Input Sourcing and Multinational Production^{*}

Stefania Garetto[†]

Boston University

April 5, 2010

Abstract

A large portion of world trade happens within firms' boundaries. This paper proposes a general equilibrium framework where firms decide whether to outsource or to integrate input manufacturing, and whether to do so domestically or abroad. By outsourcing, firms benefit from suppliers' good technologies, but pay mark-up prices. By sourcing intrafirm, they save on mark-ups and match domestic productivity with possibly lower foreign wages. Multinational corporations arise endogenously when firms integrate production abroad. The model is calibrated to match aggregate U.S. trade data, and used to quantify the welfare gains from vertical FDI and intrafirm trade. The current gains are about 1% of consumption per capita, and further FDI liberalization can increase them substantially.

Keywords: International trade, intrafirm trade, multinational firms, vertical FDI **JEL Classification** : F12, F23, L11

^{*}I am indebted to Robert E. Lucas Jr. for motivating me and helping me in this work. This paper has benefited from comments by Fernando Alvarez, Christian Broda, Thomas Chaney, José Fillat, Gene Grossman, François Gourio, Kenneth Judd, Marc Melitz, Dilip Mookherjee, Natalia Ramondo, Esteban Rossi-Hansberg, Leena Rudanko, Nancy Stokey, Mathis Wagner, and seminar participants at the University of Chicago, Arizona State University, Boston University, the Chicago Fed, the Federal Reserve Board, the New York Fed, New York University, Penn State University, Princeton University, UC Berkeley (Haas), the University of Iowa, the University of Toronto, the University of Virginia, Yale University, the 2008 SED Meeting, and the 2008 Winter ITI NBER Meeting. All errors are mine.

[†]Contact: Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215. E-mail: garettos@bu.edu.

1 Introduction

Globalization has expanded the scope of trade, as trade in finished products is being gradually outpaced by trade in intermediates, taking place both within and across the boundaries of the firm. Many studies document the growth of multistage production, in which plants in different locations contribute to the creation of value added through processing and assembly.¹ A good example is the vertical production chain of the Barbie doll quoted by Feenstra (1998), in which U.S.-produced molds cross six Asian countries before being shipped back to the U.S. where the dolls are sold. Multinational corporations play a large role in this scenario, as a substantial share of offshore production happens within their boundaries. Bernard, Jensen, and Schott (2009) report that in the year 2000 almost 50% of U.S. imports and about 30% of exports happened within firms' boundaries.

In this paper I provide a new theoretical framework to think about cost-driven, vertical multinational production and the associated flows of intrafirm trade. Firms need to acquire a set of tradeable inputs in order to produce a non-tradeable consumption good. Input production can be outsourced to unaffiliated suppliers, generating volumes of trade in intermediates, or can be integrated (or insourced) by the firm itself. When a firm decides to insource input production, it sets up a new plant, possibly in another country where factor costs are lower. This choice gives rise endogenously to the creation of multinational firms, and to vertical foreign direct investment (henceforth, FDI) in the form of integrated production abroad.² I assume that investment in an integrated facility is always associated with ownership, so that when inputs produced offshore are shipped back to the parent, we observe flows of intrafirm trade.

The novelty of this approach is the fact that the optimal sourcing strategy is achieved as a market equilibrium, while the most recent literature on this topic (notably Antràs (2003) and Antràs and Helpman (2004)) presented it as the outcome of a contracting problem. In

¹See Campa and Goldberg (1997), Yeats (2001), Hummels, Ishii, and Yi (2001), Hummels, Rapoport, and Yi (1998), and Hanson et al. (2001).

²In his survey of the literature on trade and multinational production, Helpman (2006) defines vertical FDI as "(activity done through) subsidiaries that add value to products that are not destined [...] for the host country market".

my model firms simply choose the sourcing options and the locations that minimize their production costs. Firms are heterogeneous in productivity and in the type of technology they use. Intermediate goods producers (or suppliers) have an adaptable technology with which they produce inputs that they can sell to final good producers worldwide. Final good producers are endowed with two types of technologies: a homogeneous technology to produce the consumption good and a set of heterogeneous, non adaptable technologies that they can use anywhere in the world to produce their own inputs in affiliate plants. Final good producers can either buy inputs from the suppliers or integrate input production using their "in-house" technologies. When they decide to integrate production abroad, they become the parents of a multinational corporation. Offshore integrated production takes the form of vertical FDI, and generates flows of intrafirm trade when the inputs produced offshore are shipped back to the parent.

The model delivers endogenous organizational choices without relying on the propertyrights model of the boundaries of the firm. The driving forces behind the sourcing choice are technological heterogeneity and the implications of imperfect competition on prices. By outsourcing from a supplier, a final good producer can have access to a potentially better technology, but has to pay a price which is augmented by a mark-up. On the other hand, by integrating, firms have to use their own technology, but save on the mark-ups charged by the suppliers: intrafirm trade happens between a firm and itself, and is priced at marginal cost.³ Moreover, when intrafirm sourcing takes place abroad, the parent firm is able to transfer (at least partially) its technology to the destination county, and match it with the possibly lower labor costs of the location chosen.⁴

Integrated firms increase competition in the input market: suppliers' market power and

³This way of modeling the differences between intrafirm and arm's length pricing is consistent with recent transaction-level evidence. Bernard, Jensen, and Schott (2006) document the existence of a large gap between the prices associated with arm's length transactions and the transfer prices associated with intrafirm transactions: by comparing related-party sales and arm's length sales of the same good by the same firm to the same destination market, they find that arm's length prices are significantly higher than intrafirm prices. Neiman (2010) also finds that arm's length prices are more responsive than intrafirm prices to price changes of competing firms, suggesting that intrafirm prices are set differently from arm's length prices.

⁴Grossman and Rossi-Hansberg (2008) model offshoring of intermediate goods (or "tasks") as motivated by factor cost differences. Hanson et al. (2005) document the importance of low-cost locations in vertical production networks.

prices are reduced by the possibility of integrated production. This effect is reinforced when integration takes the form of multinational production, as lower factor costs in FDI host countries make the integration option more attractive and reduce arm's length prices further. Because of the sensitivity of prices to the possibility of integration, FDI liberalization makes both outsourcing and FDI more attractive, increases competition and lowers prices.

On aggregate, the theory predicts that firms outsource mostly from suppliers located in large countries with low factor costs. Volumes of arm's length and intrafirm trade increase with cross-country heterogeneity and, while arm's length trade occurs also between identical countries, a certain degree of heterogeneity is necessary to give rise to vertical FDI and intrafirm trade.⁵ The dispersion of the cost distributions across firms also affects the sourcing pattern: the share of intrafirm transactions is larger the higher the productivity dispersion of the suppliers.

The general equilibrium structure of the model makes it suitable for welfare analysis. I calibrate the exogenous parameters (technology, endowments, and trade costs) to match aggregate moments of U.S. data, and use the calibrated economy to quantify the welfare gains arising from vertical multinational production, distinguishing them from the gains from arm's length trade.⁶ Firms' organizational choices depend on technological differences (through factor costs) and market structure (through price adjustments). I decompose the gains from multinational production to quantify the relative importance of these components: the effect of price adjustments is more relevant in scenarios where goods are more differentiated and suppliers have a significant level of market power, while the effect of productivity differences is quantitatively important across specifications. The welfare gains for the U.S. economy arising from vertical multinational production are currently about 1% of consumption per capita, and further FDI liberalization can increase them substantially: the model predicts that a 50% drop in the calibrated barrier to FDI could generate a gain of about 6% of

⁵Nocke and Yeaple (2008) obtain similar predictions when modeling the choice of vertical greenfield investment *versus* mergers and acquisitions, and report supporting empirical evidence.

⁶The model excludes horizontal FDI, *i.e.* the establishment of offshore production to serve foreign markets, and export platforms, *i.e.* the establishment of offshore production to serve third markets. Ramondo and Rodríguez-Clare (2009) present a quantitative model that captures the gains from these alternative forms of multinational production.

consumption per capita.

The rationale behind the existence of multinational firms is similar to Helpman (1984), Helpman (1985), where multinationals emerge to exploit factor cost differences across countries. In Helpman's papers, firms choose the location of their activities to minimize production costs. The incentive to become multinational arises from the existence of an immaterial factor of production that may serve product lines without being located in their plants.⁷ Similarly, I assume that firms can imperfectly transfer their productivity when they decide to integrate production abroad. In addition, the model I propose generalizes Helpman's idea to a world with heterogeneous firms, trade costs and potentially many asymmetric countries.

More recently, Antràs (2003) and Antràs and Helpman (2004) modeled the joint choice of location and organizational structure by merging existing models of trade with a contractbased theory of the firm.⁸ Their approach has the advantage of analyzing separately the two choices, and matches qualitative features of the data on intrafirm trade. My model provides a complementary analysis that emphasizes the role of technological heterogeneity and market structure in explaining organizational choices. Moreover, imperfect competition allows me to analyze optimal pricing and the interactions between pricing and organizational choices, which is absent in Antràs and Helpman's work.⁹ Finally, the general equilibrium structure of the model makes it amenable to a calibration exercise in which I evaluate the magnitude of the welfare effects deriving from multinational production and intrafirm trade.

The rest of the paper is organized as follows. Section 2 lays out the closed economy model, to isolate the choice between outsourcing and integration, without considering the location choice. Section 3 extends the model to an open economy with an arbitrary number of countries, and illustrates the properties of the general equilibrium for the two-country case. Section 4 shows the dependence of aggregate volumes of trade and FDI on the model's

⁷The idea of modeling multinational production through the existence of an immaterial factor is present also in Markusen (1984). In his setup, multinational corporations arise to increase efficiency by avoiding duplication of the control input, but this may come at the expense of higher market power and higher prices. Conversely, the structure of competition in my model implies that the presence of multinationals increases competition and reduces prices.

⁸See Helpman and Krugman (1985), Melitz (2003), Helpman, Melitz, and Yeaple (2004), Grossman and Helpman (2002), Grossman and Helpman (2004).

⁹Incomplete contracts imply that there are no explicit prices for the goods to be sourced.

parameters. Section 5 contains the calibration and the computation of the welfare gains from multinational production and intrafirm trade. Section 6 concludes.

2 The Model: the Closed Economy

In this section I present the model for a closed economy. I extend Eaton and Kortum (2002) and Alvarez and Lucas (2007) to incorporate imperfect competition and the choice of the sourcing option. Given the structure of technology heterogeneity, the organization choice simply adds one dimension to the description of goods as vectors of technology draws. This feature allows to preserve most of the tractability of the theory and to extend it to explain both trade and multinational production.

2.1 Production Technologies in a Two-Sector Economy

The economy is organized in two sectors. There is an intermediate goods sector, where a continuum of differentiated goods is produced using labor as the only input, and a final good sector, where intermediate goods and labor are combined in the production of a unique, homogeneous consumption good.

Accordingly, there are two types of producers in this economy: final good producers (or buyers) and intermediate goods producers (or suppliers). There is a continuum of intermediate goods producers, who differ in their labor productivity levels and operate in a monopolistically competitive fashion. They produce differentiated goods that are imperfect substitutes from the perspective of the buyers. Each supplier's productivity level is denoted by z, the number of units of labor needed to produce one unit of the good. z is a random draw from a common density $\psi(z)$ defined on \Re_+ . Each supplier can sell his good to any buyer, without having to incur any cost to adapt it to the buyer's specific production process. Suppliers cannot discriminate across buyers, and each supplier charges a price p(z), which depends on his cost draw, to all buyers in the market.

The final good is produced by a large number of identical producers, operating in a

perfectly competitive market. They all produce the same, homogeneous consumption good using labor and intermediate goods as inputs. For each input, a final good producer has two possible sourcing options: he can either produce it in-house or buy it from a supplier. When he decides to integrate production, his technology allows him to produce only for his own product line.¹⁰

The sourcing decision involves comparing the costs of the two options: the in-house cost of production and the outside price charged by the supplier. For each input, the final good producer has an in-house unit labor requirement x, which is a random draw from a density $\phi(x)$ defined on \Re_+ and indicates the number of units of labor needed to internalize production of one unit of input.¹¹ All the final good producers draw the in-house costs from the same distribution $\phi(\cdot)$, but they can have different cost draws for the same input. In the closed economy we can normalize the wage to one, so the unit labor requirement x is equal to the unit cost of in-house production. Hence each final good producer observes a set of input prices $\{p(z)\}$, draws a set of in-house labor requirements $\{x\}$ and then – for each intermediate good – he chooses whether he wants to buy or produce. Obviously, he buys those inputs for which the selling price p(z) is lower than the in-house unit cost of production x.

2.2 The Final Good Producer's Problem

In this framework, goods are differentiated by their unit labor requirements. I identify each intermediate good with the pair of unit labor requirements that the two types of agents need for its production: (x, z) denotes a good for which the potential buyer has unit cost x and the supplier has unit cost z and charges a price p(z). Accordingly, q(x, z) denotes the quantity produced of good (x, z). The final good producer minimizes the total cost of input

¹⁰In principle, the final good producer could aquire an adaptable technology (at some cost) to enter the intermediate goods' market and sell inputs to other final good producers. I assume that that cost is too large to be covered by the expected profits.

¹¹The cost distributions $\phi(\cdot)$ and $\psi(\cdot)$ could be equal. In the model, I use different notations to distinguish the separate effects of the two distributions on the equilibrium. In the welfare analysis, I set $\phi = \psi$ due to the scarcity of data allowing to distinguish quantitatively the two distributions. However, it is important to notice that the assumption of a continuum of goods and the unbounded support of the distributions assure that both sourcing options coexist in equilibrium regardless of whether $\phi = \psi$.

sourcing:

$$\min_{q(x,z)} \qquad \int_0^\infty \int_0^\infty \min\{x, p(z)\}q(x, z)\phi(x)\psi(z)dxdz \tag{1}$$
$$s.t. \qquad \left[\int_0^\infty \int_0^\infty q(x, z)^{1-1/\eta}\phi(x)\psi(z)dxdz\right]^{\eta/(\eta-1)} \ge q$$

where $\eta > 1$ is the elasticity of substitution across inputs, and q denotes an aggregate of intermediate goods, which the final good producer takes as given and will be determined by equilibrium conditions in the final good market.¹² The outside prices p(z) are also taken as given.

Let $B^{I} = \{(x, z) : x \leq p(z)\}$ be the set of goods that the final good producer decides to internalize and $q^{I}(x, z)$ be the solution of (1) in B^{I} . Similarly, let $B^{T} = \{(x, z) : x \geq p(z)\}$ be the set of goods that the final good producer decides to outsource and $q^{T}(x, z)$ be the solution of (1) in B^{T} . Hence:

$$q^{I}(x,z) \equiv q^{I}(x) = x^{-\eta}p^{\eta}q \qquad \forall (x,z) \in B^{I}$$
 (2)

$$q^{T}(x,z) \equiv q^{T}(p(z)) = [p(z)]^{-\eta} p^{\eta} q \quad \forall (x,z) \in B^{T}.$$
(3)

The term p is the aggregate price index for this economy:

$$p = \left[p_I^{1-\eta} + p_T^{1-\eta} \right]^{1/(1-\eta)}, \tag{4}$$

and:

$$p_{I} = \left[\int_{0}^{\infty} \int_{0}^{p(z)} x^{1-\eta} \phi(x) \psi(z) dx dz \right]^{1/(1-\eta)}$$
(5)

$$p_T = \left[\int_0^\infty p(z)^{1-\eta} \left[1 - \Phi(p(z)) \right] \psi(z) dz \right]^{1/(1-\eta)}.$$
 (6)

¹²The assumption of a continuum of goods implies that – by the law of large numbers – the aggregate q will be the same across final good producers even if they have different cost draws for each of the goods.

2.3 The Supplier's Problem

A supplier with cost draw z chooses the profit-maximizing price p(z) by trading off the higher per-unit profits given by a higher price with the possibility of capturing a larger mass of buyers with a relatively lower price:¹³

$$\max_{p(z)} [p(z) - z] \int_{p(z)}^{\infty} q^T(p(z))\phi(x)dx$$
(7)

where z is the supplier's unit cost of production, $q^T(p(z))$ is given by (3), and $\int_{p(z)}^{\infty} \phi(x) dx$ is the mass of buyers that decide to buy the good at price p(z). The first order condition for this problem is:

$$[p(z) - z] \left[q^{T}(p(z))\phi(p(z)) - \frac{\partial q^{T}(p(z))}{\partial p(z)} [1 - \Phi(p(z))] \right] = q^{T}(p(z)) \int_{p(z)}^{\infty} \phi(x) dx$$
(8)

where $\Phi(\cdot)$ denotes the c.d.f. of the in-house cost distribution.

Equation (8) summarizes the supplier's trade-off. For a given level of sales, the gain from increasing the mark-up over the marginal cost ([p(z) - z]) must be counterbalanced by the sum of the losses on both the extensive and the intensive margin. If the supplier raises the price, he is going to lose the marginal buyers (the extensive margin, captured by the term $q^T(p(z))\phi(p(z))$) and he is going to sell lower quantities to the remaining buyers (the intensive margin, captured by the term $-\frac{\partial q^T(p(z))}{\partial p(z)}[1 - \Phi(p(z))]$). Using (3), the first order condition reduces to:

$$p(z) = \frac{\hat{\eta}}{\hat{\eta} - 1} z \tag{9}$$

where:

$$\hat{\eta} \equiv \eta + \frac{\phi(p(z))}{1 - \Phi(p(z))} p(z).$$
(10)

Equation (9) shows how the buyers' possibility of integration generates a significant departure from the constant mark-up pricing rule usually implied by CES preferences associated with monopolistic competition. If integration were not possible, $\hat{\eta} = \eta$, and the optimal

¹³I rule out possibilities of price discrimination or bargaining with individual buyers.

pricing rule would depend only on the standard intensive margin, whose magnitude depends on the elasticity of substitution. The possibility of integration introduces the extra term $\frac{\phi(p(z))}{1-\Phi(p(z))}p(z)$ (the extensive margin), which depends on the hazard rate, or the probability that – after an infinitesimal price increase – the buyer switches to integrating, conditional on buying before the price increase. Since this term is non-negative, (9) implies that prices are strictly below the standard constant mark-up ones: $p(z) < \frac{\eta}{\eta-1}z$, $\forall z > 0$. Moreover, as the term $\frac{\phi(p(z))}{1-\Phi(p(z))}p(z)$ depends on the supplier's cost draw z, mark-ups are endogenous and variable across suppliers. For a wide range of parameterizations, the term $\frac{\phi(p(z))}{1-\Phi(p(z))}p(z)$ is increasing in the supplier's unit cost z, indicating that more productive suppliers charge lower prices but higher mark-ups than less productive suppliers.¹⁴

To illustrate graphically the behavior of (9), Figure 1 plots prices p(z) and mark-ups p(z)/z as functions of the unit cost z for exponentially distributed in-house costs ($\phi(x) = \lambda e^{-\lambda x}$, for $x \ge 0$).¹⁵

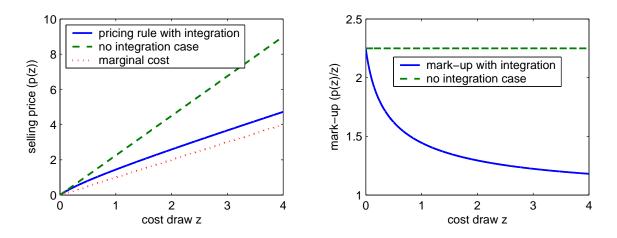


Figure 1: Prices and mark-ups, closed economy $(\eta = 1.8, \lambda = 1)$.

In the left panel, the dotted line is the supplier's marginal cost, the solid line is the price

¹⁴The result of endogenous mark-ups holds for any functional specification of the cost distribution $\phi(x)$, except for the Pareto, for which the elasticity of demand is constant and hence mark-ups are constant too (but lower than in the model without integration). Garetto (2009) provides a detailed treatment of the properties of equation (9), including conditions for monotonicity of the term $\frac{\phi(p(z))}{1-\Phi(p(z))}p(z)$ and the implications of this pricing behavior for incomplete pass-through and pricing-to-market.

¹⁵If in-house costs are exponentially distributed, (9) reduces to: $\lambda p(z)^2 + (\eta - 1 - \lambda z)p(z) - \eta z = 0$, which implies existence and uniqueness of a positive and monotonic solution $p(z; \lambda, \eta)$ such that $\frac{\partial p(z)}{\partial \lambda} < 0$, $\frac{\partial p(z)}{\partial \eta} < 0$.

p(z) (solution of equation (9)), and the dashed line is the constant mark-up pricing rule of the model without possibility of integration. Comparing the pricing strategies of the model with integration and of the standard model (with no possibility of integration) is evident that the integration option significantly reduces the profit margins of the suppliers. In the right panel of Figure 1, the solid line is the producer's mark-up (p(z)/z) as a function of the cost draw z. The dashed line is the constant mark-up of the standard model without possibility of integration. The model displays endogenous mark-ups, higher for more productive sellers, and lower for less productive ones.¹⁶

The optimal price is decreasing in η : when the degree of substitutability increases, potential buyers can more easily switch to cheaper substitutes, hence suppliers must decrease the price to keep their share of the market.¹⁷ Finally, the price is decreasing in λ . When λ decreases, the variance of the buyers' cost distribution increases, and the tail of the distribution becomes fatter: there is a larger mass of potential buyers with very high costs and suppliers can charge higher prices and mark-ups.

2.4 Equilibrium in the Final Good Market

Production of the final consumption good c is done through a constant returns to scale technology which requires the intermediate goods aggregate q and labor as inputs: $c = q^{\alpha} l_f^{1-\alpha}$, where $\alpha \in (0,1)$ and l_f is the labor force employed in the final good sector. Let L denote the country's total labor force; then $l_i = L - l_f$ is the labor force working in the intermediate good sector (for both suppliers and integrated segments of final good producers). The linearity of each intermediate good production technology implies: $q = \frac{l_i}{L}$, where k is

¹⁶Melitz and Ottaviano (2008) obtain mark-ups variability and dependence of profits on productivity by assuming linear demand systems with horizontal product differentiation. Bernard et al. (2003) obtain similar features by assuming Bertrand competition in the intermediate goods sector. In Bernard et al. (2003) though, aggregate mark-ups do not depend on country characteristics and geographic barriers, while they do in Melitz and Ottaviano (2008) and in this paper. The dependence of aggregate mark-ups on country characteristics creates a link between trade/FDI liberalization and competition, which will be clearer in the open economy section.

¹⁷Notice that, for $\eta \to \infty$, prices tend to marginal costs and the model reduces to a perfectly competitive framework where organizational choices are purely driven by productivity differences.

the number of units of labor required to produce 1 unit of the aggregate q.¹⁸ Optimality in the final good market implies that the equilibrium labor allocation and the value of q are:

$$l_f = \left(\frac{(1-\alpha)p}{(1-\alpha)p+\alpha k}\right)L \quad ; \quad l_i = \left(\frac{\alpha k}{(1-\alpha)p+\alpha k}\right)L \quad ; \quad q = \left(\frac{\alpha}{(1-\alpha)p+\alpha k}\right)L. \quad (11)$$

Finally, r denotes the zero-profit equilibrium price of the final good: $r = \alpha^{-\alpha} (1-\alpha)^{\alpha-1} p^{\alpha}$.¹⁹

3 The Open Economy

3.1 Trade Versus Domestic or Foreign Integration

I consider now producers' optimal choices in a world of N countries. Each country is a replica of the economy of the previous section, in the sense that is populated by a continuum of identical final good producers and by a continuum of specialized intermediate goods producers. A final good producer in country i ($i \in \{1, ..., N\}$) needs to source a continuum of inputs to produce a non-tradeable, homogeneous final good. As in the closed economy, each input can be either produced in an integrated facility or bought from a specialized supplier, but each of these options can be implemented domestically or in a foreign country. Suppliers in any country can sell to buyers in any country, so a final good producer can outsource from the lowest cost supplier around the world. Similarly, if a final good producer decides to integrate production, he can set up a domestic integrated plant, or engage in FDI and set up an affiliate in a foreign country. The optimal sourcing strategy is determined comparing outside prices and in-house costs of production worldwide, and is going to be affected also by trade costs and factor cost differences across countries.

When a final good producer decides to integrate production abroad, it generates flows of vertical FDI. I assume that foreign investment is associated with ownership of the foreign

$${}^{18}k = p^{\eta} \left[\int_0^\infty \int_0^{p(z)} x^{1-\eta} \phi(x) \psi(z) dx dz + \int_0^\infty z p(z)^{-\eta} \left[1 - \Phi(p(z)) \right] \psi(z) dz \right].$$

¹⁹In equilibrium, the following market clearing condition holds: $L + \int_0^\infty \pi(z) \psi(z) dz = rc$, where L denotes labor income (wages are normalized to one), $\pi(z)$ is the profit of a supplier with cost draw z, and rc denotes total expenditure in the final good.

production facility. Since the final good is non-tradeable and is assembled domestically, the inputs produced in foreign affiliate plants are shipped back to the parent, generating flows of intrafirm trade, precisely of imports from foreign affiliates. I restrict FDI to be only vertical in this economy, *i.e.* firms that decide to set a plant abroad do not serve the host country or third countries' markets. Foreign plants are only used to produce inputs for the domestic final good sector. This restriction relies on assuming that the in-house technology is not adaptable to serve other firms.

Labor is immobile, so wages may differ across countries.²⁰ I denote with w_i the wage level in country *i*. A final good producer located in country *i* has a set of technology draws $\{x_i\}$, each drawn from a country-specific distribution $\phi_i(x_i)$. If he decides to integrate production of an input, he may choose to do so in his own country or abroad. If he decides to produce at home, his unit cost is given by his technology draw times the domestic wage, $w_i x_i$. If he decides to produce abroad, he can transfer its technology draw to the foreign country and hire foreign workers at local wages. Production abroad entails other costs, like building a new plant, dealing with foreign institutions, relocate managerial know-how, in addition to the trade costs between parent and affiliates that must be paid when repatriating the produced inputs for further manufacturing. For simplicity, I model these costs as bilateral iceberg costs, implicitly assuming that they are correlated with the size of production.²¹

I denote with τ_{ij} the unit iceberg cost for a final good producer from country *i* to produce an input in an integrated facility in country *j*:

Assumption 1. $\tau_{ij} \geq 1 \quad \forall i, j, \quad \tau_{ij} = 1 \quad \forall i = j \text{ and } \tau_{ij} \leq \tau_{ik} \tau_{kj} \quad \forall i, j, k \in \{1, ...N\}.$

Hence, if a final good producer from country *i* decides to produce in country *j* an input for which he has cost draw x_i , his unit cost of production is $\tau_{ij}w_jx_i$.

We now turn to the outsourcing option. In each country j $(j \in \{1, ..., N\})$ there is a continuum of suppliers, each of whom produces a unique differentiated input with an

²⁰All producers take wages as given, and wages will be endogenously determined in equilibrium to clear the labor market in each country.

²¹In Grossman and Rossi-Hansberg (2008), the cost of offshoring is different across goods (or tasks). In the numerical exercise in Section 5, I calibrate iceberg FDI costs to match the intrafirm share of U.S. import.

adaptable technology that enables him to sell it to any buyer around the world. Each supplier in country j has a productivity draw z_j , which affects his marginal cost and the price he charges for the good. Each z_j is drawn from the country-specific distribution $\psi_j(z_j)$.²² An intermediate goods producer in country j can only hire domestic labor, hence his unit cost of production is $w_i z_i$.²³ When selling abroad, he also bears an additional cost, representing barriers to international trade, as tariffs and transportation costs. I denote with t_{ij} the iceberg trade cost for a supplier from country j that sells his good in country i:

Assumption 2. $t_{ij} \ge 1 \quad \forall \ i, j, \ t_{ij} = 1 \quad \forall i = j \ and \ t_{ij} \le t_{ik} t_{kj} \quad \forall i, j, k \in \{1, ..., N\}.$

Hence $t_{ij}w_jz_j$ is the unit cost to sell a good in country i for a supplier from country j with cost draw z_j . Given imperfect competition in the intermediate goods market, suppliers of the same input from different countries price-compete with each other to sell in a market. I denote with $p_{ij}(z_j)$ the price charged to a potential buyer in country i by a supplier in country j who has a cost draw z_j .

In this setup, an input used by a final good producer in country i is defined by the (N+1)-dimensional vector $(x_i, z) = (x_i, z_1, z_2, ..., z_N)$. I denote with $q_i(x_i, z)$ the quantity produced of an intermediate good for which a final good producer in country i has cost draw x_i and suppliers in all countries have cost draws $z = \{z_j\}_{j=1}^N$.

3.2**Organizational Choices and Location**

The analysis of the model follows basically unchanged from the previous section. A final good producer in country i observes his own set of technology draws $\{x_i\}$, a set of wages and iceberg costs $\{w_j, \tau_{ij}\}_{j=1}^N$, a set of outside C.I.F. prices (inclusive of trade costs) $\{p_{ij}(z_j)\}_{j=1}^N$, and decides whether to buy or produce, and where, each of the inputs he needs. The organizational choice is done by comparing the minimum cost across countries of producing

²²The cost distributions $\{\phi_i(\cdot)\}_{i=1}^N, \{\psi_j(\cdot)\}_{j=1}^N$ are mutually independent across countries. ²³I assume here that only final good producers can have offshore production plants, while suppliers produce only domestically. This assumption can be relaxed, and the unit cost of a supplier from country j producing in country k would take the form $\tau_{ik} w_k z_i$. This extension would come at a higher computational cost of solving for the equilibrium vector of wages.

an input and the minimum outside price of buying it. The problem is exactly as in the closed economy, but with a larger set of prices to shop for (2N instead of 2).

Let $c_i(x_i, z)$ denote the minimum unit cost of good (x_i, z) :

$$c_i(x_i, z) = \min_j \{\tau_{ij} w_j x_i, p_{ij}(z_j)\}.$$
(12)

Once decided to integrate, the location of production is determined by the interaction between iceberg costs and wages. Let m_i denote the cheapest combination of wages and FDI iceberg costs worldwide for integrated production of a final good producer from country i:²⁴

$$m_i = \min_k \tau_{ik} w_k. \tag{13}$$

On the other hand, once decided to outsource, the location of the lowest cost supplier is determined by trade costs and by the cross-country joint cost distribution of the suppliers, which affect the prices charged. A final good producer with a set of cost draws $\{x_i\}$ in country *i* solves:

$$\min_{q_i(x_i,z)} \qquad \int_{\Re^N_+} \int_0^\infty c_i(x_i,z) q_i(x_i,z) \phi_i(x_i) \psi(z) dx_i dz$$
s.t.
$$\left[\int_{\Re^N_+} \int_0^\infty q_i(x_i,z)^{1-1/\eta} \phi_i(x_i) \psi(z) dx_i dz \right]^{\eta/(\eta-1)} \ge q_i \qquad (14)$$

 $^{^{24}}$ A theory where the cost of setting up a new plant is modeled as an iceberg cost implies that a final good producer in a country will choose to locate all the integrated segments of production in the same country (or in the same set of countries in case of ties: the optimal production allocation choice in case of ties is shown in the general equilibrium section of the paper). Moreover, since the final good producers in each country are homogeneous, all firms from a country choose the same location(s) for their integrated activities. Hence the model does not necessarily generate integrated production of different goods in multiple locations by the same firm, or by firms in the same country. What it does generate, however, is the fact that producers from different countries are likely to choose different destinations for their integrated production processes, since the setup cost is bilateral. In the data, the degree of multilaterality in the operations of U.S. multinational corporations appears to be limited: the vast majority of intrafirm imports by U.S. firms happens with affiliates located in a small number of countries. Data from the Bureau of Economic Analysis show that in 2004, more than 80% of imports of U.S. parents from their foreign affiliates were shipped by only 7 countries, representing in fact 3 main areas: Canada and Mexico, the U.K. and Ireland, and Malaysia, Hong Kong and Singapore. The concentration appears even stronger if we look at data by sector. Another implication of this modeling device is that vertical FDI is going to be only unilateral, from high-wage to low-wage countries (see also Nocke and Yeaple (2008)).

where $\psi(z) = \prod_{j=1}^{N} \psi_j(z_j)$ is the density of the vector $z = (z_1, z_2, ..., z_N)$, and the intermediate goods aggregate q_i is determined by equilibrium conditions in the final good market. Let $B_i^I = \left\{ (x_i, z) \in \Re_+^{N+1} : c_i(x_i, z) = m_i x_i \right\}$ denote the set of goods that a final good producer in country *i* decides to internalize, in the location(s) with the lowest cost m_i . Let $B_{ij}^T = \left\{ (x_i, z) \in \Re_+^{N+1} : c_i(x_i, z) = p_{ij}(z_j) \right\}$ denote the set of goods that he decides to outsource from a producer in country *j*. Problem (14) is solved by:

$$q_i^I(x_i, z) = q_i^I(x_i) = (m_i x_i)^{-\eta} p_i^{\eta} q_i \quad \forall (x_i, z) \in B_i^I$$
 (15)

$$q_i^T(x_i, z) = q_i^T(\{p_{ij}(z_j)\}_{j=1}^N) = [p_{ij}(z_j)]^{-\eta} p_i^{\eta} q_i \quad \forall \ (x_i, z) \in B_{ij}^T$$
(16)

where p_i is the aggregate price index in country *i*:

$$p_i = \left[(p_i^I)^{1-\eta} + \sum_{j=1}^N (p_{ij}^T)^{1-\eta} \right]^{1/(1-\eta)}$$
(17)

and:

$$p_i^I = \left[\int_{B_i^I} (m_i x_i)^{1-\eta} \phi_i(x_i) \psi(z) dx_i dz \right]^{1/(1-\eta)}$$
(18)

$$p_{ij}^{T} = \left[\int_{B_{ij}^{T}} [p_{ij}(z_j)]^{1-\eta} \phi_i(x_i) \psi(z) dx_i dz \right]^{1/(1-\eta)}.$$
 (19)

It remains to determine the prices $\{p_{ij}(z_j)\}_{i,j=1}^N$. Markets are segmented. In the intermediate goods market, each supplier maximizes its expected profits from sales to potential buyers around the world, and may charge different prices to buyers in different countries. By assuming that no resale is possible, I study the pricing problem country by country. In choosing the optimal price to charge in a destination market, a supplier must consider competition from the producers of the same good in other countries and the fact that the potential buyers have the option of integrating production.²⁵ Each supplier observes his

 $^{^{25}}$ I assume there is only one supplier for each input in each country. This assumption can be relaxed by interpreting the production function of a supplier as the aggregate production function of a set of "lower-level" suppliers that produce the same input in a country. By assuming that the technology for producing

own marginal cost, and the aggregate parameters of the cost distributions of the potential buyers and of his competitors in other countries. Based on this information, suppliers simultaneously declare a set of prices (one for each country). In each market, the supplier that declares the lowest price sells the input to all the buyers with sufficiently high costs of insourcing. The price setting mechanism has the properties of a potentially asymmetric first-price sealed-bid auction.²⁶ Each supplier sets the price as a function of his own marginal cost in a way that, given that all the other suppliers set their price in the same way, no individual supplier could do better by choosing the price differently. The resulting equilibrium is a Bayes-Nash equilibrium, where each supplier chooses its optimal price based on his guess (correct in equilibrium) of the pricing rules followed by suppliers of the same good in other countries.²⁷

Let $b_{ij}(p_{ij}(z_j))$ be the set of technology draws of buyers in country *i* and of suppliers outside country *j* such that the buyers in country *i* decide to buy good (x_i, z) from the supplier in country *j*: $b_{ij}(p_{ij}(z_j)) = \{(x_i, \{z_k\}_{k\neq j}) \in \Re^N_+ : (x_i, z) \in B^T_{ij}\}$. A supplier in country *j* with productivity draw z_j maximizes his expected profits from sales in country *i* in the set $b_{ij}(p_{ij}(z_j))$:

$$\max_{p_{ij}(z_j)} \int_{b_{ij}(p_{ij}(z_j))} [p_{ij}(z_j) - t_{ij}w_j z_j] q_i^T(x_i, z) \phi_i(x_i) \prod_{k \neq j} \psi_k(z_k) dx_i dz_k.$$
(20)

Using (16), and due to the independence property of the cost distributions, problem (20)

an input is country-specific (*i.e.*, that z_j is constant across all producers of the same input in country j) and that each lower-level producer has a decreasing returns to scale production function and pays a fixed cost to enter the market (along the lines of Rossi-Hansberg and Wright (2007)), it can be shown that the aggregation of lower-level producers generates a constant returns to scale technology that is isomorphic to the linear technology of each supplier in the model. This is achieved by appropriately redefining the technology draw z as a function of the fixed entry cost and of the parameter ruling decreasing returns for the lower-level producers.

 $^{^{26}}$ In their survey of the auctions literature, McAfee and McMillan (1987) report that "sealed-bid tenders are [...] used by firms procuring inputs from other firms". Asymmetric auctions seem a natural tool to study pricing in international markets, "when both domestic and foreign firms submit bids and, for reasons of comparative advantage, there are systematic cost differences between domestic and foreign firms".

²⁷Appendix A provides a more extended treatment of the analogies of this problem with a first-price sealed-bid auction.

can be restated as:

$$\max_{p_{ij}(z)} \left[p_{ij}(z_j) - t_{ij} w_j z_j \right] \left(\frac{p_{ij}(z_j)}{p_i} \right)^{-\eta} q_i A_{ij}(p_{ij}(z_j))$$
(21)

where $A_{ij}(p_{ij}(z_j)) = \left[1 - \Phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\right] \cdot \prod_{k \neq j} \left[1 - F_{ik}(p_{ij}(z_j))\right]$ is the probability that – given the price $p_{ij}(z_j)$ – a final good producer in country *i* buys good (x_i, z) from the supplier in country *j*, and $F_{ik}(\cdot)$ denotes the cumulative distribution function of the prices charged by suppliers in country *k* to final good producers in country *i*.

The first order condition of problem (21) is:

$$p_{ij}(z_j)(1-\eta) + \eta t_{ij}w_j z_j - \dots$$
$$\dots p_{ij}(z_j)[p_{ij}(z_j) - t_{ij}w_j z_j] \left\{ \frac{\phi_i \left(\frac{p_{ij}(z_j)}{m_i}\right) \frac{1}{m_i}}{\left[1 - \Phi_i \left(\frac{p_{ij}(z_j)}{m_i}\right)\right]} + \sum_{k \neq j} \frac{f_{ik}(p_{ij}(z_j))}{\left[1 - F_{ik}(p_{ij}(z_j))\right]} \right\} = 0. \quad (22)$$

For each z_j , (22) is a nonlinear system of N equations in the N unknowns $p_{ij}(\cdot)$.²⁸ As each supplier competes with suppliers from other countries to sell in country i, $p_{ij}(\cdot)$ is determined by evaluating the effects of substitutability on demand, the average productivity of the potential buyers (the hazard rate term $\frac{\phi_i(p_{ij}(z_j)/m_i)/m_i}{[1-\Phi_i(p_{ij}(z_j)/m_i)]}$), and how the optimal price compares with the expected price charged by suppliers in other countries (the hazard rate term $\frac{f_{ik}(p_{ij}(z_j))}{[1-F_{ik}(p_{ij}(z_j))]}$). Notice that – since in the open economy sourcing possibilities have increased – prices are going to be lower than in the closed economy.

The system of equations (22) must be solved numerically. Here I show properties of the pricing rule for the special case in which costs are exponentially distributed $(\phi_i(x_i) = \lambda_i e^{-\lambda_i x_i},$ for $x_i \ge 0$, and $\psi_i(z_i) = \mu_i e^{-\mu_i z_i}$, for $z_i \ge 0$, $i \in \{1, ..., N\}$) and the suppliers' average productivity is constant across countries $(\mu_i = \bar{\mu}, \forall i)$. For this special case, (22) reduces to the following non-linear first order ODE:

$$p'_{ij}(z_j) = \frac{\xi_j p_{ij}(z_j) [p_{ij}(z_j) - t_{ij} w_j z_j]}{\eta t_{ij} w_j z_j + \left(1 - \eta + \frac{\lambda_i}{m_i} t_{ij} w_j z_j\right) p_{ij}(z_j) - \frac{\lambda_i}{m_i} p_{ij}(z_j)^2}$$
(23)

 $^{^{28}}$ The algorithm to solve (22) is described in Appendix A, and is available upon request to the author.

where $\xi_j = \bar{\mu} \sum_{k \neq j} \frac{t_{ij} w_j}{t_{ik} w_k}$ is a measure of relative competitiveness of suppliers in countries other than j.

For small values of z, the solution of (23) has the form: $p(z)\Big|_{z\to 0} = \frac{\eta}{\eta-1}twz + o(z)$, where the country indexes have been suppressed to simplify the notation. This indicates that the most productive suppliers are not affected by the competition of suppliers in other countries or by the possibility of integration of the buyers, since their optimal price is about the same (except for higher order terms, negligible for $z \to 0$) as in a standard monopolistically competitive model without possibility of integration. On the other hand, for very large values of z, the solution of (23) has the form: $p(z)\Big|_{z\to\infty} = \frac{tw}{\frac{\lambda}{m}tw+\xi} + twz$, which implies percentage mark-ups converging to zero, approaching the perfectly competitive case.

When $\xi_j = 0$, *i.e.* when we rule out international competition from suppliers in other countries, the problem reduces to the closed economy one (corrected for transportation costs and wages). Globally, the solution lies between the marginal cost line $t_{ij}w_jz_j$ and the closed economy pricing rule. The value of the parameter ξ_j affects the location of the curve: when international competition is tough ("high" ξ_j), the solution approaches the marginal cost line, while when international competition is low ("low" ξ_j), the solution approaches the closed economy one. Overall, prices are lower than in the closed economy, and the link between trade liberalization and prices becomes evident here: opening to trade increases the level of competition in a country (through the term ξ_j) and – as a result – prices and mark-ups shrink.²⁹

Figure 2 shows plots of the pricing rule in a world of two identical countries, for some arbitrary values of the trade barriers. The left panel shows domestic F.O.B. prices (the solid line), F.O.B. export prices (the dashed line), and marginal costs (the dotted line). Trade costs create a wedge between domestic prices and export prices: F.O.B. export prices and mark-ups are lower than the domestic ones to counteract the fact that foreign buyers must also pay the transportation cost on the imported goods (firms shrink their mark-ups to be competitive in the foreign market despite the higher costs).

²⁹Melitz and Ottaviano (2008) obtain the same qualitative result.

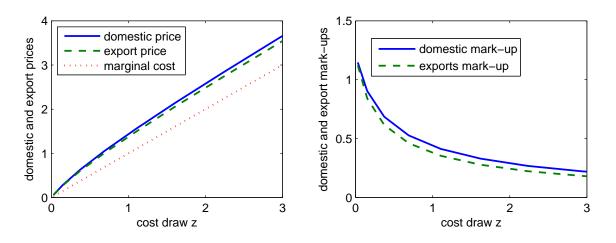


Figure 2: Open economy prices (2 symmetric countries, t = 1.3, $\tau = 1.5$).

Introducing heterogeneity in the two countries' wages and productivity distributions may create larger wedges between domestic and export prices, and may also produce export prices higher than the domestic ones, if competition in the home country is tougher than in the export market. The prices charged are affected by the number of countries in the economy: a higher number of countries generates tougher foreign competition (ξ_j increases) and prices will be consistently adjusted downwards, tending to perfectly competitive prices for $N \to \infty$. The price $p_{ij}(z_j)$ is increasing in the cost of integration m_i : a high minimum costs of integration (through wages or iceberg costs) makes the integration option less attractive, and a higher outside price still preferable for the potential buyers. Similarly, high transportation costs or low productivity in the competitors' countries increase the price charged, as foreign competition is low ($\frac{\partial p_{ij}(z_j)}{\partial t_{ik}} > 0$, $\frac{\partial p_{ij}(z_j)}{\partial \mu_k} < 0 \quad \forall k$). These properties make explicit the dependence of prices and mark-ups on country characteristics and geographic barriers. By consequence, country characteristics also affect the choice of undertaking intrafirm transactions through their effects on arm's length prices.³⁰ The analysis of the pricing strategy

³⁰The properties of the pricing rule described here depend on the assumption that suppliers cannot observe their competitors' marginal costs. This seems a reasonable assumption to make, particularly in the international context, where it may simply be too costly to monitor a foreign competitor's cost structure. Alternatively, one could remove the assumption of private information on the marginal costs and assume Bertrand competition across suppliers of the same input in different countries (as in Bernard et al. (2003)). Under this alternative scenario, the lowest cost supplier would still "win" the market, and would charge a price equal to the minimum between the second-best producer's marginal cost and the unconstrained profitmaximizing price (equal to the closed economy price in (9) corrected for transportation costs and wages). In the symmetric case, this alternative formulation implies the same average prices as (22) (by the Revenue

confirms that when a country opens to operations with other countries, both integration and trade become cheaper: integration may be relocated in lower-cost countries, and trade becomes more attractive because the higher degree of competition has the effect of lowering prices.

3.3 General Equilibrium

The final good is non-tradeable, and must be produced domestically using the intermediate goods aggregate and local labor. The final good production function in country i is: $c_i = q_i^{\alpha}(l_i^f)^{1-\alpha}$, where l_i^f is the amount of labor used in the final good sector. The labor force in each country is split in the two sectors, and the share of the labor force working in the intermediate goods sector may either work for local suppliers (serving the domestic and/or foreign markets) or for affiliates of domestic or foreign integrated firms. Labor is immobile, and the following population constraint must hold in each country:

$$L_{i} = l_{i}^{f} + \sum_{j=1}^{N} (l_{ji}^{I} + l_{ji}^{T}) \quad for \quad i = 1, \dots N,$$
(24)

where l_{ji}^{I} is the labor force of country *i* working in integrated segments of firms from country *j* and l_{ji}^{T} is the labor force of country *i* working for suppliers from country *i* selling in market *j*. Since the intermediate goods production function is linear, the labor force segments can be expressed as linear functions of the aggregate quantities $\{q_i\}_{i=1}^{N}$:

$$L_{i} = \frac{(1-\alpha)p_{i}}{\alpha w_{i}}q_{i} + \sum_{j=1}^{N} (k_{ji}^{I}q_{j} + k_{ji}^{T}q_{j}) \quad for \ i = 1, ...N$$
(25)

where the proportionality factors k_{ji}^{I} , k_{ji}^{T} are functions of the wage levels $\{w_i\}_{i=1}^{N}$ and of the model's parameters only.³¹

$${}^{31}k_{ji}^{I} = \frac{p_{j}'}{w_i} \int_{B_j^{I}} (m_j x_j)^{1-\eta} \phi_j(x_j) \psi(z) dx_j dz \text{ and } k_{ji}^{T} = p_j^{\eta} \int_{B_{ji}^{T}} t_{ji} z_i p_{ji}(z_i)^{-\eta} \phi_j(x_j) \psi(z) dx_j dz$$

Equivalence Theorem). In the asymmetric case, prices could be higher or lower than the ones implied by (22), and the discontinuity of the limit pricing rule would weaken the link between trade/FDI liberalization and prices.

Taking the wages as given, (25) is a linear system of N independent equations in the N unknowns q_i , whose solution delivers the equilibrium values of $\{q_i\}_{i=1}^N$ as functions of the wages only: $q_i^* = q_i(w_1, ..., w_N)$, for i = 1, ...N. Market clearing conditions allow to solve for the equilibrium vector of wages. In each country, total income (labor income plus the suppliers' profits) must be equal to total expenditure in the final good:

$$r_i c_i = L_i w_i + \int_0^\infty \pi_i(z_i) \psi_i(z_i) dz_i \quad for \ i = 1, ...N$$
 (26)

where r_i is the zero-profit price of the final good in country *i*: $r_i = \alpha^{-\alpha}(1-\alpha)^{(\alpha-1)}p_i^{\alpha}w_i^{1-\alpha}$, and $\pi_i(z_i)$ is the total profit of a supplier from country *i* with cost draw z_i . The market clearing condition (26) is a system of *N* equations in the *N* unknowns $\{w_i\}_{i=1}^N$ and can be solved for the equilibrium wages. In the following section, I prove existence of the equilibrium and show its qualitative properties in the two-country case.

3.4 Equilibrium Characterization: Two-Country Case

I denote the two countries with H (Home) and F (Foreign). Normalizing to one the wage in the Foreign country, the equilibrium is a relative wage w_H such that the excess demand in the Home country is equal to zero: $ED_H = L_H w_H + \int_0^\infty \pi_H(z_H) \psi_H(z_H) dz_H - r_H c_H = 0.$

Figure 3 plots the excess demand correspondence in the Home country for the symmetric case.³² Due to the discrete choice of where to locate integrated production, the correspondence has two kinks at $w_H/w_F = 1/\tau$ and $w_H/w_F = \tau$. The excess demand associated with each of these two points is an interval, and if the correspondence crosses the zero line at one of these points the corresponding relative wage does not necessarily clear the market. This happens because $w_H/w_F = 1/\tau$ and $w_H/w_F = \tau$ are the levels of the relative wage such that firms change the location of their integrated activities:

• $w_H \in (0, w_F/\tau) \Rightarrow$ firms from both countries integrate in the Home country;

³²The computation is done for parameters' values $\eta = 1.8$, $\alpha = 0.25$, t = 1.1 and $\tau = 1.2$, under exponentially distributed costs.

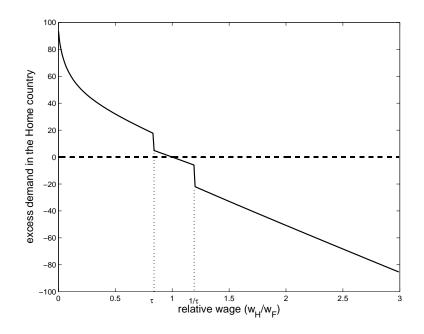


Figure 3: Excess demand correspondence in the Home country.

- $w_H \in (w_F/\tau, \tau w_F) \Rightarrow$ firms from both countries integrate domestically;
- $w_H \in (\tau w_F, \infty) \implies$ firms from both countries integrate in the Foreign country.

When $w_H = \tau w_F$, firms from H are indifferent about where to integrate production, whether domestically or abroad, while firms from F integrate domestically. The figure shows that if firms from H choose to integrate only in one country when they are indifferent, the equilibrium wage may not clear the market. Then firms from H integrate in both countries, and the allocation of labor in the integrated sectors in each country is the variable that clears the market. Similarly, when $w_H = w_F/\tau$, firms from F integrate in both countries, while firms from H integrate domestically. At these critical points, the excess demand correspondence is non-smooth because the cost structure of the firms suddenly changes: Figure 4 shows that the unit labor demand of integrated sectors of Home and Foreign firms has a kink at the point where firms switch from domestic to foreign integration and viceversa.

The following proposition establishes the existence of the equilibrium for the two-country case.

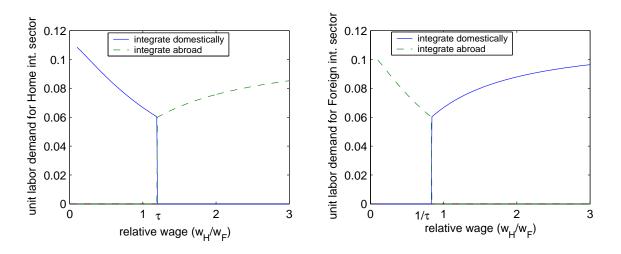


Figure 4: Labor demand in the integrated sectors.

Proposition 1. Provided that the pricing rules $\{p_{ij}(z_j)\}$ are continuous in z_j , $\forall i, j = 1, ...N$, there exists a relative wage w_H/w_F such that $ED_H = 0$.

Proof: See Appendix B.

4 Comparative Statics

In this section I use a numerical example to show the predictions of the model for the dependence of volumes of arm's length imports and intrafirm imports (vertical FDI) as functions of the exogenous parameters. Volumes of (arm's length) imports as a fraction of GDP for country i are given by:

$$(import/GDP)_i = \frac{\sum_{j \neq i} \left[(p_{ij}^T)^{1-\eta} (p_i)^{\eta} q_i \right]}{r_i c_i}$$

while volumes of vertical FDI (or intrafirm imports) as a fraction of GDP for country i are:

$$(vertical \ FDI/GDP)_{i} = \begin{cases} \frac{(p_{i}^{I})^{1-\eta}(p_{i})^{\eta}q_{i}}{r_{i}c_{i}} & ; \text{ if } w_{i} > \min_{j \neq i} \{\tau w_{j}\} \\ \frac{(1-\gamma_{i})(p_{i}^{I})^{1-\eta}(p_{i})^{\eta}q_{i}}{r_{i}c_{i}} & ; \text{ if } w_{i} = \min_{j \neq i} \{\tau w_{j}\} \\ 0 & ; \text{ otherwise} \end{cases}$$

where γ_i is the percentage of labor force hired domestically in the integrated sectors when the final good producers are indifferent about where to integrate.

4.1 Effects of Country Size

In a two-country world, I isolate the effects of size by normalizing the world population to one and denoting by s the share of world population living in the Home country. The two countries are symmetric in all the other characteristics.³³ Figure 5 plots volumes of

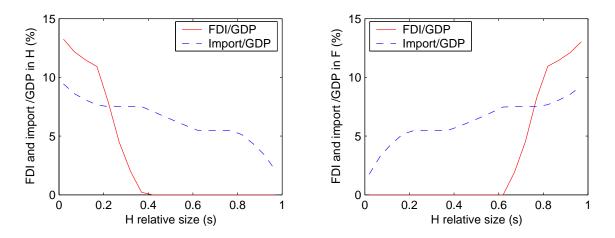


Figure 5: Volumes of arm's length import and FDI import as functions of relative size.

arm's length imports and FDI imports as fractions of GDP for the two countries, expressed as functions of s: moving to the right in the pictures means that country H accounts for

³³All the computations in this section are done for $\eta = 1.8$, $\alpha = 0.25$, t = 1.1 and $\tau = 1.2$, for exponentially distributed costs with parameters λ_i (for final good producers) and μ_i (for suppliers). I chose the symmetric case as a benchmark for this comparative statics exercise to isolate the effects of differences in one parameter at a time. In the calibration exercise in Section 5 multiple parameters need to differ across countries to give rise to the observed arm's length and intrafirm trade flows.

a bigger share of the world's labor force. Both arm's length and intrafirm imports are decreasing in the relative size of the source country, showing that smaller economies tend to source from bigger ones. Trade and FDI coexist, and while trade is bilateral, FDI is only unilateral, as it arises to exploit factor cost differentials. The smaller country (which, everything else equal, has higher labor cost) invests and sets affiliates in the larger country (where labor is more abundant, hence cheaper), and the volume of FDI is increasing in the size differential. This prediction is consistent with the the analysis of Nocke and Yeaple (2008), who find that greenfield FDI is prevalent between countries that are heterogeneous in size and other characteristics, and mostly emerges to exploit factor cost differences across countries. Hanson et al. (2001) also document a tendency for U.S. multinationals to set vertical production networks through affiliates in low-wage countries. The figures also show that, as FDI decreases faster than arm's length import as size increases, also the share of intrafirm imports over total imports is decreasing in relative size.

4.2 Effects of Productivity Dispersion

Figure 6 plots volumes of arm's length import and FDI import as functions of the productivity dispersion of final good producers, λ_H . When λ_H increases, the cost distribution of integrated production becomes less disperse, and average productivity increases, so that Home country firms find more profitable to integrate. The higher productivity translates into higher domestic wages, which push towards moving integrated production abroad (FDI). At the same time, the reliance on arm's length imports decreases. As a result, the share of intrafirm imports over total imports is increasing in λ : a higher share of intrafirm transactions is associated with a higher productivity and with a lower dispersion in the in-house costs of the final good producers. On the other hand, both types of import decrease for firms in F due to the effect of λ_H on wages, and the behavior of FDI mirrors the one in the H country. The unilaterality of FDI emerges also from this picture: is the most productive country that sets integrated production in the other country, which has lower productivity and hence lower labor costs.

Figure 7 displays volumes of arm's length import and FDI import as functions of the

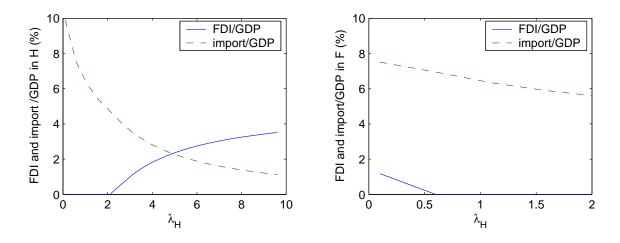


Figure 6: Volumes of arm's length import and FDI import as functions of buyers' costs dispersion.

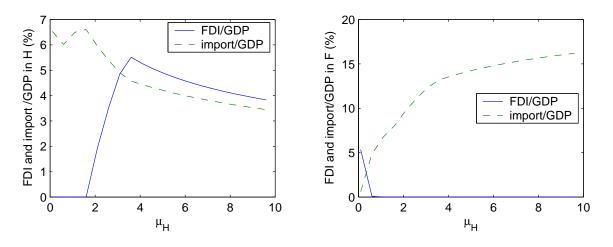


Figure 7: Volumes of arm's length import and FDI import as functions of suppliers' costs dispersion.

productivity dispersion of the intermediate goods producers, μ_H . When μ_H increases, the variance of the cost distribution of the suppliers in country H decreases: this can have different effects on the volumes of trade, because of reallocations due to the effect of productivity on wages and prices. When μ_H increases, the wage level w_H also increases, encouraging foreign sourcing (wage effect); on the other hand, more productive suppliers charge lower prices, encouraging domestic arm's length sourcing (price effect). Overall, increases in the productivity of domestic suppliers make foreign suppliers less attractive, so H's arm's length imports drop as μ_H increases. For "low" values of μ_H , w_H is low and arm's length imports

are substituted by domestic integration. As μ_H further rises, w_H rises and foreign integration becomes profitable, so arm's length imports are substituted by intrafirm imports (the wage effect dominates). As μ_H keeps rising, intrafirm imports start declining because the price effect dominates: the much higher domestic productivity makes sourcing from domestic suppliers the cheapest option, so both types of imports decline. The non-monotonicity of FDI in H is the result of these two opposite effects. Final good producers in F find profitable to set plants in H for low values of μ_H (low w_H), and overall F arm's length imports increase as μ_H increases, since the prices charged by H suppliers drop.

In summary, these calculations suggest that volumes of vertical FDI (or intrafirm import) are higher for small countries with low dispersion of the in-house costs (high λ), and the larger the differences in the fundamentals of the two countries involved (absent in case of identical countries). Volumes of arm's length imports are higher for small countries with high dispersion of the in-house costs (low λ), the larger the differences in the fundamentals of the two countries involved, but significant also between identical countries. The intrafirm share of total imports is higher the smaller the country where the parent is located and the higher the ratio $\lambda_i/\mu_{\neq i}$ in that country (*i.e.* the higher the productivity of multinational firms compared to the productivity of foreign suppliers).

5 The Gains from Multinational Production

In this section, I calibrate the two-country model to match aggregate U.S. data about volumes of trade and multinational activity.³⁴ With the calibrated model, I quantify the gains – for the U.S. economy and for the rest of the world – arising from vertical FDI (intrafirm trade). With counterfactual experiments, I show how the gains depend on the degree of competition in the market and on the extent of barriers to foreign investment. I isolate the gains from trade from the gains from multinational production, and compare them with other papers'

³⁴The U.S. Bureau of Economic Analysis publishes yearly data on U.S. multinational companies, including their intrafirm trade flows. To my knowledge, these are the only publicly available data on intrafirm trade flows. As a consequence, the calibration in this section focuses on a bilateral exercise that matches U.S. trade flows with the rest of the world.

findings.

5.1 Calibration

I start by describing the calibration of the parameters of the model. In a two-country world, I identify the Home country with the United States, and the Foreign country with an aggregate of countries that I denote as the "rest of the world" (henceforth, ROW).³⁵ Calibration of the bilateral model requires to assign values to the parameters of the production functions, α and η , to the relative productivity parameters $\lambda_{us}/\lambda_{row}$, μ_{us}/μ_{row} , to the relative size parameter L_{us}/L_{row} , and to the iceberg costs t and τ .

 $(1 - \alpha)$ represents the labor share in the final good production function. As the final good in the model is non-tradeable, Alvarez and Lucas (2007) identify $(1 - \alpha)$ with the fraction of employment in the non-tradeable sector, and compute α using data on agriculture, mining and manufacturing (defined as tradeables). Following calculations from different data sources, they choose $\alpha = 0.25$ as a reasonable value for industrialized countries. Since I want to match features of U.S. trade data, I use their calibrated value in this computation.

The exponential parameterization of the cost distributions restricts the elasticity of substitution between intermediates to $\eta \in (1,2)$ to assure an elasticity of demand larger than one and convergence of the price integrals.³⁶ In the model, η is a measure of product differentiation and market power, and it has a large effect on the computation of the welfare gains and on their decomposition. In the baseline calibration I choose a value of $\eta = 1.8$,

³⁵The ROW in the calibration is composed by 157 countries, representing 98% of the world GDP.

³⁶In Eaton and Kortum (2002) and Alvarez and Lucas (2007), the Weibull parameterization of the cost distribution implies a wider interval for the value of η . The exponential distribution corresponds to a Weibull distribution with shape parameter $\vartheta = 1$. Eaton and Kortum (2002) and Alvarez and Lucas (2007) use values of $\vartheta \in [0.1, 0.3]$, much below the value I use. I choose $\vartheta = 1$ because it assures existence and uniqueness of a continuous and increasing pricing rule in the closed economy and in the symmetric version of the open economy model, which I use to initialize the algorithm to solve for the prices in the asymmetric case, as described in Appendix A. In Eaton and Kortum (2002) and Alvarez and Lucas (2007), the value of ϑ matches across-goods heterogeneity in countries' relative efficiences, and hence it is more relevant in a multi-country framework. I argue that – for the purpose of the exercise in this paper – cross-country variability is not key for the results, since the calibration is bilateral. Moreover, since Eaton and Kortum (2002) show that the gains from trade are an increasing function of ϑ , the results reported in this section can be interpreted as an upper bound for the actual gains.

but I also present the results for a lower value of η ($\eta = 1.2$) to show how the gains from multinational production depend on this aspect of competition in the market.³⁷

The remaining parameters are calibrated jointly to match a set of relevant moments in the data. I identify the ratio μ_{us}/μ_{row} with the relative average productivity of U.S. firms with respect to ROW firms. Bernard, Jensen, and Schott (2009) report that multinational corporations appear to be on average more productive than non-multinational firms. Due to the scarcity of available data to quantify this productivity differential, I assume that final good producers (the potential multinational corporations) draw their productivity from the same distribution as local firms, and hence have the same relative average productivity: $\lambda_{us}/\lambda_{row} = \mu_{us}/\mu_{row}$. As in Alvarez and Lucas (2007), L_{us}/L_{row} represents labor in efficiency units in the US relative to the ROW.³⁸ t and τ are average iceberg costs of trade and FDI. I choose the four parameters μ_{us}/μ_{row} , L_{us}/L_{row} , t, and τ to match:

- 1. the intrafirm share of imports of U.S. multinational corporations from their foreign affiliates,
- 2. U.S. total imports as a fraction of GDP,
- 3. U.S. share of world GDP, and
- 4. U.S. GDP per worker relative to an average of the ROW.

All matched data are for the year 2004. The intrafirm share of imports of U.S. parents from their foreign affiliates was 13.5% in 2004, and almost constant over the last decade.³⁹ From

³⁷The value $\eta = 1.8$ implies mark-ups ranging from 125% to zero. Average mark-ups depend on productivity parameters and trade barriers.

 $^{^{38}}$ As explained in Alvarez and Lucas (2007), the size parameter *L* cannot be measured directly. In the model, size and productivity parameters are directly linked to GDP and GDP per worker in each country, variables that I include in the set of moments to be matched.

³⁹I construct the share of intrafirm imports in U.S. total imports by merging data published by the U.S. Bureau of Economic Analysis (U.S. Direct Investment Abroad: Financial and Operating Data for U.S. Multinational Companies, available at http://www.bea.gov/international/) with Census data on U.S. imports. The share I construct is smaller than the ones reported by other papers: Antràs (2003) and Bernard, Jensen, and Schott (2009) report an intrafirm share of imports of about 40%. This discrepancy is due to the fact that I consider only the portion of intrafirm imports that the model explains: imports of U.S. parents from their foreign affiliates. I am excluding imports of U.S.-located affiliates from foreign parents (because the model does not support bilateral intrafirm transactions, more common when talking about horizontal FDI), and transactions between affiliates.

Census data, U.S. imports were 13.3% of U.S. GDP. The share of U.S. GDP in the world GDP was 30% (from WDI's GDP data), and U.S. GDP per worker relative to an average of the ROW was 2.22.⁴⁰

Let $P = [\mu_{us}/\mu_{row} \quad L_{us}/L_{row} \quad t \quad \tau]$. The vector of calibrated parameters is a vector $P^* = \arg \min_P \sum [mom - \widehat{mom}(P)]^2$, where mom is the vector of moments from the data, and $\widehat{mom}(P)$ is the vector of moments generated by the model as function of the vector of parameters P. The baseline calibrated model (for $\eta = 1.8$) implies $L_{us}/L_{row} = 0.16$, t = 1.1, $\tau = 2.65$, $\mu_{us}/\mu_{row} = \lambda_{us}/\lambda_{row} = 1.92$. Table 1 summarizes the calibrated parameters. The values in parentheses are the ones implied by the alternative calibration, with $\eta = 1.2$.

PARAMETER	VALUE	DEFINITION	SOURCE
α	0.25	1 - labor share in non-tradeables	Alvarez and Lucas (2007)
η	1.8(1.2)	elasticity of substitution	model restrictions
t	1.1	iceberg trade cost)
τ	2.65(2.66)	iceberg FDI cost	
L_{us}/L_{row}	0.16	relative labor in efficiency units	to match data
μ_{us}/μ_{row}	1.92(1.77)	suppliers' relative productivity	
$\lambda_{us}/\lambda_{row}$	1.92(1.77)	buyers' relative productivity	

Table 1: Summary of calibrated parameters.

Given the nonlinearities introduced in the model through the shape of the pricing functions and the discrete choice of location, it is hard to talk about identification of the parameters. The calibrated parameters must be determined jointly, as each or them affects all the four matched moments. This said, sensitivity analysis reveals that the computed intrafirm

$$(GDP \ per \ worker)_{row} = \sum_{i \neq us} \left(\frac{GDP_i}{labor \ force_i} \times \frac{imports_{us,i}}{imports_{us,row}} \right)$$

Import data are from the U.S. Census, while GDP and labor force data are from the World Bank's WDI.

⁴⁰I compute the average GDP per worker in the ROW as a weighted average of each country's GDP per worker, with the shares of US imports from that country as weights:

share of imports is extremely sensitive to the choice of the value of the iceberg cost τ . To be able to match the share of intrafirm import from the data, the calibrated value of τ implies that producing one unit of input abroad almost triplicates its unit costs. I believe that the necessity of this high cost to match the data depends on the fact that the model does not consider other types of transaction costs, fixed costs of entering the foreign market, or legal restrictions to intrafirm activities. These frictions – that the model does not consider explicitly – are reflected in the results of the calibration.

Conversely, the calibrated trade iceberg cost t is on the low side. Particularly, it is significantly lower than the ones implied by Eaton and Kortum (2002) estimates. This depends on the fact that in my model (unlike in Eaton and Kortum (2002) and Alvarez and Lucas (2007)) firms have also the option of integrating production domestically, and this option is very attractive in the calibrated economy.⁴¹ For this reason, a high trade cost would generate substitution away from arm's length imports into domestic integrated production. Hence a low iceberg trade cost is necessary to match the share of import that we observe in the data.

5.2 Gains from Multinational Production: Decomposition and Policy Experiments

I compute the welfare gains that the theory implies by comparing the calibrated economy and a counterfactual world without possibility of integration and multinational production. I compute the gain in consumption per capita as:

welfare gain =
$$\left(\frac{\text{consumption p.c. in calibrated model}}{\text{consumption p.c. in model without integration}} - 1\right) \times 100$$

where the term in the denominator is obtained by computing the model with the calibrated parameters, but shutting down the possibility of in-house production. In the model without integration, firms can still trade intermediates at arm's length, so the difference in con-

 $^{^{41}\}mathrm{In}$ the next subsection I show that in the calibrated economy the majority of integrated production is done domestically.

sumption per capita is purely due to the possibility of integration. In the model though, integrated production can happen domestically or abroad, so it is necessary to disentangle how much of the welfare gain comes from domestic activity and how much of it comes from foreign investment. The calibrated model implies that U.S. firms decide to integrate both domestically and abroad, and delivers the equilibrium share of labor that U.S. integrated firms hire in each country.⁴² Due to the linearity of the intermediate goods production technology, the share of labor hired abroad is equal to the share of integrated production that is done abroad and to the share of gains arising from foreign integration. On the other hand, in equilibrium, the ROW economy integrates only domestically. Hence the welfare gains for the ROW consumers come from two sources: the possibility of (domestic) integrated activity and the upward pressure on wages determined by the entry of U.S. firms.

The results are shown in the first column of Table 2. The calibrated economy implies a gain in U.S. consumption per capita of 4.87% with respect to a world economy with no possibility of integration. Multinational activity (foreign integration) accounts for a gain of 0.71% only, while the rest is due to integrated activity at home. The magnitude of this number should not be underestimated. What this exercise says is that the change from a world without vertical FDI to a world where vertical FDI amounts to 13% of total imports generates an increase of 0.71% in consumption per capita. We are opening the economy only to match a volume of FDI that is relatively small in the data. Nonetheless, the induced welfare gain is sizeable.⁴³

Table 2 also shows another decomposition of the gains, which quantifies the fact that the model produces gains from integration through two different channels. On one hand, there is a "productivity effect": with integration, a share of inputs is sourced at marginal cost. With foreign integration, the parent's productivity is matched to lower wages in the host country, further increasing the gains for the parent's country.⁴⁴ In the host country, entry of foreign

⁴²The calibrated parameters imply that in equilibrium $w_{us} = \tau w_{row}$. Hence U.S. firms set plants both domestically and abroad, while ROW firms integrate only domestically.

⁴³Ramondo and Rodríguez-Clare (2009) compute the welfare gains from horizontal multinational production. They find welfare gains of about 3% for the U.S., much larger than in this exercise, and consistent with the observation that horizontal FDI flows are larger than vertical ones: they report that horizontal multinational sales of U.S. companies are equal to 16% of GDP.

 $^{^{44}}$ The productivity effect is similar to the one analyzed by Grossman and Rossi-Hansberg (2008).

	Baseline calibration $(\eta = 1.8, \tau = 2.65)$	FDI reform $(\eta = 1.8, \tau' = 1.82)$	Higher market power $(\eta' = 1.2, \tau = 2.66)$
U.S.			
welfare gains $(\%)$	4.87	6.95	13.29
domestic integration	4.16	0	11.35
foreign integration	0.71	6.95	1.94
productivity effect	4.63	6.41	4.41
competition effect	0.24	0.54	8.88
implied share of U.S. intrafirm import (%)	13.7	62.52	13.7
ROW			
welfare gains $(\%)$	12.39	14.87	23.1
domestic integration	11.84	11.84	22.58
foreign integration	0.55	3.33	0.52
productivity effect	11.76	14.35	14.37
competition effect	0.63	0.52	8.73

Table 2: Gains from multinational production.

firms increases labor demand, and as a result the relative wage increases. On the other hand there is a "competition effect": suppliers in both countries shrink their mark-ups in response to the possibility of integration on the side of their potential buyers. This produces an overall decrease in the price of inputs. To disentangle these two effects, I compute the model for a hypothetical world where firms do not have the possibility of integrating, but suppliers do reduce their prices as if there were the possibility of integrating. The difference between this instrumental model and the model without integration isolates the competition effect. The residual gain is to be attributed to the productivity effect.⁴⁵

In the calibrated economy, the competition effect accounts for a small portion (4.93%) of the total welfare gain for the U.S., which is hence mostly attributable to the large productivity

⁴⁵The welfare gain induced by the productivity effect can also be interpreted as the welfare gain of the "competitive version" of the model. If suppliers had no market power, arm's length transactions would be priced at marginal costs, and sourcing patterns and welfare gains would be driven purely by productivity differences.

differential between the two regions.⁴⁶ The bottom portion of the table presents the same calculations for the ROW economy. The gain from incoming foreign firms is small (0.55% of consumption per capita) because the limited entry of U.S. integrated firms has only a small effect on ROW's wages. As for the decomposition of the gains, the competition effect is small, widely dominated by the large productivity gain of the (domestically) integrated firms.

The second column of Table 2 reports the same calculations performed in a world where the unit cost of integrated production abroad drops of 50%. I refer to this experiment as to an "FDI reform", like an institutional liberalization, that in the model takes the form of a reduction in the parameter τ .⁴⁷ As expected, a drop in the cost of FDI increases the welfare gains: compared to the baseline calibration, consumption levels are significantly higher in both countries. The drop in τ generates a shift in the world allocation of production: all integrated activity of U.S. firms is now happening abroad, and the 6.95% gain in consumption per capita is entirely due to foreign production.⁴⁸ Notice also the sensitivity of the computed intrafirm import share to changes in τ : a 50% drop in τ increases the share of intrafirm imports almost five-fold. The larger extent of foreign investment in ROW countries also increases ROW's relative wage and induces the 3.03% gain in consumption per capita attributable to the entry of foreign firms. The decomposition of the gains in competition and productivity effects is basically unchanged with respect to the baseline scenario, with the large majority of the gains coming from the productivity differential across the two regions.

The third column of Table 2 reports the results of the alternative calibration, with a lower value of the elasticity of substitution: $\eta' = 1.2$. This version of the calibration corresponds to a world where the degree of differentiation across intermediates is higher. As a

⁴⁶Notice that the magnitude of the competition effect also depends on the number of countries involved, so a multilateral calibration exercise, where the ROW is composed by many countries, could generate larger gains through this channel.

⁴⁷An example of such liberalization could be any legislative action "to increase intellectual property protection and to provide the legal conditions for the participation of transnational corporations in the privatization of state industries" (see UNCTAD (1993)). It is true that most episodes of liberalization involve measures designed to facilitate both trade and FDI. Nonetheless, there have been examples of liberalizations explicitly targeting FDI, like for example the inclusion of FDI-related issues in the Uruguay Round agreement and the Multilateral Agreement on Investment in the OECD (see UNCTAD (1996)).

⁴⁸After the drop in τ , $w'_{us} > \tau' w_{row}$.

result, competition is lower and suppliers have more market power. In this setup, gains from opening to intrafirm trade are higher than in the baseline calibration, because the possibility of integration reduces more significantly the suppliers' market power and boosts competition in the economy.⁴⁹ More importantly, under this scenario the second decomposition of the gains looks significantly different: a much larger share is due to the competition effect (66.82% in the U.S., 37.79% in the ROW), because suppliers reduce their mark-ups more than in the previous cases.⁵⁰ We can interpret this result as describing how the decomposition of the gains varies across sectors characterized by more/less substitutability across intermediates. The model predicts that sectors where inputs are more differentiated gain from opening to FDI mostly due to the effect of increased competition on prices. Conversely, those sectors where goods are less differentiated gain mostly from the improved use of less costly technologies.

5.3 Gains from Trade and Multinational Production: from Autarky to Free Trade and FDI

In this section, I compare the gains implied by the model with what other authors have found using different underlying theories. Using the calibrated parameters, I compute consumption per capita in the two countries in the autarky case, in which barriers to trade and FDI are prohibitively high and there is no foreign sourcing. I allow for domestic integration in the autarky economy, so that the gains from domestic integration are not reflected in the calculations. I normalize the results, and compute the welfare gain (expressed in terms of increases in consumption per capita) arising from opening the economy to free trade (t = 1), but not allowing foreign integration $(\tau \to \infty)$. The results for the two countries are displayed in the second row of Table 3.

The gain for the U.S. is 13% of consumption per capita, close to estimates from other

⁴⁹This result is consistent with Rauch (1999), who finds that the impact of trade barriers is lower on commodities that on differentiated goods. Accordingly, in my model the effect of the removal of barriers to FDI is larger, the larger the degree of differentiation across goods.

 $^{^{50}}$ In the baseline calibration, the possibility of integrated production generates a reduction in average domestic mark-ups of 8.2%, while under this scenario mark-ups shrink of 42.33%.

	U.S.	ROW
autarky	1	1
costless trade and no FDI	1.13	1.05
costless trade and costless FDI	1.23	1.06

Table 3: Gains from autarky to costless trade and FDI.

papers: Alvarez and Lucas (2007) estimate a gain of 10%, while Eaton and Kortum (2002) obtain a gain of 17%. The gains in the ROW are smaller, but still significant. The calibration exercise in this paper is bilateral, and less ambitious in scope compared to the ones in Eaton and Kortum (2002) and Alvarez and Lucas (2007). Nonetheless, the computed welfare gains for the U.S. economy are of a similar order of magnitude. The fact that the model matches the order of magnitude of the gains from trade produced by somehow different quantitative exercises raises one's confidence on the model's ability to generate reasonable gains from FDI.

The third row of the table computes consumption in a frictionless world ($t = \tau = 1$). Opening to costless FDI implies an additional increase in consumption per capita of about 9% for the U.S. and 1% for the ROW, for a total gain with respect to autarky of 23% and 6% respectively. Rodríguez-Clare (2007) estimates that the combined gains from trade and diffusion of ideas across countries can reach about 200% of consumption, depending on the relative importance of a country's research intensity. Given their large role in total world research, the gains for the U.S. are much lower than this upper bound, reaching about 10% of consumption. Compared to Rodríguez-Clare's analysis, my model concentrates the attention on a very specific channel of diffusion – vertical FDI –, nonetheless the computed total gains for the U.S. are larger. This feature depends on the different source of the gains I consider. In Rodríguez-Clare (2007), countries profit from openness because they can get access to ideas generated in other countries, so the gains are limited for a country, like the U.S., that accounts for the majority of world research. In my model, the gains arise from the match of "good ideas" (high productivity draws) with low labor costs, so even a country that accounts for the totality of world research can benefit from opening.

Burstein and Monge-Naranjo (2009) compute the welfare gains from relocating firms' managerial know-how to control factors of production abroad. Their computed gains (for an average host country) are about 5% and 3% of consumption per capita for unilateral and multilateral liberalizations respectively, while the source countries lose from the relocation. The relocation of managerial ability to control foreign factors of production in Burstein and Monge-Naranjo (2009) may be compared to the technology transfer in this paper. In their model though, the relocation of managerial ability abroad entails a loss of productivity at home, which explains the smaller gains, especially in the case of multilateral liberalizations.

6 Conclusions

This paper proposes a new general equilibrium framework aimed to explain the decisions of firms to fragment their production processes across national borders, both in terms of location and organizational structure, through the choice of outsourcing *versus* insourcing input production. Firms' optimal sourcing strategies are the outcome of a market equilibrium, where the major role is played by technology heterogeneity and by the implications of imperfect competition on prices. Multinational corporations arise endogenously when firms decide to integrate production in foreign countries.

I study optimal pricing in presence of multinational firms: imperfect competition creates a wedge between trade prices and transfer prices. The possibility of integration induces downward pressure on arm's length prices, establishing a link between trade and FDI liberalization and equilibrium prices. The model has predictions for the dependence of aggregate trade volumes on the economy's fundamentals: volumes of imports and FDI are inversely related to the parent's country size and increase with cross-country heterogeneity. Trade occurs among identical countries, while a certain degree of heterogeneity is necessary to give rise to vertical FDI. The dispersion of the cost distributions across firms also affects the choice of the sourcing strategy, in the sense that the prevailing sourcing strategy in a country is the one associated with the lowest cost dispersion. The general equilibrium structure of the model makes it amenable to perform welfare calculations. I calibrate the model to match aggregate U.S. data and compute the implied gains from multinational production and intrafirm trade. The welfare gains are currently about 1% of consumption per capita, and the model shows that further liberalization would substantially increase them.

Extensions of the model should be devoted to a more flexible characterization of the FDI technology, able to reproduce multilateral patterns that we observe in the data. Nonetheless, I believe the analysis conducted here is a useful starting point to get a deeper understanding of the role of technology and market structure in shaping firms' sourcing decisions, and of the welfare consequences of this aspect of globalization.

Appendix

A On Optimal Pricing in the Open Economy

In this section I show some results on the properties of the pricing rule (solution of equation (22)), drawing from the analogy with the auction theory literature.

I. Symmetric Countries, Frictionless World

The symmetric version of the pricing problem (for $m_i = w_i = w \ \forall i, t = \tau = 1$) satisfies all the conditions describing a symmetric first-price sealed-bid auction: each supplier observes his own cost (his "valuation"), but not the ones of his competitors, and costs are independently distributed according to a common density $\psi(\cdot)$ (independent private values). Moreover, the revenues in case of "win" are equal to the optimal price declared (the "bid"). McAfee and McMillan (1987) illustrate that this problem admits a Bayes-Nash equilibrium, where optimal prices are equal across countries and are given by:

$$p(z) = wz + \frac{\int_{z}^{\infty} wp(z)^{-\eta} \left[1 - \Phi\left(\frac{p(z)}{w}\right)\right] \left[1 - \Psi(\xi)\right]^{N-1} d\xi}{p(z)^{-\eta} \left[1 - \Phi\left(\frac{p(z)}{w}\right)\right] \left[1 - \Psi(z)\right]^{N-1}}.$$
 (A.1)

For exponentially distributed draws $(\phi(x) = \lambda e^{-\lambda x} \text{ for } x \ge 0 \text{ and } \psi(z) = \mu e^{-\mu z} \text{ for } z \ge 0)$, the optimal symmetric pricing rule p(z) is the positive root of:

$$\eta w z + p(z)(1-\eta) + -p(z)[p(z) - wz] \left(\frac{\lambda}{w} + (N-1)\mu\right) = 0.$$
 (A.2)

II. Symmetric Productivities, Positive Trade and FDI Costs

The structure of this problem allows to extend the result of the previous paragraph to the case in which trade and FDI costs are positive and heterogeneous across countries, provided that the suppliers' cost distribution is common across countries: $\Psi_i(\cdot) = \Psi(\cdot) \forall i$. This introduces differences in the distribution of the marginal costs across suppliers, but these differences are deterministic, and preserve the tractability of the problem. To understand this, notice the following: in selling to market *i*, a supplier from country *j* must charge a price $p_{ij}(z_j)$ which is below the prices charged by suppliers in other countries. This happens with probability equal to:

$$prob\{p_{ij}(z_j) = \min_k p_{ik}(z_k)\} = \prod_{k \neq j} prob\{p_{ij}(z_j) \le p_{ik}(z_k)\}$$
$$= \prod_{k \neq j} prob\{t_{ij}w_jz_j \le t_{ik}w_kz_k\}$$
$$= \prod_{k \neq j} \left[1 - \Psi\left(\frac{t_{ij}w_jz_j}{t_{ik}w_k}\right)\right].$$

Hence the optimal pricing rule can be written as:

$$p_{ij}(z_j) = t_{ij}w_j z_j + \frac{\int_{z_j}^{\infty} t_{ij}w_j p_{ij}(z_j)^{-\eta} \left[1 - \Phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\right] \prod_{k \neq j} \left[1 - \Psi\left(\frac{t_{ij}w_j \xi}{t_{ik}w_k}\right)\right] d\xi}{p_{ij}(z_j)^{-\eta} \left[1 - \Phi\left(\frac{p_{ij}(z_j)}{m_i}\right)\right] \prod_{k \neq j} \left[1 - \Psi\left(\frac{t_{ij}w_j z_j}{t_{ik}w_k}\right)\right]}$$
(A.3)

which, under exponentially distributed draws, is equivalent to equation (23).

III. Asymmetric Countries

The results of the previous section cannot be extended to the case of asymmetric

cost distributions $\Psi_i(\cdot)$. Proofs of existence of a Bayes-Nash equilibrium in first-price sealed-bid asymmetric auctions require the cost distributions $\Psi_i(\cdot)$ to be defined on a compact interval.⁵¹ For tractability reasons, I prefer not to introduce bounds for the cost distributions. For each destination country *i*, I solve numerically the system of equations (22) with the following algorithm:

- 1. Given an initial vector of productivities z_j for each country, guess an initial pricing rule $p_{ij}^0(z_j)$ for each origin country j to destination country i.
- 2. Compute $F_{ij}^{0}(p_{ij}^{0}(z_{j}))$.
- 3. Given $F_{ij}^0(\cdot)$, find the pricing rules $p_{ij}^1(z_j)$ (one for each origin country j) that solve (22).
- 4. If $p_{ij}^1(z_j) p_{ij}^0(z_j) < \varepsilon$, $\forall j$, STOP. If $p_{ij}^1(z_j) - p_{ij}^0(z_j) \ge \varepsilon$, $\forall j$, compute $F_{ij}^1(p_{ij}^1(z_j))$, and continue iterating until convergence.

If a solution exist, it must lie between the marginal cost and the closed-economy pricing rule. I use the closed-economy pricing rule to initialize the algorithm.

B Existence of the Equilibrium

This section contains the proof of Proposition 1. To show the existence of the equilibrium in the two-country case, it is sufficient to show that the excess demand correspondence ED_h is continuous and that $\exists \underline{w}_h, \overline{w}_h$ such that $ED_h(\underline{w}_h) > 0$ and $ED_h(\overline{w}_h) < 0$.

It is clear from the construction of the model that – provided that the pricing rules are continuous – the excess demand is differentiable (hence continuous) almost everywhere. The only two points where the excess demand correspondence is not differentiable are $w_h = \tau$ and $w_h = 1/\tau$. At these wage levels, firms switch the location of production, and labor demand is not differentiable. The labor demand for integrated segments of firms from country H is:⁵²

⁵¹See Maskin and Riley (2000).

⁵²The labor demand for integrated segments of firms from country F is constructed in the same way, with the non-differentiability at $w_h = 1/\tau$.

$$l_{h}^{I} \begin{cases} = \frac{p_{h}^{\eta}q_{h}}{w_{h}} \int_{B_{h}^{I}} (w_{h}x_{h})^{1-\eta} \phi_{h}(x_{h}) \psi(z) dx_{h} dz & ; \text{ if } w_{h} < \tau \\ \in \left[\frac{p_{h}^{\eta}q_{h}}{w_{h}} \int_{B_{h}^{I}} (w_{h}x_{h})^{1-\eta} \phi_{h}(x_{h}) \psi(z) dx_{h} dz, \frac{p_{h}^{\eta}q_{h}}{w_{s}} \int_{B_{h}^{I}} (\tau w_{f}x_{h})^{1-\eta} \phi_{h}(x_{h}) \psi(z) dx_{h} dz \right] & ; \text{ if } w_{h} = \tau \\ = \frac{p_{h}^{\eta}q_{h}}{w_{s}} \int_{B_{h}^{I}} (\tau w_{f}x_{h})^{1-\eta} \phi_{h}(x_{h}) \psi(z) dx_{h} dz & ; \text{ if } w_{h} > \tau \end{cases}$$

where l_h^I takes values in a closed interval for $w_h = \tau$, and $\lim_{w_h^- \to \tau} l_h^I \in l_h^I(\tau)$, $\lim_{w_h^+ \to \tau} l_h^I \in l_h^I(\tau)$. This is sufficient to ensure continuity of ED_h at $w_h = \tau$ (and similarly at $w_h = 1/\tau$).

On the second point, it is sufficient to compute the limits of the excess demand correspondence for $w_H \to 0$ and $w_H \to \infty$. Using the definition of profits, the first order conditions in the final good market and the population constraint, the excess demand correspondence can be written as:

$$ED(w_h) = L_h w_h + (p_{hh}^T)^{1-\eta} p_h^{\eta} q_h + (p_{fh}^T)^{1-\eta} p_f^{\eta} q_f - w_h (l_{hh}^T + l_{fh}^T) - r_h c_h$$

= $\left[\frac{(p_{hh}^I)^{1-\eta} + (p_{hh}^T)^{1-\eta}}{p_h^{1-\eta}} - 1 \right] p_h q_h + \left[\frac{(p_{fh}^I)^{1-\eta} + (p_{fh}^T)^{1-\eta}}{p_f^{1-\eta}} \right] p_f q_f$

where, given the definition of the price indexes:

$$\left[\frac{(p_{hh}^{I})^{1-\eta} + (p_{hh}^{T})^{1-\eta}}{p_{h}^{1-\eta}} - 1\right] \in (-1,0) \text{ and } \left[\frac{(p_{fh}^{I})^{1-\eta} + (p_{fh}^{T})^{1-\eta}}{p_{f}^{1-\eta}}\right] \in (0,1).$$

Since prices are increasing in wages:

$$\lim_{w_h \to 0} p_h = \lim_{w_h \to 0} p_f = 0 \text{ and } \lim_{w_h \to \infty} p_h = \lim_{w_h \to \infty} p_f = \infty.$$

It remains to determine the limits of q_h , q_f . The term q_h can be rewritten as:

$$q_h = \frac{\alpha w_h \left[(1-\alpha) p_f + \alpha w_f k_{ff} \right] L_h - \alpha^2 w_h w_f k_{fh} L_f}{\left[(1-\alpha) p_h + \alpha w_h k_{hh} \right] \left[(1-\alpha) p_f + \alpha w_f k_{ff} \right] - \alpha^2 w_h w_f k_{fh} k_{hf}}$$

$$= \left\{ \frac{(1-\alpha)p_{h}}{\alpha w_{h}L_{h}} + \frac{k_{hh}}{L_{h}} - \frac{\alpha w_{f}k_{fh}k_{hf}}{[(1-\alpha)p_{f} + \alpha w_{f}k_{ff}]L_{h}} \right\}^{-1} \dots \\ \dots - \left\{ \frac{[(1-\alpha)p_{h}/w_{h} + \alpha k_{hh}][(1-\alpha)p_{f} + \alpha w_{f}k_{ff}]}{\alpha^{2}w_{f}k_{fh}L_{f}} - \frac{k_{hf}}{L_{f}} \right\}^{-1}$$

When $w_n \to 0$, the term $\frac{(1-\alpha)p_h}{\alpha w_h L_h} + \frac{k_{hh}}{L_h}$ is a positive constant; the term $\frac{\alpha w_f k_{fh} k_{hf}}{[(1-\alpha)p_f + \alpha w_f k_{ff}]L_h}$ is a $\frac{0}{0}$ indeterminacy; the term $\frac{[(1-\alpha)p_h/w_h + \alpha k_{hh}]}{\alpha^2 w_f k_{fh} L_f}$ is a positive constant, and the term $[(1-\alpha)p_f + \alpha w_f k_{ff}]$ tends to zero, so $\frac{[(1-\alpha)p_h/w_h + \alpha k_{hh}][(1-\alpha)p_f + \alpha w_f k_{ff}]}{\alpha^2 w_f k_{fh} L_f}$ tends to zero; the term $\frac{k_{hf}}{L_f}$ tends to zero. As a result: $\lim_{w_h \to 0} q_h = -\infty$. Similarly, one can show that $\lim_{w_h \to \infty} q_h = \infty$, $\lim_{w_h \to 0} q_f = \infty$, $\lim_{w_h \to \infty} q_f = -\infty$. It is immediate to conclude that:

$$\lim_{w_h \to 0} ED(w_h) = \infty \text{ and } \lim_{w_h \to \infty} ED(w_h) = -\infty.$$

References

- Alvarez, Fernando, and Robert E. Lucas. 2007. "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade." Journal of Monetary Economics 54 (6): 1726–1768.
- Antràs, Pol. 2003. "Firms, Contracts, and Trade Structure." Quarterly Journal of Economics 118 (4): 1375–1418.
- Antràs, Pol, and Elhanan Helpman. 2004. "Global Sourcing." Journal of Political Economy 112 (3): 552–580.
- Bernard, Andrew B., Jonathatan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003.
 "Plants and Productivity in International Trade." *The American Economic Review* 93 (4): 1268–1290.
- Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. 2006. "Transfer Pricing by U.S.-Based Multinational Firms." NBER Working Paper 12493.
 - ——. 2009. "Importers, Exporters, and Multinationals: a Portrait of Firms in the U.S. that Trade Goods." In T. Dunne, J.B. Jensen and M.J. Roberts (eds.), *Producer Dynamics: New Evidence from Micro Data.* Chicago: University of Chicago Press.

- Burstein, Ariel, and Alexander Monge-Naranjo. 2009. "Foreign Know-How, Firm Control, and the Income of Developing Countries." Quarterly Journal of Economics 124 (1): 149–195.
- Campa, José, and Linda S. Goldberg. 1997. "The Evolving External Orientation of Manufacturing Industries: Evidence from Four Countries." Federal Reserve Bank of New York Economic Policy Review 3 (2): 53–81.
- Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography and Trade." *Econometrica* 70 (5): 1741–1779.
- Feenstra, Robert C. 1998. "Integration of Trade and Disintegration of Production in the Global Economy." The Journal of Economic Perspectives 12 (4): 31–50.
- Garetto, Stefania. 2009. "Firms' Heterogeneity and Incomplete Pass-Through." Mimeo, Boston University.
- Grossman, Gene M., and Elhanan Helpman. 2002. "Integration Versus Outsourcing in Industry Equilibrium." *Quarterly Journal of Economics* 117 (1): 85–120.
- ———. 2004. "Managerial Incentives and the International Organization of Production." Journal of International Economics 63:237–262.
- Grossman, Gene M., and Esteban Rossi-Hansberg. 2008. "Trading Tasks: A Simple Theory of Offshoring." The American Economic Review 98 (5): 1978–1997.
- Hanson, Gordon H., Raymond J. Mataloni, , and Matthew J. Slaughter. 2001. "Expansion Strategies of U.S. Multinational Firms." In D. Rodrik and S.M. Collins (eds.), Brookings Trade Forum 2001.
- ———. 2005. "Vertical Production Networks in Multinational Firms." Review of Economics and Statistics 87 (4): 664–678.
- Helpman, Elhanan. 1984. "A Simple Theory of International Trade with Multinational Corporations." Journal of Political Economy 92 (3): 451–471.
- ———. 1985. "Multinational Corporations and Trade Structure." *Review of Economic Studies* 52 (3): 443–457.
- ——. 2006. "Trade, FDI, and the Organization of Firms." NBER Working Paper 12091.

- Helpman, Elhanan, and Paul Krugman. 1985. Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. Cambridge, MA: MIT Press.
- Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple. 2004. "Exports Versus FDI with Heterogeneous Firms." *American Economic Review* 94 (1): 300–316.
- Hummels, David, Jun Ishii, and Kei-Mu Yi. 2001. "The Nature and Growth of Vertical Specialization in World Trade." Journal of International Economics 54:75–96.
- Hummels, David, Dana Rapoport, and Kei-Mu Yi. 1998. "Vertical Specialization and the Changing Nature of World Trade." Federal Reserve Bank of New York Economic Policy Review 4 (2): 79–99.
- Markusen, James R. 1984. "Multinationals, Multi-Plant Economies, and the Gains from Trade." Journal of International Economics 16:205–226.
- Maskin, Eric, and John Riley. 2000. "Equilibrium in Sealed High Bid Auctions." *Review* of *Economic Studies* 67 (3): 439–454.
- McAfee, Preston, and John McMillan. 1987. "Auctions and Bidding." Journal of Economic Literature 25 (2): 699–738.
- Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–1725.
- Melitz, Marc J., and Gianmarco I.P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies* 75 (1): 295–316.
- Neiman, Brent. Forthcoming 2010. "Stickiness, Synchronization, and Passthrough in Intrafirm Trade Prices." *Journal of Monetary Economics*.
- Nocke, Volker, and Stephen Yeaple. 2008. "An Assignment Theory of Foreign Direct Investment." *Review of Economic Studies* 75 (2): 529–557.
- Ramondo, Natalia, and Andrès Rodríguez-Clare. 2009. "Trade, Multinational Production, and the Gains from Openness." Mimeo, Pennsylvania State University.
- Rauch, James E. 1999. "Networks Versus Markets in International Trade." Journal of International Economics 48:7–35.

- Rodríguez-Clare, Andrès. 2007. "Trade, Diffusion, and the Gains from Openness." Mimeo, Pennsylvania State University.
- Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2007. "Establishment Size Dynamics in the Aggregate Economy." *The American Economic Review* 97 (5): 1639–1666.
- UNCTAD. 1993. World Investment Report; Transnational Corporations and Integrated International Production. New York: United Nations.
- ——. 1996. World Investment Report; Investment, Trade, and International Policy Arrangements. New York: United Nations.
- Yeats, Alexander J. 2001. "Just How Big is Global Production Sharing?" In Sven W. Arndt and Henryk Kierzkowski (eds.), Fragmentation: New Production Patterns in the World Economy. Oxford: Oxford University Press.