Input Sourcing and Multinational Production

Stefania Garetto†
University of Chicago
Department of Economics
(JOB MARKET PAPER)
12th November 2007

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. Outsourcing allows to benefit from suppliers’ good technologies, but entails the cost of a mark-up price. Intrafirm sourcing allows to save on mark-ups and to match a firm’s productivity with possibly lower foreign wages. The pricing implications of the theory unveil a positive relationship between the intrafirm share of imports and the degree of differentiation across inputs in a sector, for which I find strong support in the data. Moreover, 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). As a result, there is complementarity between trade and foreign investment. 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 small (about 1% of consumption per capita) but 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. I wish to thank Fernando Alvarez, Christian Broda, Thomas Chaney, José Fillat, Kenneth Judd, Natalia Ramondo, Nancy Stokey, Robert Shimer, Chad Syverson, Mathis Wagner, and participants at the University of Chicago Capital Theory Working Group and Money and Banking Workshop for helpful comments. All errors are mine.

†Contact: Department of Economics, The University of Chicago, 1126 East 59th Street, Chicago, IL 60637. E-mail: garettos@uchicago.edu. http://home.uchicago.edu/~garettos.
1 Introduction

Globalization has expanded the scope of trade. Portuguese wine trading for English cloth was a good example in Ricardo’s days, but does only half the job in describing today’s international economics. 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 firms’ boundaries. Bernard, Jensen and Schott (2005) 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 these facts, particularly 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. Trade and FDI are substitutes at the level of the single input, since a firm decides the optimal sourcing strategy good by good, but are likely to coexist on aggregate, since it may be optimal for a firm to outsource some inputs while integrating production of others.

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

---

\(^1\)See Yeats (2001).


\(^3\)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”.

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, 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 allow 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, it allows firms to 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, the integration option allows to save on the mark-ups charged by the suppliers: intrafirm trade happens between the firm and itself, and is priced at marginal cost\(^5\).

The two determinants behind the optimal sourcing choice imply testable predictions at the firm and sector level. At the firm level, due to technology heterogeneity, we expect the most productive firms to be the ones opting for insourcing and multinational production. Bernard, Jensen and Schott (2005) report evidence on the productivity advantage of multinational firms: in the data\(^6\) they describe, firms engaged in transactions with foreign affiliates appear to be larger and more productive than non-multinational firms. At the sector level, imperfect competition implies that suppliers’ prices and mark-ups are going to be higher in sectors where inputs are less substitutable with each other, generating a stronger incentive to insource in these sectors. The relationship

\(^4\)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 (2006). 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.

\(^6\)Bernard, Jensen and Schott (2005) merge transaction-level customs data with a longitudinal database tracking almost all U.S. private sector firms. The merged dataset allows them to measure how much of each firm’s trade takes place at arm’s length versus intrafirm.
between within-sector differentiation and intrafirm trade is unexplored in the literature. I provide empirical support for this prediction and show the statistical relevance of the degree of input differentiation in explaining sourcing patterns.

In summary, technology heterogeneity and imperfect competition are the two main features of the theory. The imperfectly competitive market structure establishes a link between trade liberalization, competition and optimal pricing in the presence of multinational firms. Multinationals add a new margin of competition among the 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 possibility of integrated production on the side of the buyers. FDI liberalization makes both trade and FDI more profitable, increases competition and lowers prices. As a result, the model predicts complementarity between trade and vertical FDI7. 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, the suppliers optimally decide to do pricing-to-market, i.e. to charge different prices in different countries.

On aggregate, the theory predicts that small countries rely more on sourcing abroad, both from unaffiliated suppliers and from integrated plants. 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 trade8. The dispersion of the cost distributions across firms also affects the sourcing strategy, with the prevailing sourcing mode being the one associated with the lowest cost dispersion.

The calibrated version of the model allows to quantify the welfare gains resulting from vertical multinational production9 in addition to the gains from trade, and also allows to disentangle 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 characterized by a significant level of market power on the side of the suppliers, while the effect of productivity and technology drives most of the results. The welfare gains arising from vertical multinational production are currently quite small (about 1% of consumption per capita), but further FDI liberalization would increase them substantially: the model predicts that a 50% drop in the calibrated barrier to

---

7The data show a positive correlation of arm’s length and intrafirm import by country, driven in large part by Canada and Mexico, which accounted for 56% of intrafirm import and 28% of total imports in 2000.
8Nocke and Yeaple (2005) have similar results when modeling the choice of vertical, greenfield investment versus mergers and acquisitions, and report empirical evidence in support of this fact.
9The model excludes horizontal FDI, i.e. the establishment of offshore production to serve foreign markets.
FDI would imply a gain of about 8% 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\(^{10}\) of production that may serve product lines without being located in their plants. I assume instead that firms can relocate 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, Antrás (2003) and Antrás and Helpman (2004) modeled the joint choice of location and organizational structure by merging existing models of trade\(^ {11}\) with a contract-based theory of the firm\(^ {12}\). Their approach has the advantage of analyzing separately the two choices, and matches qualitative features of the data on intrafirm trade. My model is a complementary analysis that uses firms’ heterogeneity and market structure to explain the same choices. While disregarding the theory of the firm aspects, my approach unveils a systematic relationship between intrafirm trade and inputs differentiation. Moreover, my theory has the advantage of providing a detailed analysis of optimal pricing, which is absent in Antrás and Helpman’s work\(^ {13}\), and allows a relatively simple calibration of the model in order 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 properties

---

\(^{10}\)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.


\(^{12}\)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. Antrás (2003) and Antrás 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 possibility that they allow.

\(^{13}\)Incomplete contracts imply that there are no explicit prices for the goods to be sourced.
of the model in explaining the dependence of aggregate volumes of trade and FDI on the economy’s fundamentals. Section 5 provides supporting empirical evidence on the relationship between within-sector differentiation and intrafirm trade. Section 6 contains the calibration of the model and the computation of the gains from multinational production and intrafirm trade. Section 7 concludes.

2 The Model: the Closed Economy

This section develops the analysis of the closed economy model, to isolate the choice between outsourcing and integration, setting aside the location choice. The model is based on recent models of trade, particularly on Eaton and Kortum (2002) and Alvarez and Lucas (2007). Precisely, I adopt Alvarez and Lucas’ setup and extend it 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 in the same model.

2.1 Production Technologies in a Two–Sector Economy

I introduce the setup of the model with a simple example. Consider two big producers of sport shoes, say Nike and Adidas. Shoes, despite being quite a “simple” good to produce, incorporate many different parts, like soles, laces, uppers, embroideries. Nike and Adidas can produce all the separate parts and then assemble them into the shoe that they sell, or they can subcontract various stages of the production process to outside producers and take care of the final assembly only, or of the distribution only. For example, Adidas can decide to produce his own laces, with his own technology, but I assume it is good only at producing specific Adidas laces: it knows how long and thick they have to be, and how to adapt them to his own different models of shoes. However, it does not know exactly how to produce the laces Nike uses, and acquiring that capability is too costly to be worthwhile. So if Adidas decides to produce shoelaces, it will produce only its own, and Nike will have a similar decision to take. A specialized firm that produces shoelaces, on the other hand, has a technology (like a versatile machine for example) that can be switched at no cost to produce either Nike or Adidas laces, hence it can easily serve both buyers.

I formalize this problem in my model. 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 are heterogeneous 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 $z$, which is a random
draw from a common distribution $\psi(z) : \mathbb{R}_+ \to [0, 1]$ and indicates the number of units of labor
needed to produce one unit of the good. Moreover, the intermediate goods producers’ technology
is freely adaptable, in the sense that each intermediate goods producer 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 sector is populated by a continuum of identical producers, who operate in a per-
fectly 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 acquire 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 is taken 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 distribution $\phi(x) : \mathbb{R}_+ \to [0, 1]$ and
indicates the number of units of labor needed to internalize production of one unit of producer-
specific input. All the final good producers face the same cost distribution $\phi(\cdot)$, but they can have
different cost draws. 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, as in many recent models of trade with firms’ heterogeneity\(^{14}\), goods are differentiated by their unit labor requirements. Consistently, 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 \( (x, z) \). Let \( q \) denote the aggregate quantity of intermediate goods\(^{15}\) that each final good producer needs for the production of the final good:

\[
q = \left[ \int_0^\infty \int_0^\infty q(x, z) x^{1-1/\eta} \phi(x) \psi(z) dx dz \right]^{\eta/(\eta-1)}
\]  

where \( \eta > 1 \) is the elasticity of substitution among the single inputs. The aggregate \( q \) will be determined by equilibrium conditions in the final good market, and the final good producer takes it as a given while deciding about whether producing or buying the single inputs.

The final good producer minimizes the total cost of the input aggregate \( q \):

\[
\min_{q(x,z)} \int_0^\infty \int_0^\infty \min\{x, p(z)\} q(x, z) \phi(x) \psi(z) dx dz \\
\text{s.t. } \left[ \int_0^\infty \int_0^\infty q(x, z) x^{1-1/\eta} \phi(x) \psi(z) dx dz \right]^{\eta/(\eta-1)} \geq q
\]

where the outside price \( p(z) \) is taken as given.

Problem (2) may be rewritten as:

\[
\min_{q(x,z)} \left[ \int_0^\infty \int_0^{p(z)} x q(x, z) \phi(x) \psi(z) dx dz + \int_0^{\infty} p(z) \int_{p(z)}^\infty q(x, z) \phi(x) \psi(z) dx dz \right]
\]


\(^{15}\)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.
\[ s.t. \quad \left[ \int \int q(x, z)^{1-1/\eta} \phi(x) \psi(z) dx dz \right]^{\eta/(\eta-1)} \geq q. \quad (3) \]

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 (3) 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 (3) in \( B^T \). Hence:

\[ q^I(x, z) \equiv q^I(x) = \left( \frac{x}{p} \right)^{-\eta} q \quad \forall \ (x, z) \in B^I \quad (4) \]
\[ q^T(x, z) \equiv q^T(p(z)) = \left( \frac{p(z)}{p} \right)^{-\eta} q \quad \forall \ (x, z) \in B^T. \quad (5) \]

Notice that when a final good producer decides to integrate production of a good, the optimal quantity does not depend on the outside price \( (q^I(x, z) \equiv q^I(x)) \). Similarly, when it is optimal to buy a good, the demand for it does not depend on the in-house cost, but only on the price charged \( (q^T(x, z) \equiv q^T(p(z))) \). The term \( p \) is the aggregate price index for this economy:

\[ p = \left[ p_I^{1-\eta} + p_T^{1-\eta} \right]^{1/(1-\eta)} \quad (6) \]

and:

\[ p_I = \left[ \int_0^\infty \int_0^{p(z)} x^{1-\eta} \phi(x) \psi(z) dx dz \right]^{1/(1-\eta)} \quad (7) \]
\[ p_T = \left[ \int_0^\infty p(z)^{1-\eta} \left[ 1 - \Phi(p(z)) \right] \psi(z) dz \right]^{1/(1-\eta)}. \quad (8) \]

The economy-wide price index is a CES aggregate of two sub-indexes: the aggregate price of integrated goods \( p_I \) and the aggregate price of outsourced (or traded) goods \( p_T \).

2.3 The Supplier’s Problem

While taken as a given by the potential buyer, each selling price \( p(z) \) is optimally set by an intermediate goods producer. 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:

$$\max_{p(z)} [p(z) - z]q^T(p(z)) \int_{p(z)}^{\infty} \phi(x)dx$$  \hspace{1cm} (9)

where $q^T(p(z))$ is given by equation (5), $\int_{p(z)}^{\infty} \phi(x)dx$ is the mass (or the percentage) of buyers that decide to buy the good at price $p(z)$, and $z$ is equal to the unit cost of production since wages are normalized to 1. The first order condition for this problem is:

$$[p(z) - z] \left[ q^T(p(z))\phi(p(z)) - \frac{\partial q^T(p(z))}{\partial p(z)} [1 - \Phi(p(z))] \right] = q^T(p(z)) \int_{p(z)}^{\infty} \phi(x)dx. \hspace{1cm} (10)$$

Equation (10) summarizes the monopolist’s trade-off. Keeping constant its total quantity sold (the right hand side of (10)), the gain from increasing the mark-up over the marginal cost (the first term on the left hand side of (10)) must be counterbalanced by the sum of the losses on both the extensive and the intensive margin. In fact, if the monopolist raises the price, he is going to lose the marginal buyers (this is the extensive margin, captured by the term $q^T(p(z))\phi(p(z))$) and he is going to sell lower quantities to the remaining buyers (this is the intensive margin, captured by the term $-\frac{\partial q^T(p(z))}{\partial p(z)} [1 - \Phi(p(z))]$).

Using (5), problem (9) becomes:

$$\max_{p(z)} [p(z) - z] \left( \frac{p(z)}{p} \right)^{-\eta} q[1 - \Phi(p(z))]$$  \hspace{1cm} (11)

and the first order condition implies:

$$p(z) = \left[ 1 - \frac{1}{\eta + \frac{\phi(p(z))}{1 - \Phi(p(z))}p(z)} \right]^{-1} z. \hspace{1cm} (12)$$

Equation (12) shows how the possibility of integration on the side of the buyers 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 the cost draw distribution $\phi(x)$ to be exponential with parameter $\lambda$:

**Assumption 1.** $\phi(x) = \lambda e^{-\lambda x}$ \hspace{1cm} for $x \geq 0$. 


Under Assumption 1, (12) reduces to:
\[
\lambda p(z)^2 + (\eta - 1 - \lambda z)p(z) - \eta z = 0.
\]

where the parameter \( \lambda \) is an inverse measure of variation of the buyers’ cost distribution\(^{16}\). Equation (13) 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 \( \eta \) and \( \lambda \). 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.

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 \( \eta \) and \( \lambda \).

The dashed line is the producer’s marginal cost, while the dash-dotted line is the constant mark-up pricing rule that would be optimal if the buyers did not have the possibility of integrating production. Comparing the pricing strategy of the model with integration with the price in a standard monopolistically competitive model (with no possibility of integration) is evident that the integration option significantly reduces the profit margins of the monopolistic sellers.

Figure 2 shows the producer’s mark-up \( (p - z)/z \) as a function of the cost draw \( z \) (for the same values of the parameters \( \eta \) and \( \lambda \) as in the previous picture). The dash-dotted line depicts the constant mark-up of the standard model without possibility of integration. The model displays endogenous mark-ups\(^{17}\), higher for the most productive sellers, whose productivity advantage allows

\(^{16}\)1/\( \lambda \) is equal to the standard deviation of the buyers’ cost distribution. In other words, when \( \lambda \) 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 (12) 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 an example of empirical productivity distribution, see Syverson (2004)).

\(^{17}\)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). I chose
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 stay in the market, but in percentage terms the prices they

the exponential specification also because it preserves the feature of mark-ups variability while keeping the solution relatively simple.
charge are closer to their marginal costs, so that the profit margin they actually get is very small. 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:

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

(14)

![Figure 3: Mark-up as a function of market share, closed economy ($\eta = 1.8, \lambda = 1$).](image)

These predictions are aligned with the ones obtained by Melitz and Ottaviano (2005), which achieve mark-ups variability and dependence of profits on productivity by assuming a specific functional form for the demand that firms face\textsuperscript{18}. Bernard, Eaton, Jensen and Kortum (2003) achieve similar features by assuming Bertrand competition in the intermediate goods sector\textsuperscript{19}. 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\textsuperscript{20}.

\textsuperscript{18}Melitz and Ottaviano (2005) use the linear demand system with horizontal product differentiation developed by Ottaviano, Tabuchi, and Thisse (2002).

\textsuperscript{19}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 (2005) 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.

\textsuperscript{20}As explained in Bernard, Eaton, Jensen and Kortum (2003), measured productivity for one unit of input is given
Regarding the dependence of the pricing rule on the model’s parameters, comparative statics of equation (13) implies that the optimal price is decreasing in $\eta$: when the degree of substitutability increases, competition in the market increases and potential buyers can more easily switch to cheaper substitutes, hence the intermediate goods producers must decrease the price charged to keep their share of the market. Moreover, the price is decreasing in $\lambda$. When $\lambda$ 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

Given $p(z)$, I can compute the price indexes $p^I$, $p^T$ and $p$. To close the model, I need to impose equilibrium conditions on the final good market and solve for the intermediate goods aggregate $q$. 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:

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

(15)

where $\alpha \in (0, 1)$ and $l_f$ is the labor force employed in the final good sector. Let $L$ denote the country’s total labor force; then $l_i = L - l_f$ is the labor force working in the intermediate good sector (for both suppliers and integrated segments of final good producers). The linearity of each intermediate good production technology implies:

$$q = \frac{l_i}{k}$$

(16)

where $k$ is the number of units of labor required to produce 1 unit of the aggregate $q$:

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

(17)

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

14
of \( q \) are:

\[
    l_f = \left( \frac{(1 - \alpha)p}{(1 - \alpha)p + \alpha k} \right) L \quad (18)
\]

\[
    l_i = \left( \frac{\alpha k}{(1 - \alpha)p + \alpha k} \right) L \quad (19)
\]

\[
    q = \left( \frac{\alpha}{(1 - \alpha)p + \alpha k} \right) L. \quad (20)
\]

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

\[
    r = \alpha^{-\alpha}(1 - \alpha)^{\alpha-1}p^\alpha. \quad (21)
\]

### 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. As in the previous section, a final good producer in country \( i \) \((i \in \{1,...,N\})\) needs to source a continuum of inputs to produce a final good, and for each of these inputs he can either decide to produce it or to buy it from a specialized seller. In an open economy though, he also has to decide where to implement the preferred option, as before comparing outside prices and costs of production, but taking into account also other variables like wages and trade costs.

Before introducing the notation, I recall the example of the previous section, and adapt it to the open economy setup. Nike and Adidas have to decide how to source the intermediate goods needed for shoes manufacturing, but their available options have increased. If they decide to buy shoelaces, they do not necessarily have to rely on a domestic supplier, but they can shop around and eventually locate a supplier in another country who can supply the same, adaptable shoelaces at a lower cost than the domestic supplier. In terms of the in-house option, both firms can integrate production of shoelaces by extending their facilities to incorporate the input production as well. So they can enlarge their domestic plants or — equivalently — build another plant in another location to produce shoelaces which are then shipped back to the headquarters of the firm for final assembly and distribution. For example, Nike can build a facility in Malaysia and exploit factor
costs differentials to produce the laces, and then ship them back to the U.S. On the other hand, Adidas may find favorable another country (being the firm German, a good option could be an Eastern European country where factor costs are presumably lower and the transportation costs are relatively low).

Based on productivity and factor cost differentials, each final good producer can optimally decide to integrate production abroad, generating flows of FDI. If — as I assume — the foreign investment realizes in ownership of the foreign production facility, when the inputs produced abroad are shipped back to the parents, 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.

To summarize, Nike and Adidas will choose the optimal sourcing option among outsourcing from the cheapest supplier around the world and integrating production in the most convenient country.

To formalize this idea in the framework of my model, I assume labor is immobile, so that wages\textsuperscript{21} 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 technology draws $\{x_i\}$, drawn from a country-specific distribution $\phi_i(x_i)$. If he decides to internalize an input, he may choose to do so in his own country or abroad, setting a production facility elsewhere. 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 carries its technology draw with him, but faces 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 $\tau_{ij}$ the unit iceberg cost for a final good producer from country $i$ to setup production of an input in country $j$:

\textsuperscript{21}Wages are going to be endogenously determined in equilibrium.
Assumption 2. $\tau_{ij} \geq 1 \ \forall \ i, j, \ \tau_{ij} = 1 \ \forall i = j \ \text{and} \ \tau_{ij} \leq \tau_{ik}\tau_{kj} \ \forall i, j, k.$

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

Let’s now turn to the outsourcing option. In the open economy, each country $j$ ($j \in \{1, ..., N\}$) is populated by a continuum of intermediate goods producers, each of whom produces a unique differentiated input for which he has an adaptable technology that allows 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 $\psi_j(z_j)$. An intermediate goods producer in country $j$ can only hire labor from country $j$, hence his marginal cost of production is $w_jz_j$. On the other hand, he can sell to final good producers worldwide, and to do so he is price competing with the producers of the same good in other countries. Finally, intermediate goods producers can do pricing-to-market, i.e. they can charge different prices to buyers in different countries. I denote with $p_{ij}(z_j)$ the price charged to a potential buyer in country $i$ by a monopolist in country $j$ who has a cost draw $z_j$.

When a final good producer decides to buy an input, he faces $N$ outside prices charged by sellers in different locations. In the open economy though, the effective price that a buyer ends up by paying is actually augmented by the existence of barriers to international trade, as tariffs and transportation costs. Hence, when a final good producer in country $i$ decides to buy an input from a seller in country $j$, the actual price he pays is $t_{ij}p_{ij}(z_j)$, where $t_{ij}$ represents the extent of barriers to trade between countries $i$ and $j$:

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

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 used by a final good producer in country $i$ who has cost draw $x_i$ for that good and also faces a set of suppliers with draws $z = \{z_j\}_{j=1}^N$ charging prices $\{p_{ij}(z_j)\}_{j=1}^N$.

\[\text{22} \text{All the productivity distributions } \{\phi_i(\cdot)\}_{i=1}^N, \{\psi_j(\cdot)\}_{j=1}^N \text{ are mutually independent across countries.}\]

\[\text{23} \text{This assumption is motivated by the fact that I think about the final good producers in the model as the (potential) multinational corporations, and about the intermediate goods producers as national suppliers.}\]
3.2 Organizational Choices and Location

The analysis of the model follows basically unchanged from the previous section. A final good producer in country $i$ observes his own set of technology draws $\{x_i\}$, a set of wages and iceberg costs $\{w_j, \tau_{ij}\}_{j=1}^N$, a set of outside C.I.F. prices $\{t_{ij}p_{ij}(z_j)\}_{j=1}^N$, and then 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 outside price of buying it. The problem is exactly as in the closed economy, but with a larger set of prices to shop for ($2N$ instead of 2).

Let $c_i(x_i, z)$ be the unit cost of obtaining good $(x_i, z)$:

$$c_i(x_i, z) = \min_j \{\tau_{ij}w_jx_i, t_{ij}p_{ij}(z_j)\}.$$  \hfill (22)

Notice that, 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\(^{25}\) of wages and iceberg costs worldwide for integrated production of a final good producer from country $i$:

$$m_i = \min_k \tau_{ik}w_k.$$  \hfill (23)

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.

---

\(^{24}\)Inclusive of trade costs.

\(^{25}\)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: I will show the optimal production allocation choice in case of ties 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 this version of 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. This representation fails to capture multilateral investment patterns that we observe in the data, but the cost of capturing those patterns would come at the price of further complications in the model. I do not think that this limitation invalidates the results of the paper, since the degree of multilaterality in the operations of U.S. multinational corporations appears to be limited: the vast majority of intrafirm trade by U.S. firms happens with affiliates located in a small number of countries. Data from the Bureau of Economic Analysis show that in 2004, more than 80% of imports of U.S. parents from their foreign affiliates were shipped by only 7 countries, representing in fact 3 main areas: Canada and Mexico, the U.K. and Ireland, and Malaysia, Hong Kong and Singapore. The concentration appears even stronger if we look at data by sector.
A final good producer with a set of cost draws \( \{x_i\} \) in country \( i \) solves:

\[
\min_{q_i(x_i,z)} \int_{\mathbb{R}^N_+} \int_{0}^{\infty} c_i(x_i, z) q_i(x_i, z) \phi_i(x_i) \psi(z) dx_i dz
\]

s.t. \[
\left[ \int_{\mathbb{R}^N_+} \int_{0}^{\infty} q_i(x_i, z)^{1-\eta} \phi_i(x_i) \psi(z) dx_i dz \right] \eta/(\eta-1) \geq q_i
\]  

where \( \psi(z) \) is the density of the vector \( z = (z_1, z_2, ... z_N) \) :

\[
\psi(z) = \prod_{j=1}^{N} \psi_j(z_j)
\]

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 country or in the set of countries with the lowest cost \( m_i \)) and let \( B_T^{ij} \) denote the set of goods that he decides to outsource from a producer in country \( j \):

\[
B_I^i = \left\{(x_i, z) \in \mathbb{R}^{N+1}_+ : c_i(x_i, z) = m_i x_i \right\}
\]

\[
B_T^{ij} = \left\{(x_i, z) \in \mathbb{R}^{N+1}_+ : c_i(x_i, z) = t_{ij} p_{ij}(z_j) \right\}.
\]

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

\[
\min_{q_i(x_i,z)} \left[ \int_{B_I^i} m_i x_i q_i(x_i, z) \phi_i(x_i) \psi(z) dx_i dz + \sum_{j=1}^{N} \int_{B_T^{ij}} t_{ij} p_{ij}(z_j) q_i(x_i, z) \phi_i(x_i) \psi(z) dx_i dz \right]
\]

s.t. \[
\left[ \int_{\mathbb{R}^{N+1}_+} q_i(x_i, z)^{1-\eta} \phi_i(x_i) \psi(z) dx_i dz \right] \eta/(\eta-1) \geq q_i
\]  

Problem (26) is solved by:

\[
q_I^i(x_i, z) = q_I^i(x_i) = (m_i x_i)^{-\eta} p_i^I q_i \quad \forall (x_i, z) \in B_I^i
\]

\[
q_T^i(x_i, z) = q_T^i(z_j) = [t_{ij} p_{ij}(z_j)]^{-\eta} p_i^T q_i \quad \forall (x_i, z) \in B_T^{ij}
\]  

\[26\] Notice that \( B_I^i \cup (\cup_j B_T^{ij}) = \mathbb{R}^{N+1}_+ \).
where $p_i$ is the aggregate price index in country $i$:

$$p_i = \left[ (p_i^I)^{1-\eta} + \sum_{j=1}^{N} (p_{ij}^T)^{1-\eta} \right]^{1/(1-\eta)} \quad (29)$$

and:

$$p_i^I = \left[ \int_{B_i^I} (m_ix_i)^{1-\eta}\phi_i(x_i)\psi(z)dx_idz \right]^{1/(1-\eta)} \quad (30)$$

$$p_{ij}^T = \left[ \int_{B_{ij}^T} [t_{ij}p_{ij}(z_j)]^{1-\eta}\phi_i(x_i)\psi(z)dx_idz \right]^{1/(1-\eta)} \quad (31)$$

It remains to determine the outside prices $\{p_{ij}(z_j)\}$. In the intermediate goods market, each supplier maximizes its profits from sales to potential buyers around the world, and may charge different prices to potential buyers in different countries. In choosing the optimal price to charge in a country, he must consider the competition from the producers of the same good in the other countries and the fact that the potential buyers have always the option to integrate production. Pricing-to-market from the side of the intermediate goods producers allows to study the pricing problem country by country. Each supplier in each country chooses $N$ prices to charge, one for each country it enters. Given that there are no fixed costs in the model, each seller will sell in every country.

Let $b_{ij}(p_{ij}(z_j))$ be the set of technology draws of buyers in country $i$ and of sellers outside country $j$ such that the buyers in country $i$ decide to buy the good from the seller in country $j$:

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

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

$$\max_{p_{ij}(z_j)} \int_{b_{ij}(p_{ij}(z_j))} [p_{ij}(z_j) - w_jz_j]q_i^T(z_j)\phi_i(x_i)\prod_{k \neq j} \psi_k(z_k)dx_idz_k. \quad (32)$$

Using (28), and due to the independence property of the draws’ distributions, problem (32) can be restated as:

$$\max_{p_{ij}(z)} [p_{ij}(z_j) - w_jz_j] \left( \frac{t_{ij}p_{ij}(z_j)}{p_i} \right)^{-\eta}q_iA_{ij}(p_{ij}(z_j)) \quad (33)$$
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 \( j \):

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

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

The solution of the pricing rule in the open economy is algebra-intensive, and I relegate it to the Appendix. Consistently with the closed economy section, I assume the cost draw distributions of both sectors in each country \( i \), \( \phi_i(x_i) \) and \( \psi_i(z_i) \) (\( i = 1, ... N \)), to be exponentials with parameters \( \lambda_i \) and \( \mu_i \) 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 \( z_i \geq 0 \).

Using Assumption 4, the solution of problem (33) is characterized by the following non-linear first order differential equation:

\[
p_{ij}'(z_j) = \frac{\xi_j p_{ij}(z_j) - w_j z_j}{\eta w_j z_j + \left( 1 - \eta + \frac{\lambda_i t_{ij}}{m_i} w_j z_j \right) p_{ij}(z_j) - \frac{\lambda_i t_{ij}}{m_i} p_{ij}(z_j)^2}
\] (34)

where \( \xi_j = \sum_{k \neq j} \mu_k \) is a measure of competitiveness of sellers in countries other than \( j \).

### 3.3 Properties of the Pricing Rule

Equation (34) does not have a closed form solution. To provide intuition about the properties of the pricing rule \( p_{ij}(z_j) \), I provide a semi-parametric approximation of the solution based on the parametric path method developed in Judd (2002).

I approximate the pricing function with a parameterization that respects the properties of the solution for \( z = 0 \) and asymptotically for \( z \to \infty \), and builds a smooth link between the conditions at the limits. For small values of \( z \), the solution of (34) has the form:

\[
p(z) \bigg|_{z \to 0} = \frac{\eta}{\eta - 1} wz + o(z)
\] (35)

which indicates that the most productive suppliers are not affected by the competition of suppliers
in other countries or by the possibility of integration of the buyers, since their optimal price is about the same (except for higher order terms, negligible for $z \rightarrow 0$) as in a standard monopolistically competitive model without possibility of integration on the side of the buyers. On the other hand, for very large values of $z$, the solution of (34) has the form:

$$p(z) \bigg|_{z \rightarrow \infty} = \frac{w}{m w + \xi} + wz$$

(36)

which implies percentage mark-ups converging to zero, approaching the perfectly competitive case.

Based on these conditions, I choose the following parameterization\(^{27}\), where the country indexes $i, j$ are suppressed to ease the notation:

$$\tilde{p}(z; a, \rho) = \left( a_1 z + a_2 z^2 + \sum_{l=3}^{M} a_l z^l \right) e^{-\rho z} + \left( \frac{w}{m w + \xi} + wz \right) \left( 1 - e^{-\rho z} \right).$$

(37)

The parameterization (37) reduces to an ordinary polynomial for $z \rightarrow 0$, and tends to the asymptotic solution (36) for $z \rightarrow \infty$, with the free parameter $\rho$ controlling the speed of convergence to the asymptote. The parameters $a_1$ and $a_2$ are computed algebraically (see Appendix), while $\{a_l\}_{l=3}^{M}$ are free parameters chosen to minimize the sum of squared residuals obtained by substituting the parameterization (37) and its derivative into equation (34). The order of the polynomial $M$ is set high enough to get a precise characterization of the solution (here I use $M = 10$).

Concerning the properties of the solution, equation (34) implies that 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)\(^{28}\). Globally, the solution lies between the marginal cost line $w_j z_j$ and the closed economy pricing rule. The value of the parameter $\xi_j$ affects the location of the curve in the sense that when international competition is tougher ("high" $\xi_j$), the solution approaches the marginal cost line, while when international competition is low ("low" $\xi_j$), 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 —

\(^{27}\)The choice of parameterization (37) is explained in the Appendix.

\(^{28}\)The equilibrium pricing rule must have positive first derivative (higher marginal costs result in higher prices). When $\xi_j = 0$, the only way in which the condition $\tilde{p}'_{ij}(z_j) > 0$ can be satisfied is when the denominator of the right hand side of (34) is equal to zero. The resulting expression is analogous to the quadratic expression characterizing the closed economy pricing rule (13).
prices and mark-ups shrink \(^{29}\).

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 prices (the solid line), 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 \(^{30}\) 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 \(\xi_j\) 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 the prices charged domestically and the export prices.

![Figure 4: Pricing-to-market (2 symmetric countries, \(t=1.3, \tau=1.5\).](image)

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 \(^{31}\), 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 country generates tougher foreign competition (\(\xi_j\) increases) and prices will be consistently adjusted downwards. Figure 5 shows the effect of international competition on prices through an increase in the number of countries in the economy.

\(^{29}\)Melitz and Ottaviano (2005) obtain the same qualitative result.

\(^{30}\)Exclusive of trade costs.

\(^{31}\)If for example the productivity dispersion in the foreign country is higher than in the home country, domestic sellers face potential buyers and competitors that are likely to be less productive, so — even charging higher prices — domestic sellers are still a better option for the foreign buyers.
In terms of the dependence of prices on the other parameters of the model, the price $p_{ij}(z_j)$ is increasing in the local wage $w_j$ and 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. Higher transportation costs and tariffs make the “gross” outside price less attractive for the potential buyers, so the sellers must decrease the net price they charge to be competitive in spite of these higher costs. The comparative statics with respect to $\lambda$ and $\eta$ is unchanged with respect to the closed economy section. 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 incentivate higher volumes of production in each country.

3.4 General Equilibrium

In the open economy, the final good is non-tradeable, and must be produced domestically using the intermediate goods aggregate and some (local) labor. The final good production function in
country $i$ $(i = 1, ... N)$ is:

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

(38)

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:

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

(39)

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 $i$ targeting market $j$. Let $l_i(x_i, z)$ denote the labor force producing good $q_i(x_i, z)$. The linearity of the intermediate goods production process allows to express the labor force segments as linear functions of the quantities \{$q_i$\}$_{i=1}^{N}$:

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

(40)

where, given the pricing rule in equation (34), 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. Taking the wages as given, (40) is a linear system of $N$ independent equations in the $N$ unknowns $q_i$: this system can be solved, delivering the equilibrium values of \{$q_i$\}$_{i=1}^{N}$ as functions of the wages only:

$$q_i^* = q_i(w_1, ... w_n)$$

(41)

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

\footnote{Notice that firms from many countries may decide to integrate production in the same countries, and there may be countries where integrated production does not occur. Hence it may happens that the labor force segments $l_{ji}^I$ are zero for some countries $i$.}

\footnote{Explicit expressions for the proportionality factors $k_{ji}^I$, $k_{ji}^T$ are derived in the Appendix.}
be equal to total expenditure in the final good:

$$r_i c_i = L_i w_i + \int_0^\infty \pi_i(z_i) \psi_i(z_i) dz_i \quad \text{for } i = 1, \ldots, N$$  \hspace{1cm} (42)

where $r_i$ is the zero-profit price of the final good in country $i$:

$$r_i = \alpha^{-\alpha} (1 - \alpha)^{(\alpha-1)} p_i \alpha^{1-\alpha}$$

and $\pi_i(z_i)$ is the total profit of an intermediate goods producer from country $i$ with cost draw $z_i$:

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

The market clearing condition (42) is a system of $N$ equations in the $N$ unknowns $\{w_i\}_{i=1}^{N}$ and can be solved for the wages’ equilibrium values.

The fact that the model must be solved numerically makes difficult to investigate the conditions for the existence of the equilibrium. In the following section, I compute the equilibrium and show its qualitative properties in the two-country case. This allows a simple graphical proof of the existence and uniqueness of the equilibrium.

### 3.5 Equilibrium Characterization: 2–Country Case

In a world of two countries, the equilibrium vector of wages can be expressed as one relative wage, and this allows to show the equilibrium as the wage that sets the excess demand in one country equal to zero (Walras’ law assures that the excess demand is zero in the other country as well).

I denote the two countries with $H$ (Home) and $F$ (Foreign). Normalizing to 1 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\(^{34}\). The excess demand

\(^{34}\)In the computation, the elasticity of substitution parameter and the input share in the production process are kept fixed at $\eta = 1.8$ and $\alpha = 0.3$; the trade barriers are set at $t = 1.1$ and $\tau = 1.2$. The figure represents the
correspondence is continuous and decreasing in the relative wage \( w_H \). For \( w_H \to 0 \), production is costless, so total production and profits diverge, and the excess demand is positive. Conversely, for \( w_H \to \infty \), both income and expenditure terms diverge, and the net effect is that the excess demand diverges to \(-\infty\). These features guarantee the existence of a unique finite equilibrium wage.

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

- \( w_H \in (0, w_F/\tau) \) \( \Rightarrow \) firms from both countries integrate in the Home country;
- \( w_H \in (w_F/\tau, \tau w_F) \) \( \Rightarrow \) firms from both countries integrate domestically;
- \( w_H \in (\tau w_F, \infty) \) \( \Rightarrow \) 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 symmetric case of two identical countries, hence the excess demand function crosses the zero line at \( w_H/w_F = 1 \).

\[ \text{Recall that a firm from country } i \text{ locates integrated activities in the country } j \text{ such that } \tau_{ij} w_j = \min_k \{ \tau_{ik} w_k \}. \]
domestically or abroad, while firms from $F$ integrate domestically. The picture shows that if firms from $H$ choose to integrate in only one country when they are indifferent, the equilibrium wage does 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 $F$ are indifferent about where to integrate production, while firms from $H$ integrate domestically. In this case are the firms from $F$ which will have integrated facilities in both countries.

![Diagram of labor demand in the integrated sectors](image)

Figure 7: Labor demand in the integrated sectors (symmetric countries, $t = 1.1$, $\tau = 1.2$).

At these critical points, the excess demand correspondence is non-smooth because the cost structure of the firms suddenly changes: Figure 7 shows the unit labor demand of integrated sectors of Home and Foreign firms. There is a kink in the unit labor demand at the point where firms switch from domestic to foreign integration and vice versa. When the equilibrium wage falls in a point corresponding to a kink, firms from one country are indifferent about where to integrate, and they integrate in both countries. The labor force in the integrated sectors in each country is endogenously determined to clear the labor market.

### 4 Comparative Statics: Numerical Analysis

In this section I explain the predictions of the model about the determinants of volumes of trade and foreign direct investment/intrafirm trade\(^{36}\). Since the aim of the theory is to explain sourcing

\(^{36}\)As I mentioned in the Introduction, the model does not distinguish between different ownership structures, since every decision is taken on a purely technological basis. I assume that vertical FDI always entails ownership, and I
strategies, and in the model vertical FDI takes the form of intrafirm imports, I show the predictions for volumes of intrafirm imports and arm’s length imports 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:

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

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

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

where $\gamma_i$ 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 2-country case. I isolate the effects of size by normalizing the world population to 1 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\(^\text{37}\).

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 the relative labor force of country $H$: 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 size, showing that smaller economies rely more on sourcing abroad. Trade and FDI coexist, and while trade is bilateral, FDI is only unilateral, since in the model 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

---

\(^{37}\)The elasticity of substitution parameter and the input share in the production process are $\eta=1.8$ and $\alpha=0.3$, while the trade barriers are $t = 1.1$ and $\tau = 1.2$. The productivity parameters are the same in the two countries: $\lambda_H = \lambda_F = \mu_H = \mu_F = 1$. 

---
(where labor is more abundant, hence cheaper), and the volume of FDI is increasing in the size differential. This prediction is consistent with the analysis of FDI in Nocke and Yeaple (2005), 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. 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.

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

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

Nocke and Yeaple report data from the Bureau of Economic Analysis 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 (and low-wage) countries.
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 equilibrium effect of size on equilibrium wages, and on the implications of different relative wages for the location choice: when country $H$ is “small” ($s \in (0, 0.2)$), 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: for $s \in (0.2, 0.4)$, $w_H = \tau w_F$ 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 $s \in (0.4, 0.6)$, 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 $s \in (0.6, 1)$.

### 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 for integrated production of Home country firms, $\lambda_H$.

![Figure 10: Volumes of arm’s length import and FDI import as functions of buyers’ costs dispersion.](image)

When $\lambda_H$ increases, the cost distribution of integrated production becomes less disperse, and 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 $\lambda_H$ increases). As a result, the share of intrafirm imports over total imports is increasing in $\lambda$: 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 $\lambda_H$ 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.

I repeated the experiment plotting volumes of trade as functions of the productivity dispersion of the intermediate goods producers. The results are displayed in Figure 11. The variable on the horizontal axis is the dispersion parameter of the suppliers’ cost distribution in the Home country, $\mu_H$. When $\mu_H$ 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 $\mu_H$ 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 (productivity effect). Overall, increases in the productivity of domestic suppliers make foreign suppliers less attractive, so $H$’s arm’s length imports drop as $\mu_H$ increases. For “low” values of $\mu_H$, ($\mu_H \in (0,1.8)$), $w_H$ is low and arm’s length imports are substituted by domestic integration. As $\mu_H$ further rises, $w_H$ rises and foreign integration becomes profitable, so arm’s length imports are substituted by intrafirm imports (the wage effect dominates). As $\mu_H$ keeps rising ($\mu_H \geq 4$), intrafirm imports start declining because

![Figure 11: Volumes of arm’s length import and FDI import as functions of sellers’ costs dispersion.](image-url)
the productivity 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 \( \mu_H \) (low \( w_H \)), and overall \( F \) arm’s length imports increase as \( \mu_H \) 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. For this computation, the values of the fixed parameters are the same as in the previous sections (\( \eta = 1.8, \alpha = 0.3, t = 1.1 \) and \( \tau = 1.2 \)). Since in a two-country world only relative variables matter, I also fix the labor force of the Home country \( (L_H = 100) \) and the parameters of the cost distributions of the Foreign country \( (\lambda_F = \mu_F = 1) \).

<table>
<thead>
<tr>
<th>( L_F = L_H, ) ( \lambda_H = 1; \mu_H = 2 ) ( (\gamma_H = .8) )</th>
<th>( L_F = L_H, ) ( \lambda_H = \mu_H = 2 ) ( (\gamma_H = .74) )</th>
<th>( L_F = 2L_H, ) ( \lambda_H = \mu_H = 2 ) ( (\gamma_H = .64) )</th>
<th>( L_F = 4L_H, ) ( \lambda_H = \mu_H = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>symmetric countries</td>
<td>1.03</td>
<td>7.7</td>
<td>0</td>
</tr>
<tr>
<td>( L_F = L_H, ) ( \lambda_H = 1; \mu_H = 2 ) ( (\gamma_H = .8) )</td>
<td>1.16</td>
<td>7.29</td>
<td>20.01</td>
</tr>
<tr>
<td>( L_F = L_H, ) ( \lambda_H = \mu_H = 2 ) ( (\gamma_H = .74) )</td>
<td>1.25</td>
<td>5.27</td>
<td>42.8</td>
</tr>
<tr>
<td>( L_F = 2L_H, ) ( \lambda_H = \mu_H = 2 ) ( (\gamma_H = .64) )</td>
<td>1.29</td>
<td>5.27</td>
<td>70.17</td>
</tr>
<tr>
<td>( L_F = 4L_H, ) ( \lambda_H = \mu_H = 4 )</td>
<td>1.76</td>
<td>3.36</td>
<td>85.32</td>
</tr>
</tbody>
</table>

Table 1: Computation results, 2-countries economy.

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 \( \gamma_H \) 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 (first row of the table) wages are the same and there is no incentive to foreign investment, nonetheless the model predicts a significant volume of arm’s length trade. Productivity heterogeneity on the side of the suppliers (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). Higher productivity on the side of the buyers (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 $\lambda$), 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 $\lambda$), 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 Input Differentiation and Intrafirm Import

In this section, I provide some evidence on the role of product differentiation and market structure in shaping firms’ sourcing decisions. To introduce the empirical analysis of the prediction I am going to test, I consider a slightly modified version of the model. Consider a setup where the final consumption good is produced using labor and two types of intermediate aggregates: $c = q_h q_l^{\beta} l^{1-\alpha-\beta}$, where $q_h$ and $q_l$ only differ in the degree of substitutability across intermediate goods: $q_h$ ($q_l$) is characterized by a higher (lower) elasticity of substitution $\eta_h$ ($\eta_l$). The suppliers of the inputs composing $q_l$ are going to charge on average higher prices and mark-ups with respect to their counterparts producing goods entering $q_h$, so there is a higher incentive for the final good producers to integrate input production in sector $l$.

In the words of the model, this is equivalent to say that the intrafirm share of imports is inversely
related to the elasticity of substitution $\eta$. The intrafirm share of imports of country $i$ from country $j$ can be expressed as:

$$intrafirm\ share\ of\ import = \left[ 1 + \frac{\int_{B_i^T}^{B_i^I} [t_{ij}p_{ij}(z_j)]^{1-\eta} \phi_i(x_i) \psi(z) dx_i dz}{\int_{B_i^I}^{B_i^I} \phi_i(x_i) \psi(z) dx_i dz} \right]^{-1}.$$ \hspace{1cm} (43)

When $\eta$ increases, import prices $p_{ij}(z_j)$ drop, and arm’s length import expenditure increases (since both $[t_{ij}p_{ij}(z_j)]^{1-\eta}$ and $B_i^T$ increase). On the other side, intrafirm import expenditure drops, since the set $B_i^I$ shrinks, and as a result $\frac{\partial \text{share}}{\partial \eta} < 0$. Figure 12 shows this result numerically\(^{39}\). For a set of calibrated parameters, the intrafirm share of imports decreases monotonically from 24% to 8% as the parameter $\eta$ increases from 1 to 2.

Figure 12: Intrafirm share of import as a function of $\eta$.

This comparative statics can be generalized to say that, in a multi–sector framework, the model predicts that the share of intrafirm imports over total imports is going to be higher, the higher the degree of differentiation (the lower the elasticity of substitution) across goods in that sector.

Figure 13 shows some preliminary evidence about the empirical validity of this prediction. The variable on the horizontal axis is a measure of within-sector product differentiation, while the variable on the vertical axis is the volume of imports of U.S. parents from their foreign affiliates.

\(^{39}\)The plot in Figure 12 is done computing the model for the values of the parameters I calibrate in the next section: $\alpha = 0.25$, $L_i = 100$, $L_j = 692.7176$, $\lambda_i = \mu_i = 2.22$, $\lambda_j = \mu_j = 1$, $t_{ij} = 1.1$ and $\tau_{ij} = 3.0184$. 

35
as a share of U.S. total imports. Data are plotted for 29 manufacturing industries and 4 years. The construction of both variables is discussed in detail in the following subsection. The positive correlation in the plot clearly shows how intrafirm imports are a prevailing sourcing channel in sectors where goods are highly differentiated.

5.1 Specification and Data

The hypothesis I test is that the share of intrafirm imports of U.S. parents from their foreign affiliates is higher, the higher the degree of differentiation across goods in a sector. In a recent influential paper, Antràs (2003) has shown that the share of intrafirm imports is positively related to the capital share of the exporting industry. Even if the theory presented in this paper has nothing to say about the role of factor intensity for intrafirm trade, it is necessary to include this variable in the analysis to disentangle its effect from the effect of product differentiation.

In the next subsection I report estimates from regressions of the form:

$$ln(S_{jt}) = \beta_0 + \beta_1 D_{jt} + \beta_2 \ln(K_{jt}/L_{jt}) + \varepsilon_{jt}$$  \hspace{1cm} (44)

where $S_{jt}$ is equal to imports of U.S. parents from their foreign affiliates as a fraction of total U.S.
imports, by sector and year, $D_j^q$ is the degree of product differentiation in sector $j$, $K_{jt}/L_{jt}$ is the capital/labor ratio, by sector and year, and $\varepsilon_{jt}$ is an orthogonal error term.

The left-hand side variable is constructed combining data on intrafirm imports from the Bureau of Economic Analysis$^{40}$ and data on total U.S. imports available from the Center for International Data at UC Davis$^{41}$. Since the model does not predict bilateral intrafirm trade flows, I identify intrafirm imports with imports of U.S. parents from their foreign affiliates$^{42}$. Data on intrafirm import by industry are available for the 34 BEA manufacturing sectors. Data on total imports are available at the 5-digit SITC level, and I aggregate them to the BEA classification level with a conversion table available on the same website. I construct the share of intrafirm imports for 4 years (1983–1987–1992–1996) and exclude natural resources industries (Primary Metal Industries, Ferrous and Nonferrous) and those sectors for which data are missing (Tobacco Products, Leather and Leather Products, Other Paper and Allied Products). I end up with 29 industries for 4 years, for a total of 116 observations. The share of intrafirm imports varies from a minimum of 0.29% for Apparel and Other Textile Products in 1987 to a maximum of 79.39% for Electronic Components in 1983.

Like in Antràs (2003), the capital intensity variable $K_{jt}/L_{jt}$ is taken from the NBER Manufacturing Industry Productivity Database$^{43}$, and is measured as the ratio of total capital stock to total employment in the corresponding industry.

I now turn to the description of the degree of differentiation variable. In the model, the degree of differentiation within a sector is captured by the elasticity of substitution parameter. Broda and Weinstein (2006) estimate sectoral elasticities of substitution from price and volume data on U.S. consumption of imported goods. By using data at the 10-digit Harmonized System, they estimate how much demand shifts between 10-digit varieties when relative prices vary, within each 3-digit SITC sector. In order to overcome the measurement error problems associated with using their estimates as regressors, I build a dummy that takes value 1 (0) when the estimated sectoral elasticity is below (above) the median value (2.54):

$$D_{SITC}^\eta = \begin{cases} 1 & \text{if } \eta_i < \text{median}(\eta_i) ; \\ 0 & \text{if } \eta_i \geq \text{median}(\eta_i) . \end{cases}$$

$^{40}$Henceforth, BEA. Sector-level data on the operations of U.S. multinational companies are available at: http://www.bea.gov/international/di1usdop.htm.


$^{42}$Antràs includes in his analysis also imports shipped to U.S. affiliates by their foreign parents.

$^{43}$http://www.nber.org/nberces/nbprod96.htm.
In order to aggregate this variable at the BEA classification level, I weight each value of $D^n_{SITC_i}$ by the import share of the 3-digit SITC sector in the corresponding BEA sector:

$$D^n_{BEA_j} = \sum_{SITC_i \in BEA_j} \frac{\text{import}_{SITC_i}}{\text{import}_{BEA_j}} \times D^n_{SITC_i}.$$  

The resulting variable assigns values between 0 and 1 to each of the 34 BEA industries. A low value of $D^n_{BEA_j}$ is associated to commodities, while a high value is associated to highly differentiated sectors. Table 2 reports the value of $D^n_{BEA_j}$ for each of the industries I include in my calculations.

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>$D^n_{BEA_j}$</th>
<th>Code</th>
<th>Description</th>
<th>$D^n_{BEA_j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Grain, Mill and Bakery Products</td>
<td>0</td>
<td>21</td>
<td>Construction, Mining, etc.</td>
<td>0.4979</td>
</tr>
<tr>
<td>2</td>
<td>Beverages</td>
<td>0</td>
<td>22</td>
<td>Computer and Office Equipment</td>
<td>0.4092</td>
</tr>
<tr>
<td>4</td>
<td>Other Food and Kindred Products</td>
<td>0.4303</td>
<td>23</td>
<td>Other Non Electric Machinery</td>
<td>0.6464</td>
</tr>
<tr>
<td>5</td>
<td>Apparel and Other Textile Products</td>
<td>0.4093</td>
<td>24</td>
<td>Household Appliances</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Pulp, Paper and Boardmills</td>
<td>0</td>
<td>25</td>
<td>Household Audio and Video, etc.</td>
<td>0.8958</td>
</tr>
<tr>
<td>9</td>
<td>Printing and Publishing</td>
<td>1</td>
<td>26</td>
<td>Electronic Components</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Drugs</td>
<td>1</td>
<td>27</td>
<td>Other Electrical Machinery</td>
<td>0.3919</td>
</tr>
<tr>
<td>11</td>
<td>Soaps, Cleaners and Toilet Goods</td>
<td>0.7006</td>
<td>28</td>
<td>Motor Vehicles and Equipment</td>
<td>0.6503</td>
</tr>
<tr>
<td>12</td>
<td>Agricultural Chemicals</td>
<td>0.2304</td>
<td>29</td>
<td>Other Transportation Equipment</td>
<td>0.2657</td>
</tr>
<tr>
<td>13</td>
<td>Industrial Chemicals and Syntetics</td>
<td>0.4891</td>
<td>30</td>
<td>Lumber, Wood, Furniture, etc.</td>
<td>0.5101</td>
</tr>
<tr>
<td>14</td>
<td>Other Chemicals</td>
<td>1</td>
<td>31</td>
<td>Glass Products</td>
<td>0.5518</td>
</tr>
<tr>
<td>15</td>
<td>Rubber Products</td>
<td>0.8827</td>
<td>32</td>
<td>Stone, Clay, Concrete, Gypsum, etc.</td>
<td>0.8330</td>
</tr>
<tr>
<td>16</td>
<td>Miscellaneous Plastic Products</td>
<td>0</td>
<td>33</td>
<td>Instruments and Apparatus</td>
<td>0.9650</td>
</tr>
<tr>
<td>19</td>
<td>Fabricated Metal Products</td>
<td>0.4464</td>
<td>34</td>
<td>Other Manufacturing</td>
<td>0.5131</td>
</tr>
<tr>
<td>20</td>
<td>Farm and Garden Machinery</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Sector differentiation variable.

To check whether the results are sensible to the methodology chosen to construct $D^n_{BEA_j}$, I construct the same variable and run the regressions also based on another classification. Rauch (1999) classified goods at the 4-digit SITC level in three categories: homogeneous goods or commodities (if traded in organized exchanges), reference priced goods (if they have price quotations reported in trade journals), and differentiated goods (if they could not be priced in either of these ways). To construct my differentiation variable, I define a dummy that assigns value 0 to homogeneous goods, value 0.5 to reference priced goods and value 1 to differentiated products:

$$D^n_{Rauch\ SITC_i} = \begin{cases} 1 & \text{if } i \in \text{differentiated goods} \\ 0.5 & \text{if } i \in \text{reference priced goods} \\ 0 & \text{if } i \in \text{commodities} \end{cases}.$$
The variable $D^{η}_{Rauch\ SITC}$ is aggregated to the BEA classification level following the same weighting scheme as in (45).

5.2 Results

The upper panel of Table 3 reports OLS estimates of equation (44), with the product differentiation variable based on Broda and Weinstein (2006)’s estimates. Columns 1 and 2 report the results of the regression isolating each regressor; column 3 inserts both independent variables into the analysis. In column 4, I cluster the standard errors by sector, to control for potential within-sector serial correlation of the error term. Column 5 adds year fixed effects, to control for any year-specific shock that may affect the results. The standard errors are reported in parentheses, and *** indicates statistic significance at the 1% significance level.

The coefficient on $D^{η}_{j}$ has the expected positive sign, and is statistically significant at the 1% significance level across all columns. Clustering the standard errors and including year fixed effects does not change the estimated coefficients\(^{44}\). The coefficient $\hat{\beta}_1 = 1.478$ in column 5 means that a change in $D^{η}_{j}$ from 0 to 1 is associated to a 338% increase in the intrafirm share of imports, i.e. highly differentiated goods ($D^{η}_{j} = 1$) are associated to a share of intrafirm imports on average 3.38 times higher than the one associated to commodities ($D^{η}_{j} = 0$).

The lower panel of Table 3 reports the results of the same regressions with the variable $D^{η}_{j}$ based on the classification in Rauch (1999). The coefficients have different magnitudes with respect to the previous estimates, reflecting the different construction of the variable, but are of the expected sign and statistically significant at the 1% level in all the columns.

6 The Gains from Multinational Production

In this section, I provide a calibrated version of the two-country model that matches some relevant facts about trade and volumes of multinational activity for U.S. firms. With the calibrated model, I quantify the gains — for the U.S. economy and for the rest of the world\(^ {45}\) — arising from the possibility of FDI through the channel of foreign integration. I perform counterfactual experiments to show how the gains depend on the degree of competition in the market and on the extent of

\(^{44}\) None of the year dummies is statistically significant.

\(^{45}\) Henceforth, ROW.


<table>
<thead>
<tr>
<th>Dep. Var: $\ln(S_{jt})$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Broda and Weinstein classification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D^\eta_j$</td>
<td>1.102***</td>
<td>1.474***</td>
<td>1.474***</td>
<td>1.478***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.293)</td>
<td>(0.251)</td>
<td>(0.434)</td>
<td>(0.442)</td>
<td></td>
</tr>
<tr>
<td>$\ln(K_{jt}/L_{jt})$</td>
<td>0.829***</td>
<td>1.005***</td>
<td>1.005***</td>
<td>1.018***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.159)</td>
<td>(0.143)</td>
<td>(0.246)</td>
<td>(0.255)</td>
<td></td>
</tr>
<tr>
<td>clustered errors</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>year fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.102</td>
<td>0.185</td>
<td>0.370</td>
<td>0.381</td>
<td>0.385</td>
</tr>
<tr>
<td>No. of obs.</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
</tr>
<tr>
<td>No. of clusters</td>
<td>29</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rauch classification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D^\eta_j$</td>
<td>0.881</td>
<td>3.528***</td>
<td>3.528***</td>
<td>3.579***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.593)</td>
<td>(0.559)</td>
<td>(0.899)</td>
<td>(0.954)</td>
<td></td>
</tr>
<tr>
<td>$\ln(K_{jt}/L_{jt})$</td>
<td>0.829***</td>
<td>1.408***</td>
<td>1.408***</td>
<td>1.440***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.159)</td>
<td>(0.165)</td>
<td>(0.257)</td>
<td>(0.272)</td>
<td></td>
</tr>
<tr>
<td>clustered errors</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>year fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.010</td>
<td>0.185</td>
<td>0.392</td>
<td>0.403</td>
<td>0.410</td>
</tr>
<tr>
<td>No. of obs.</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
</tr>
<tr>
<td>No. of clusters</td>
<td>29</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Intrafirm Share, Factor Intensity and Product Differentiation.

the barriers to foreign investment. I isolate the gains from trade from the gains from multinational production, and compare them with other papers’ findings.

6.1 Calibration

I want to provide a calibrated version of the two-country model, with the Home country being the United States, and the Foreign country being an aggregate of the rest of the world. In this calibration, I adopt some parameters from the literature, I compute directly the productivity parameters, and — for those parameters for which I do not have sufficient information — I write an algorithm that selects the values that best match some relevant moments of the data.

The parameter $(1 - \alpha)$ 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 \(\alpha\) 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 model restricts the elasticity of substitution between intermediate goods to \(\eta \in (1, 2)\) to assure an elasticity of demand larger than 1 and convergence of the price integrals. In my model, \(\eta\) 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 \(\eta (\eta = 1.2)\) to show how the gains from multinational production depend on the degree of competition in the market\(^{46}\).

In a two-country world only relative variables matter, so I need values for the relative average productivity of firms. I identify the ratio \(\mu_{\text{us}}/\mu_{\text{row}}\) with the relative average productivity of U.S. firms with respect to an average of their counterparts in the ROW. I compute this ratio as:

\[
\mu_{\text{us}}/\mu_{\text{row}} = \frac{(GDP \text{ per worker})_{\text{us}}}{(GDP \text{ per worker})_{\text{row}}} = 2.22
\]

where the ROW’s labor productivity is a weighted average of each country’s labor productivity, with the shares of US imports from that country as weights:

\[
(GDP \text{ per worker})_{\text{row}} = \sum_{i \neq \text{us}} \left( \frac{GDP_i}{\text{labour force}_i} \times \frac{\text{imports}_{\text{us},i}}{\text{imports}_{\text{us,row}}} \right).
\]

To construct the ROW entry I aggregate data for 157 countries, representing 98% of the world GDP. The data on U.S. imports by country are from the U.S. Census\(^{47}\), while data on GDP and labor force by country are from the World Bank’s World Development Indicators\(^{48}\) (WDI). All the data used in this calibration are for the year 2004.

Recent empirical research about the relative productivity of MNCs versus non-multinational firms seems to agree on the fact that MNCs tend to be more productive than domestic, non-multinational firms. This fact is documented in Bernard, Jensen and Schott (2005) for U.S. firms,\(^{46}\)

---

\(^{46}\)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.

\(^{47}\)http://www.census.gov/foreign-trade/balance/.

and in Criscuolo and Martin (2005) for U.K. firms. Not having a precise measure of this productivity differential, I assume that final good producers (the potential MNCs) draw their productivity from the same distribution of local firms, and hence have the same relative average productivity: \( \lambda_{us}/\lambda_{row} = \mu_{us}/\mu_{row} = 2.22 \).

As in Alvarez and Lucas (2007), the parameter \( L_i \) represents labor in efficiency units in country \( i \), and as such it cannot be observed directly. Similarly, the calibration of the average iceberg costs \( t \) and \( \tau \) is problematic because it requires to aggregate non-U.S. countries’ entries and because I do not have reliable measures of the costs of intrafirm transactions. In order to recover values for these parameters, I calibrate them jointly to match three magnitudes in the data that are relevant for the model to fit the facts I am interested into.

The first magnitude I match is the ratio of intrafirm imports of U.S. parents from their foreign affiliates over total U.S. imports: this share was 13.5% in 2004, and almost constant over the last decade (oscillating between 13 and 15% from 1992 to 2004). 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\(^{49}\) (BEA) with the Census data on U.S. imports. The share I construct is smaller than the ones reported by other papers\(^{50}\), 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. The second magnitude I match is the volume of U.S. import as a fraction of GDP. Using Census data for the year 2004, I obtain a value of 13.3%. The third magnitude I match is the share of U.S. GDP in world GDP, which was 30% in 2004 (from WDI’s GDP data). The model calibrated to match these magnitudes implies \( L_{us}/L_{row} = 0.14, \ t = 1.1 \) and \( \tau = 3.02 \).\(^{51}\)

Table 4 summarizes the calibrated parameters I use in the computation of the welfare gains. The values in parentheses for \( \tau \) and \( L_{us}/L_{row} \) are the ones used when choosing \( \eta = 1.2 \).


\[^{50}\]Antràs (2003) and Bernard, Jensen and Schott (2005) report an intrafirm share of total imports of about 40%.

\[^{51}\]The calibrated value of the iceberg cost of intrafirm import implies that producing one unit of input abroad almost triplicates its unit costs. I believe that the necessity of this high cost to match the data depends on the fact that the model does not consider other types of transaction costs, fixed costs of entering the foreign market, or legal restrictions to intrafirm activities. These frictions — that the model does not consider explicitly — are reflected in the calibration of the unit cost of foreign integration.
<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
<th>definition</th>
<th>source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.25</td>
<td>labor share in non-tradeables</td>
<td>Alvarez and Lucas (2007)</td>
</tr>
<tr>
<td>$\eta$</td>
<td>1.8 (1.2)</td>
<td>elasticity of substitution</td>
<td>model restrictions</td>
</tr>
<tr>
<td>$\mu_{us}/\mu_{row}$</td>
<td>2.22</td>
<td>suppliers’ relative productivity</td>
<td>U.S. Census, WDI (2004)</td>
</tr>
<tr>
<td>$\lambda_{us}/\lambda_{row}$</td>
<td>2.22</td>
<td>buyers’ relative productivity</td>
<td>U.S. Census, WDI (2004)</td>
</tr>
<tr>
<td>$t$</td>
<td>1.1</td>
<td>iceberg trade cost</td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>3.02 (3.23)</td>
<td>iceberg FDI cost</td>
<td>to match data</td>
</tr>
<tr>
<td>$L_{us}/L_{row}$</td>
<td>0.14 (0.137)</td>
<td>relative labor in efficiency units</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Summary of calibrated parameters.

6.2 Gains from Multinational Production: Decomposition and Policy Experiments

With the calibrated model, I compute the welfare gains (in terms of percentage changes in consumption per capita) that the theory implies compared to a counterfactual world without the possibility of integration and multinational production.

The gain in consumption per capita is computed as follows:

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

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.

In the model though, integrated production can happen domestically or abroad, so I need to disentangle how much of the welfare gain comes from domestic activity and how much of it comes from foreign investment. The calibrated model implies that U.S. firms decide to integrate both domestically and abroad, and delivers the equilibrium share of labor 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 5. The calibrated economy implies a gain in U.S. consumption per capita of 4.84% with respect to a world where there is no possibility of integration. Multinational activity (foreign integration) accounts for a gain of 0.72% only, while the rest is due to integrated activity at home. The small gain from multinational production is due to the fact that the model is calibrated to match the share of intrafirm import of U.S.-based MNCs over total imports, which is small compared to the size of the economy\textsuperscript{52} (13.5% of imports, where imports are about 13% of U.S. GDP).

\begin{table}[h]
\centering
\begin{tabular}{lccc}
\hline
 & \textit{baseline calibration} & \textit{FDI reform} & \textit{higher market power} \\
 & \textit{(}\eta = 1.8, \tau = 3.02\text{)} & \textit{(}\eta = 1.8, \tau' = 2\text{)} & \textit{(}\eta' = 1.2, \tau = 3.23\text{)} \\
\hline
\textbf{U.S.} & & & \\
\text{welfare gains (%)} & 4.84 & 7.37 & 12.68 \\
of which: & & & \\
	ext{domestic integration} & 4.12 & 0 & 10.8 \\
foreign integration & 0.72 & 7.37 & 1.88 \\
of which: & & & \\
predictivity effect & 4.64 & 6.85 & 6.15 \\
competition effect & 0.2 & 0.52 & 6.53 \\
\hline
\text{implied share of U.S.} & 13.69 & 63.32 & 14.28 \\
\text{intrafirm import (%)} & & & \\
\hline
\textbf{ROW} & & & \\
\text{welfare gains (%)} & 12.41 & 15.14 & 22.39 \\
of which: & & & \\
	ext{domestic integration} & 11.82 & 11.82 & 21.78 \\
foreign integration & 0.59 & 3.32 & 0.61 \\
of which: & & & \\
predictivity effect & 11.82 & 14.67 & 14.74 \\
competition effect & 0.59 & 0.47 & 7.65 \\
\hline
\end{tabular}
\caption{Gains from multinational production: decompositions and policy experiments.}
\end{table}

\textsuperscript{52}Since both the model and the calibration disregard other channels driving intrafirm imports, the computed gains should be interpreted as a lower bound on the total gains from foreign investment and intrafirm activity.
Table 5 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 productivity 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 input price level.

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 (4.13% for the U.S. and 4.75% for the ROW) of the total welfare gain, 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 (0.59% 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 5 reports the same calculations performed in a world where the unit cost of integrated production abroad drops of 50% (from the calibrated value $\tau = 3.02$ to $\tau' = 2$). Lowering this cost increases the gains, and the levels of consumption are significantly higher in both countries. Notice that this experiment still involves a very high unit cost of foreign production, so we can argue that dropping $\tau$ at an even lower level will increase further the welfare gains. The drop in $\tau$ generates a shift in the world allocation of production: all integrated activity of U.S. firms is now happening abroad, and the 7.37% 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 3.32% 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 5 reports the same calculations for a lower value of the elasticity of substitution \( \eta \) \((\eta' = 1.2)\). This version of the calibration corresponds to a world where the degree of differentiation across intermediates is lower. 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. 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. Moreover, under this scenario the second decomposition of the gains looks significantly different: a much larger share is due to the competition effect (51% in the U.S., 34% in the ROW), because suppliers reduce their high mark-ups more than in the previous cases.

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

In this section, I use the calibrated model to perform an experiment that can be used to compare the gains that this model produces with what other authors have found using different underlying theories. Using the parameters calibrated in the previous section, I compute consumption in the two countries in the autarky case, in which the barriers to trade and FDI \( t \) and \( \tau \) 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 6. The gain for the U.S. is equal to 26% of consumption per capita. This number is significantly higher than other estimates (for the U.S., Alvarez and Lucas (2007) estimate a gain of 10%, while Eaton and Kortum (2002) obtain a gain of 17%), because imperfect competition on the side of the suppliers implies that both domestic and foreign suppliers shrink their mark-ups when facing tougher competition from abroad. The gains in the ROW are smaller, but still significant. The difference between the two countries is due to the productivity differential: U.S. firms have higher exports and profits than ROW firms.

46
<table>
<thead>
<tr>
<th></th>
<th>U.S.</th>
<th>ROW</th>
</tr>
</thead>
<tbody>
<tr>
<td>autarky</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>costless trade and no FDI</td>
<td>1.26</td>
<td>1.08</td>
</tr>
<tr>
<td>costless trade and FDI</td>
<td>1.47</td>
<td>1.22</td>
</tr>
</tbody>
</table>

Table 6: Gains from moving from autarky to costless trade and FDI.

The third row of the table computes consumption in the two zones allowing costless trade and also costless foreign integration ($t = \tau = 1$). This possibility implies an additional increase in consumption of about 17% for the U.S. firms, and of about 13% for the ROW, for a total gain with respect to autarky of 47% and 22% 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\(^{53}\). Given their large role in total world research\(^{54}\), the gains for the U.S. are much lower than this upper bound (about 10% of consumption). Compared to Rodríguez-Clare’s analysis, my model concentrates the attention on a very specific channel of diffusion — vertical FDI —, nonetheless the computed gains are larger. This feature depends on the different source of the gains I consider. In Rodríguez–Clare (2007), openness allows countries to profit from ideas generated in other countries, so there is not much to gain for the countries that account 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 (2007) compute the welfare gains from allowing firms to relocate their 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, their computed gains are smaller because in their model the relocation of managerial ability abroad entails a loss of productivity at home.

\(^{53}\)Precisely, he concludes that a country accounting for 1% of the total world research experiences gains of 183% when moving from autarky to free trade and is open to diffusion of new ideas to/from other countries.

\(^{54}\)Rodríguez-Clare reports that the scientists and engineers engaged in research that reside in the U.S. are about 41% of the world total.
7 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.

Imperfect competition creates a wedge between trade prices and transfer prices. In sectors where inputs are more differentiated, this wedge will be larger, and intrafirm transactions will prevail. I find strong support for this prediction in the data: the regression analysis shows that input differentiation explains about 20% of the observed intrafirm share of imports.

Moreover, imperfect competition allows to study optimal pricing in presence of multinationals: 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 competitiveness.

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 calibrated the model to match aggregate U.S. data and computed the implied gains from multinational production and intrafirm trade. The welfare gains are currently quite small (about 1% of consumption per capita), but the model shows that further liberalization would substantially increase them.

Extensions of the model should be devoted to a better 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

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 $i$. The first order condition of problem (33) is:

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

Dividing each term by $[t_{ij}p_{ij}(z_j)]^{-\eta-1}$ and collecting common terms:

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

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

$$
[p_{ij}(z_j) - \eta(p_{ij}(z_j) - w_jz_j)] - [p_{ij}(z_j) - w_jz_j)p_{ij}(z_j) \frac{\phi_i \left( \frac{t_{ij}p_{ij}(z_j)}{m_i} \right) \frac{t_{ij}}{m_i} \prod_{k \neq j} \left[ 1 - F_{ik} \left( \frac{t_{ij}p_{ij}(z_j)}{t_{ik}} \right) \right]} {1 - \Phi_i \left( \frac{t_{ij}p_{ij}(z_j)}{m_i} \right)} + \sum_{l \neq j} f_{il} \left( \frac{t_{ij}p_{ij}(z_j)}{t_{il}} \right) \frac{t_{ij}}{t_{il}} = 0.
$$
Since $p'_{ij}(z_j) > 0$:

\[
F_{ij}(p_{ij}(z_j)) = \Psi_j(p_{ij}(z_j)) \quad \text{(A.1)}
\]

\[
f_{ij}(p_{ij}(z_j)) = \psi_j(p_{ij}(z_j)) \quad \text{(A.2)}
\]

and the first order condition becomes:

\[
p_{ij}(z_j)p_{ij}(z_j) - w_j z_j \sum_{l \neq j} \frac{\psi_l \left( \frac{t_{ij} p_{ij}(z_j)}{m_l} \right)}{1 - \Phi_l \left( \frac{t_{ij} p_{ij}(z_j)}{m_l} \right)} - w_j z_j p_{ij}(z_j)^2 = \xi_j \quad \text{(A.3)}
\]

Assuming that the cost draws are exponentially distributed, with country-specific and sector-specific parameters:

\[
\phi_i(x_i) = \lambda_i e^{-\lambda_i x_i} \quad \text{for } x_i \geq 0 \quad \text{(A.4)}
\]

\[
\psi_i(z_i) = \mu_i e^{-\mu_i z_i} \quad \text{for } z_i \geq 0 \quad \text{(A.5)}
\]

the pricing rule in (A.3) reduces to:

\[
p'_{ij}(z_j) = \frac{\xi_j p_{ij}(z_j)[p_{ij}(z_j) - w_j z_j]}{\eta w_j z_j + (1 - \eta + \frac{\lambda_i}{\mu_i} w_j z_j) p_{ij}(z_j) - \frac{\lambda_i}{\mu_i} p_{ij}(z_j)^2} \quad \text{(A.6)}
\]

where $\xi_j = \sum_{l \neq j} \mu_l$ is a measure of worldwide competition that a firm from country $j$ faces while selling to country $i$ ($\xi_{ij}$ is the sum of the inverses of the standard deviations of the cost distributions of all other producers of the same good, and is increasing in the number of countries in the economy).

**B Choice of the Parameterization to Approximate the Open Economy Pricing Rule**

The parameterization (37), chosen to compute numerically the solution of equation (34), is an application of Judd (2002)'s parametric path method. Equation (34) is problematic because of a
bifurcation point at $z = 0$, which makes shooting methods unstable. The behavior of the solution is known at both end points (zero and $+\infty$), and the economics behind it give a reasonable idea of the qualitative behavior of the function we are looking for. The parametric path method seems suitable because it constrains the shape of the function to a set of candidate solutions that respect the economically sensible qualitative behavior of the function, and — in this set — it selects the one that is the closest to the exact solution of the differential equation.

The parametric path method (in its more common formulation) uses a polynomial approximation of the solution at the initial value of the path, and imposes convergence to the steady state. In most ODEs, the convergence rate can be derived analytically by log-linearizing the differential equation around the steady state. In equation (34) this is not possible, so the term $\rho$ in the parameterization cannot be computed separately, but must be included in the non-linear least squares problem that solves for the undetermined coefficients.

To escape from the bifurcation point, I compute algebraically the first and second order polynomial coefficients to match the first and second derivatives of the solution for $z = 0$:

$$
p'(0) = \frac{\eta w}{\eta - 1}
$$

$$
p''(0) = \frac{-2w}{(\eta - 1)^2} \left[ \xi + \frac{\lambda t}{m} \frac{w}{\eta - 1} \right]
$$

which implies that $a_1$ and $a_2$ are:

$$
a_1 = \frac{\eta w}{\eta - 1} - \frac{\rho w}{m w + \xi}
$$

$$
a_2 = \frac{w}{\eta - 1} \left[ \rho - \frac{1}{\eta - 1} \left( \xi + \frac{\lambda t}{m} \frac{w}{\eta - 1} \right) \right] - \frac{\rho^2 w}{2 (m w + \xi)}.
$$

I impose the analytical solution of $a_1$ and $a_2$ into (37) and compute numerically the remaining polynomial coefficients ($a_3,...a_M$) and the convergence parameter $\rho$ by solving the nonlinear least squares problem constructed by plugging the parameterization (37) into equation (34). For the computation, I chose a 10-th order polynomial, and I constructed the sum of the squared residuals with Gauss-Legendre integration over ordinary polynomials, with 20 nodes.

To evaluate the goodness of the approximation, Figure 14 shows the computed residuals (i.e., the difference between the two sides of (37)).
In this section, I explicitly derive the formulas for the labor force segments in each sector. Let $l_{ji}^I$ denote the labor force of country $i$ working in integrated segments of firms from country $j$. Then:

$$l_{ji}^I = \begin{cases} 
0 & \text{if } m_j \neq \tau_j w_i \\
 k_{ji}^I q_j & \text{if } m_j = \tau_j w_i 
\end{cases} \quad (A.7)$$

where:

$$k_{ji}^I = \left( \frac{m_j}{p_j} \right)^{-\eta} \int_{B_j^I} x_j^{-\eta} \phi_j(x_j) \psi(z) \, dx_j \, dz. \quad (A.8)$$

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

$$l_{ji}^T = k_{ji}^T q_j \quad (A.9)$$

where:

$$k_{ji}^T = \left( \frac{t_{ji}}{p_j} \right)^{-\eta} \int_{B_j^T} z_i p_{ji}(z_i)^{-\eta} \phi_j(x_j) \psi(z) \, dx_j \, dz. \quad (A.10)$$
References


