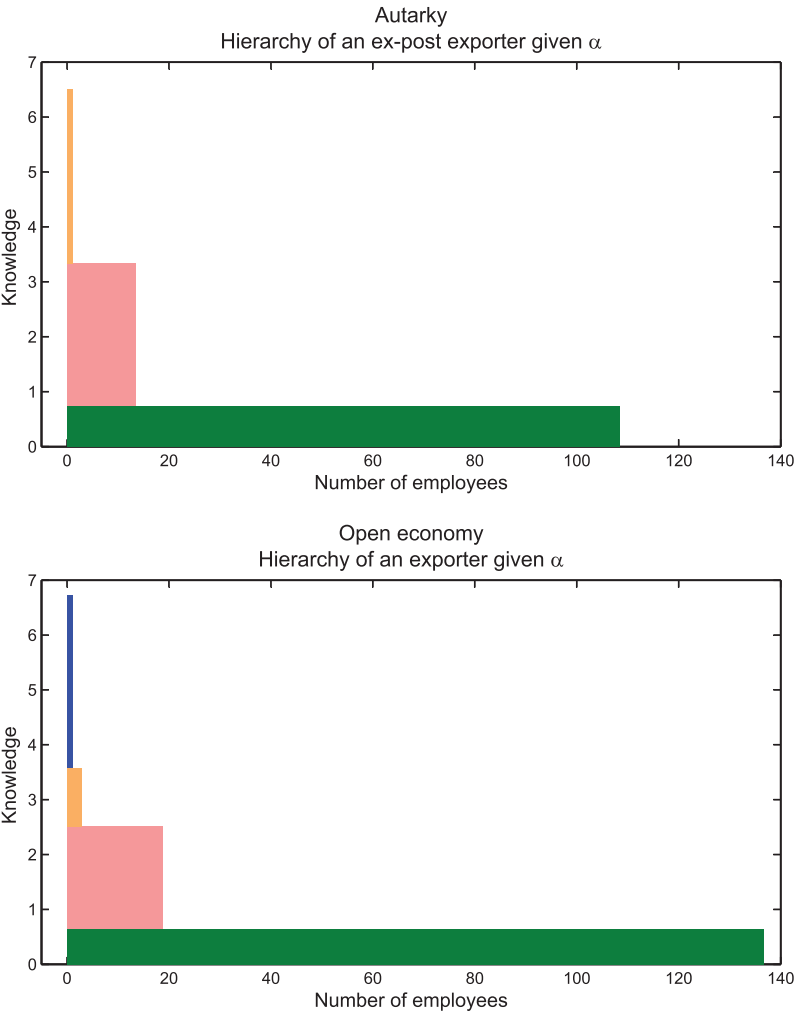


FIGURE VIII  
Continued

We plot rectangles in higher layers above rectangles in lower ones. The effects are now determined by Propositions 2 and 8.

Compare first the two panels in the figure, which represent a firm with given demand level,  $\alpha$ , that does not export as a result of the trade liberalization. The change in the characteristics of the firm is dramatic. The firm reduces the number of layers from

three to two. This implies a reduction in the number of workers (agents in layer zero) from 21 to 12. Note that the main reason to have an intermediate layer is to increase the span of control of the entrepreneur. Since in autarky the firm also hired 3 agents in the second layer, the total employment size of the firm went from 25 to 13. The levels of knowledge also changed. In the open economy, employees at each comparable layer of the firm know more. Worker's knowledge increases by 109%, and the knowledge of managers in the first-layer increases by 182%. Since the firm has one less layer, the firm as a whole solves fewer of the problems faced in production as the cumulative knowledge of the firm decreases by 10%.

These implications reverse when we look at the exporter represented in the third and fourth panels. Now, the number of layers increases as we open to trade, and so does the number of employees in the firm. However, the amount of knowledge of agents in each comparable layer decreases. Workers's knowledge decreases by 12%, that of first-layer managers decreases by 28%, and that of second-layer managers decreases by 67%. Still, since the firm has one more layer, its cumulative knowledge increases by 4%. The firm decides to decrease the knowledge of agents at all comparable layers to economize on the knowledge of the many workers and lower-level managers in the firm. The organizational changes described in Figure VIII are present for all firms that adjust the number of layers. Of course, these changes in the levels of knowledge, are mapped directly into changes in the distribution of income within the firm.

Figure VIII presents an example of changes in organization but does not allow us to describe these changes across the universe of firms. To do so, we use the concept of autonomy that we introduced in Propositions 1 and 2. In a model of management by exception, like ours, autonomy of a class of workers can be measured by the set of the decisions they make relative to the total set of decisions made by the firm. That is, we can measure the set of problems solved by agents in a given layer, relative to the total amount of knowledge in the firm. So  $\frac{z_L^t}{Z_L^t}$  measures the share of knowledge of the entrepreneur, or its autonomy. Of course, the larger the autonomy of the entrepreneur, the more centralized the decision making in the firm.

Figure IX presents the measure of autonomy of entrepreneurs. Propositions 1 and 2 imply that  $\frac{z_L^t}{Z_L^t}$  decreases with  $q$  and

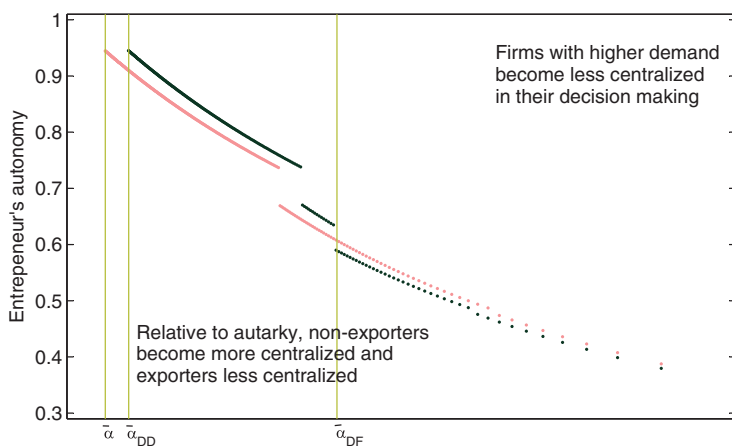


FIGURE IX

Changes in Autonomy before and after a Bilateral Trade Liberalization

drops discontinuously at the  $q$ 's where firms change layers. Proposition 4 then implies that the autonomy of entrepreneurs should fall with respect to  $\alpha$ . This is clearly the case in Figure IX. Furthermore, a bilateral trade liberalization increases the level of centralization for all nonexporters and decreases it for all exporters. So the model delivers a clear testable implication for this key dimension of the organization of the firm. Bloom et al. (2010) present evidence that centralization tends to fall with firm size, as predicted by our theory. As far as we know, the empirical literature has not studied the effect of a trade liberalization on centralization.<sup>30</sup>

#### IV.B. Export Costs, the Costs of Knowledge, and Information Technology

In the previous sections we studied the effect of reductions in variable trade costs. We can also use numerical simulations to understand the effect of changes in the fixed cost of exporting on productivity and welfare. From Proposition 7 we know that if the fixed costs of exporting are small enough so that all firms that produce also export, in a world with two symmetric countries,

30. Marin and Verdier (2008) present evidence that the degree of concentration as a function of sales per employee exhibits an inverted U-shape pattern in data from German and Austrian firms. These pattern is also generated by our theory.



firms produce identical quantities with and without trade. Thus, only the number of firms and wages adjust. So for low enough transport costs, each firm has a productivity identical to its productivity in autarky. When  $f_{ij} = 0$ , it is also the case that exporting does not increase the fixed cost of production, and so the production threshold  $\bar{\alpha}$  does not change. Hence, when  $f_{ij} = 0$  the productivity gain from trade liberalization is equal to zero. Similarly, as  $f_{ij} \rightarrow \infty$ , no firm exports and so the economy is in autarky. Hence, productivity gains are zero as well. For the intermediate fixed cost of exporting, the productivity gains are positive.

Figure X presents average productivity gains and welfare gains relative to autarky for a variety of exporting costs and for two levels of iceberg costs,  $\tau$ . As can be seen in the figure, there is a range of fixed costs for which all firms export. Note that, in contrast to Melitz (2003) with two symmetric countries, in our theory even if  $f_{ij} > f_{ii}$  all firms in the economy might export. This is never the case in Melitz (2003), since there the profits from foreign sales (which are always smaller than the profits from domestic sales) cannot compensate for the fixed cost of the marginal producers. In our theory the profits from domestic and foreign sales are not separable, since it is the total quantity produced that determines the organizational structure (as can be seen in equations (24) and (25)). So exporting increases the profits from domestic sales by changing the organizational structure of the firm. This can be appreciated in Figure X where for the case  $\tau = 1$ , the largest  $f_{ij}$  for which all firms export is about 80% larger than  $f_{ii}$ . Of course, as we increase  $\tau$ , the sales in the export market decrease and the ordering of the threshold is reversed, although this effect is still present.

As we argued, as long as all firms export, firms' quantities do not change, and therefore, the productivity of all firms remains constant. However, average productivity increases as a result of selection. The firms with the lowest demand thresholds exit, which increases average productivity. Average productivity increases due to selection until some firms decide to enter only to the domestic market. At that point further increases in fixed cost decrease average productivity, since the productivity of non-exporters tends to decrease relative to autarky. Increasing  $f_{ij}$  further leads to fluctuations in average productivity as the production threshold moves to the left and the export threshold to the right, changing their position relative to the levels of  $\alpha$  at which firms change layers. In any case, as  $f_{ij}$  increases further

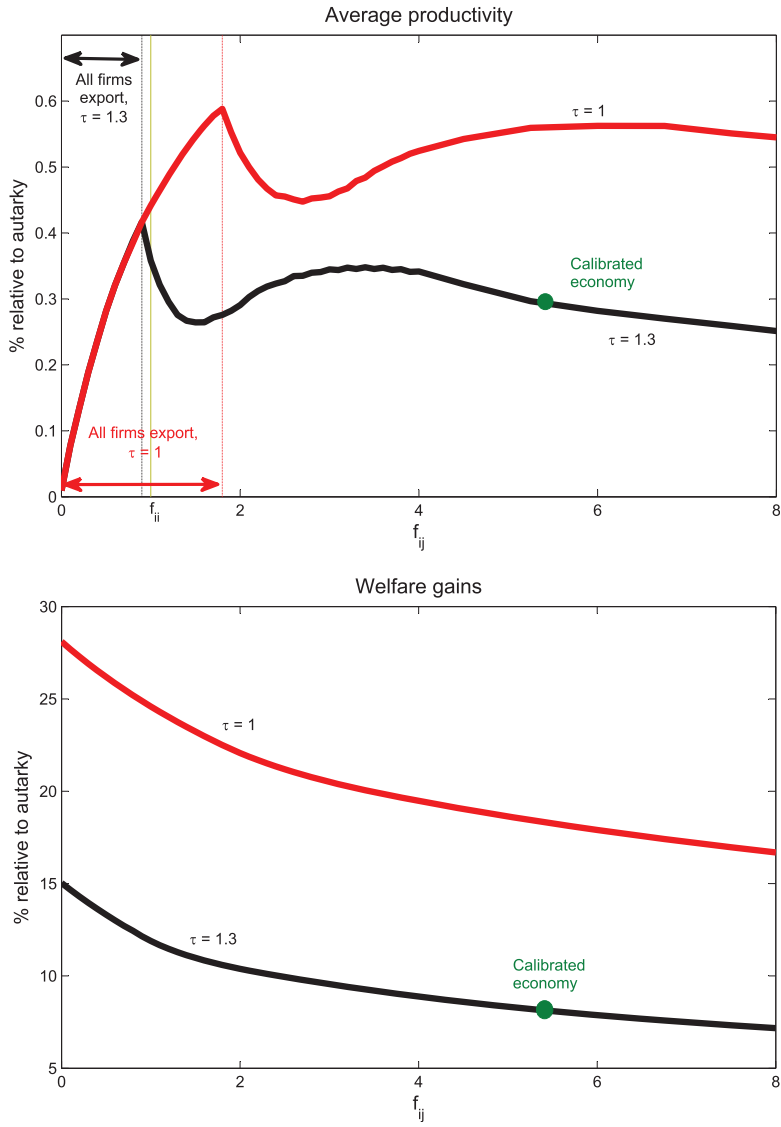


FIGURE X

Welfare and Productivity Gains from Trade as a Function of  $f_{ij}$

the productivity gains tend to zero. In contrast to the effect on average productivity, the welfare gains from trade decrease monotonically with  $f_{ij}$ , since the technology to export becomes progressively worse. Furthermore, a reduction in  $\tau$  increases welfare for any finite level of fixed exporting costs.

We now turn to the impact of changes in the cost of knowledge,  $c$ . As we increase  $c$  leveraging knowledge using a complex organization becomes more important. So all firms tend to have taller organizations with more layers, where decision making is more concentrated. Hence, the increase in the quantity produced of exporters associated with international trade leads to more important changes in organization when  $c$  is relatively large. In contrast, the organizational structure of nonexporters, which is already relatively simple, does not change much. The result is a monotonically increasing relationship between average productivity and  $c$ . Of course, as  $c$  increases, the technology to produce knowledge deteriorates and welfare (given by wages since  $P_D = 1$ ) decreases both in autarky and the open economy. However, the decrease in welfare in autarky dominates, leading to larger gains from trade the higher are learning costs. Organization is more important if  $c$  is large, because its sole purpose is to economize on knowledge acquisition. The higher  $c$ , the more important is organization, and the increase in scale triggered by trade for exporters. The results of this exercise are presented in Figure XI.

We finish this section with a discussion of the effect of communication technology on productivity, welfare, and the gains from trade, presented in Figure XII. A higher parameter  $h$  decreases the span of control of managers. Hence, creating new layers in an organization becomes less attractive. The knowledge of managers can be leveraged less, since the spans of control are smaller given that communication costs are larger. The economy uses less organization and the productivity gains associated with trade become less important. The increase in the quantities produced of exporters changes the organization of exporters by less, leading to smaller average productivity gains in the economy. Because a higher  $h$  implies a worse communication technology, it decreases welfare both under autarky and in the open economy. However, the decline for the open economy is larger and so the gains from trade decrease with  $h$ . An economy with worse communication technology makes the improvements in the organization of exporters less valuable.

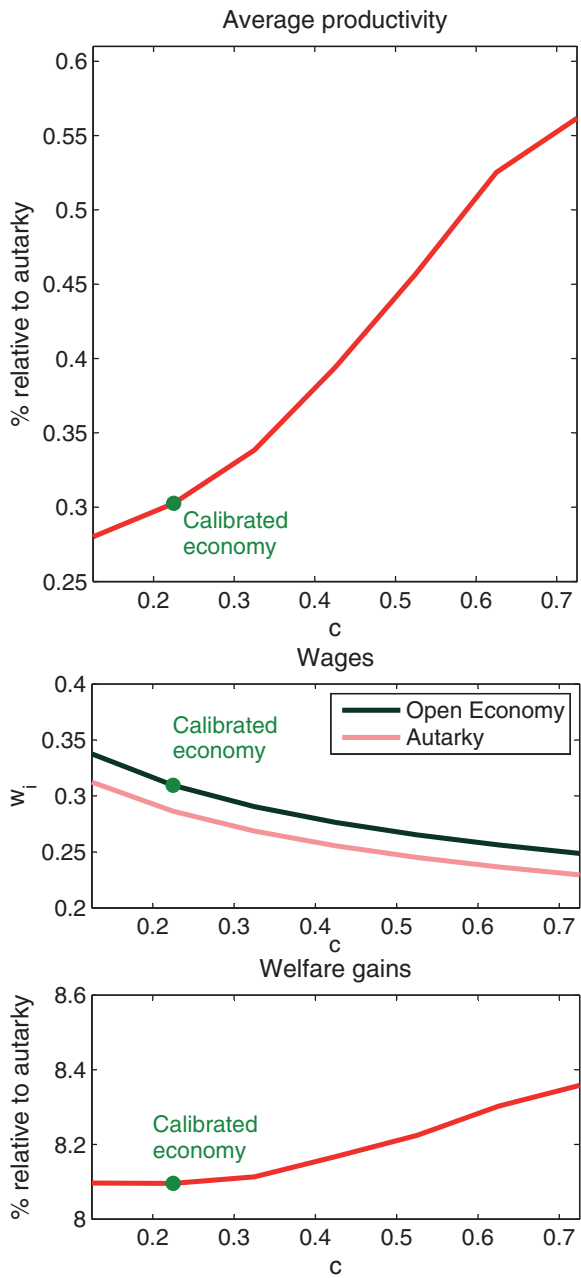


FIGURE XI  
Comparative Statics in  $c$

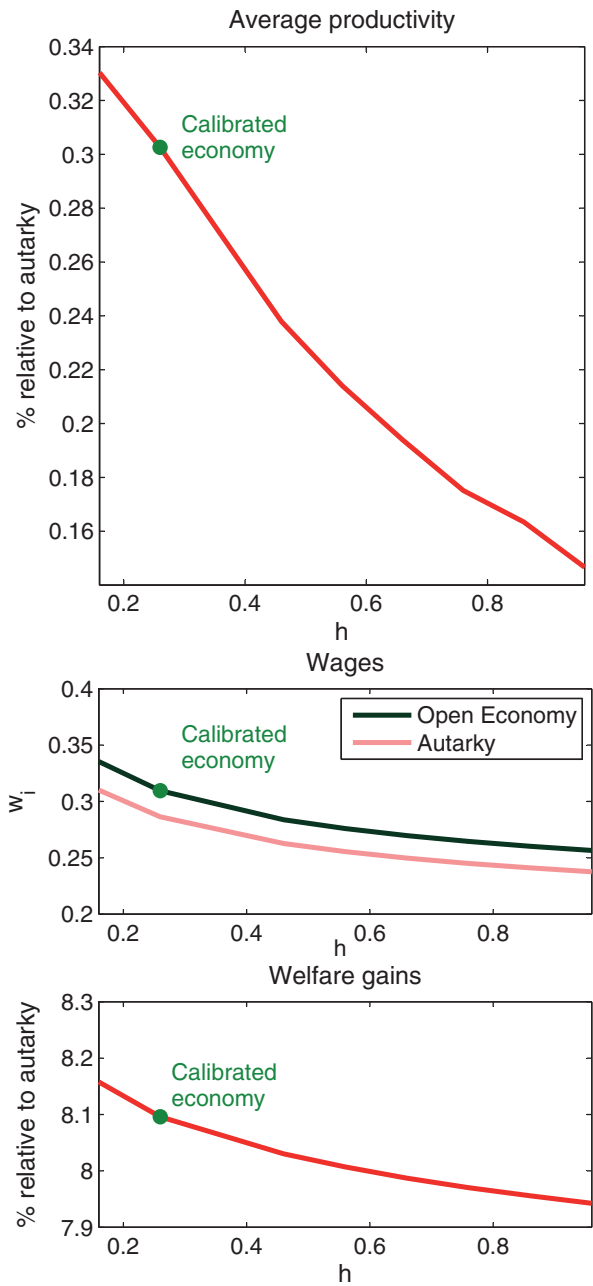


FIGURE XII  
Comparative Statics in  $h$

## V. CONCLUSIONS

The goal of this article is to understand how the ability of an economy to trade internationally impacts the organization of production. For this purpose we propose a general equilibrium theory in which firms produce differentiated products using labor and knowledge. The organization decisions of firms determine the classes of employees used in production, their knowledge, as well as the number of each of them.

Our theory was built out of two main components. First, we have adapted the theory of organization introduced in Garicano (2000) to derive and characterize the variable cost function for any quantity produced. The construction and characterization of this function is one of the contributions of this article. Its properties are reminiscent of other costs curves proposed previously in the literature, some of them as early as Viner (1931). However, as far as we know, this is the first article to microfound them using a fully specified model of organization. The benefits of such a microfoundation, are twofold. First, each point in this curve can be associated to a full description of the organization of the firm. Second, we can relate the shape and characteristics of this function to the parameters that determine trade, as well as the cost of acquiring and communicating knowledge.

We can then embed the cost function into a general equilibrium international trade model with firm heterogeneity to link changes in the ability to trade with changes in within firm characteristics. This can be done in multiple ways, with multiple demand systems, and using different sources of exogenous heterogeneity. We purposely have chosen the most standard of models to do so. Hence, we use a CES demand system in which firms produce goods that differ in their level of demand. Of course, some of our quantitative predictions are specific to this particular structure. However, the main qualitative implications of the model are robust to introducing other exogenous forms of heterogeneity (e.g., management ability) and other demand systems. The main qualitative implication of trade for within-firm characteristics are the result of movements along the cost function generated by changes in the quantities produced. So as long as the general equilibrium structure generates increases in the quantity produced for exporters and decreases for nonexporters, the within-firm effects will be qualitatively similar. Adding endogenous organization seems to matter quantitatively, too. Our

numerical simulations imply that the gains from trade are about 40% larger when we account for organizational changes.

The robustness of these implications is important because they can be contrasted with data. Bloom and van Reenen (2007) and Bloom et al. (2011) have presented evidence that management practices matter for a variety of firm' outcomes. Our theory can rationalize these predictions as well as more nuance ones like, for example, the relationship between size and autonomy (see Bloom et al. 2010).

The theory we propose also has strong and detailed implications for the effect of changes in demand, or access to markets via trade, on organizational change. To verify empirically these implications, one needs data on firm-level outcomes, as in Bloom and Van Reenen (2007), as well as on the occupations and wages of all employees within firms. Occupation data can be used to classify employees into layers. Wage data can be used as a measure of the component of knowledge that is relevant in the labor market. This match between the data and the concepts in the model can then be used to test the implications of the model among several dimensions. First, the theory implies that a trade liberalization increases the layers of management of exporters and decreases the layers of management of nonexporters. For exporters that increase the number of layers, we should observe decreases in the knowledge and wages of employees in each layer, but increases in the number of employees in each of them. In contrast, exporters that do not change the number of layers should exhibit higher knowledge, wages, and numbers of employees in each layer. The inverse implications apply for nonexporters. The theory also links these implications to effects on prices, productivity, revenue productivity, as well as aggregate implications on the distribution of knowledge, income, productivity, and firm sizes. The increased availability of large national employer-employee data sets that include information on occupations and wages gives us hope that our theory will serve as a guide for this type of empirical analysis in the future.

## APPENDIX I: OTHER MEASURES OF PRODUCTIVITY

In the main text we discussed the implications of our model for the effect of a bilateral trade liberalization on productivity. We used a measure of productivity given by  $\frac{q(\alpha)}{C(q(\alpha);1)}$ . Even though

this measure is theoretically appealing, in many empirical studies data on quantities and average costs at constant prices are not readily available. Hence, we now discuss the predictions of our model for other commonly used measures of productivity. In particular, we study how trade affects revenue productivity (measured as  $\frac{r(\alpha)}{C(q(\alpha);1)} = p(\alpha) \frac{q(\alpha)}{C(q(\alpha);1)}$ , labor productivity (measured as  $q(\alpha)/n(\alpha)$  where  $n(\alpha)$  denotes the total number of employees in the firm without including education workers or workers used to cover fixed costs), and revenue labor productivity (measured as  $\frac{r(\alpha)}{n(\alpha)}$ ). These measures use progressively more easily available data. Figure A1 presents all of these measures as a function of the demand draw,  $\alpha$ .

Note first a set of common characteristics of all plots in Figure A1. First, on average exporters are more productive than nonexporters, independently of the measure used or whether we use weighted averages by employment size or revenue, or an unweighted average. As is evident in Figure A1, not all exporters are more productive than nonexporters, but on average this is always the case by between 2% and 5%. Second, all measures of productivity increase discontinuously for firms that export. So the marginal exporter always gains in productivity, independently of the measure we use. Of course, the jump in revenue productivity or revenue labor productivity for the marginal firm is larger compared to the case when we use only quantities, because, as we have shown, the prices of the marginal firm increase discontinuously. The effect of prices also implies that using revenue productivity leads to overall much larger changes in productivity as a result of international trade. In fact, all firms, including the nonexporters, increase their revenue productivity and their revenue labor productivity. Remember that the increase in firms' prices is the result of the increase in wages that results from the extra varieties available to consumers in the open economy. Hence, using revenue and not quantity-based productivity measures overestimates the gains in productivity that result from international trade. This is consistent with the findings in De Loecker (2011).

Perhaps the more commonly used measure of productivity in the empirical literature that uses quantity-based measures of productivity is what we call labor productivity. The reason is that it is hard to properly adjust empirically for the knowledge of workers within the firm. This measure yields a very clear empirical implication for the changes in productivity. All



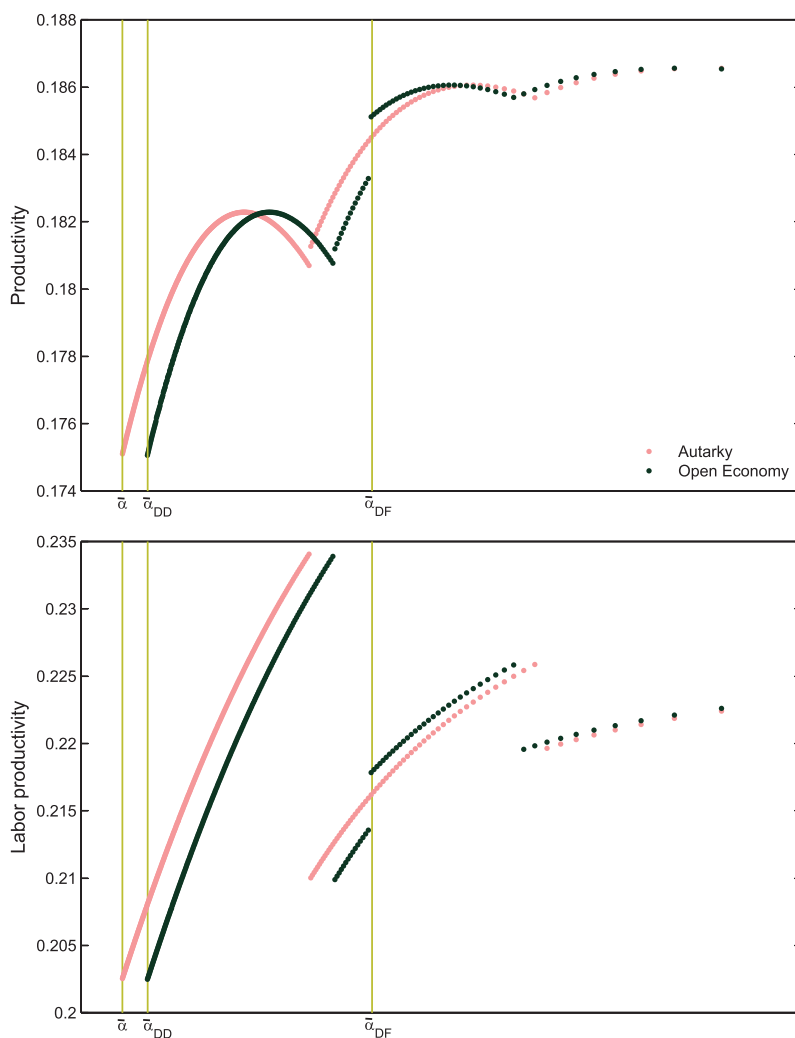


FIGURE A1

Different Measures of Productivity

nonexporters with the same number of layers should decrease their labor productivity and all exporters with the same number of layers should increase their labor productivity as a result of a trade liberalization. Furthermore, firms with smaller demand draws experience larger changes (positive or negative) in labor productivity than firms with high  $\alpha$ 's. The results are

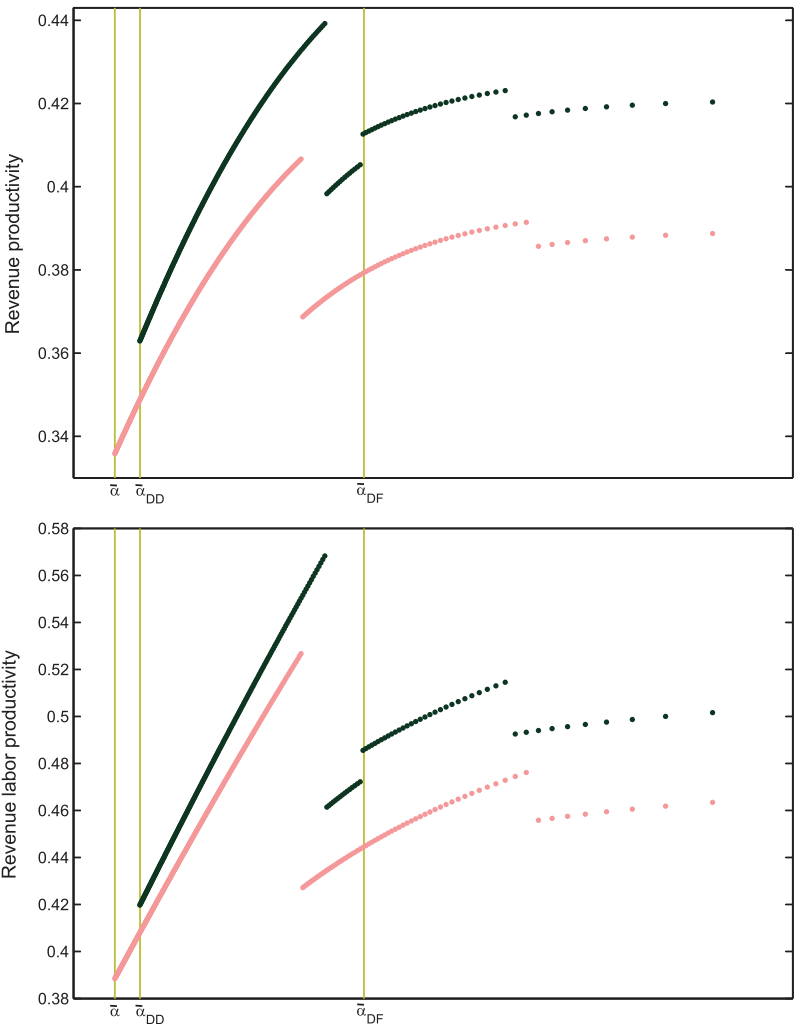


FIGURE A1

Continued

reversed for firms that add or drop layers. Nonexporters that decrease their layers gain labor productivity and all exporters that increase their number of layers lose productivity. Table A1 presents the average changes in all of these measures of productivity for all firms, exporters, nonexporters, and the marginal exporter using either no weights (the column labeled 1) or the

TABLE A1  
PRODUCTIVITY GAINS RELATIVE TO AUTARKY

Weight	Productivity				Revenue productivity			
	1	$n(\alpha)$	$r(\alpha)$	$q(\alpha)$	1	$n(\alpha)$	$r(\alpha)$	$q(\alpha)$
All firms	0.03%	0.30%	0.27%	0.22%	8.16%	8.63%	8.52%	8.47%
Exporters	0.10%	0.04%	0.05%	0.05%	8.33%	8.22%	8.21%	8.22%
Nonexporters	-0.08%	-0.18%	-0.20%	-0.21%	7.95%	7.87%	7.92%	7.89%
Marginal firm		1.00%				1.82%		

Weight	Labor productivity				Revenue labor productivity			
	1	$n(\alpha)$	$r(\alpha)$	$q(\alpha)$	1	$n(\alpha)$	$r(\alpha)$	$q(\alpha)$
All firms	0.08%	0.35%	0.30%	0.28%	8.21%	8.65%	8.54%	8.53%
Exporters	0.33%	0.13%	0.12%	0.13%	8.63%	8.30%	8.29%	8.29%
Nonexporters	-0.03%	0.02%	0.09%	0.08%	8.00%	8.10%	8.25%	8.21%
Marginal firm		2.00%				2.83%		

number of employees, revenue, or quantities as weights to calculate the average. All results are consistent across weights except for the change in labor productivity for nonexporters. In this case, the unweighted average decreases while the weighted ones increase. This is natural given that a large number of small firms experience a decrease in productivity and a few larger ones experience a substantial gain, as can be seen in the second panel of Figure A1. The table shows that the increase in labor productivity for the marginal exporter is about 2%.

## APPENDIX II: PROOFS OF PROPOSITIONS

Throughout we impose:

ASSUMPTION 1. The parameters  $\lambda$ ,  $c$ , and  $h$  are such that  $\frac{c}{\lambda} \leq \frac{h}{1-h}$ .

Assumption 1 is sufficient to guarantee that, for any organization with  $L \neq 1$  management layers, the entrepreneur never sets knowledge to zero at any layer.<sup>31</sup> Before we state and

31. The case of  $L = 1$  is the only one where a simple parameter restriction cannot guarantee that  $z_1^0 > 0$ . The reason is that an organization with zero management layers is constrained to produce  $q < A$ . Hence, even if workers do not learn anything, an organization with one layer of managers can dominate the

prove Proposition 1 we prove the following result:

**LEMMA 1.** Given the number of layers  $L$ , under Assumption 1, as the multiplier on the output constraint,  $\phi$ , decreases, the constraints over  $z_L^l \geq 0$  for all  $l$  bind such that the first to hit zero is  $z_L^{l-1}$ , then  $z_L^{l-2} \dots$  until  $z_L^0$  and finally,  $z_L^l$ .

*Proof.* Without loss of generality, consider that all knowledge in the organization is positive; therefore,  $\theta_L^l = 0$  for all  $l$ . From the first-order conditions of the cost minimization problem we obtain the following system of  $L + 1$  equations relating all the  $z$ 's.

$$(27) \quad e^{-\lambda z_L^L} = \frac{hwc}{\phi \lambda A},$$

$$(28) \quad e^{-\lambda z_L^l} = \frac{ch}{c + \lambda(cz_L^0 + 1)},$$

$$(29) \quad e^{-\lambda z_L^0} = \frac{c/h}{c + \lambda(cz_L^1 + 1)},$$

$$(30) \quad e^{-\lambda z_L^{l-1}} = \frac{c}{c + \lambda(cz_L^l + 1)} \text{ for all } 1 < l < L.$$

The constraint on quantity is always binding, therefore  $\phi > 0$ . We will now solve sequentially, starting from the lowest layer for the value of the Lagrange multiplier  $\phi$  such that it makes each of the  $z$ 's zero. Let  $\phi$  be large enough. First note from (27) that for any value of  $\phi > \phi_L^*$ ,  $z_L^L > 0$ , where  $\phi_L^* = \frac{hwc}{\lambda A}$ . Consider now layer 0 and assume that all other  $z$ 's are positive. Using condition (27) and (28) we can solve for  $z_L^0$  as a function of  $\phi$ . In particular  $z_L^0 = \frac{\phi A - w}{wc} - \frac{1}{\lambda}$ . This immediately implies that there exists a unique  $\phi^* > 0$  given by  $\phi_0^* = \frac{w}{A}(\frac{c}{\lambda} + 1)$  such that  $z_L^0 = 0$ . Note that  $\phi_0^* > \phi_L^*$ . This implies that the 0 layer constraint binds before the  $L$  layer constraint. Now consider layer 1 and solve for  $\phi_1^*$  such that  $z_L^1 = 0$  and all other  $z$ 's are positive. From the first-order conditions with respect to  $z_L^1$  we get  $z_L^1(\phi) = \frac{1}{\lambda h} e^{\lambda(\frac{\phi A - w}{wc} - \frac{1}{\lambda})} - \frac{1}{\lambda} - \frac{1}{c}$ . It immediately follows that there exists a unique  $\phi_1^*$  such that  $z_L^1(\phi_1^*) = 0$ . Let  $\Theta \equiv 1 + \frac{\lambda}{c}$ , then  $\phi_1^* = \frac{w}{A}(1 + \frac{c}{\lambda}(1 + \ln(h\Theta)))$ . Note that under Assumption 1

---

self-employed firm when the quantity produced is close enough to  $A$  or larger. However, all the results below hold for the case of  $L = 1$  as well.

$\phi_1^* \geq \phi_0^*$  since  $\phi_1^* = \phi_0^* + \frac{1}{A} \left( \frac{wc}{\lambda} \ln(h\Theta) \right)$  and  $\ln(h\Theta) \geq 0$ . Let's consider now layer 2. We solve for  $\phi_2^*$  such that  $z_L^2 = 0$  and consider that all other  $z$ 's are positive, we get  $z_L^2(\phi) = \frac{z_L^1(\phi)}{\lambda} - \frac{1}{\lambda} - \frac{1}{\lambda}$ . Note that there exists a  $\phi_2^*$  such that  $z_L^2(\phi_2^*) = 0$ ,  $\phi_2^* = \phi_0^* + \frac{1}{A} \left( \frac{wc}{\lambda} \ln(h \ln(\Theta) + h\Theta) \right)$ . Note that  $\phi_2^* > \phi_1^*$  since  $\phi_2^* - \phi_1^* = \frac{1}{A} \frac{wc}{\lambda} (\ln(\ln(\Theta) + \Theta) - \ln(\Theta)) > 0$ . Continuing in a recursive way we will find that  $\phi_{L-1}^* > \phi_{L-2}^*$  since  $\phi_{L-1}^* - \phi_{L-2}^* = \frac{1}{A} \frac{wc}{\lambda} (\ln(\ln(\Theta) + \Theta) - \ln(\Theta)) > 0$ . Therefore,  $\phi_{L-1}^* > \phi_{L-2}^* \dots > \phi_1^* \geq \phi_0^* > \phi_L^*$ . Hence, as the multiplier over the output constraint,  $\phi$ , decreases, the constraints over  $z_L^l \geq 0$  bind in a descending order starting from  $z_L^{L-1}$  up to  $z_L^0$  and then  $z_L^L$ . ■

**PROPOSITION B.1.** Under Assumption 1, for all  $L \neq 1$  and any output level  $q$ , the knowledge of agents at all layers is positive ( $z_L^l \geq 0$  never binds).

*Proof.* Let  $L > 1$  and consider two organizations, an unconstrained organization  $L$  with  $L$  layers, and a constrained organization  $L'$  with  $L'$  layers, where  $L' = L + 1$  and constrained to  $z_{L'}^{L'-1} = 0$ . Consider the problem of the constrained organization  $L'$ . The cost function to minimize is

$$\frac{w \left( cz_{L'}^0 + 1 + h \sum_{l=0}^{L'-3} e^{-\lambda Z_{L'}^l} (cz_{L'}^{l+1} + 1) + h e^{-\lambda Z_{L'}^{L'-2}} (cz_{L'}^{L'} + 1) + h e^{-\lambda Z_{L'}^{L'-2}} \right)}{h e^{-\lambda Z_{L'}^{L'-2}}}.$$

The first-order condition of the cost minimization problem for  $z_{L'}^0$  is

$$c + \lambda (cz_{L'}^0 + 1) - \frac{\phi}{w} \lambda A + \frac{\theta_{L'}^0}{w n_{L'}^0} \leq 0,$$

for  $z_{L'}^l$ ,  $0 < l < L' - 1$

$$h e^{-\lambda Z_{L'}^{l-1}} c + \lambda (cz_{L'}^0 + 1) + \lambda h \sum_{k=0}^{l-1} e^{-\lambda Z_{L'}^k} (cz_{L'}^{k+1} + 1) - \lambda A \frac{\phi}{w} + \frac{\theta_{L'}^l}{w n_{L'}^0} \leq 0,$$

and for  $z_{L'}^{L'}$

$$c - \frac{\phi}{w} \frac{\lambda A}{h} e^{-\lambda Z_{L'}^{L'}} + \frac{\theta_{L'}^{L'}}{w} \leq 0.$$

Note that these first-order conditions are exactly the same as the ones of the  $L$  layer firm, equations (7) to (9), since  $n_L^0 = \frac{1}{h} e^{\lambda Z_L^{L-1}}$  and  $n_{L'}^0 = \frac{1}{h} e^{\lambda Z_{L'}^{L'-2}}$ . Therefore, the firms will choose the same  $z$ 's.

However, organization  $L'$  will have more employees than organization  $L$  since it has one more layer. Therefore, the total cost of organization  $L'$  is strictly larger than organization  $L$ , by  $w$ . Note that the number of managers at layer  $L' - 1$  that the entrepreneur hires is  $n_{L'}^{L'-1} = n_L^0 h e^{-\lambda Z_{L'}^{L'-2}} = 1$ . Hence, the firm will never choose to have an organization with  $z_{L'}^{L'-1} = 0$ . Lemma 1 then implies that it will never choose to have  $z_l^l = 0$  for  $l < L - 1$ .

The case for  $L=0$  is trivial, since  $z_0^0 = 0$  implies that the output of the firm is equal to zero. ■

We now present the proofs of the propositions in the text:

*Proof of Proposition 1.* Consider first part 1. From the first-order conditions of the cost minimization problem we have that the marginal cost is strictly increasing in  $q$ ,

$$\frac{d\phi}{dq} = \frac{\frac{hwc}{A\lambda} e^{\lambda z_L^L}}{A \left[ e^{\lambda Z_L^L} \left( 1 + n_L^0 \sum_{l=1}^{L-1} (1/n_L^l) \right) + \frac{1}{h} e^{-\lambda z_L^L} \right]} > 0,$$

where we used that

$$(31) \quad \frac{dz_L^l}{dz_L^L} = \begin{cases} \frac{dz_L^0}{dz_L^L} = h e^{\lambda z_L^L} > 0 & \text{for } l = 0 \\ \frac{dz_L^0}{dz_L^L} \prod_{v=1}^l \frac{dz_L^v}{dz_L^{v-1}} = e^{\lambda z_L^L} e^{\lambda Z_L^{l-1}} > 0 & \text{for } l > 0, \end{cases}$$

and

$$(32) \quad \frac{dz_L^L}{d\phi} = \frac{A}{hwce^{\lambda z_L^L}} > 0.$$

Therefore, the knowledge of all employees,  $z_L^l$  for all  $l$ , increase with  $q$ . From this result it immediately follows that the number of employees at all  $l < L$ , and the span of control of managers at all  $l$ , are increasing in  $q$ , given that both are an increasing function of knowledge; namely

$$(33) \quad n_L^0 = \frac{e^{\lambda Z_L^{L-1}}}{h}, n_L^l = e^{\lambda(Z_L^{L-1} - Z_L^{l-1})} = e^{\lambda \sum_{v=1}^{L-l} z_L^{L-v}}, \frac{n_L^0}{n_L^1} = \frac{e^{\lambda z_L^0}}{h}, \frac{n_L^l}{n_{L+1}^l} = e^{\lambda z_L^l}.$$

We now show that given  $L$  the autonomy of the entrepreneur decreases with  $q$ ,  $\frac{d(z_L^L/Z_L^L)}{dq} < 0$ . Note that

$$\frac{d\left(\frac{z_L^L}{Z_L^L}\right)}{dq} = \frac{\frac{dz_L^L}{dq}}{(Z_L^L)^2} \sum_{l=0}^L \left( z_L^l - \frac{dz_L^L}{dz_L^L} z_L^l \right).$$

For the autonomy of the entrepreneur to decrease it suffices to show that  $z_L^l - \frac{dz_L^l}{dz_L^0} z_L^0 < 0$  for all  $l$ . Given Proposition B.1, we specialize to the case of positive knowledge. If we allow for zero knowledge at some  $l$ , then  $z_L^l - \frac{dz_L^l}{dz_L^0} z_L^0 = 0$  for all  $l$  such that  $z_L^l = 0$ , and then if  $z_L^l - \frac{dz_L^l}{dz_L^0} z_L^0 < 0$  for all  $l$  such that  $z_L^l > 0$  it immediately implies that  $d\left(\frac{z_L^l}{z_L^0}\right)/dq \leq 0$ . Consider first the case of  $l=0$ . From equation (31)  $z_L^0 - \frac{dz_L^0}{dz_L^0} z_L^0 = z_L^0 - h e^{\lambda z_L^0} z_L^0$ . Using the first-order conditions (27) to (30) we can rewrite  $z_L^0$  as a function of  $z_L^0$ , namely  $z_L^0 = \frac{1}{\lambda} \ln\left(\frac{c+\lambda(cz_L^0+1)}{hc}\right)$ . Then  $z_L^0 - h e^{\lambda z_L^0} z_L^0 = \left(\frac{c+\lambda}{c\lambda} + z_L^0\right)(f(z_L^0) - g(z_L^0))$ , where  $f(z_L^0) = \frac{\lambda c z_L^0}{c+\lambda(cz_L^0+1)}$  and  $g(z_L^0) = \ln\left(\frac{1}{h} + \frac{\lambda}{h}(z_L^0 + \frac{1}{c})\right)$ . These functions have the following properties,  $f(0)=0$ ,  $g(0) = \ln\left(\frac{c+\lambda}{hc}\right) > 0$ ,  $f'(z_L^0) > 0$ ,  $g'(z_L^0) > 0$  and  $\frac{g'(z_L^0)}{f'(z_L^0)} > 1$ . Therefore  $g(z_L^0) > f(z_L^0)$  for all  $z_L^0 > 0$ . Hence

$$(34) \quad e^{\lambda z_L^0} z_L^0 > z_L^0.$$

Now consider the case of  $l=1$ . From equation (31)  $z_L^1 - \frac{dz_L^1}{dz_L^0} z_L^0 = z_L^1 - e^{\lambda z_L^0} e^{\lambda z_L^1} z_L^0$ . Since  $e^{\lambda z_L^0} z_L^0 > z_L^0/h$ , it is sufficient to show that  $z_L^1 - (z_L^0/h) e^{\lambda z_L^0} < 0$ . Using again the first-order conditions in equations (27) to (30) we can rewrite  $z_L^0$  as a function of  $z_L^1$ , namely  $z_L^0 = \frac{1}{\lambda} \ln\left(\frac{c+\lambda(cz_L^1+1)}{c/h}\right)$ . Therefore  $z_L^1 - (z_L^0/h) e^{\lambda z_L^0} = \left(\frac{c+\lambda}{c\lambda} + z_L^1\right)(f(z_L^1) - g(z_L^1) - 2\ln h)$ , where  $f(\cdot)$  and  $g(\cdot)$  are defined as before. Note that  $f(0) - g(0) - 2\ln h = -\ln\left(\frac{c+\lambda}{c/h}\right) < 0$  by Assumption 1. Since  $g(z_L^1) > f(z_L^1)$  for all  $z_L^1 > 0$ ,  $z_L^1 - \frac{dz_L^1}{dz_L^0} z_L^0 < 0$ . Moreover

$$(35) \quad e^{\lambda z_L^0} e^{\lambda z_L^1} z_L^0 > z_L^1.$$

To show that  $z_L^l - \frac{dz_L^l}{dz_L^0} z_L^0 < 0$  for all  $l > 0$ , we first show that  $e^{\lambda z_L^{l-1}} z_L^{l-1} > z_L^l$  for all  $1 < l < L$ . From the first-order conditions (27) to (30) we can rewrite  $z_L^{l-1}$  as a function of  $z_L^l$ , namely  $z_L^{l-1} = \frac{1}{\lambda} \ln\left(\frac{c+\lambda(cz_L^l+1)}{c}\right)$ . Then  $z_L^l - z_L^{l-1} e^{\lambda z_L^{l-1}} = \left(\frac{c+\lambda}{c\lambda} + z_L^l\right)(f(z_L^l) - g(z_L^l) - 2\ln h)$ , where  $f(\cdot)$  and  $g(\cdot)$  are defined as before. Note that  $f(0) - g(0) - 2\ln h = -\ln\left(\frac{c+\lambda}{c}\right)$  and given the properties of  $f(\cdot)$  and  $g(\cdot)$ , then  $g(z_L^l) > f(z_L^l)$  for all  $z_L^l > 0$ . Therefore

$$(36) \quad e^{\lambda z_L^{l-1}} z_L^{l-1} > z_L^l, \text{ for all } 1 < l < L.$$

We now show that  $z_L^l - \frac{dz_L^l}{dz_L^l} z_L^l < 0$  for all  $1 < l < L$ . From equation (31)  $\frac{dz_L^l}{dz_L^l} = e^{\lambda z_L^l} e^{\lambda Z_L^{l-1}}$ . Then

$$e^{\lambda z_L^l} z_L^l e^{\lambda Z_L^{l-1}} = e^{\lambda z_L^0} e^{\lambda z_L^l} z_L^l e^{\lambda (Z_L^{l-1} - z_L^0)} > e^{\lambda z_L^1} z_L^1 e^{\lambda (Z_L^{l-1} - Z_L^1)} \text{ by (35)}$$

$$e^{\lambda z_L^1} z_L^1 e^{\lambda (Z_L^{l-1} - Z_L^1)} > e^{\lambda z_L^k} z_L^k e^{\lambda (Z_L^{l-1} - Z_L^k)} \text{ for } 2 \leq k \leq l-2 \text{ by (36),}$$

$$e^{\lambda z_L^{l-2}} z_L^{l-2} e^{\lambda z_L^{l-1}} > z_L^{l-1} e^{\lambda z_L^{l-1}} > z_L^l \text{ by (36).}$$

Therefore,  $z_L^l - \frac{dz_L^l}{dz_L^l} z_L^l < 0$  for all  $1 < l < L$ . For the case in which  $l = L$ , then  $z_L^l - \frac{dz_L^l}{dz_L^l} z_L^l = 0$ . Hence  $\sum_{l=0}^L \left( z_L^l - \frac{dz_L^l}{dz_L^l} z_L^l \right) < 0$ , the autonomy of the entrepreneur,  $\frac{z_L^l}{Z_L^l}$  decreases with  $q$ .

Now consider part 2. To prove (i) note that since  $MC_L(q; w) = \frac{hwc}{\lambda A} e^{\lambda z_L^l} > 0$  for any  $z_L^l$ , then the cost function  $C_L(q; w)$  is strictly increasing. Since this is the case for all  $L$ ,  $C(q; w)$  is strictly increasing.

The proof of (ii) is immediate from equation (11) and (iii).

The claim in (iii) follows from the first-order conditions of the cost minimization problem. Note that equations (7) to (9), with  $\theta_L^l = 0$  for all  $l$  and  $L$  depend on  $w$  only through the term  $\phi/w$ . Thus, if the system of first-order conditions is satisfied with  $w$ , and  $\phi$ . If  $w' = tw$  for  $t > 0$ , the system is satisfied with  $t\phi$  and the rest of the allocation does not change. Hence, factor demands do not change with  $w$ .

(iv). consider first the case of a self-employed entrepreneur,  $L = 0$ . The average cost is given by

$$AC_0(q; w) = w \left( \frac{c}{\lambda q} \ln \left( \frac{A}{A - q} \right) + \frac{1}{q} \right).$$

The partial derivative of the  $AC_0(q; w)$  with respect to  $q$  is

$$\frac{\partial AC_0(q; w)}{\partial q} = - \frac{AC_0(q; w) - MC_0(q; w)}{q}.$$

Note that the average cost is decreasing if  $AC_0(q; w) > MC_0(q; w)$  and increasing if  $AC_0(q; w) < MC_0(q; w)$ . At  $AC_0(q; w) = MC_0(q; w)$  it has a unique stationary point  $q^*$  that solves  $\frac{c}{\lambda} \ln \left( \frac{A}{A - q^*} \right) + w = \frac{c}{\lambda A} \frac{q^*}{A - q^*}$ . This stationary point is a minimum since

$$\frac{\partial^2 AC_0(q; w)}{\partial q^2} = \frac{c}{\lambda} \frac{1}{q^2} + \left( \frac{2AC_0(q; w) + \frac{c}{\lambda A} (A - q)}{q^2} \right) > 0 \text{ since } q < A,$$



and

$$\lim_{q \rightarrow 0} AC_0(q; w) = +\infty, \lim_{q \rightarrow A} AC_0(q; w) = +\infty,$$

which shows that the average cost function is strictly convex. The marginal cost is

$$MC_0(q; w) = \frac{wc}{\lambda} \frac{1}{A - q} = \frac{wc}{\lambda A} e^{\lambda z_0^L}.$$

Now consider the cost function of an organization with  $L$  layers producing  $q$

$$C_L(q; w) = w \left( \frac{ch}{\lambda A} e^{\lambda z_L^L} q + \frac{c}{\lambda} \left( 1 - e^{\lambda z_L^{L-1}} + \lambda z_L^L + \frac{\lambda}{c} \right) \right).$$

Recall that the marginal cost is given by  $MC_L(q; w) = \frac{wch}{\lambda A} e^{\lambda z_L^L}$ . This implies that

$$(37) \quad \Rightarrow \left( \frac{ch}{A} e^{\lambda z_L^L} \frac{\partial z_L^L}{\partial q} + \frac{c}{q} \left( \frac{\partial z_L^L}{\partial q} - e^{\lambda z_L^{L-1}} \frac{\partial z_L^{L-1}}{\partial q} \right) \right) q = 0.$$

Now consider the average cost of this organization,

$$AC_L(q; w) = w \left( \frac{ch}{\lambda A} e^{\lambda z_L^L} + \frac{c}{\lambda} \frac{1 - e^{\lambda z_L^{L-1}} + \lambda z_L^L + \lambda/c}{q} \right).$$

Some properties of this function are

$$\lim_{q \rightarrow 0} AC_L(q; w) = +\infty, \lim_{q \rightarrow +\infty} AC_L(q; w) = +\infty,$$

where we use the fact that  $\frac{\partial z_L^L}{\partial q} > 0$  and  $\frac{\partial z_L^{L-1}}{\partial q} > 0$ , which follows from part 1 of the proof. The partial derivative of the  $AC_L(q; w)$  with respect to  $q$  is

$$\frac{\partial AC_L(q; w)}{\partial q} = -\frac{1}{q} (AC_L(q; w) - MC_L(q; w)),$$

where we use equation (37). Note that the average cost is decreasing if  $AC_L(q; w) > MC_L(q; w)$  and increasing if  $AC_L(q; w) < MC_L(q; w)$ . It has a stationary point where they are equal. In particular, there is a unique stationary point  $q^*$  such that  $\frac{\partial AC_L(q^*; w)}{\partial q} = 0$  and  $q^*$  solves

$$\Rightarrow AC_L(q^*; w) = \frac{wch}{\lambda A} e^{\lambda z_L^L(q^*)}.$$

By evaluating the second derivative of the average cost function, we show that the stationary point is a minimum. That is

$$\frac{\partial^2 AC_L(q; w)}{\partial q} = \frac{2}{q^2} \left( AC_L(q; w) - \frac{wch}{\lambda A} e^{\lambda z_L^L} \right) + \frac{wch}{qA} e^{\lambda z_L^L} \frac{\partial z_L^L}{\partial q},$$

and so at  $q^*$ ,

$$\frac{\partial^2 AC_L(q^*; w)}{\partial q} = \frac{wch}{qA} e^{\lambda z_L^L(q^*)} \frac{\partial z_L^L}{\partial q} > 0.$$

This implies that the average cost function is convex and it reaches a minimum when it intersects with the marginal cost function. Also note that the MES is implicitly given by

$$e^{-\lambda z_L^{L-1}(q^*)} = \frac{c}{c + \lambda(cz_L^L(q^*) + 1)}. \quad \blacksquare$$

*Proof of Proposition 2.* We first prove that the minimum average cost decreases with the number of layers. Consider an organization with  $L$  layers producing  $q_L^*$  at the MES. From the first-order conditions of the cost minimization problem, together with the fact that the firm is producing at its MES, the  $L+1$  knowledge levels of the organization solve the following set of  $L+1$  equations

$$\begin{aligned} \lambda z_L^L &= \ln(\lambda z_L^0 + \Theta) - \ln h, \\ \lambda z_L^0 &= \ln(\lambda z_L^1 + \Theta) + \ln h, \\ \lambda z_L^{l-1} &= \ln(\lambda z_L^l + \Theta) \text{ for all } 1 < l \leq L, \end{aligned}$$

where  $\Theta \equiv 1 + \frac{\lambda}{c}$ . Define the function  $g(x) = \ln x + \Theta$  and abusing notation we will refer to  $g^L(v)$  to the operator applied to the function such that it evaluates the function  $L$  times over itself, namely,  $g^L(v) = g(g(\dots g(v)))$  where we are applying the  $g$  function  $L$  times.

Some properties of the function  $g$  are:

Property 1: The function  $g^l(v)$  evaluated at  $\Theta$ ,  $g^l(\Theta)$ , is increasing in  $l$ . Namely  $g^L(\Theta) > g^{L-1}(\Theta) > \ln \Theta + \Theta > 1$  for all  $L$ .

Property 2: There exists a unique  $v^*$ , such that  $v^* > \Theta$  and  $g^L(v^*) = g^{L+1}(v^*)$ . Note that the function  $g(v)$  is continuously increasing. The  $\lim_{v \rightarrow \infty} g(v) = +\infty$ ,  $\lim_{v \rightarrow 0} g(v) = -\infty$ ,  $g'(v) = \frac{1}{v}$ .  $g'(v) \geq 1$  for all  $v \leq 1$  and  $g'(v) \leq 1$  for all  $v \geq 1$  and also

$\lim_{v \rightarrow \infty} g'(v) = 0$ . Note that at  $v = \Theta$ ,  $g(\Theta) = \ln \Theta + \Theta > \Theta$ . By the intermediate value theorem, the function  $g(v)$  has two stationary points. In particular, there exists a unique  $v^* > \Theta$  such that  $g^L(v^*) = g^{L+1}(v^*)$ .

Property 3:  $\lim_{v \rightarrow \infty} g(v) = +\infty$  and this implies that  $\lim_{v \rightarrow \infty} g^l(v) = +\infty$  for all  $l$ .

Property 4: The partial derivative of  $g^l(v)$  with respect to  $v$  is given by

$$\frac{\partial g^L(v)}{\partial v} = g'(v) \prod_{i=1}^{L-1} g'(g^i(v)) = g'(g^{L-1}(v)) \frac{\partial g^{L-1}(v)}{\partial v}.$$

Property 5:  $g'(v) = \frac{1}{v} \Rightarrow g'(v) < 1$  for all  $v > 1$ .  $g'(g(v)) = \frac{1}{g(v)} < 1$  for all  $v > \exp(1 - \Theta)$  since  $g(v) > 1$  for all  $v > \exp(1 - \Theta)$ .

These five properties of  $g$  are going to be used in the remaining part of the proof. We define the following two functions  $f(v) = \frac{v - \Theta + \ln h}{\lambda}$ , and  $G(v, L) = \frac{1}{\lambda} \ln(g^L(v) + \ln h)$ . Note that the solution for  $z_L^L$  at the MES solves  $f(\lambda z_L^L + \Theta) = G(\lambda z_L^L + \Theta, L)$ . We will now show that there is a unique solution to this equation and that the solution decreases with  $L$ . Consider the following change of variables,  $v = \lambda z + \Theta$ . Note that when  $z = 0$ ,  $v = \Theta$ ; and when  $z \rightarrow \infty$ ,  $v \rightarrow \infty$ . First note that at  $v = \Theta$ ,  $f(\Theta) = \frac{\ln h}{\lambda} < 0$ , and  $G(\Theta, L) = \frac{1}{\lambda} \ln(g^L(\Theta) + \ln h) > \frac{1}{\lambda} \ln(\ln \Theta + \Theta + \ln h)$  by Property 1 and  $> 0$  by Assumption 1. Note that both functions converge to infinity as  $v$  converges to infinity,  $\lim_{v \rightarrow +\infty} f(v) = +\infty$ , and  $\lim_{v \rightarrow +\infty} G(v, L) = \frac{1}{\lambda} \ln(g^L(+\infty) + \ln h) = +\infty$  by Property 3. Now note that both functions are strictly increasing,

$$f'(v) = \frac{1}{\lambda} > 0,$$

and

$$G'(v, L) = \frac{1}{\lambda} \frac{1}{g^L(v) + \ln h} \frac{\partial g^L(v)}{\partial v} = \left( \prod_{i=1}^{L-1} g'(g^i(v)) \right) \frac{g'(v)}{\lambda} > 0 \text{ by Property 4,}$$

and also

$$G'(v, L) < \frac{1}{\lambda} \text{ for all } z > 0 \text{ by Property 5.}$$

We have just shown that the functions cross only once. This follows because, at  $z=0$ ,  $f(\Theta) < G(\Theta, L)$  and both functions are continuously increasing where  $f$  increases at a higher rate than  $G$ . By the intermediate Value theorem, this immediately implies that there exists a unique  $z_L^L$  that solves  $f(\lambda z_L^L + \Theta) = G(\lambda z_L^L + \Theta, L)$ . We now characterize how  $z_L^L$  changes as we change  $L$ .

We show that the value of  $z_L^L$  such that  $f(\lambda z_L^L + \Theta) = G(\lambda z_L^L + \Theta, L)$  decreases as we increase  $L$ . To do so, we first show that the function  $G(v, L)$  increases in  $L$  at  $z=0$  ( $v=\Theta$ ). Then we show that there exists a unique  $z^* > 0$  ( $v^* > \Theta$ ) such that  $G(v^*, l)$  has the same value for all  $l$ . Then we show that the  $z_L^L$  that solves  $f(\lambda z_L^L + \Theta) = G(\lambda z_L^L + \Theta, L)$  has to be larger than  $z^*$ . Finally we show that  $z_L^L$  decreases with  $L$ .

First note that  $G(\Theta, l)$  increases with  $l$ , since  $G(\Theta, l) = \frac{1}{\lambda} \ln(g^l(\Theta) + \ln h) > \frac{1}{\lambda} \ln(g^{l-1}(\Theta) + \ln h) = G(\Theta, l-1)$  by Property 1. From property 2, there exists a unique  $v^* > \Theta$  such that  $g^l(v^*) = g^{l+1}(v^*)$  and this implies that  $G(v^*, l) = G(v^*, l+1)$ , for all  $l > 2$ . Note that at  $v^*$ ,  $G(v^*, L) > f(v^*)$ . Since  $G(v^*, L) = \frac{1}{\lambda} \ln(v^* + \ln h)$  and  $f(v^*) = \frac{v^* - \Theta + \ln h}{\lambda} = \frac{\ln v^* + \ln h}{\lambda}$  using that  $v^* = \ln v^* + \Theta$  and so this implies that  $G(v^*, L) > f(v^*)$  for all  $1 > h > 1/v^*$ . Note that  $h > 1/\Theta > 1/v^*$  by Assumption 1 and Property 2.

Finally note that  $G'(v, L-1) > G'(v, L)$ . Since

$$\begin{aligned} \frac{G'(v, L)}{G'(v, L-1)} &= \frac{g^{L-1}(v) + \ln h}{g^L(v) + \ln h} g'(g^{L-1}(v)) \\ &= \left( \frac{1}{g^L(v) + \ln h} \right) \left( \frac{g^{L-1}(v) + \ln h}{g^{L-1}(v)} \right) < 1. \end{aligned}$$

Therefore, the knowledge of the entrepreneur,  $z_L^L$ , at the MES,  $q_L^*$ , decreases with  $L$ . Since at the MES the average cost is equal to the marginal cost and the marginal cost is an increasing function of  $z_L^L$ , we have shown that the minima of the average cost decrease with the number of layers,  $L$ .

To show that the level of output that achieves the MES increases with the number of layers consider an organization with  $L$  layers producing at the MES,  $q_L^*$ . Given  $q_L^*$ , this organization has no incentives to modify its knowledge level at any layer. To see this, consider adding a top layer to this organization with a span of control of one. Assigning positive knowledge to the new top layer is suboptimal because it has the same cost and benefit than increasing the knowledge of the manager at layer  $L$ . Given

that the knowledge of this manager was already optimal and that the cost of acquiring knowledge is linear, a new top manager with span of control of one will increase the total cost by  $w$  but will not change production.

Now consider an organization with  $L + 1$  layers constrained to produce  $q_L^*$  where the span of control of the manager is greater than one, as in our model. Consider the case where the knowledge of the top manager at layer  $L + 1$  is zero. First note that this organization, from layer  $L$  down the hierarchy, is a replica of the organization with  $L$  layers. The difference is that there are  $n_{L+1}^L$  “managers at the top” replicas of these  $L$ -suborganizations. This immediately implies that in order to produce  $q_L^*$  each of these  $L$ -suborganizations has to produce less than  $q_L^*$ . That is, the cost of the  $L + 1$  organization with zero knowledge at the top that produces  $q_L^*$  is equal to

$$w + AC_L\left(\frac{q_L^*}{n_{L+1}^L}; w\right)q_L^* > AC_L(q_L^*; w)q_L^*,$$

where the inequality is the result of the convexity of  $AC(\cdot)$  (Proposition 1) and that  $q_L^*$  achieves the MES with  $L$  layers. Note that if each of these  $L$ -suborganizations produces at its MES,  $q_L^*$ , then the  $L + 1$  organization will be producing more than  $q_L^*$ , namely  $n_{L+1}^L q_L^*$ . Such an organization will have an average cost of  $AC_L(q_L^*; w)$  which is lower than  $AC_{L+1}(q_L^*; w)$  with  $z_{L+1}^{L+1} = 0$ , but it will be producing strictly more than  $q_L^*$  violating the constraint of producing  $q_L^*$ .

In general,  $z_{L+1}^{L+1} > 0$ , since the span of control of the manager at the top is larger than one, which results in suboptimal knowledge acquisition in the  $L$ -suborganization. However, this adjustment is just the result of the inability of the organization to have  $L$ -suborganizations that produce  $q_L^*$ . Hence, the  $L + 1$  organization necessarily incurs a higher average cost when producing  $q_L^*$ . Because organizations with higher layers have lower minimum average costs, this immediately implies that the  $L + 1$  organization is not at its MES at  $q_L^*$ . Moreover, we also just argued that if we relaxed the constraint over the quantity produced, the  $L + 1$  organization is able to produce at a lower average cost but with a higher scale, for example, by producing  $n_{L+1}^L q_L^*$ . Since the average cost is a convex function, and when producing more than  $q_L^*$  the  $L + 1$  layer organization has strictly lower average costs, this

implies that  $q_{L+1}^* > q_L^*$ . Thus, the MES of the  $L + 1$  organization has a higher scale of production than the  $L$  layer organization.

We now show that the marginal cost ( $\phi$ ) decreases at the crossing where the number of layers ( $L$ ) increases. Consider an organization with  $L$  layers producing  $q$  and let  $\phi_L$  be the multiplier under organization  $L$ . Now suppose that we are at the  $q$  such that the entrepreneur finds it optimal to increase the number of layers, say, to  $L' = L + 1$  layers. Note that the cost minimization problem of organization  $L'$  has an extra first-order condition compared to organization  $L$ , the first-order condition with respect to  $z_{L'}^{L'-1}$ . Let the multiplier under organization  $L'$  be  $\phi_{L'}$ . First note  $\phi_{L'} = \phi_L$  is inconsistent with fixing  $q$ . The reason is that if  $\phi_{L'} = \phi_L$  then from the first-order conditions we have that  $z_{L'}^l = z_L^l$  for  $l = 0, \dots, L - 1$ ,  $z_{L'}^L = z_L^L$  and  $z_{L'}^{L'-1} > 0$  by Proposition B.1, since it was optimal to increase the number of layers. This implies that  $Z_{L'}^{L'-1} > Z_L^{L-1}$  and the organization will produce  $\frac{A}{h} [e^{\lambda Z_{L'}^{L'-1}} - e^{-\lambda z_{L'}^L}] > q$ . Following the same argument,  $\phi_{L'} > \phi_L$  is inconsistent with fixing  $q$ , because the organization will also produce more than  $q$  given that each  $z$  is an increasing function of  $\phi_{L'}$ . Hence, the only way  $L'$  is optimal and can produce  $q$  is if  $\phi_{L'} < \phi_L$ .

The result that the knowledge of all employees  $z_L^l$  for all  $l$  decrease discontinuously at the crossing where the number of layers ( $L$ ) increases, follows from equations (7) to (9) since each  $z$  is an increasing function of  $\phi$  and  $\phi$  decreases discontinuously as  $L$  increases.

From equation (33) it follows that the span of control of managers at layer  $l + 1$  are an increasing function of knowledge at layer  $l$ . Because the knowledge of all employees  $z_L^l$  for all  $l$  decrease discontinuously at the crossing where the number of layers ( $L$ ) increases, the span of control of managers at layer  $l$ ,  $\frac{n_L^l}{n^{l+1}}$ , decrease discontinuously at the crossing where the number of layers ( $L$ ) increases.

We now show that the autonomy of the entrepreneur decreases discontinuously at the crossing where the number of layers ( $L$ ) increases. At  $q$  such that the cost of producing with  $L$  layers of management is equal to the cost of producing with  $L + 1$  layers of management,

$$\frac{A}{h} (e^{\lambda Z_L^{L-1}} - e^{-\lambda z_L^L}) = \frac{A}{h} (e^{\lambda Z_{L+1}^L} - e^{-\lambda z_{L+1}^{L+1}}).$$

Since the marginal cost ( $\phi$ ) decreases discontinuously as the number of layers ( $L$ ) increases, the knowledge of the

entrepreneur decreases discontinuously,  $z_L^L > z_{L+1}^{L+1}$ . Therefore, at  $q$ ,  $Z_L^{L-1} < Z_{L+1}^L$ . This immediately results in  $\frac{Z_L^{L-1}}{z_L^L} < \frac{Z_{L+1}^L}{z_{L+1}^{L+1}}$ , and so adding one on both sides,  $\frac{Z_L^L}{z_L^L} < \frac{Z_{L+1}^{L+1}}{z_{L+1}^{L+1}}$  which implies  $\frac{z_L^L}{Z_L^L} > \frac{z_{L+1}^{L+1}}{Z_{L+1}^{L+1}}$ .

Finally, we show that the number of employees at all layers,  $n_L^l$  for all  $l$  increase discontinuously at  $q$  such that the number of layers ( $L$ ) increases by first showing that the number of workers, employees at layer 0, increase at  $q$ . By definition,  $n_L^0 = \frac{e^{\lambda Z_L^{L-1}}}{h}$ , and  $n_{L+1}^0 = \frac{e^{\lambda Z_{L+1}^L}}{h}$ , and since at  $q$ ,  $Z_L^{L-1} < Z_{L+1}^L$ , it immediately implies that  $n_{L+1}^0 > n_L^0$  at  $q$ . Now consider the case of managers at layer  $l \leq L-1$ . By definition,  $n_L^l = hn_L^0 e^{-\lambda Z_L^{l-1}}$  and  $n_{L+1}^l = hn_{L+1}^0 e^{-\lambda Z_{L+1}^{l-1}}$ , and since at  $q$ ,  $n_{L+1}^0 > n_L^0$  and  $Z_L^{l-1} > Z_{L+1}^{l-1}$ , it immediately implies that  $n_{L+1}^l > n_L^l$  for all  $l \leq L-1$  at  $q$ . At layer  $l=L$ ,  $n_{L+1}^L = hn_{L+1}^0 e^{-\lambda Z_{L+1}^{L-1}} = e^{\lambda Z_{L+1}^L} > 1 = n_L^L$ . ■

*Proof of Proposition 3.* Consider an organization with  $L=1$ , fix  $c$  and let  $\lambda \rightarrow \infty$  so that  $\frac{c}{\lambda} \rightarrow 0$ . We guess that  $\lim_{\lambda \rightarrow \infty} \lambda z_1^1 = \infty$  and verify our guess. Note that this guess combined with the first-order condition implies that

$$\lim_{\lambda \rightarrow \infty} \lambda z_1^1 = \infty \Rightarrow \lim_{\lambda \rightarrow \infty} \lambda z_1^0 = \ln \frac{qh}{A}.$$

Then from the first-order condition with respect to  $z_1^0$

$$(38) \quad \lim_{\lambda \rightarrow \infty} \lambda z_1^1 = \lim_{\lambda \rightarrow \infty} \ln \frac{1}{h} \left( 1 + \ln \frac{qh}{A} + \frac{\lambda w}{c} \right) = \infty,$$

which verifies our guess. This also implies that  $\lim_{\lambda \rightarrow \infty} z_1^1 = 0$ . Note that from equation (38) we get

$$\lim_{\lambda \rightarrow \infty} \frac{\ln \frac{1}{h} \left( 1 + \ln \frac{qh}{A} + \frac{\lambda w}{c} \right)}{\lambda} = \lim_{\lambda \rightarrow \infty} \frac{\frac{w}{c}}{\left( 1 + \ln \frac{qh}{A} + \frac{\lambda w}{c} \right)} = 0.$$

Then as  $\lambda \rightarrow \infty$

$$\lim_{\lambda \rightarrow \infty} MC_1(q; w) = \lim_{\lambda \rightarrow \infty} \frac{hc}{A\lambda e^{\ln h / (1 + \ln \frac{qh}{A} + \frac{\lambda w}{c})}} = \lim_{\lambda \rightarrow \infty} \frac{c + c \ln \frac{qh}{A} + \lambda w}{A\lambda} = \frac{w}{A},$$

$$\lim_{\lambda \rightarrow \infty} AC_1(q; w) = \lim_{\lambda \rightarrow \infty} \frac{w + \frac{A}{q}w + \frac{A}{q}cz_1^1}{A} = w \left( \frac{1}{A} + \frac{1}{q} \right),$$

$$\lim_{\lambda \rightarrow \infty} C_1(q; w) = w \left( \frac{q}{A} + 1 \right) = w \frac{q}{A} + w.$$

In the limit economy, when  $\lambda \rightarrow \infty$ , the entrepreneur is able to solve all problems without any knowledge. Therefore, the problems that the workers generate are all converted into output but  $z_1^0 = 0$ . Since no knowledge is needed for production, there are no costs incurred in acquiring knowledge. The only costs are the wages that the manager has to pay to the workers and himself. The number of workers in a limit economy is given by  $n_1^0 = \frac{q}{A}$ ; therefore, the cost of production includes a variable cost  $\frac{wq}{A}$  plus a fixed cost  $w$  which is the wage of the entrepreneur. Removing this extra fixed cost generated in the limit economy, the variable cost function is given by  $C(q; w) = q \frac{w}{A}$ .

The marginal cost of increasing production is the extra cost of hiring more workers,  $\frac{w}{A}$ . Note that there are no gains from adding an extra layer of problem solvers. This will only increase the cost of the organization, since it will have to pay more salaries. In the limit economy, the average variable cost is equal to the marginal cost,  $AC(q; w) = \frac{w}{A}$ .

Productivity as defined in the text for the limiting case is given by

$$\alpha(q) = \frac{q}{\lim_{\lambda \rightarrow 0} C(q; 1)} = A. \quad \blacksquare$$

*Proof of Proposition 4.* From Proposition 1 we know that the marginal cost increases with  $q$ , and at the  $q$  such that the entrepreneur finds it optimal to switch layers, the marginal cost decreases (Proposition 2), since the price set by the entrepreneur is a constant markup over marginal cost. The properties of the marginal cost schedule are reflected in the pricing rule of the monopolist, namely, that prices increase within a number of layers and they jump down as we switch layers. We now need to show that  $q$  increases with  $\alpha$ . Within a number of layers,

$$\frac{\partial q(\alpha)}{\partial \alpha} = \frac{\frac{q(\alpha)}{\alpha}}{1 + \frac{\sigma q(\alpha)}{MC(q(\alpha); w)} \frac{\partial MC(q(\alpha); w)}{\partial q}} = \frac{\frac{q(\alpha)}{\alpha}}{1 + \sigma \lambda q(\alpha) \frac{\partial z^L}{\partial q}} > 0,$$

which follows from  $\frac{\partial MC(q(\alpha); w)}{\partial q(\alpha)} > 0$ . Note that when the organization increases layers,  $q(\alpha)$  does not move continuously with  $\alpha$ . This follows from the dependence of the production supplied on the marginal cost. In particular,  $q$  is a decreasing function of the marginal cost,

$$\frac{\partial q(\alpha)}{\partial MC(q(\alpha); w)} = -\sigma \frac{q(\alpha)}{MC(q(\alpha); w)}.$$



Therefore, when the marginal cost decreases as a consequence of adding a new layer, the new quantity supplied goes up discontinuously. The elasticity of the increase in  $q(\alpha)$  with respect to the marginal cost is given by  $\sigma$ .

Because  $r(\alpha) = p(\alpha)q(\alpha)$ , the elasticity of  $r(\alpha)$  with respect to  $MC(q(\alpha); w)$  is equal to  $1 - \sigma < 0$ , and so revenues jump up when  $MC(q(\alpha); w)$  jumps down as a result of an increase in the number of layers. ■

*Proof of Proposition 5.* The result follows immediately from

$$\frac{\partial^2 \pi(\alpha; w)}{\partial q} = -\left(\frac{\sigma - 1}{\sigma^2}\right) q(\alpha)^{-1/\sigma-1} (\alpha R)^{1/\sigma} P^{(\sigma-1)/\sigma} - \frac{\partial MC(q(\alpha); w)}{\partial q} < 0,$$

$$\text{since } \frac{\partial MC(q(\alpha); w)}{\partial q} = \begin{cases} \frac{wc}{A} e^{\lambda z_0^0} \frac{\partial z_0^0}{\partial q} > 0 & \text{for } L = 0 \\ \frac{wch}{A} e^{\lambda z_L^L} \frac{\partial z_L^L}{\partial q} > 0 & \text{for } L > 0 \end{cases}.$$

Note also that since the profit function is given by

$$\pi(\alpha) = \alpha R \left( \frac{\sigma}{\sigma - 1} MC(q(\alpha); w) \right)^{1-\sigma} - C(q(\alpha); w) - wf,$$

and the cost and marginal cost functions are not directly affected by  $\alpha$ ,  $\pi(\alpha)$  is increasing in  $\alpha$ . Since at a change in layers profits have to be identical,  $\pi(\alpha)$  is continuous. ■

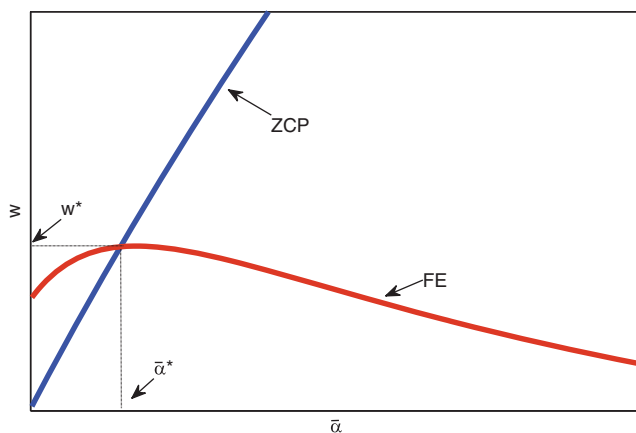


FIGURE B1

Equilibrium  $w^*$  and  $\bar{\alpha}^*$

*Proof of Proposition 6.* Define the following functions,

$$\Delta^{ZCP}(\alpha, w) = \pi(\alpha) = 0,$$

$$\Delta^{FE}(\alpha, w) = \int_{\alpha}^{\infty} \pi(x)g(x)dx - \delta wf^E = 0.$$

An example with the properties that we prove is presented in Figure B1.

First we characterize the zero cut-off profit (ZCP) schedule. Note that the schedule starts at  $(0, 0)$  because  $\Delta^{ZCP}(0, 0) = 0$ . Then note that it is increasing in the  $(\alpha, w)$  plane, since  $\frac{\partial w}{\partial \alpha} = -\frac{\Delta_{\alpha}^{ZCP}}{\Delta_w^{ZCP}} > 0$ , which follows from  $\Delta_{\alpha}^{ZCP} = \frac{\partial \pi(\alpha)}{\partial \alpha} = \left(\frac{\partial r(\alpha)}{\partial \alpha} - MC(\alpha; w)\right) \frac{\partial q(\alpha)}{\partial \alpha} + \frac{\partial r(\alpha)}{\partial \alpha} = \frac{1}{\sigma} \frac{r(\alpha)}{\alpha} > 0$ , and  $\Delta_w^{ZCP} = \frac{\partial \pi(\alpha)}{\partial w} = \frac{1-\sigma}{\sigma} \frac{r(\alpha, w)}{w} < 0$ .

Implicitly define  $\tilde{\alpha}(w)$  by  $\Delta^{ZCP}(\tilde{\alpha}(w), w) = 0$ . Then we have proven that  $\tilde{\alpha}(0) = 0$  and that  $\tilde{\alpha}(w)$  is strictly increasing.

Let's focus now on the  $\Delta^{FE}(\alpha, w)$  schedule. Note that it starts at  $(0, \hat{w})$  where  $\hat{w} > 0$ , since  $\Delta^{FE}(0, \hat{w}) = 0$ , where  $\hat{w}$  is defined as

$$\hat{w} = \left(\frac{\sigma-1}{\sigma}\right)^{\frac{1}{1-\sigma}} \left(\frac{\tilde{C} + (f + \delta f^E)}{\tilde{N}(\frac{\sigma}{\sigma-1})^{-\sigma} \widetilde{MC}}\right)^{\frac{1}{1-\sigma}} > 0,$$

and where  $\tilde{C} = \int_0^{\infty} \frac{C(\alpha; w)}{w} g(\alpha) d\alpha$  and  $\widetilde{MC} = \int_0^{\infty} \alpha \left(\frac{MC(\alpha; w)}{w}\right)^{1-\sigma} g(\alpha) d\alpha$  are constants.

Also note that in the limit, as  $\alpha \rightarrow \infty, w \rightarrow 0$ , so  $\Delta^{FE}(\infty, 0) = 0$ . This implies, by the intermediate value theorem, that there exists an  $\tilde{\alpha}^*$  and  $w^*$  such that  $\Delta^{ZCP}(\tilde{\alpha}^*, w^*) = 0$  and  $\Delta^{FE}(\tilde{\alpha}^*, w^*) = 0$ .

We now show that the  $\Delta^{FE}(\alpha, w)$  schedule increases until it reaches the  $\Delta^{ZCP}(\alpha, w)$  schedule and it declines thereafter. Note that

$$\frac{\partial w}{\partial \alpha} = -\frac{\Delta_{\alpha}^{FE}}{\Delta_w^{FE}} > 0 \quad \text{if } \alpha < \tilde{\alpha}(w)$$

$$< 0 \quad \text{if } \alpha > \tilde{\alpha}(w),$$

which follows from

$$\Delta_{\alpha}^{FE} = -\pi(\alpha)g(\alpha) > 0 \quad \text{if } \alpha < \tilde{\alpha}(w)$$

$$< 0 \quad \text{if } \alpha > \tilde{\alpha}(w),$$

and

$$\Delta_w^{FE} = \int_{\alpha}^{\infty} \frac{\partial \pi(x)}{\partial w} g(x) dx - \delta f^E = \frac{1-\sigma}{\sigma} \int_{\alpha}^{\infty} \frac{r(x)}{w} g(x) dx < 0.$$

Hence, the pair  $(\bar{\alpha}^*, w^*)$  that solves  $\Delta^{ZCP}(\bar{\alpha}^*, w^*) = 0$  and  $\Delta^{FE}(\bar{\alpha}^*, w^*) = 0$  is unique. Equations (18) and (19) can then be directly solved to obtain  $M$  and  $R$ . This implies that an equilibrium of this economy exists and is unique. ■

*Proof of Proposition 7.* Let  $(\alpha^*, w^*)$ , be the solution under  $\tilde{N}$ . We now show that  $(\alpha^*, (\frac{\tilde{N}'}{\tilde{N}})^{\frac{1}{\sigma-1}} w^*)$  is the solution under  $\tilde{N}'$ , and  $q(\alpha)$  are the same under  $\tilde{N}$  than under  $\tilde{N}'$ . Note that if  $(\alpha^*, w^*)$  are the solution under  $\tilde{N}$ , then

$$\Delta^{ZCP}(\alpha^*, w^*, \tilde{N}) = q(\alpha^*)^{1-\frac{1}{\sigma}} (\alpha^* w^* \tilde{N})^{\frac{1}{\sigma}} - w^* C(\alpha^*) - w^* f = 0,$$

which implies

$$\frac{(\tilde{N})^{\frac{1}{\sigma}}}{w^{*1-\frac{1}{\sigma}}} = \frac{C(\alpha^*) + f}{q(\alpha^*)^{1-\frac{1}{\sigma}} (\alpha^*)^{\frac{1}{\sigma}}},$$

and

$$\begin{aligned} \Delta^{FE}(\alpha^*, w^*, \tilde{N}) &= \int_{\alpha^*}^{\infty} \left( q(\alpha)^{1-\frac{1}{\sigma}} (\alpha w^* \tilde{N})^{\frac{1}{\sigma}} - w^* C(\alpha) - w^* f \right) g(\alpha) d\alpha \\ &\quad - \delta w^* f^E = 0, \end{aligned}$$

which implies

$$\frac{(\tilde{N})^{\frac{1}{\sigma}}}{w^{*1-\frac{1}{\sigma}}} = \frac{f[1 - G(\alpha^*)] + \delta f^E + \int_{\alpha^*}^{\infty} C(\alpha) g(\alpha) d\alpha}{\int_{\alpha^*}^{\infty} q(\alpha)^{1-\frac{1}{\sigma}} (\alpha)^{\frac{1}{\sigma}} g(\alpha) d\alpha}.$$

Now change  $\tilde{N}$  to  $\tilde{N}'$ . Let  $q(\alpha)$  be the same under  $\tilde{N}'$  as under  $\tilde{N}$ . Set the same  $\alpha^*$  and find  $w$  such that  $\Delta^{ZCP}(\alpha^*, w, \tilde{N}') = 0$  where  $\tilde{N}' \neq \tilde{N}$ . Note that  $w^{*'} = \left(\frac{\tilde{N}'}{\tilde{N}}\right)^{\frac{1}{\sigma-1}} w^*$  solves the ZCP condition,

$$\begin{aligned} \Delta^{ZCP}(\alpha^*, w^{*'}, \tilde{N}') &= \left( \frac{q(\alpha^*)^{1-\frac{1}{\sigma}} (\alpha^* w^{*'} \tilde{N}')^{\frac{1}{\sigma}} P^{\frac{\sigma-1}{\sigma}}}{w^{*'}} - C(\alpha^*) - f \right) w^{*'}, \\ &= \Delta^{ZCP}(\alpha^*, w^*, \tilde{N}) \frac{w^{*'}}{w^*} = 0. \end{aligned}$$

Note that this solution also solves the free entry condition,

$$\begin{aligned} \int_{\alpha^*}^{\infty} \pi(\alpha, w^{*'}) g(\alpha) d\alpha &= \delta w^{*'} f^E, \\ \int_{\alpha^*}^{\infty} \frac{\pi(\alpha, w^{*'})}{w^{*'}} g(\alpha) d\alpha &= \delta f^E + f(1 - G(\alpha^*)), \\ \int_{\alpha^*}^{\infty} \left( q(\alpha)^{1-\frac{1}{\sigma}} (\alpha)^{\frac{1}{\sigma}} P^{\frac{\sigma-1}{\sigma}} \frac{(\tilde{N}')^{\frac{1}{\sigma}}}{w^{*'} 1-\frac{1}{\sigma}} - c(\alpha) \right) g(\alpha) d\alpha &= \delta f^E + f(1 - G(\alpha^*)), \\ \frac{(\tilde{N}')^{\frac{1}{\sigma}}}{w^{*'} 1-\frac{1}{\sigma}} &= \frac{\delta f^E + f(1 - G(\alpha^*)) + \int_{\alpha^*}^{\infty} c(\alpha) g(\alpha) d\alpha}{\int_{\alpha^*}^{\infty} q(\alpha)^{1-\frac{1}{\sigma}} (\alpha)^{\frac{1}{\sigma}} P^{\frac{\sigma-1}{\sigma}} g(\alpha) d\alpha} = \frac{(\tilde{N}')^{\frac{1}{\sigma}}}{w^{*'} 1-\frac{1}{\sigma}}. \end{aligned}$$

We still need to show that our guess does not change  $q(\alpha)$ . Use equation (14) to conclude that for any  $\alpha$

$$\begin{aligned} q(\alpha) &= \alpha(w^*)^{1-\sigma} \tilde{N} \left( \frac{\sigma}{\sigma-1} MC(q(\alpha); 1) \right)^{-\sigma} \\ &= \alpha(w^{*'})^{1-\sigma} \tilde{N}' \left( \frac{\sigma}{\sigma-1} MC(q(\alpha); 1) \right)^{-\sigma}. \end{aligned}$$

Hence  $q(\alpha)$  remains unchanged after the change in  $\tilde{N}$ . Equation (13) implies that given  $q(\alpha)$ , prices are proportional to wages. Then, an increase in the price with  $\tilde{N}$ , together with constant quantity, increases revenues. Because revenues change proportionally to the wage, and so do costs, operating profits change proportionally as well. ■

*Proof of Proposition 8.* Consider first the case in which  $f_{DF} = f_{FD} = 0$ . Then, all firms export in equilibrium. In this case a trade liberalization is equivalent to an increase in the population size of a closed economy. Proposition 7 then implies that wages increase but the selection threshold and quantities produced remain unchanged. Now consider an economy in which  $f_{DF} = f_{FD} > 0$  and some firms do not export. Since not all firms expand their market,  $w_D$  then increases by less than  $\left( \frac{(\tau_{ij} \tilde{N}_j + \tilde{N}_i)}{\tilde{N}_i} \right)^{\frac{1}{\sigma-1}}$  (the increase that would result if all firms were to export), because nonexporters are not selling to foreign firms and they are not demanding labor accordingly. As a result the effect of wages and population sizes that affect production

according to equations (23), (24), and (25) goes from  $\bar{w}_D^{1-\sigma} \tilde{N}_i$  in the closed economy to  $w_D^{1-\sigma} (\tau_{ij} \tilde{N}_j + \tilde{N}_i)$  (where  $\bar{w}_D$  denotes the wage in the closed economy). The new term is larger in the open economy, since wages adjust less than what Proposition 7 predicts. This implies that the total production of exporters has to necessarily increase. Nonexporters do not get access to the foreign market so the relevant effect on sales is necessarily decreasing in  $w_D$  by equation (22). Toward a contradiction, suppose that opening to trade leads to a decline in  $w_D$ . Then, by the reasoning above, production of nonexporters and exporters increases, which by Proposition 1 increases their labor demands since by Proposition 3 the cost function is homogeneous of degree 1. This contradicts the labor market clearing condition in equation (26). Hence,  $w_D$  strictly increases, and the production of nonexporters decreases and that of exporters increases with trade liberalization. Of course, because our choice of a domestic country is arbitrary, this applies to the other country as well. ■

YALE UNIVERSITY

PRINCETON UNIVERSITY

## REFERENCES

- Antràs, Pol, Luis Garicano, and Esteban Rossi-Hansberg, "Offshoring in a Knowledge Economy," *Quarterly Journal of Economics*, 121 (2006), 31–77.
- , "Organizing Offshoring: Middle Managers and Communication Costs," In *The Organization of Firms in a Global Economy*, Elhanan Helpman, Dalia Marin, and Thierry Verdier, ed. (Cambridge, MA: Harvard University Press, 2008).
- Antràs, Pol, and Esteban Rossi-Hansberg, "Organizations and Trade," *Annual Review of Economics*, 1 (2009), 43–64.
- Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models Same Old Gains?" *American Economic Review*, 102 (2012), 94–130.
- Atkeson, Andrew, and Ariel Burstein, "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 118 (2010), 433–484.
- Aw, Bee Yan, Sukkyn Chung, and Mark J. Roberts, "Productivity and Turnover in the Export Market: Micro Evidence from Taiwan (China) and the Republic of Korea," *World Bank Economic Review*, 14 (2000), 65–90.
- Aw, Bee Yan, Mark J. Roberts, and Tor Winston, "Export Market Participation, Investments in R&D and Worker Training, and the Evolution of Firm Productivity," *World Economy*, 14 (2007), 83–104.
- Aw, Bee Yan, Mark J. Roberts, and Daniel Y. Xu, "R&D Investments, Exporting, and the Evolution of Firm Productivity," *American Economic Review: Papers and Proceedings*, 98 (2008), 451–456.
- Baldwin, John R., and Wulong Gu, "Participation in Export Markets and Productivity Performance in Canadian Manufacturing," *Canadian Journal of Economics*, 36 (2003), 634–657.
- Bernard, Andrew B., and Bradford J. Jensen, "Exceptional Exporter Performance: Cause Effect, or Both?" *Journal of International Economics*, 47 (1999), 1–25.

- , “Exporting and Productivity in the U.S,” *Oxford Review of Economic Policy*, 20 (2004), 343–357.
- Bernard, Andrew B., Jonathan Eaton, Bradford J. Jensen, and Samuel Kortum, “Plants and Productivity in International Trade,” *American Economic Review*, 93 (2003), 1268–1290.
- Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott, “Firms in International Trade,” *Journal of Economic Perspectives*, 21 (2007), 105–130.
- Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott, “Multiple-Product Firms and Product Switching,” *American Economic Review*, 100 (2010), 70–97.
- Bloom, Nicholas, and John Van Reenen, “Measuring and Explaining Management Practices across Firms and Countries,” *Quarterly Journal of Economics*, 122 (2007), 1351–1408.
- Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John van Reenen, “The Distinct Effects of Information Technology and Communication Technology on Firm Organization,” (Mimeo, London School of Economics, 2010).
- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, “Does Management Matter: Evidence from India,” (Mimeo, Stanford University, 2011).
- Bustos, Paula, “Trade Liberalization, Exports and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms,” *American Economic Review*, 101 (2011), 304–350.
- Clerides, Sofronis, Saul Lach, and James Tybout, “Is Learning-by-Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico and Morocco,” *Quarterly Journal of Economics*, 113 (1998), 903–947.
- Conconi, Paola, Patrick Legros, and Andrew F. Newman, “Trade Liberalization and Organizational Change,” (CEPR Discussion Paper No. DP7028, 2008).
- Costantini, James A., and Marc J. Melitz “The Dynamics of Firm-Level Adjustment to Trade Liberalization,” In *The Organization of Firms in a Global Economy*, Elhanan Helpman, Dalia Marin, and Thierry Verdier, ed. (Cambridge, MA: Harvard University Press, 2008).
- Delgado, Miguel A., Jose C. Fariñas, and Sonia Ruano, “Firm Productivity and Export Markets: A Non-parametric Approach,” *Journal of International Economics*, 57 (2002), 397–422.
- De Loecker, Jan, “Do Exports Generate Higher Productivity? Evidence from Slovenia,” *Journal of International Economics*, 73 (2007), 69–98.
- , “Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity,” *Econometrica*, 79 (2011), 1407–1451.
- Eaton, Jonathan, and Samuel Kortum, “Technology, Geography, and Trade,” *Econometrica*, 70 (2002), 1741–1779.
- Fernandes, Ana M., and Alberto E. Isgut, “Learning-by-Exporting Effects: Are They for Real?” (Mimeo, World Bank and Institute for Competitiveness and Prosperity, 2006).
- Garicano, Luis, “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 108 (2000), 874–904.
- Garicano, Luis, and Esteban Rossi-Hansberg, “Inequality and the Organization of Knowledge,” *American Economic Review*, 94 (2004), 197–202.
- , “Organization and Inequality in a Knowledge Economy,” *Quarterly Journal of Economics*, 121 (2006), 1383–1435.
- , “Organizing Growth,” *Journal of Economic Theory*, 147 (2012), 593–629.
- Ghironi, Fabio, and Marc J. Melitz, “International Trade and Macroeconomic Dynamics with Heterogeneous Firms,” *Quarterly Journal of Economics*, 120 (2005), 865–915.
- Guadalupe, Maria, and Julie M. Wulf, “The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies,” *American Economic Journal: Applied Economics*, 2 (2010), 105–127.
- Hallward-Driemeier, Mary, Giuseppe Iarossi, and Kenneth L. Sokoloff, “Exports and Manufacturing Productivity in East Asia: A Comparative Analysis with Firm-Level Data,” (NBER Working Paper No. w8894, 2002).

- Krugman, Paul, "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70 (1980), 950–959.
- Lileeva, Alla, and Daniel Treffer, "Improved Access to Foreign Markets Raises Plant-Level Productivity... For Some Plants," *Quarterly Journal of Economics*, 125 (2010), 1051–1099.
- Marin, Dalia, and Thierry Verdier "Competing in Organizations: Firm Heterogeneity and International Trade," In *The Organization of Firms in a Global Economy*, Elhanan Helpman, Dalia Marin, and Thierry Verdier, ed. (Cambridge, MA: Harvard University Press, 2008).
- , "Corporate Hierarchies and International Trade: Theory and Evidence," (Mimeo, University of Munich, 2010).
- Mayer, Thierry, Marc J. Melitz, and Gianmarco I.P. Ottaviano, "Market Size, Competition and the Product Mix of Exporters," (NBER Working Paper No. w16959, 2011).
- Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71 (2003), 1695–1725.
- Park, Albert, Dean Yang, Xinzhen Shi, and Yuan Jiang, "Exporting and Firm Performance: Chinese Exporters and the Asian Financial Crisis," (NBER Working Paper No. w14632, 2009).
- Rubini, Loris, "Innovation and the Elasticity of Trade Volumes to Tariff Reductions," (EFIGE Working Paper No. 31, 2011).
- Van Biesebroeck, Johannes, "Exporting Raises Productivity in Sub-Saharan African Manufacturing Firms," *Journal of International Economics*, 67 (2004), 373–391.
- Viner, Jacob, "Cost Curves and Supply Curves," *Zeitschrift für Nationalökonomie*, 3 (1931), 23–46.
- Yeaple, Stephen R., "A Simple Model of Firm Heterogeneity, International Trade, and Wages," *Journal of International Economics*, 65 (2005), 1–20

This page intentionally left blank