Input Sourcing and Multinational Production*

Stefania Garetto†

Boston University and Princeton University

November 17, 2008

Abstract

A large portion of world trade happens within firms' boundaries. This paper proposes a new general equilibrium framework where firms decide whether to outsource to unaffiliated suppliers or to integrate input manufacturing. Multinational corporations and intrafirm trade arise endogenously when firms integrate production in foreign countries. By outsourcing, firms benefit from suppliers' good technologies, but pay a mark-up price. Intrafirm sourcing allows to save on mark-ups and to match a firm's productivity with possibly lower foreign wages. Imperfect competition establishes a link between FDI liberalization and optimal pricing: suppliers find optimal to reduce their prices in response to the possibility of insourced production (the "pro-competition effect" of multinationals). The model is calibrated to match aggregate U.S. trade data, and used to quantify the gains arising from vertical multinational production and intrafirm trade. The computed gains are currently about 1% of consumption per capita, and the model shows that further 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 School of Business), the University of Iowa, the University of Toronto, and the 2008 SED Meeting. All errors are mine.

†Contact: Department of Economics, International Economics Section, Princeton University, Fisher Hall, Princeton, NJ 08540. E-mail: sgaretto@princeton.edu.

1 Introduction

Globalization has expanded the scope of trade, in the sense that trade in finished products is being gradually outpaced by trade in intermediates[1], taking place both within and across the boundaries of the firm. Many studies document the growth of multistage production[2], in which plants in different locations contribute to the creation of value added through processing and assembly. 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 (Forthcoming) 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 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[3] (henceforth, FDI) in the form of integrated production abroad. 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 my model, firms simply choose the sourcing options and the locations that minimize their production costs. The driving force behind this choice is technology: firms are heterogeneous in productivity and in the type of technology they use. Some of them have an adaptable technology with which they produce an input that can be sold to any other firm in the world (I call these firms the intermediate goods producers,

[1]See Yeats (2001).

[2]See Campa and Goldberg (1997), Hummels, Ishii and Yi (2001), Hummels, Rapoport and Yi (1998), and Hanson, Mataloni and Slaughter (2001, 2005).

[3]In 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"*.

or suppliers). Other firms are endowed with two types of technologies: a homogeneous technology to produce the final 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. I call these firms the final good producers, and give them the option of buying inputs from the suppliers or to produce them with their non adaptable "in-house" technologies. When these firms decide to produce abroad, they become the parents of a multinational corporation. Offshore production takes the form of vertical FDI, and – when the inputs produced offshore are shipped back to the parent – flows of intrafirm trade.

The incentive to insource production is two-fold: on the one hand, firms can exploit a technology that may be better than the one of the suppliers, by matching the potentially high productivity of the parent with the possibly lower labor costs of the location chosen[4]. On the other hand, in an imperfectly competitive market, by integrating firms save on the mark-ups charged by the suppliers: intrafirm trade happens between a firm and itself, and is priced at marginal cost[5].

In summary, technology heterogeneity and imperfect competition are the two main features of the theory. Imperfect competition establishes a link between trade liberalization, competition and optimal pricing in presence of multinational firms. Multinationals add a new margin of competition among suppliers, who need to adjust down their prices in order to survive in a global market where they have to compete with other suppliers around the world and with the buyers' possibility of integrated production abroad. FDI liberalization makes both trade and FDI more profitable, increases competition and lowers prices. Firms' heterogeneity implies that the price adjustments vary depending on the productivity (or the size) of the single firms. In response to trade competition and multinationals' competition, suppliers optimally decide to do pricing-to-market, *i.e.* to charge different prices in different countries.

On aggregate, the theory predicts that firms outsource mostly from suppliers located in large countries. Volumes of arm's length and intrafirm imports increase with cross-country heterogeneity and, while trade occurs also between identical countries, a certain degree of heterogeneity is necessary to give rise to vertical FDI and intrafirm trade[6]. The dispersion of the cost distributions across

[4]Hanson, Mataloni and Slaughter (2005) document the importance of low-cost locations in vertical production networks. The choice of offshore locations based on factor cost differences is present also in Grossman and Rossi-Hansberg (Forthcoming). While not exploring the insourcing *versus* outsourcing choice, they focus on modeling the growth of the international division of labor and the associated growth of trade in intermediates (or trade in "tasks", to use their terminology).

[5]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.

[6]Nocke and Yeaple (2008) have similar results when modeling the choice of vertical, greenfield investment *versus*

firms also affects the sourcing strategy: the share of intrafirm transactions is larger the higher the productivity dispersion of the suppliers.

The calibrated version of the model quantifies the welfare gains resulting from vertical multinational production[7] in addition to the gains from trade, and distinguishes the relative importance of productivity and technology *versus* market structure and competition for the results. The effect of competition on prices is more relevant in scenarios where the suppliers have a significant level of market power, while the effect of productivity and technology drives most of the results. The welfare gains arising from vertical multinational production are currently about 1% of consumption per capita, and further FDI liberalization would increase them substantially: the model predicts that a 50% drop in the calibrated barrier to FDI would imply a gain of about 7% of consumption per capita.

The rationale behind the existence of multinational firms is similar to Helpman (1984, 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, and the incentive to do so comes from the existence of an immaterial factor[8] of production that may serve product lines without being located in their plants. Similarly, I assume that firms can transfer their productivity when they decide to integrate production abroad. In addition, the model I propose in this paper generalizes Helpman's idea to a world with heterogeneous firms and potentially many countries, adds trading costs and the analysis of optimal pricing to his setup.

More recently, Antràs (2003) and Antràs and Helpman (2004) modeled the joint choice of location and organizational structure by merging existing models of trade[9] with a contract-based theory of the firm[10]. 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

mergers and acquisitions, and report supporting empirical evidence.

[7]The model excludes horizontal FDI, *i.e.* the establishment of offshore production to serve foreign markets.

[8]The idea of modeling multinational production through the existence of an immaterial factor is present also in Markusen (1984). In his setup, multinational corporations (henceforth, MNCs) 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 MNCs increases competition and reduces prices.

[9]See Helpman and Krugman (1985), Melitz (2003) and Helpman, Melitz and Yeaple (2004).

[10]Grossman and Helpman (2002) model the choice of integration *versus* outsourcing, disregarding location, as trading-off the higher costs associated with an integrated firm with the search costs and potential hold-up problems associated with dealing with an outside agent who supplies a good whose quality is non-verifiable. Antràs (2003) and Antràs and Helpman (2004) extend their approach to one where even the integration option does not solve completely the hold-up problem. Grossman and Helpman (2004) examine jointly the organizational form and location choice with a model where is the effort of the agent that is non-observable, and integration and outsourcing differ in the monitoring possibilities that they allow.

analysis that uses firms' heterogeneity and market structure to explain the same choices. Moreover, my theory has the advantage of providing a detailed analysis of optimal pricing, which is absent in Antràs and Helpman's work[11], and can be calibrated to 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 presents the open economy model for the general case of an arbitrary number of countries and characterizes the general equilibrium for the two-country case. Section 4 shows the dependence of aggregate volumes of trade and FDI on the model's parameters. Section 5 contains the calibration of the model and the computation of the gains from multinational production and intrafirm trade. Section 6 concludes.

2 The Model: the Closed Economy

In this section I present the closed economy model, to isolate the choice between outsourcing and integration, setting aside the location choice. 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 final 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

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

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.

The final good is produced by a continuum of identical producers, operating in a perfectly competitive market. They all produce the same, homogeneous consumption good using labor and producer-specific intermediate goods as inputs. For each input, the 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 insource production, his technology allows him to produce only for his own product line. In principle, the final good producer could aquire an adaptable technology (at some cost) to enter also the intermediate goods' market. I assume that that cost is too large to be covered by the expected profits.

The sourcing decision involves comparing the costs of the two options: the in-house cost of production and the outside price. For each of the inputs, 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 producer-specific input. All the final good producers draw from the same cost distribution $\phi(\cdot)$, but they can have different cost draws for each input. Since in the closed economy the wage is normalized to one, the unit labor requirement x is also equal to the unit cost of production. Notice that for both kinds of producers, a low draw implies a low marginal cost of production, so that the "low x" and "low z" producers are the most productive ones.

The outsourcing option is given by the outside price of the good, which I denote $p(z)$ since it is a function of the supplier's cost draw. The buyer takes this price as given, while the intermediate goods producer sets the price based on his marginal cost and the demand function he faces.

Hence each final good producer sees 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 sourcing:

$$\begin{aligned} \min_{q(x,z)} \quad & \int_0^\infty \int_0^\infty \min\{x, p(z)\} q(x,z)\phi(x)\psi(z)dxdz \\ s.t. \quad & \left[\int_0^\infty \int_0^\infty q(x,z)^{1-1/\eta}\phi(x)\psi(z)dxdz\right]^{\eta/(\eta-1)} \geq q \end{aligned} \tag{1}$$

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

$$\begin{aligned} \min_{q(x,z)} \quad & \left[\int_0^\infty \int_0^{p(z)} xq(x,z)\phi(x)\psi(z)dxdz + \int_0^\infty p(z) \int_{p(z)}^\infty q(x,z)\phi(x)\psi(z)dxdz\right] \\ s.t. \quad & \left[\int_0^\infty \int_0^\infty q(x,z)^{1-1/\eta}\phi(x)\psi(z)dxdz\right]^{\eta/(\eta-1)} \geq q. \end{aligned} \tag{2}$$

This problem maps the goods – previously defined on a bi-dimensional space – on a one-dimensional space where they are denoted by their minimum cost. 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 (2) 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 buy and $q^T(x, z)$ be the solution of (2) in B^T. Hence:

$$q^I(x,z) \equiv q^I(x,p(z)) = \left(\frac{x}{p}\right)^{-\eta} q \qquad \forall\ (x,z) \in B^I \tag{3}$$

$$q^T(x,z) \equiv q^T(x,p(z)) = \left(\frac{p(z)}{p}\right)^{-\eta} q \quad \forall\ (x,z) \in B^T. \tag{4}$$

[12] Assuming a continuum of goods implies that – by the law of large numbers – the aggregate q will be the same across final good producers even allowing them to have different cost draws for each of the goods.

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

$$p = \left[p_I^{1-\eta} + p_T^{1-\eta} \right]^{1/(1-\eta)} \tag{5}$$

and:

$$p_I = \left[\int_0^\infty \int_0^{p(z)} x^{1-\eta} \phi(x) \psi(z) dx dz \right]^{1/(1-\eta)} \tag{6}$$

$$p_T = \left[\int_0^\infty p(z)^{1-\eta} \Big[1 - \Phi(p(z)) \Big] \psi(z) dz \right]^{1/(1-\eta)} \tag{7}$$

where p_I is the aggregate price of integrated goods and p_T is the aggregate price of outsourced (or traded) 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. An intermediate goods producer with random draw z solves:

$$\max_{p(z)} \; [p(z) - z] \int_{p(z)}^\infty q^T(x, p(z)) \phi(x) dx \tag{8}$$

where $q^T(x, p(z))$ is given by equation (4), $\int_{p(z)}^\infty \phi(x) dx$ is the mass of buyers that decide to buy the good at price $p(z)$, and z is the unit cost of production since wages are normalized to one. The first order condition for this problem is:

$$[p(z) - z] \left[q^T(x, p(z)) \phi(p(z)) - \frac{\partial q^T(x, p(z))}{\partial p(z)} [1 - \Phi(p(z))] \right] = q^T(x, p(z)) \int_{p(z)}^\infty \phi(x) dx. \tag{9}$$

Equation (9) summarizes the supplier's trade-off. For a given level of sales (the right hand side of (9)), 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 (this is the extensive margin, captured by the term $q^T(x, p(z))\phi(p(z))$) and he is going to sell lower quantities to the remaining buyers (this is the intensive margin, captured by the term $-\frac{\partial q^T(x,p(z))}{\partial p(z)}[1 - \Phi(p(z))]$).

Using (4), problem (8) becomes:

$$\max_{p(z)} [p(z) - z] \left(\frac{p(z)}{p}\right)^{-\eta} q[1 - \Phi(p(z))] \tag{10}$$

and the first order condition reduces to:

$$p(z) = \left[1 - \frac{1}{\eta + \frac{\phi(p(z))}{1-\Phi(p(z))} p(z)}\right]^{-1} z. \tag{11}$$

Equation (11) 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. In order to characterize the pricing rule and to describe its properties, I assume that the cost draw distribution $\phi(x)$ is exponential with parameter λ:

Assumption 1. $\phi(x) = \lambda e^{-\lambda x} \qquad for \quad x \geq 0.$

Under Assumption 1, (11) reduces to:

$$\lambda p(z)^2 + (\eta - 1 - \lambda z)p(z) - \eta z = 0. \tag{12}$$

where the parameter λ is an inverse measure of variation of the buyers' cost distribution[13]. Equation (12) admits only one positive solution, which is the profit maximizing price $p(z)$ expressed as a function of the technology draw z and of the model's parameters η and λ. Since $p(z) > z \ \forall z$, each seller gets positive profits: with a continuum of potential buyers, even if a seller's cost draw z is very high, there is always going to be some buyer who has a higher cost and is willing to buy the good from him. Also, since I assume that there are no fixed costs of production, the entire distribution of sellers will be in the market in equilibrium. As expected, profits are higher for the more productive sellers, while they tend to zero as the unit labor requirement z increases.

[13] $1/\lambda$ is equal to the standard deviation of the buyers' cost distribution: when λ increases, the cost distribution becomes less disperse. Even if the exponential law does not allow to disentangle the effects of the mean and the variance of the distribution, the fact that here only the variance matters can be observed by solving equation (11) with a generalized exponential law with both a location and a shape parameter: the location parameter cancels out, showing that the mean of the distribution plays no role in affecting the pricing strategy. I use the standard, one-parameter exponential law because it implies a roughly symmetric and unimodal distribution of log-productivity ($\log(1/x)$), consistent with many empirical studies of productivity variation at the industry level (for example Syverson (2004)).

2.4 Properties of the Pricing Rule

Figure 1 plots the producer's profit-maximizing price as a function of his own draw z for some arbitrary values of the parameters η and λ. The dashed line is the producer's marginal cost, while the dash-dotted 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.

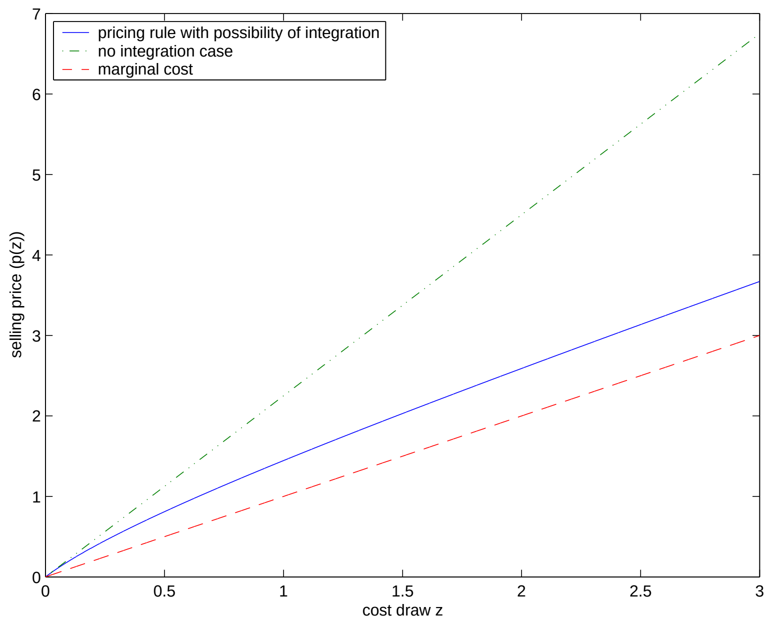

Figure 1: Pricing strategy, closed economy ($\eta = 1.8$, $\lambda = 1$).

Figure 2 shows the producer's mark-up ($(p - z)/z$) as a function of the cost draw z. The dash-dotted line depicts the constant mark-up of the standard model without possibility of integration. The model displays endogenous mark-ups[14], higher for the most productive sellers, whose productivity advantage allows them to keep low prices but to have larger profit margins, and lower for the least productive sellers, who have to charge higher prices to cover their costs, but only get small profit margins. This feature also implies that mark-ups are increasing in the firms' market share and size. Figure 3 shows the firm's mark-up as a function of its market share, where the share of an intermediate goods producer with cost z is given by:

[14]The result of endogenous mark-ups holds for any functional specification of the distribution $\phi(x)$, except for the Pareto law. When the costs of integrated production are distributed according to a Pareto law, the price elasticity of demand is constant and hence mark-ups are constant too (but lower than in the model without integration).

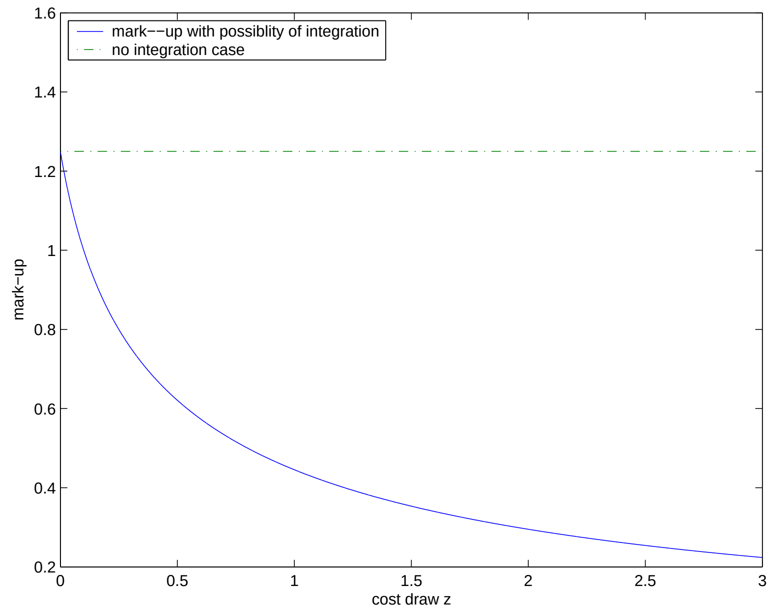

Figure 2: Mark-up over marginal cost, closed economy (η=1.8, λ=1).

$$s(z) \equiv \frac{p(z)q^T(x,p(z))\int_{p(z)}^{\infty}\phi(x)dx}{pq} = \left(\frac{p(z)}{p}\right)^{1-\eta}[1-\Phi(p(z))]\ .$$

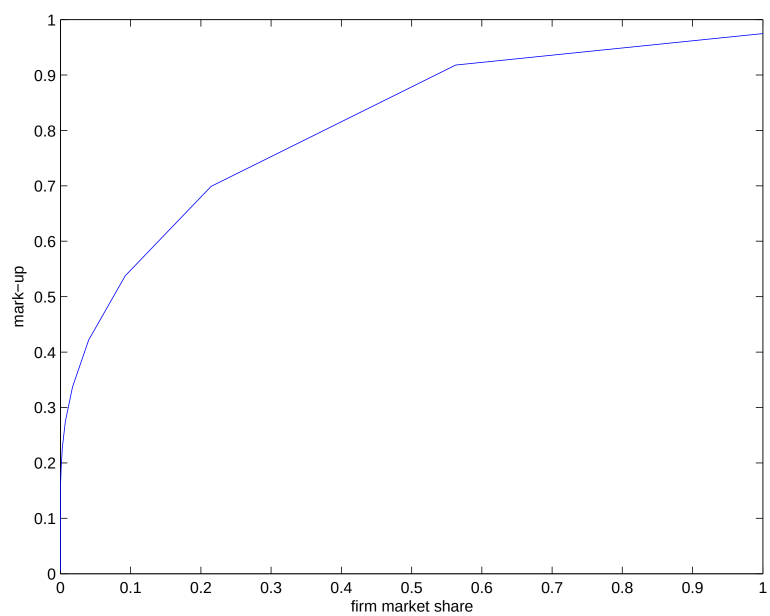

Figure 3: Mark-up as a function of market share, closed economy (η =1.8, λ = 1).

These predictions are aligned with the ones obtained by Melitz and Ottaviano (2008), which achieve mark-ups variability and dependence of profits on productivity by assuming linear demand systems with horizontal product differentiation. Bernard, Eaton, Jensen and Kortum (2003) achieve similar features by assuming Bertrand competition in the intermediate goods sector[15]. The nice feature of this result is that the assumed productivity heterogeneity is not absorbed by variation in prices, but also translates into heterogeneity in measured productivity, expressed as value of output per unit of input[16]. Comparative statics of equation (12) implies that the optimal price is decreasing in η: when the degree of substitutability increases, potential buyers can more easily switch to cheaper substitutes, hence the suppliers must decrease the price to keep their share of the market. 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 the sellers act on the intensive margin charging higher prices and mark-ups.

2.5 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:

$$c = q^{\alpha} l_f^{1-\alpha} \tag{13}$$

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 that $q = \frac{l_i}{k}$, where k is the number of units of labor required to produce 1 unit of the aggregate q:

$$k = p^{\eta} \left[\int_0^{\infty} \int_0^{p(z)} x^{1-\eta} \phi(x) \psi(z) dx dz + \int_0^{\infty} z p(z)^{-\eta} \Big[1 - \Phi(p(z)) \Big] \psi(z) dz \right]. \tag{14}$$

Optimality in the final good market implies that the equilibrium labor allocation and the value

[15] In Bernard, Eaton, Jensen and Kortum (2003), the distribution of mark-ups does not depend on country characteristics and geographic barriers, while it does in Melitz and Ottaviano (2008) and in this paper. The dependence of mark-ups on country characteristics creates a link between trade liberalization and competition, which will be clearer in the open economy section.

[16] As explained in Bernard, Eaton, Jensen and Kortum (2003), measured productivity for one unit of input is given by $p(z)/z$. In a model with constant mark-ups, this magnitude is also constant (independent on the single firm productivity $1/z$), while in a model with variable mark-ups the term $p(z)/z$ is decreasing in z, showing that low-cost, high productivity firms also exhibit high measured productivity.

of q are:

$$l_f = \left(\frac{(1-\alpha)p}{(1-\alpha)p+\alpha k}\right)L \quad ; \quad l_i = \left(\frac{\alpha k}{(1-\alpha)p+\alpha k}\right)L \quad ; \quad q = \left(\frac{\alpha}{(1-\alpha)p+\alpha k}\right)L. \tag{15}$$

Finally, r denotes the zero-profit equilibrium price of the final good:

$$r = \alpha^{-\alpha}(1-\alpha)^{\alpha-1}p^{\alpha}. \tag{16}$$

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 final good, and each of these inputs can be either produced in an integrated facility or bought from a specialized seller. In an open economy, a final good producer can integrate production domestically or abroad, and buy imputs from suppliers located in any country. The optimal sourcing strategy is going to be determined comparing outside prices and costs of production, but taking into account also other variables like wages and trade costs.

When a final good producer decides to integrate production abroad, it generates flows of FDI. I assume that the foreign investment realizes in ownership of the foreign production facility, so when inputs produced abroad are shipped back to the parent, these flows appear in the data as intrafirm trade, precisely as imports from foreign affiliates. I assume that FDI is only vertical in this economy, *i.e.* firms that decide to set a plant abroad do not serve the host country market, but use the foreign facility only to produce inputs for the domestic final good sector. This restriction relies on assuming that the technology of integrated final good producers is not adaptable to serve other firms.

I assume labor is immobile, so that wages[17] may differ across countries. I denote with w_i the wage level in country i ($i = 1, ...N$). A final good producer located in country i has a set of

[17] Wages are going to be endogenously determined in equilibrium.

technology draws $\{x_i\}$, each drawn from the country-specific distribution $\phi_i(x_i)$. If he decides to produce an input himself, he may choose to do so in his own country or abroad. If he decides to produce at home, his marginal 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 pay local wages. Also, it is reasonable to think that production abroad entails some other costs due to the necessity of building a new facility: for simplicity, I model these costs as iceberg costs, implicitly assuming that they are correlated with the size of production. Iceberg costs are bilateral, reflecting characteristics as proximity, common language, religion or past colonization. Also, there may be transportation costs due to the necessity of repatriating the produced inputs for further manufacturing, and legal restrictions to foreign investment, which are also modeled as part of these additional costs. I denote with τ_{ij} the unit iceberg cost for a final good producer from country i to setup production of an input in country j:

Assumption 2. $\tau_{ij} \geq 1 \;\; \forall\, i, j, \;\; \tau_{ij} = 1 \;\; \forall i = j$ *and* $\tau_{ij} \leq \tau_{ik}\tau_{kj} \;\; \forall i, j, k.$

Hence, if a final good producer from country i decides to produce an input for which he has a draw x_i in country j, his unit cost of production is $\tau_{ij} w_j x_i$. As in the closed economy, integrated production is producer's specific, *i.e.*, the final good producer cannot enter the intermediate goods market and sell the internalized good to other firms.

We now turn to the outsourcing option. In the open economy, each country j ($j \in \{1, ...N\}$) has a continuum of suppliers, each of whom produces a unique differentiated input with an adaptable technology that enables him to sell it to any buyer around the world. Each intermediate goods producer 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 $\psi_j(z_j)$[18]. An intermediate goods producer in country j can only hire labor from country j[19], hence his marginal cost of production is $w_j z_j$. On the other hand, he can sell to final good producers worldwide. 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 its good in country i:

Assumption 3. $t_{ij} \geq 1 \;\; \forall\, i, j, \;\; t_{ij} = 1 \;\; \forall i = j$ *and* $t_{ij} \leq t_{ik}t_{kj} \;\; \forall i, j, k.$

[18] The productivity distributions $\{\phi_i(\cdot)\}_{i=1}^N$, $\{\psi_j(\cdot)\}_{j=1}^N$ are mutually independent across countries.

[19] This assumption is motivated by the fact that I think about the final good producers in the model as the (potential) multinational corporations, and about the intermediate goods producers as national suppliers.

Hence, $t_{ij}w_jz_j$ is the marginal 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 from different countries price-compete with each other to sell in a market, and can charge different prices to buyers in different countries (pricing-to-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 $(x_i, z) = (x_i, z_1, z_2, ...z_N)$, which includes the final good producer's technology draw and the draws of the N suppliers of that input around the world. 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 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, by comparing the minimum cost (across countries) of producing an input with the minimum 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 $c_i(x_i, z)$ be the minimum unit cost of good (x_i, z):

$$c_i(x_i, z) = \min_j \{\tau_{ij}w_jx_i \ , \ p_{ij}(z_j)\}. \tag{17}$$

Once decided to integrate, the location of production is determined by the interaction between iceberg costs and market wages. Let m_i denote the cheapest combination[20] of wages and iceberg costs worldwide for integrated production of a final good producer from country i:

$$m_i = \min_k \tau_{ik}w_k. \tag{18}$$

[20]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 will 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 processes, due to the fact that the setup cost is bilateral.

On the other hand, the choice of the location of the trade partner is determined by trade costs and by the cross-country joint productivity distribution of the suppliers, which affects the prices charged.

A final good producer with a set of cost draws $\{x_i\}$ in country i solves:

$$\begin{aligned} \min_{q_i(x_i,z)} \quad & \int_{\Re_+^N}\int_0^\infty c_i(x_i,z)q_i(x_i,z)\phi_i(x_i)\psi(z)dx_idz \\ s.t. \quad & \left[\int_{\Re_+^N}\int_0^\infty q_i(x_i,z)^{1-1/\eta}\phi_i(x_i)\psi(z)dx_idz\right]^{\eta/(\eta-1)} \geq q_i \end{aligned} \tag{19}$$

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 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) and let B_{ij}^T denote the set of goods that he decides to outsource from a producer in country j[21]:

$$\begin{aligned} B_i^I &= \left\{(x_i,z)\in\Re_+^{N+1}: c_i(x_i,z)=m_ix_i\right\} \\ B_{ij}^T &= \left\{(x_i,z)\in\Re_+^{N+1}: c_i(x_i,z)=p_{ij}(z_j)\right\}. \end{aligned}$$

The final good producer's problem may be rewritten as:

$$\begin{aligned} \min_{q_i(x_i,z)} \quad & \left[\int_{B_i^I} m_ix_iq_i(x_i,z)\phi_i(x_i)\psi(z)dx_idz+\sum_{j=1}^{N}\int_{B_{ij}^T}p_{ij}(z_j)q_i(x_i,z)\phi_i(x_i)\psi(z)dx_idz\right] \\ s.t. \quad & \left[\int_{\Re_+^{N+1}} q_i(x_i,z)^{1-1/\eta}\phi_i(x_i)\psi(z)dx_idz\right]^{\eta/(\eta-1)} \geq q_i. \end{aligned} \tag{20}$$

Problem (20) is solved by:

$$q_i^I(x_i,z) = (m_ix_i)^{-\eta}p_i^\eta q_i \quad \forall\ (x_i,z)\in B_i^I \tag{21}$$

$$q_i^T(x_i,z) = [p_{ij}(z_j)]^{-\eta}p_i^\eta q_i \quad \forall\ (x_i,z)\in B_{ij}^T \tag{22}$$

[21]Notice that $B_i^I\cup(\cup_jB_{ij}^T)=\Re_+^{N+1}$.

where p_i is the aggregate price index in country i:

$$p_i = \left[(p_i^I)^{1-\eta} + \sum_{j=1}^{N} (p_{ij}^T)^{1-\eta} \right]^{1/(1-\eta)} \tag{23}$$

and:

$$p_i^I = \left[\int_{B_i^I} (m_i x_i)^{1-\eta} \phi_i(x_i) \psi(z) dx_i dz \right]^{1/(1-\eta)} \tag{24}$$

$$p_{ij}^T = \left[\int_{B_{ij}^T} [p_{ij}(z_j)]^{1-\eta} \phi_i(x_i) \psi(z) dx_i dz \right]^{1/(1-\eta)} . \tag{25}$$

It remains to determine the prices $\{p_{ij}(z_j)\}_{i,j=1}^N$. 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 potential buyers in different countries. By assuming that no resale is possible, I can study the pricing problem country by country. Each supplier in each country hence chooses N prices to charge, one for each country. In choosing the optimal price to charge in a country, a supplier must consider competition from the producers of the same good in the other countries[22] and the fact that the potential buyers have the option of integrating production. Each supplier can observe his own marginal cost, and the parameters of the cost distributions of its potential buyers and of its competitors in other countries. Based on this information, suppliers simultaneously declare a set of prices (one for each country). The price setting mechanism works as a sealed bid auction, where each supplier cannot observe the prices set by its competitors. The resulting equilibrium is a Bayesian Nash equilibrium, as each supplier sets its optimal price based on incomplete information on the cost structure (and hence the payoffs) of its competitors.

Let $b_{ij}(p_{ij}(z_j))$ be the set of technology draws of buyers in country i and of sellers outside

[22] 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 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.

country j such that the buyers in country i decide to buy good (x_i, z) from the seller in country j:

$$b_{ij}(p_{ij}(z_j)) = \{(x_i, \{z_k\}_{k \neq j}) \in \Re_+^N : (x_i, z) \in B_{ij}^T\}. \tag{26}$$

A supplier in country j with productivity draw z_j maximizes its expected profits from sales in country i in the set $b_{ij}(p_{ij}(z_j))$:

$$\max_{p_{ij}(z_j)} \int_{b_{ij}(p_{ij}(z_j))} [p_{ij}(z_j) - t_{ij} w_j z_j] q_i^T(x_i, z) \phi_i(x_i) \prod_{k \neq j} \psi_k(z_k) dx_i dz_k. \tag{27}$$

Using (22), and due to the independence property of the draws' distributions, problem (27) can be restated as:

$$\max_{p_{ij}(z)} [p_{ij}(z_j) - t_{ij} w_j z_j] \left(\frac{p_{ij}(z_j)}{p_i}\right)^{-\eta} q_i A_{ij}(p_{ij}(z_j)) \tag{28}$$

where $A_{ij}(p_{ij}(z_j))$ is the probability that – given the price $p_{ij}(z_j)$ – a final good producer in country i will buy good (x_i, z) from the seller in country j:

$$A_{ij}(p_{ij}(z_j)) = \left[1 - \Phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\right] \cdot \prod_{k \neq j} [1 - F_{ik}(p_{ij}(z_j))] \tag{29}$$

and $F_{ik}(\cdot)$ denotes the cumulative distribution function of the prices charged by sellers in country k to final good producers in country i.

As in the closed economy section, I assume that the cost draw distributions of both sectors in each country i, $\phi_i(x_i)$ and $\psi_i(z_i)$ $(i = 1, ...N)$, are exponentials with parameters λ_i and μ_i respectively:

Assumption 4. $\phi_i(x_i) = \lambda_i e^{-\lambda_i x_i}$ *for* $x_i \geq 0$, $\psi_i(z_i) = \mu_i e^{-\mu_i z_i}$ *for* $z_i \geq 0$.

Under Assumption 4, the first order condition of problem (28) is:

$$p_{ij}(z_j)(1-\eta) + \eta t_{ij} w_j z_j - p_{ij}(z_j)[p_{ij}(z_j) - t_{ij} w_j z_j] \left\{\frac{\lambda_i}{m_i} + \sum_{k \neq j} \frac{f_{ik}(p_{ij}(z_j))}{[1 - F_{ik}(p_{ij}(z_j))]}\right\} = 0. \tag{30}$$

For each z_j, (30) is a nonlinear system[23] of N equations in the N unknowns $p_{ij}(\cdot)$, $j = 1, ...N$. As each supplier competes with suppliers from other countries to sell in country i, $p_{ij}(\cdot)$ is determined

[23] The algorithm to solve (30) is available upon request to the author.

by evaluating the usual effects of substitutability on demand, the average productivity of the potential buyers (the term λ_i/m_i), and how the optimal price compares with the expected price charged by suppliers in other countries (the hazard rate term $\dfrac{f_{ik}(p_{ij}(z_j))}{[1-F_{ik}(p_{ij}(z_j))]}$ is the probability that a final good producer in i buys good (x_i, z) from a supplier in k following an infinitesimal increase in the price charged by the supplier in j, conditional on wanting to buy from the supplier in j before the price increase).

3.3 Properties of the Pricing Rule, $\mu_i = \bar{\mu}, \ \forall i$

The system of equations (30) must be solved numerically, but it is possible to provide some intuition about the behavior of the pricing rule $p_{ij}(z_j)$ by analyzing the solution of the problem for the special case in which the suppliers' average productivity is constant across countries: $\mu_i = \bar{\mu}, \ \forall i$.

Proposition 1. *If the suppliers' average productivity is constant across countries, the solution of (28) is characterized by the following non-linear first order ODE:*

$$p'_{ij}(z_j) = \frac{\xi_j p_{ij}(z_j)[p_{ij}(z_j) - t_{ij}w_j z_j]}{\eta t_{ij} w_j z_j + \left(1 - \eta + \frac{\lambda_i}{m_i} t_{ij} w_j z_j\right) p_{ij}(z_j) - \frac{\lambda_i}{m_i} p_{ij}(z_j)^2} \tag{31}$$

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

Proof: See Appendix A.

For small values of z, the solution of (31) has the form:

$$p(z)\Big|_{z \to 0} = \frac{\eta}{\eta - 1} twz + o(z) \tag{32}$$

where the country indexes have been suppressed to simplify the notation. Equation (32) indicates that the most productive suppliers are not affected by 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 (31) has the form:

$$p(z)\Big|_{z \to \infty} = \frac{tw}{\frac{\lambda}{m} tw + \xi} + twz \tag{33}$$

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 sellers in other countries, the problem reduces to the closed economy one (with the correction for transportation costs and wages). Globally, the solution lies between the marginal cost line $t_{ij} w_j z_j$ and the closed economy pricing rule[24]. The value of the parameter ξ_j affects the location of the curve: when international competition is tougher ("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 and – as a result – prices and mark-ups shrink[25].

Figure 4 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[26] (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: the fact that in this economy mark-ups are endogenous implies that 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 on the foreign market despite the higher costs, as shown in the right panel of the figure). The parameter ξ_j does not affect the pricing-to-market behavior (*i.e.*, the wedge between domestic and export prices), since international competition affects in the same way both prices charged domestically and export prices.

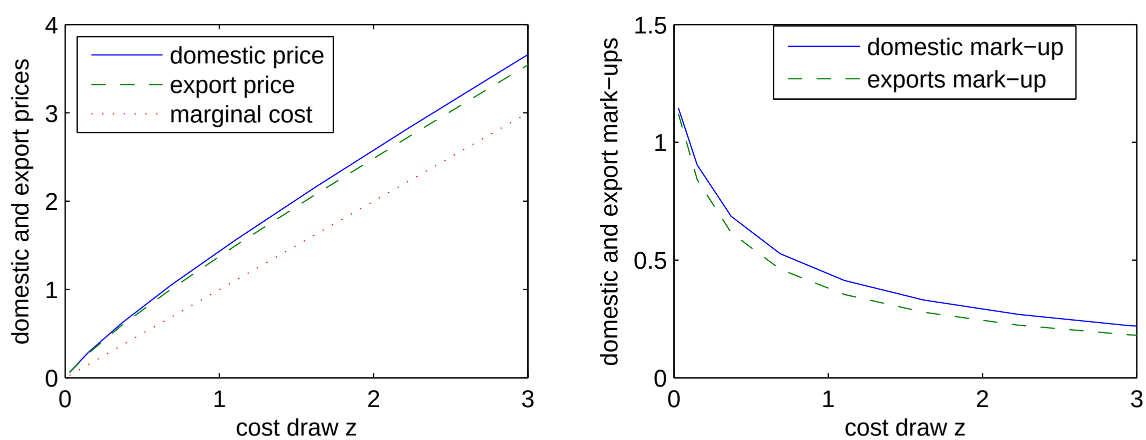

Figure 4: Pricing-to-market (2 symmetric countries, t=1.3, τ=1.5) .

[24]More precisely, the solution lies below the closed economy pricing rule adjusted for transportation costs and wage differences, that can be obtained by equating to zero the denominator of the right hand side of (31).

[25]Melitz and Ottaviano (2008) obtain the same qualitative result.

[26]Exclusive of trade costs.

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 are consistently adjusted downwards, tending to perfectly competitive prices for $N \to \infty$. Figure 5 shows the effect of international competition on prices through an increase in the number of countries in the economy.

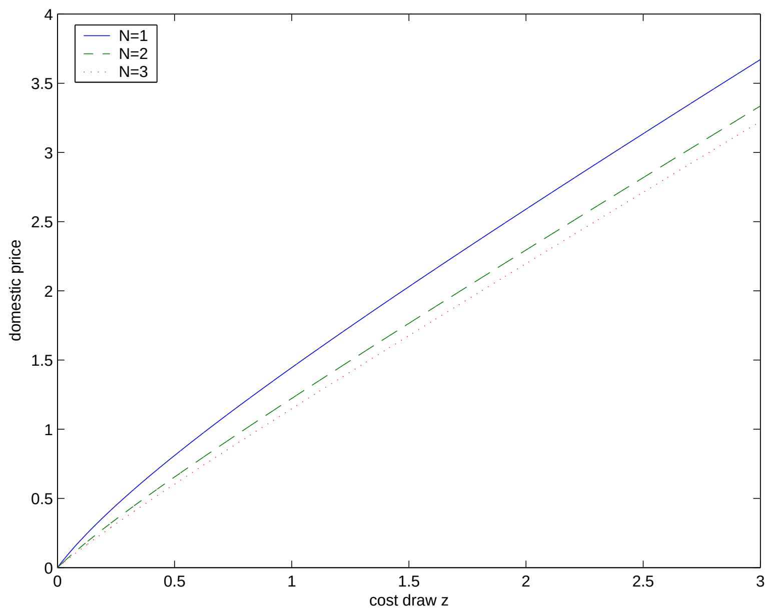

Figure 5: Domestic prices, open economy (N symmetric countries).

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 $(\frac{\partial p_{ij}(z_j)}{\partial t_{ik}} > 0 \ \ , \ \frac{\partial p_{ij}(z_j)}{\partial \mu_k} < 0 \ \ \forall k)$. These properties make explicit the dependence of prices and mark-ups on country characteristics and geographic barriers. By consequence, country characteristics will also affect the choice of undertaking intrafirm transactions through their effects on arm's length prices. The analysis of the pricing strategy confirms that when a country opens to operations with other countries, both the integration and the trade options become cheaper for domestic firms: 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. Lower prices lead to higher volumes of production in each country.

3.4 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 ($i = 1, ...N$) is:

$$c_i = q_i^{\alpha}(l_i^f)^{1-\alpha} \tag{34}$$

where l_i^f is amount of labor used in the final good sector in country i.

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 the foreign market) or for affiliates of domestic or foreign integrated firms. Given that labor is immobile, the following population constraint must hold in each country:

$$L_i = l_i^f + \sum_{j=1}^{N}(l_{ji}^I + l_{ji}^T) \quad for \;\; i = 1, ...N \tag{35}$$

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 specialized intermediate goods producers 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 quantities $\{q_i\}_{i=1}^N$:

$$L_i = \frac{(1-\alpha)p_i}{\alpha w_i}q_i + \sum_{j=1}^{N}(k_{ji}^I q_j + k_{ji}^T q_j) \quad for \;\; i = 1, ...N \tag{36}$$

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[27]. Taking the wages as given, (36) 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:

$$q_i^* = q_i(w_1, ...w_n). \tag{37}$$

Market clearing conditions in each country allow to solve for the equilibrium vector of wages. In each country, total income (labor income plus the profits of the intermediate goods producers) must

[27] Explicit expressions for the proportionality factors k_{ji}^I, k_{ji}^T are derived in Appendix B.

be equal to total expenditure in the final good:

$$r_i c_i = L_i w_i + \int_0^\infty \pi_i(z_i)\psi_i(z_i)dz_i \quad for \;\; i = 1, ...N \tag{38}$$

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 an intermediate goods producer from country i with cost draw z_i:

$$\pi_i(z_i) = \sum_{j=1}^{N} \left[[p_{ji}(z_i) - t_{ji} w_i z_i] \left(\frac{p_{ji}(z_i)}{p_j} \right)^{-\eta} q_j \left[1 - \Phi_j \left(\frac{p_{ji}(z_i)}{m_j} \right) \right] \prod_{k \neq i} [1 - F_{jk}(p_{ji}(z_i))] \right].$$

The market clearing condition (38) 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.

3.5 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 6 plots the excess demand correspondence in the Home country for the symmetric case[28]. Due to the discrete choice of where to locate[29] 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:

[28] The computation is done for parameters values $\eta = 1.8$, $\alpha = 0.25$, $t = 1.1$ and $\tau = 1.2$.

[29] Recall that a firm from country i locates integrated activities in the country j such that $\tau_{ij} w_j = \min_k\{\tau_{ik} w_k\}$.

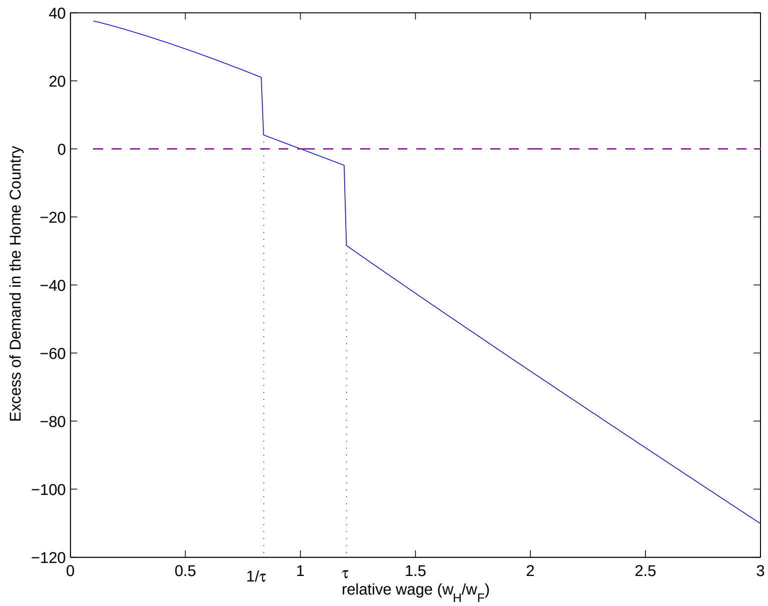

Figure 6: Excess demand correspondence in the Home country.

- $w_H \in (0, w_F/\tau) \quad \Rightarrow$ firms from both countries integrate in the Home country;
- $w_H \in (w_F/\tau, \tau w_F) \Rightarrow$ firms from both countries integrate domestically;
- $w_H \in (\tau w_F, \infty) \quad \Rightarrow$ firms from both countries integrate in the Foreign country.

Hence, 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 in only 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 7 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.

Proposition 2. *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 $ED_H = 0$.*

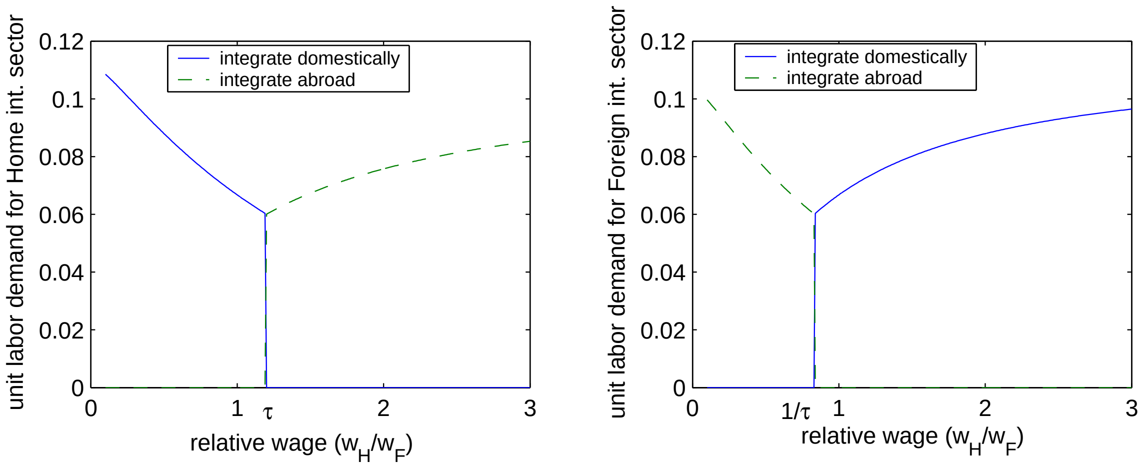

Figure 7: Labor demand in the integrated sectors.

Proof: See Appendix C.

4 Comparative Statics: Numerical Analysis

In this section I show the predictions of the model for the volumes of arm's length imports and intrafirm imports (vertical FDI) as fractions of GDP, and for the share of total imports that is performed intrafirm.

Volumes of (arm's length) imports over GDP for country i are given by:

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

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

$$(vertical\ FDI/GDP)_i = \begin{cases} \dfrac{(p_i^I)^{1-\eta}(p_i)^\eta q_i}{r_i c_i} & ; \text{if } w_i > \min\limits_{j \neq i}\{\tau w_j\} \\ \dfrac{(1-\gamma_i)(p_i^I)^{1-\eta}(p_i)^\eta q_i}{r_i c_i} & ; \text{if } w_i = \min\limits_{j \neq i}\{\tau w_j\} \\ 0 & ; \text{otherwise} \end{cases}$$

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

Figure 8 shows the effects of relative size on volumes of imports and foreign integration in the two-country case. I isolate the effects of size by normalizing the world population to one and denoting by s the share of the world population living in the Home country. The two countries are perfectly symmetric in all the other characteristics[30].

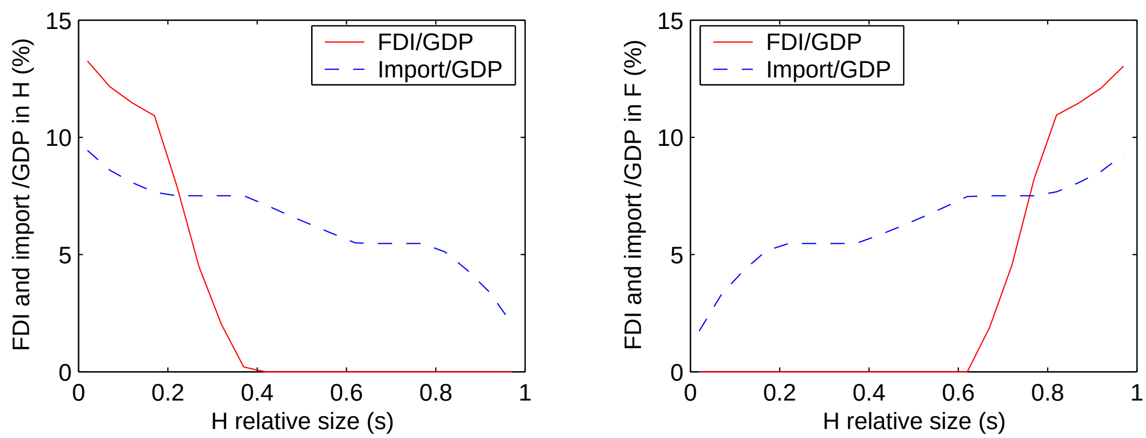

Figure 8: Volumes of arm's length import and FDI import as functions of relative size.

The figure plots volumes of arm's length imports and of 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 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 FDI in 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[31]. This prediction is also empirically supported by Hanson, Mataloni and Slaughter (2001), who document a tendency for U.S. multinationals to set vertical production networks through affiliates in low-wage countries.

[30] All the computations in this section are done for $\eta = 1.8$, $\alpha = 0.3$, $t = 1.1$ and $\tau = 1.2$.

[31] Nocke and Yeaple (2008) report data on the operations of U.S. multinationals suggesting that the share of greenfield FDI in total foreign investment is decreasing in the host country level of development: this seems to suggest that greenfield FDI is the preferred strategy to enter low-income countries.

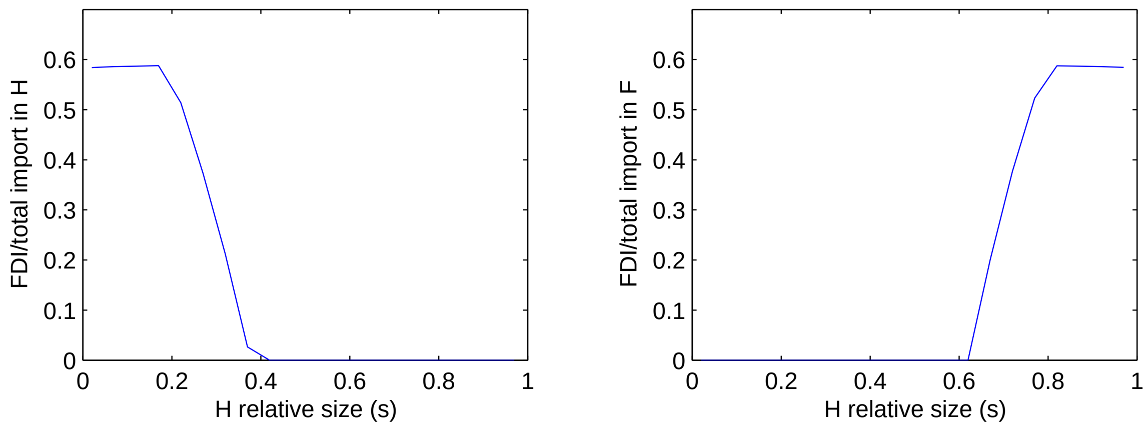

Figure 9: Share of FDI import over total import as a function of relative size.

Figure 9 shows that also the share of intrafirm imports over total import drops with increases in relative size. When a country relative size increases, the profitability of foreign labor force drops, and so the return from setting affiliates abroad. The slope of the plot depends on the effect of size on equilibrium wages, and on the implications of different relative wages for the location choice: when country H is "small", the scarce labor force in H has the effect of pushing up equilibrium wages, $w_H > \tau w_F$ and firms from H integrate in F only. Under this scenario, increases in size that do not change the equilibrium location choice affect arm's length and FDI imports in the same way, and the ratio is flat. As H's size keeps increasing, the equilibrium location choice changes, $w_H = \tau w_F$ and H firms integrate both domestically and in F. Under this scenario, increases in size have the effect that FDI imports are substituted by domestic integration, so FDI imports drop more drastically than arm's length imports. For s approaching 0.5, cross-country differences are not large enough to justify foreign investment, so all integrated activity is domestic, and the share reaches zero. The pattern is reversed for country F.

4.2 Effects of Productivity Dispersion

Figure 10 shows the effects of the productivity dispersion of final good producers on volumes of import and FDI in the two countries. The variable on the horizontal axis is the dispersion parameter of integrated production of Home country firms, λ_H.

When λ_H increases, the cost distribution of integrated production becomes less disperse, and

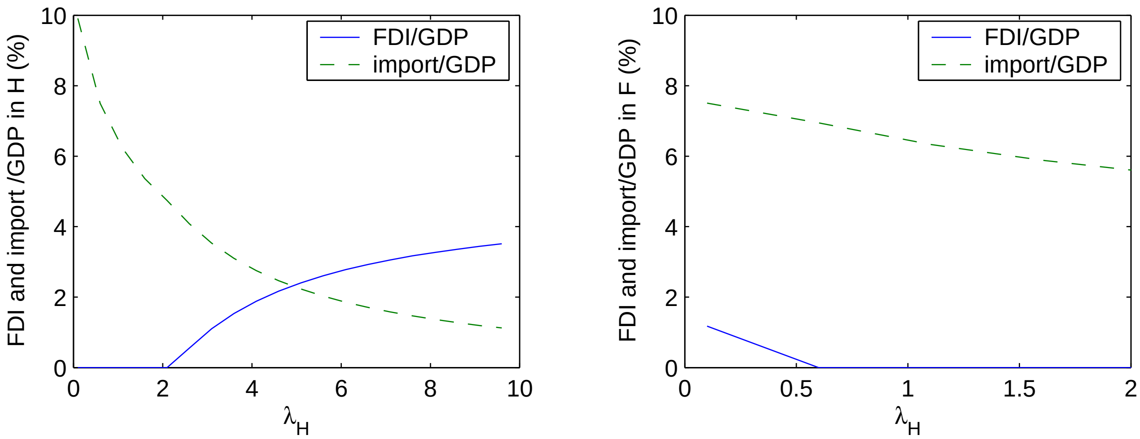

Figure 10: Volumes of arm's length import and FDI import as functions of buyers' costs dispersion.

the average productivity increases, so that Home country firms find more profitable to integrate. The higher productivity reflects into higher domestic wages, which push towards moving integrated production abroad (FDI). At the same time, the reliance on arm's length imports decreases (outside suppliers are relatively less productive when λ_H increases). 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 one, which has lower productivity and hence lower labor costs.

Figure 11 displays volumes of trade as functions of the productivity dispersion of the intermediate goods producers, μ_H. The variable on the horizontal axis is the dispersion parameter of the suppliers' cost distribution in the Home country, 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

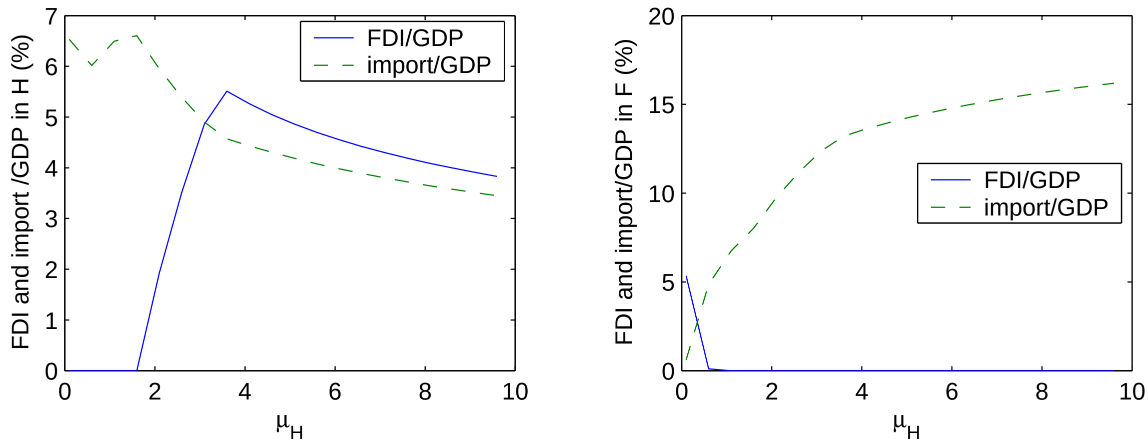

Figure 11: Volumes of arm's length import and FDI import as functions of sellers' costs dispersion.

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. 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.

4.3 Effects of Higher Heterogeneity Across Countries

In the last two sections I have shown the effects of changes in relative size and relative productivity on the volumes of intrafirm and arm's length imports. Each effect has been isolated by keeping all the other variables fixed. In this section I show the results of the model for combined changes in size and productivity.

Table 1 displays the results. Since in a two-country world only relative variables matter, I fix the labor force of the Home country ($L_H = 100$) and the parameters of the cost distributions of the Foreign country ($\lambda_F = \mu_F = 1$). The table reports equilibrium relative wage, consumption per capita, volumes of intrafirm and arm's length imports as fractions of GDP, and the share of intrafirm imports over total imports. Each entry reports the values for the two countries. The value γ_H which appears in parenthesis with some equilibrium wages is the percentage of labor force hired domestically in H's integrated sectors when firms are indifferent about the location of production.

In a symmetric world wages are the same and there is no incentive to foreign investment, nonetheless the model predicts a significant volume of arm's length trade. Increasing the produc-

	equilibrium relative wage	*consumption per capita*	*FDI import/ GDP (%)*	*arm's length import/GDP (%)*	*intrafirm share of import (%)*
symmetric countries	1	1.03	0	7.7	0
		1.03	0	7.7	0
$L_F = L_H$,	1.2	1.16	1.83	7.29	20.01
$\lambda_H = 1$; $\mu_H = 2$	$(\gamma_H = .8)$	1.08	0	11.3	0
$L_F = L_H$,	1.2	1.25	3.94	5.27	42.8
$\lambda_H = \mu_H = 2$	$(\gamma_H = .74)$	1.07	0	11.3	0
$L_F = 2L_H$,	1.2	1.29	12.4	5.27	70.17
$\lambda_H = \mu_H = 2$	$(\gamma_H = .64)$	1.07	0	11.3	0
$L_F = 4L_H$,	2.01	1.76	19.55	3.36	85.32
$\lambda_H = \mu_H = 4$		1.09	0	12.53	0

Table 1: Comparative statics, two-country economy.

tivity of the suppliers in one country (second row) distorts the equilibrium wages and creates a rationale for foreign investment. Under this scenario, domestic integrated production dominates FDI at a ratio of 4 to 1, and the higher productivity in country H increases the volume of exports (F's imports). Increasing the productivity of the buyers as well (third row) increases the incentive to integrate abroad, reduces imports and boosts consumption. Additional size differences (last two rows) tilt the sourcing strategy of the smaller and more productive country (here H) towards FDI import.

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 domestic MNCs productivity compared with the foreign suppliers' one).

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. 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' findings.

5.1 Calibration

In the two-country model, let the Home country be the United States, and the Foreign country be an aggregate of the rest of the world (henceforth, ROW). $(1-\alpha)$ 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 (which they define as tradeables). Following calculations from different data sources, they choose a value of $\alpha = 0.25$ as a reasonable value for the industrialized countries. Since I am trying to match features of U.S. trade data, I use their calibrated value in this computation. The exponential parameterization for the productivity 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. 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 $\eta = 1.8$, 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 a market[32].

I use the model to calibrate the remaining parameters. I identify the ratio μ_{us}/μ_{row} with the relative average productivity of U.S. firms with respect to ROW firms. As in Alvarez and Lucas (2007), L_{us}/L_{row} represents labor in efficiency units in the US relative to the ROW. t and τ are average iceberg costs of trade and FDI. I calibrate these parameters jointly to match four relevant moments of the data. For simplicity, I assume that final good producers (the potential multinational

[32]The value $\eta = 1.8$ implies mark-ups ranging from 125% (for the most productive firms) to zero (for the least productive ones). Average mark-ups depend on productivity parameters and trade barriers. With the set of calibrated parameters, average (sales-weighted) mark-ups are 57% on U.S. domestic sales and 55% on U.S. exports.

corporations) draw their productivity from the same distribution of local firms, and hence have the same relative average productivity: $\lambda_{us}/\lambda_{row} = \mu_{us}/\mu_{row}$.

I calibrate the four parameters τ, t, L_{us}/L_{row}, and μ_{us}/μ_{row} to match the intrafirm share of U.S. imports, U.S. total imports as a fraction of GDP, U.S. share of world GDP, and 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[33]. From Census data, I find that 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[34] was 2.22. Let $P = [\mu_{us}/\mu_{row} \;\; L_{us}/L_{row} \;\; t \;\; \tau]$. The vector of calibrated parameters is a vector $P^* = \arg\min \sum [mom - \widehat{mom}(P)]^2$, 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 set of parameters P. The model calibrated to match these magnitudes implies $L_{us}/L_{row} = 0.16$, $t = 1.1$, $\tau = 2.65$, $\mu_{us}/\mu_{row} = \lambda_{us}/\lambda_{row} = 1.92$. Table 2 summarizes the calibrated parameters I use in the computation of the welfare gains. The values in parentheses for τ, L_{us}/L_{row}, μ_{us}/μ_{row} and $\lambda_{us}/\lambda_{row}$ are the ones used when choosing $\eta = 1.2$.

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 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

[33] I construct the share of intrafirm imports in U.S. total imports by merging the 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 (Forthcoming) report an intrafirm share of imports of about 40%. This because 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 inter-affiliates transactions.

[34] I 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:

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

Data are for 157 countries, representing 98% of the world GDP. Import data are from the U.S. Census, while GDP and labor force data are from the World Bank's WDI.

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*	to match data
τ	2.65 (2.66)	*iceberg FDI cost*	to match data
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*	to match data
$\lambda_{us}/\lambda_{row}$	1.92 (1.77)	*buyers' relative productivity*	to match data

Table 2: Summary of calibrated parameters.

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 calibration.

5.2 Gains from Multinational Production: Decomposition and Policy Experiments

I compute the welfare gains that the theory implies compared to a counterfactual world without possibility of integration and multinational production. The gain in consumption per capita is computed as follows:

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

where the term "consumption in model without integration" is obtained by computing the model with the calibrated parameters and shutting down the possibility of in-house production for the final good producers. The model without integration still allows firms to trade in intermediates, so the difference in consumption per capita is purely due to the possibility of integration[35].

In the model though, integrated production can happen domestically or abroad, so I need to

[35]The gains computed in this section take the model as a correct description of reality, and are meant to illustrate its workings for a reasonable parameterization. As such, the magnitude of the gains depends on the specific market structure assumed. Having imperfect competition in the final good market could reduce the welfare gains, as some of the cost savings associated with integration would not be passed through to the consumers due to the inefficiencies associated with market power. Nonetheless, the model would still exhibit positive gains.

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 in U.S. integrated sectors that is hired in each country. Due to the linearity of the intermediate goods production technology, the share of labor hired in integrated sectors abroad is also the share of integrated production that is done abroad and the share of gains that come from foreign integration. On the other hand, in equilibrium, the ROW economy integrates only domestically and the gains in consumption come from two sources: the possibility of (domestic) integrated activity and the upward pressure on wages determined by the entry of U.S. firms.

These results are shown in the first column of Table 3. The calibrated economy implies a gain in U.S. consumption per capita of 4.87% with respect to a world where there is 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.

	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
of which:			
domestic integration	4.16	0	11.35
foreign integration	0.71	6.95	1.94
of which:			
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
of which:			
domestic integration	11.84	11.84	22.58
foreign integration	0.55	3.33	0.52
of which:			
productivity effect	11.76	14.35	14.37
competition effect	0.63	0.52	8.73

Table 3: Gains from multinational production: decompositions and policy experiments.

Table 3 also shows another decomposition of the gains, which quantifies the fact that – according to the model – gains from integration arise from two different channels. On one hand, there is a "productivity effect": with integration, a share of inputs is sourced at marginal cost, which can be lower than the one of the suppliers due to potentially higher productivity of the integrated firms. With foreign integration, this potentially high productivity can be matched to low wages, further increasing the gains for the parent's country. In the host country, entry of foreign firms increases labor demand, and as a result the wage increases: the producvitity gain for the host country depends on the upward pressure on wages induced by the entry of foreign firms. On the other hand there is a "competition effect" due to the fact that suppliers in both countries shrink their mark-ups to respond to the higher competition arising from 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 model and the model without integration isolates the competition effect. The residual gain is to be attributed to the productivity effect. 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 differential between the two regions. 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 3 reports the same calculations performed in a world where the unit cost of integrated production abroad drops of 50% . As expected, lowering this cost increases the gains, and 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 comes entirely from foreign production. The larger extent of foreign investment in ROW countries also increases ROW's relative wage and causes the 3.03% gain in consumption per capita attributable to the entry of foreign firms. The decomposition of the gains in competition and integration effect is basically unchanged with respect to the baseline scenario, with a large majority coming from the productivity differential across the two regions.

The third column of Table 3 reports the same calculations for a lower value of the elasticity of substitution η ($\eta' = 1.2$). This version of the calibration corresponds to a world where the degree of differentiation across intermediates is higher. As a 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[36]. 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 high mark-ups more than in the previous cases[37].

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

In this section, I compare the gains that the model produces with what other authors have found using different underlying theories. Using the calibrated parameters, I compute consumption in the two countries in the autarky case, in which the barriers to trade and FDI t and τ are prohibitively high and there is no foreign sourcing. I normalize the results to one, and compute the welfare gain 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 4. The gain for the U.S. is equal to 13% of consumption per capita, close to estimates from other 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 third row of the table computes consumption in a frictionless world ($t = \tau = 1$). This possibility implies an additional increase in consumption of about 9% for the U.S. firms 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 (about 10% of consumption). Compared to Rodríguez-Clare's analysis, my

[36] This 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.

[37] In the baseline calibration, the possibility of integrated production generates a reduction in 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 4: Gains from moving from autarky to costless trade and FDI.

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 (Forthcoming) compute the welfare gains from relocating firms' managerial know-how to control factors of production abroad. Their computed gains are about 9% and 6% of consumption per capita for unilateral and multilateral liberalizations respectively. Even if the thought experiment is similar to this paper, in their model the relocation of managerial ability abroad entails a loss of productivity at home (which explains the smaller gains associated to a multilateral liberalization).

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 imperfectly competitive market structure. I study optimal pricing in presence of multinationals: imperfect competition creates a wedge between trade prices and transfer prices. The possibility of integration of input production induces downward pressure on the prices charged by the suppliers, establishing a link between trade and FDI liberalization and equilibrium prices. Due to firms' heterogeneity, the price responses vary across firms, generating prices and mark-ups distributions that are sensitive to changes in technology and trade costs. Firms' heterogeneity also implies the optimality of pricing-to-market behavior: firms charge lower prices and mark-ups in foreign markets to counteract the negative effects of trade costs on their competitivity.

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 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.

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. This would allow to calibrate the multicountry version of the model to match bilateral trade and FDI facts. 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 Derivation of the Open Economy Pricing Rule, $\mu_i = \bar{\mu},\ \forall i$

In this section I derive the optimal pricing rule of a supplier in country j with cost draw z_j who sells his good to a final good producer in country i, under the assumption that suppliers' average productivity is constant across countries: $\mu_i = \bar{\mu},\ \forall i$. The first order condition of problem (28) is:

$$\left[[p_{ij}(z_j)]^{-\eta} - \eta[p_{ij}(z_j) - t_{ij}w_j z_j][p_{ij}(z_j)]^{-\eta-1}\right] \cdot$$

$$...\left[1-\Phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\right]\prod_{k\neq j}[1-F_{ik}(p_{ij}(z_j))]-$$

$$...[p_{ij}(z_j)-t_{ij}w_jz_j][p_{ij}(z_j)]^{-\eta}\Bigg\{\phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\frac{1}{m_i}\prod_{k\neq j}[1-F_{ik}(p_{ij}(z_j))]-$$

$$...\sum_{l\neq j}\left[1-\Phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\right]\prod_{k\neq j,l}[1-F_{ik}(p_{ij}(z_j))]\cdot f_{il}(p_{ij}(z_j))\Bigg\}=0.$$

Dividing each term by $[p_{ij}(z_j)]^{-\eta-1}\left[1-\Phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\right]\prod_{k\neq j}[1-F_{ik}(p_{ij}(z_j))]$ and collecting common terms:

$$[p_{ij}(z_j)-\eta[p_{ij}(z_j)-t_{ij}w_jz_j]]-$$

$$...p_{ij}(z_j)[p_{ij}(z_j)-t_{ij}w_jz_j]\left\{\frac{\phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\frac{1}{m_i}}{\left[1-\Phi_i\left(\frac{p_{ij}(z_j)}{m_i}\right)\right]}+\sum_{l\neq j}\frac{f_{il}(p_{ij}(z_j))}{[1-F_{il}(p_{ij}(z_j))]}\right\}=0.$$

Since $p'_{ij}(z_j)>0$:

$$\begin{aligned}[1-F_{il}(p_{ij}(z_j))] &= \left[1-\Psi_l\left(\frac{t_{ij}w_jz_j}{t_{il}w_l}\right)\right] && \text{(A.1)}\\ f_{il}(p_{ij}(z_j)) &= \psi_l\left(\frac{t_{ij}w_jz_j}{t_{il}w_l}\right)\frac{t_{ij}w_j}{t_{il}w_l}\cdot\frac{1}{p'_{ij}(z_j)}. && \text{(A.2)}\end{aligned}$$

Assuming that the cost draws are exponentially distributed, with country-specific parameters for the cost of integration, and with a common parameter for the costs of the suppliers:

$$\begin{aligned}\phi_i(x_i) &= \lambda_i e^{-\lambda_i x_i} \quad for \;\; x_i\geq 0\\ \psi_i(z_i) &= \bar{\mu}e^{-\bar{\mu}z_i} \quad for \;\; z_i\geq 0\end{aligned}$$

the first order condition characterizing the pricing rule reduces to the following ODE:

$$p'_{ij}(z_j)=\frac{\xi_j p_{ij}(z_j)[p_{ij}(z_j)-t_{ij}w_jz_j]}{\eta t_{ij}w_jz_j+(1-\eta+\frac{\lambda_i}{m_i}t_{ij}w_jz_j)p_{ij}(z_j)-\frac{\lambda_i}{m_i}p_{ij}(z_j)^2} \tag{A.3}$$

where $\xi_j=\bar{\mu}\sum_{l\neq j}\frac{t_{ij}w_j}{t_{il}w_l}$ is a measure of worldwide competition that a firm from country j faces while selling to country i.

B Equilibrium Labor Force Segments

Let l_{ji}^I denote the labor force of country i working in integrated segments of firms from country j. Then:

$$l_{ji}^I = \begin{cases} 0 & ; \text{if } m_j \neq \tau_{ji} w_i \\ k_{ji}^I q_j & ; \text{if } m_j = \tau_{ji} w_i \end{cases} \tag{B.1}$$

where:

$$k_{ji}^I = \frac{p_j^\eta}{w_i} \int_{B_j^I} (m_j x_j)^{1-\eta} \phi_j(x_j) \psi(z) dx_j dz. \tag{B.2}$$

Similarly, let l_{ji}^T denote the labor force of country i working for specialized intermediate goods producers from i targeting market j:

$$l_{ji}^T = k_{ji}^T q_j \tag{B.3}$$

where:

$$k_{ji}^T = p_j^\eta \int_{B_{ji}^T} t_{ji} z_i p_{ji}(z_i)^{-\eta} \phi_j(x_j) \psi(z) dx_j dz. \tag{B.4}$$

C Existence of the Equilibrium

This section contains the proof of Proposition 2. 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 $\exists \underline{w}_H, \bar{w}_H$ such that $ED_H(\underline{w}_H) > 0$ and $ED_H(\bar{w}_H) < 0$.

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. Particularly, the labor demand for integrated segments of firms from country H is[1]:

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

$$
l_h^I \begin{cases} = \dfrac{p_h^\eta q_h}{w_h} \displaystyle\int_{B_h^I} (w_h x_h)^{1-\eta} \phi_h(x_h)\psi(z) dx_h dz & ; \text{if } w_h < \tau \\ \in \left[\dfrac{p_h^\eta q_h}{w_h} \displaystyle\int_{B_h^I} (w_h x_h)^{1-\eta} \phi_h(x_h)\psi(z) dx_h dz, \dfrac{p_h^\eta q_h}{w_s} \displaystyle\int_{B_h^I} (\tau w_f x_h)^{1-\eta} \phi_h(x_h)\psi(z) dx_h dz \right] & ; \text{if } w_h = \tau \\ = \dfrac{p_h^\eta q_h}{w_s} \displaystyle\int_{B_h^I} (\tau w_f x_h)^{1-\eta} \phi_h(x_h)\psi(z) dx_h dz & ; \text{if } w_h > \tau \end{cases}
$$

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 rewrite the excess demand correspondence and compute its limits 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:

$$
\begin{aligned}
ED(w_h) &= L_h w_h + (p_{hh}^T)^{1-\eta} p_h^\eta q_h + (p_{fh}^T)^{1-\eta} p_f^\eta q_f - w_h(l_{hh}^T + l_{fh}^T) - r_h c_h \\
&= \left[\frac{(p_{hh}^I)^{1-\eta} + (p_{hh}^T)^{1-\eta}}{p_h^{1-\eta}} - 1 \right] p_h q_h + \left[\frac{(p_{fh}^I)^{1-\eta} + (p_{fh}^T)^{1-\eta}}{p_f^{1-\eta}} \right] p_f q_f
\end{aligned}
$$

where, given the definition of the price indexes:

$$
\begin{aligned}
\left[\frac{(p_{hh}^I)^{1-\eta} + (p_{hh}^T)^{1-\eta}}{p_h^{1-\eta}} - 1 \right] &\in (-1, 0) \\
\left[\frac{(p_{fh}^I)^{1-\eta} + (p_{fh}^T)^{1-\eta}}{p_f^{1-\eta}} \right] &\in (0, 1).
\end{aligned}
$$

As prices are increasing in wages:

$$
\begin{aligned}
\lim_{w_h \to 0} p_h &= \lim_{w_h \to 0} p_f = 0 \\
\lim_{w_h \to \infty} p_h &= \lim_{w_h \to \infty} p_f = \infty.
\end{aligned}
$$

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

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

$$
\begin{aligned}
&= \left\{ \frac{(1-\alpha)p_h}{\alpha w_h L_h} + \frac{k_{hh}}{L_h} - \frac{\alpha w_f k_{fh} k_{hf}}{\left[(1-\alpha)p_f + \alpha w_f k_{ff}\right] L_h} \right\}^{-1} \dots \\
&\quad \dots - \left\{ \frac{\left[(1-\alpha)p_h/w_h + \alpha k_{hh}\right]\left[(1-\alpha)p_f + \alpha w_f k_{ff}\right]}{\alpha^2 w_f k_{fh} L_f} - \frac{k_{hf}}{L_f} \right\}^{-1}.
\end{aligned}
$$

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 $\frac{\alpha w_f k_{fh} k_{hf}}{\left[(1-\alpha)p_f + \alpha w_f k_{ff}\right] L_h}$ is a $\frac{0}{0}$ indeterminacy; the term $\frac{\left[(1-\alpha)p_h/w_h + \alpha k_{hh}\right]}{\alpha^2 w_f k_{fh} L_f}$ is a positive constant, and the term $\left[(1-\alpha)p_f + \alpha w_f k_{ff}\right]$ tends to zero, so $\frac{\left[(1-\alpha)p_h/w_h + \alpha k_{hh}\right]\left[(1-\alpha)p_f + \alpha w_f k_{ff}\right]}{\alpha^2 w_f k_{fh} L_f}$ 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 $\lim_{w_h \to \infty} q_h = \infty$, $\lim_{w_h \to 0} q_f = \infty$, $\lim_{w_h \to \infty} q_f = -\infty$. Given these results, it is immediate to conclude that:

$$
\begin{aligned}
\lim_{w_h \to 0} ED(w_h) &= \infty \\
\lim_{w_h \to \infty} ED(w_h) &= -\infty.
\end{aligned}
$$

References

[1] Alvarez, Fernando, and Robert E. Lucas, Jr. (2007), "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade", *Journal of Monetary Economics*, Vol. 54, No. 6, pp. 1726-1768.

[2] Antràs, Pol (2003), "Firms, Contracts, and Trade Structure", *Quarterly Journal of Economics*, Vol. 118, No. 4, pp. 1375-1418.

[3] Antràs, Pol, and Elhanan Helpman (2004), "Global Sourcing", *Journal of Political Economy*, Vol. 112, No. 3, pp. 552–580.

[4] Bernard, Andrew B., Jonathatan Eaton, J. Bradford Jensen and Samuel Kortum (2003), "Plants and Productivity in International Trade", *The American Economic Review*, Vol. 93, No. 4, pp. 1268-1290.

[5] Bernard, Andrew B., J. Bradford Jensen and Peter K. Schott (2006), "Transfer Pricing by U.S.-Based Multinational Firms", NBER Working Paper 12493.

[6] Bernard, Andrew B., J. Bradford Jensen and Peter K. Schott (Forthcoming) "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*, University of Chicago Press.

[7] Burstein, Ariel, and Alexander Monge-Naranjo (Forthcoming), "Foreign Know–How, Firm Control, and the Income of Developing Countries", *Quarterly Journal of Economics*.

[8] 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 4, pp. 79-99.

[9] Eaton, Jonathan, and Sam Kortum (2002), "Technology, Geography and Trade", *Econometrica*, Vol. 70, No. 5, pp. 1741-1779.

[10] Feenstra, Robert C. (1998), "Integration of Trade and Disintegration of Production in the Global Economy", *The Journal of Economic Perspectives*, Vol. 12, No. 4, pp. 31-50.

[11] Grossman, Gene M., and Elhanan Helpman (2002), "Integration Versus Outsourcing in Industry Equilibrium", *Quarterly Journal of Economics*, Vol. 117, No.1, pp. 85-120.

[12] Grossman, Gene M., and Elhanan Helpman (2004), "Managerial Incentives and the International Organization of Production", *Journal of International Economics*, No. 63, pp. 237-262.

[13] Grossman, Gene M., and Esteban Rossi-Hansberg (Forthcoming), "Trading Tasks: A Simple Theory of Offshoring", *The American Economic Rewiew*.

[14] 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*.

[15] Hanson, Gordon H., Raymond J. Mataloni, and Matthew J. Slaughter (2005), "Vertical Production Networks in Multinational Firms", *Review of Economics and Statistics*, Vol. 87, No. 4, pp. 664-678.

[16] Helpman, Elhanan (1984), "A Simple Theory of International Trade with Multinational Corporations", *Journal of Political Economy*, Vol. 92, No. 3, pp. 451-471.

[17] Helpman, Elhanan (1985), "Multinational Corporations and Trade Structure", *Review of Economic Studies*, Vol. 52, No. 3, pp. 443-457.

[18] Helpman, Elhanan (2006), "Trade, FDI, and the Organization of Firms", NBER Working Paper 12091.

[19] Helpman Elhanan, and Paul Krugman (1985) "Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy", MIT Press.

[20] Helpman, Elhanan, Marc J. Melitz and Stephen R. Yeaple (2004), "Exports Versus FDI with Heterogeneous Firms", *The American Economic Review*, Vol. 94, No.1, pp. 300-316.

[21] Hummels, David, Jun Ishii and Kei-Mu Yi (2001), "The Nature and Growth of Vertical Specialization in World Trade", *Journal of International Economics*, Vol. 54, pp. 75-96.

[22] 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.

[23] Markusen, James R. (1984), "Multinationals, Multi-Plant Economies, and the Gains from Trade", *Journal of International Economics*, Vol. 16, pp. 205-226.

[24] Melitz, Marc J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, Vol.71, No.6, pp.1695-1725.

[25] Melitz, Marc J., and Gianmarco I.P. Ottaviano (2008), "Market Size, Trade, and Productivity", *Review of Economic Studies*, Vol. 75, No.1, pp.295-316.

[26] Nocke, Volker, and Stephen Yeaple (2008), "An Assignment Theory of Foreign Direct Investment", *Review of Economic Studies*, Vol. 75, No. 2, pp. 529-557.

[27] Rauch, James E. (1999), "Networks versus markets in international trade", *Journal of International Economics*, Vol. 48, pp. 7-35.

[28] Rodríguez-Clare, Andrès (2007), "Trade, Diffusion, and the Gains from Openness", mimeo, Pennsylvania State University.

[29] Rossi-Hansberg, Esteban, and Mark L. J. Wright (2007), "Establishment Size Dynamics in the Aggregate Economy", *The American Economic Review*, Vol. 97, No. 5, pp. 1639-1666.

[30] Syverson, Chad (2004), "Market Structure and Productivity: A Concrete Example", *Journal of Political Economy*, Vol. 112, No. 6, pp. 1181-1222.

[31] 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 University Press.