# Do Contractual Frictions Shape Global Sourcing? Evidence from Spanish Firm-Level Data

by

Wilhelm Kohler
*Tübingen University*
and
Marcel Smolka
*Tübingen University*

This version as of October 26th, 2009

**Abstract**

We investigate Spanish manufacturing firms' global sourcing decisions with regard to their ownership structure and their locational choice of intermediate input production by exploring a firm-level panel data set covering the years 2000-2007. We distinguish four sourcing categories (foreign integration, foreign outsourcing, domestic integration, domestic outsourcing) thus providing a much more comprehensive picture about the global production organization than was the case with earlier studies. Our findings are: first, there is a crucial interplay between a firm's capital intensity and total factor productivity which is shaping the probability of integrating intermediate input production into the firm boundaries of control, both at home and abroad. In particular, a productivity-increase yields significantly higher probabilities of sourcing inputs from a vertically integrated supplier if and only if the firm's capital intensity lies above a well-identified cut-off intensity level. Second, as regards the locational choice of production organization, offshoring firms are on average more capital-intensive than purely domestic firms, which suggests that there remain unexplored channels through which the factor content of production affects the firms' global sourcing decisions. Third, there is a productivity-dependent sorting pattern of firms into the four sourcing categories according to which foreign-integration firms feature the highest and domestic-outsourcing firms the lowest total factor productivity. There is also evidence that foreign-outsourcing firms are slightly more productive than domestic-integration firms.

*Keywords:* intra-firm trade, foreign direct investment, incomplete contracts, outsourcing, productivity, factor intensity, firm-level data.

*JEL classification:* F14, F23, L22, L23

---

We thank seminar participants at Tübingen University, the 5th Danish International Economics Workshop, Aarhus School of Business, the International Workshop on Firm and Product Heterogeneity, H.U. Brussels, and the ETSG 2009, University of Rome, for useful comments. We are especially grateful to Dieter Urban for helpful suggestions.
**Correspondence:** Chair in International Economics, University of Tübingen, Nauklerstraße 47, 72074 Tübingen, Germany. E-mail: wilhelm.kohler@uni-tuebingen.de; marcel.smolka@uni-tuebingen.de

# 1 Introduction

During the past decade, trade theory has turned *micro*: The individual firm has entered the stage of analysis. Of course, this is of interest only to the extent that firms are not entirely symmetric. We have ample evidence indicating that firms exhibit substantial heterogeneity in the degree, as well as the organizational form of internationalization. Importantly, this heterogeneity is not random. There are characteristic patterns; see Bernard et al. (2007). Meanwhile, we also have a wealth of theories that help us understand why such patterns might come about; see Helpman (2006) for a survey. However, we still have relatively few empirical studies that directly test specific predictions about firms' strategies of internationalization that are derived from such micro-models of trade. Very often, the models are not sufficiently geared towards such predictions, and to the extent that they are the data required are often not available. In this paper we test the key predictions derived from a theoretical model of *global sourcing* that was developed by Antràs & Helpman (2004), based on the pioneering contribution of Antràs (2003). We do so using a survey-based micro data set for Spanish firms that allows us to test the model predictions in a direct and comprehensive way, improving in several ways upon existing literature.

The issue of global sourcing strategies has caught particular attention in the economic globalization debate. Over the past two decades, improvements in transport and communication have fostered *fragmentation* of production processes. Increasingly, firms consider sourcing parts of their value added chain from independent contractors, and in many cases also from foreign countries that offer a cost advantage, say in terms of lower wages. Such *offshoring* of production steps has repeatedly been at the forefront of the economic policy debate in the US and Western European Economies, on the grounds that it might increase inequality in earnings and employment perspectives; see Krugman (2008). Sometimes this debate has struck a protectionist tone, although up to this point protectionist policy measures have not been targeted at offshoring. From a theoretical perspective, it is quite obvious that offshoring might change the factor content of a country's trade and, thus, affect its domestic factor prices and unemployment. But the empirical significance of this effect has turned out difficult to pin down; see OECD (2007).

Identifying the effects of offshoring requires knowledge of its determinants. In this paper, we therefore address the determinants of firms' sourcing behavior. More specifically, we want to test the empirical significance of *"hold-up" problems* for the choice of sourcing strategies, whereby such strategies have two dimensions: *location* and *organizational form*. The literature originating in Antràs (2003) and Antràs & Helpman (2004) emphasizes that in manufacturing production the provision of essential inputs is often plagued by imperfect, or complete lack of, contractibility. Firms and input suppliers may not be able to write ex ante enforceable contracts, or enforceable

contingent contracts, because the nature of the input negates third party verification of the ex post quality and/or quantity of the inputs provided by the two parties. Moreover, the inputs may be relationship-specific, so that once produced they have zero value outside the production relationship considered. As a result, the firm and the supplier are pitted against each other in negotiations about sharing the surplus of the relationship, and provision of both inputs is plagued by a "hold-up" problem. Production is distorted by inefficient input levels.

Drawing on Grossman & Hart (1986), Antràs (2003) argues that firms may respond to this problem by acquiring property rights in the input provided by the supplier, if this is possible in the given institutional environment. Under certain assumptions about cost and bargaining, this generates a prediction on whether a firm organizes input provision through *integrated production*, i.e., acquiring property rights, or through *outsourcing*, whereby the firm relies on a contract with an independent supplier. The key determinant according to Antràs (2003) is the elasticity of the output level with respect to the input provided by the supplier, as opposed to the input provided by firm, which is called a "headquarter service" in Antràs & Helpman (2004). The larger this elasticity, the more important is the supplier-input, the less attractive therefore the acquisition of property rights, as this reduces the supplier's incentive. Existing empirical literature has mostly measured this elasticity through a factor-intensity variable, thus assuming that the headquarter-input is unambiguously more capital- or skill-intensive than the supplier-input. Our empirical results indicate that coefficients on factor-intensity variables reflect confounding factors: The headquarter effect (Antràs effect) and a comparative advantage effect which derives from factor price differences across countries.

However, as indicated above, the sourcing strategy involves a second dimension which relates to the location, as opposed to the organizational mode, of input provision. In Antràs & Helpman (2004), the headquarter service must by definition be provided domestically. Thus, if integrated production involves foreign sourcing of the supplier-input, then what we observe is often called *intra-firm trade*, or cross-border vertical integration.[1] But foreign sourcing may also take place through independent suppliers, in which case we observe trade between *unrelated parties*. Existing empirical work mostly looks only at the organizational choice of offshoring, thus ignoring transactions that entail domestic sourcing. We argue that any full-fledged test of the model predictions requires that we look at both dimensions, the organizational mode and the location of input provision. Our data allow us to do so.

[1]It will become evident later in this paper that the term "intra-firm" trade is somewhat misleading in this context. In particular, it should not be interpreted as cross-border transactions under complete firm *hierarchy* (command). Instead, it is trade between *related parties* in the sense that the firm has acquired property rights in the traded inputs, which affects its bargaining power. But it is still *bargaining*, not *command*, that governs this trade.

Antràs & Helpman (2004) extend the model to include *firm heterogeneity* in terms of productivity along the lines pioneered by Melitz (2003). This generates a richer set of predictions that also address the question why, even within a given industry, some firms choose vertical integration of input provision, while others rely on independent suppliers (outsourcing), and why some decide to obtain the "supplier-input" domestically, while others obtain it from abroad (offshoring). These predictions include a direct effect of *productivity* on the firm's sourcing choice, which is relatively straightforward. More interestingly, and less straightforwardly, the model predicts that firm productivity also affects the relationship between the aforementioned supplier-input-elasticity and the strategy choice. In other words, there are interaction effects between firm productivity and the supplier-input-elasticity (or its counterpart, the headquarter elasticity) in the determination of a firm's sourcing strategy. Our data allow us to also test for such interaction effects.

Antràs & Helpman (2008) propose a further generalization of the model that allows them to consider variations in the *degree of contractibility*. Such variations may occur across industries and countries, as well as through time. It turns out that an increase in the fraction of supplier-inputs (headquarter inputs) that are contractible enhances the case for vertical integration (outsourcing). This seems counterintuitive at first sight, but is relatively easy to understand, once the above mentioned strategic implication of the supplier-input-elasticity is realized. In fundamental terms, what matters is the output elasticity with respect to non-contractible inputs. Therefore, a firm should be more willing to accept a lower incentive for the supplier that comes along with integration, if the share of noncontractible supplier-inputs (and thus the output elasticity with respect to these inputs) is small. In this paper, we summarize the results from some approaches to test predictions derived from Antràs & Helpman (2008), identifying weaknesses of existing literature, but for reasons of space we leave detailed empirical analysis of a varying degree of contractibility to future work.

We thus have a rich set of theoretical predictions on whether or not firms rely on trade in obtaining their inputs, as well as on the organizational form that they choose for input provision. Why are these interesting predictions? Why would we want to know about their empirical validity? The common thread running through all these predictions is that input provision is plagued by a "hold-up" problem. Whether or not such problems play a significant role in input trade is practically important, for two reasons. The first has to do with the *factor content* of a country's trade, which is an important determinant for a country's wages or, more generally, factor prices; see Deardorff & Staiger (1988) and Deardorff (2000). Traditional trade theory holds that the factor content of a country's foreign trade is determined by its *factor endowment*, relative to its trading partners. The recent tendency of increased fragmentation and offshoring of value-added-stages obviously affects the factor content of trade. From conventional theory, we would expect that this is driven by factor price differences; see Jones (2000). Thus, in a

capital abundant economy it would be the more labor intensive slices of value-added within each industry that are likely to fall victim to global sourcing strategies. However, models in the spirit of Antràs & Helpman (2004) now introduce a new distinction between headquarter-inputs and supplier-input.[2] What offshoring does to the factor content of trade then crucially depends on whether the supplier-input happens to involve a relatively labor- or capital intensive activity.[3] Moreover, the models suggest that much also depends on whether foreign provision of such inputs is plagued by a more serious "hold-up" problem than domestic provision. Thus, a country's *property rights institutions* play an important role, in addition to its factor endowment, in shaping the factor content of its trade in the ongoing process of fragmentation and offshoring. In this paper, we do not test any prediction about the relationship between the factor-intensity of an input and the severity of the "hold-up" problem involved. That must be left for future work. But investigating the general significance of such "hold-up" problems for firms' sourcing decisions is an important first step.

There is a second reason why it seems important to establish the empirical significance of "hold-up" problems in input trade which has to do with trade policy. If "hold-up" problems that are differently severe for traded and non-traded inputs, this generates a genuine trade distortion that may be addressed by means of trade policy. For instance Ornelas & Turner (2008) look at a case where a "hold-up" problem is present in domestic input provision, but not in foreign input provision. As a result, input provision is distorted, with too much offshoring, which establishes a case for a tariff on the imported input. Conversely, in Antràs & Staiger (2008) the "hold-up" problem is present only in foreign input provision, in which case the optimal response is an import subsidy. What a country's trade policy, if pursued along these lines, does to the factor content of its trade thus depends, not only on the relative severity of "hold-up" problems in domestic and foreign input provision, but also on the relative factor intensities of these inputs. This is an area of trade policy research that is only just developing, and we do not take up trade policy issues in this paper. But it is obvious that such policy questions further enhance the relevance of the empirical work that we present in this paper.

The paper is structured as follows. In section 2, we present a stripped down version of a model in the spirit of Antràs & Helpman (2004), in order to specify the exact predictions that we want to test. Section 3 introduces related empirical work, focusing on similarities and differences in terms of data structure and estimation methodologies. Subsequently, in section 4, we describe our data set and provide some descriptive statistics that give a flavor of the evidence that we use in order to test our predictions. In section 5, we present econometric results from

[2]The reader will notice that this is reminiscent of the first general equilibrium models of the multinational firm proposed by Helpman (1984); see also Markusen (1984).

[3]Notice again that in Antràs & Helpman (2004), by definition, the headquarter service is a domestic input.

two alternative approaches. Our first choice is Tobit estimation of equations, derived from our theoretical model, that explain the extent to which firms rely on certain sourcing strategies. As a sort of consistency check, we also run regressions where dummy variables for sourcing strategies explain the productivity and size of our firms. Section 6 concludes and indicates lines for future research.

# 2 A theory of sourcing without contractibility

Stripped down to its bare bones, what are the essentials of the "hold-up model" of global sourcing? A key tenet of all models in the spirit of Antràs (2003) is that production requires essential inputs that the firm cannot generate within its firm (or command) hierarchy, but must be obtained from a supplier. The line separating "firm-inputs" and "supplier-inputs" is exogenous to the model. What is endogenous is the organizational mode and location of the "supplier-input". Alternative modes of organization differ in their implications for the incentive structure for the firm and the input supplier. By choosing a specific organizational mode the firm may thus alleviate the problem of distorted input levels. The choice is modeled in the following theoretical framework.

## 2.1 The framework

The model assumes that there are two types of agents, firm and the input suppliers. Firms have access to some market where they raise revenue from selling final goods which are produced from two essential inputs that are provided, respectively, by the firms themselves (the so-called "headquarter service") and by the input suppliers. Input-suppliers may either be located in the domestic economy or in the foreign economy, but firms must locate in the domestic economy. Firms are unable to produce non-headquarter inputs and must, therefore, always resort to input suppliers in order to generate revenue from the final output. In turn, input provision requires primary factors which must be purchased in the factor market of the country where a provider is located. For simplicity, one may assume that labor (domestic or foreign) is the only primary input involved, with a unit of labor required per unit of the input provided.

Firms produce final goods from the two types of inputs according to some linearly homogeneous production function, whereby firms engage in product differentiation which affords market power along the lines of Dixit & Stiglitz (1977). Importantly, each firm produces its own brand, and inputs must be provided in a unique specification for each brand. We assume extreme specificity, so that any input, once provided, has economic value only from production and sale of the corresponding final good. This is often called relationship-specificity. Moreover, we assume that the quality and/or quantity of input provision cannot be verified by a third party, so that firms

and input suppliers cannot write enforceable contracts for input provision. Therefore, firms and input suppliers must ensure corresponding specifications, and when deciding about the amounts of labor to invest into generating their respective inputs, they expect to be pitted against each other in a negotiation about how to share the revenue from selling the output generated by their specific inputs. Negotiation is assumed to take the form of Nash-bargaining, with a bargaining power equal to $\beta$ for the firm, and $1-\beta$ for the input supplier. In addition to the variable cost of input provision, raising revenue from the production relationship requires an organizational effort which causes a fixed cost. As we shall detail below, the magnitude of this cost depends on the organizational form and location of input provision.

For simplicity, we assume that in both countries input suppliers are in infinitely elastic supply, with zero outside options. Thus, any input supplier is willing to engage in a production relationship with a firm, if and only if she may expect to obtain a share of the revenue that is at least equal to the cost of input provision. We assume perfect competition among input suppliers which implies that in equilibrium the profit that an input supplier may expect from participating in production must be zero. Hence, any firm may attract a suitable input supplier by charging a participation fee which is equal to the surplus of the supplier's expected revenue-share over the her wage cost for input provision. This implies zero expected profits for the input supplier. The wage to be paid by the input supplier will be the domestic or the foreign wage rate, depending on whether a firm chooses a domestic or a foreign input supplier. We assume a foreign locational advantage for the supplier-input, which means that the foreign wage rate is lower than the domestic wage rate.[4] In the domestic economy, there is an infinitely elastic supply of potential firms that may enter the market, with an uncertain total productivity of the relationship between the two inputs. Firm entry evolves as in Melitz (2003), meaning that entry is governed by a sunk entry cost and zero expected profits from a potential relationship with an input supplier.

As regards sourcing strategies, the key difference between firms and input suppliers is that firms may "purchase", ex-ante, an outside option for the ex-post Nash-bargaining. This takes the form of choosing an organizational form that involves *property rights* in the supplier-input. Following Antràs & Helpman (2004), we call this *vertical integration*, but it is important to bear in mind that it does not mean the firm has command over the entire input provision. It simply means that, if bargaining breaks down, the firm may invoke a property right on the supplier-input and still sell the output. However, this involves two types of cost. First, there

[4]Considering the empirical test of the model predictions on Spanish data, one may question this assumption. However, in the empirical estimation the model should be interpreted in more general terms, with many factors. Then, the foreign locational advantage for input-provision derives from relative factor price discrepancies, which may involve a relatively *lower* domestic wage rate. The model predictions tested below are robust to this modification.

is an iceberg-type cost of invoking property rights in that the revenue accruing to the firm is only a fraction $\delta$ of market revenue.[5] And secondly, the organizational fixed cost to the firm is higher with vertical integration than without. We call an organizational form of input provision that does not involve any property rights *outsourcing*. Moreover, we assume that the organizational cost is also larger if the firm relies on a foreign input supplier. Antràs & Helpman (2004) make further assumptions leading to following ranking of the fixed organizational cost: $F^{FI} > F^{FO} > F^{DI} > F^{DO}$, whereby $I$ and $O$ indicate vertical integration and outsourcing, respectively, while $F$ and $D$ indicate a foreign or domestic input supplier. Organizational cost involve the use of domestic labor and are expressed in terms of domestic labor. One may also assume that $\delta$ is larger for an input provided domestically: $\delta^D > \delta^F$. Notice that this share is always zero, if the firm chooses outsourcing. It is obvious that in a multi-country framework $\delta^F$ may vary across trading partners.

With this legal environment of property rights, the firm may influence the expected pay-off from ex-post Nash-bargaining through an ex-ante choice of its organizational form of sourcing. More specifically, with vertical integration the share of revenue that accrues to the firm is equal to $\beta^{\ell I} := \delta^\ell + \beta\left(1-\delta^\ell\right)$, and with outsourcing it is equal to $\beta^{FO} = \beta^{DO} = \beta < \beta^{\ell\kappa}$. Notice that $1-\delta^\ell$ times the revenue to be generated from the market represents the "gains from trade", trade meaning completion of the production relationship through successful Nash-bargaining.

## 2.2 Decision making

Decision making within this framework takes place in a multi-stage manner. In a first step, firms enter the market in order to find out about their productivity level $\theta$ (TFP). In stage 2, conditional upon $\theta$ being large enough to avoid losses from the subsequent stages,[6] each firm decides about its optimal sourcing strategy, which means the simultaneous choice of a location $\ell = F$ or $\ell = D$ and an organizational form $\kappa = I$ or $\kappa = O$. In the third stage, the two input providers (i.e., the firm and the supplier) choose levels of their inputs that maximize the difference between their expected shares of revenue, to be generated from the relationship, and their respective wage cost. Expected shares depend on the anticipated Nash-bargaining power and the outside options of the two parties. Importantly, revenue depends on the levels of *each* input chosen. In stage 4, output is produced, and the revenue is shared according to the Nash-bargaining solution. When contemplating entry, a potential firm anticipates its optimal choice of an organizational form in stage 2, which in turn will anticipate the optimal input levels

[5] Antràs & Helpman (2004) assume that the iceberg-cost parameter applies to the quantity of the final good produced. But this translates into an iceberg-cost parameter for revenue in a straightforward way, depending on the details of the revenue function that are not important for our purpose.

[6] Details of this first stage are as in Melitz (2003), and will not be looked at in detail here.

emerging in stage 3. Thus, the sourcing strategy chosen is an integral part of a subgame-perfect equilibrium. In the following, we focus on stages 2 through 4 of decision making, employing the usual backward induction.[7]

**Revenue (stage 4):** Under the above assumptions, revenue $R$ is concave in $x$, the quantity of the final good produced:

$$R(x) = A^{1-\alpha}x^{\alpha} \quad \text{with} \quad x = \theta y(h, m) \tag{1}$$

The first expression follows from Dixit-Stiglitz-type product differentiation, with an elasticity of substitution equal to $1/(1-\alpha)$, whereby $0 < \alpha < 1$. In the second expression, $y(h, m)$ is a linearly homogeneous production function for the final good, with input levels $h$ and $m$ for the headquarter service and the supplier-input, respectively, and $\theta$ as the firm-specific TFP-level. In the sequel, we denote the elasticity of $x$ with respect to $h$ by $\eta$. Given linear homogeneity, the elasticity with respect to $m$ is equal to $1-\eta$.

**Input levels (stage 3):** Input-providers simultaneously choose input levels $h$ and $m$ that maximize $\gamma R(x) - w^D h$ and $(1-\gamma)\,R(x) - w^\ell m$, respectively, where $l$ denotes the location of the input supplier and $\gamma$ denotes the revenue share accruing to the firm in ex-post Nash-bargaining. As explained above, this depends on organizational choice $\kappa \in \{I, O\}$. However, for the time being we ignore this dependency, in order to better grasp the strategic role of this share. It can be shown that the optimal input levels (indicated by an asterisk) satisfy

$$h^* = \gamma\alpha R^*\eta / w^D \quad \text{and} \quad m^* = (1-\gamma)\,\alpha R^*(1-\eta) / w^\ell \tag{2}$$

where $R^*$ denotes the relationship-specific total revenue deriving from input levels $h^*$ and $m^*$, determined simultaneously in this stage of decision making. The lack of contractibility introduces a distortion, relative to an environment with ex-ante enforceable contracts. Both inputs are provided in less than optimal amounts, whereby the share parameter $\gamma$ determines the distortion (underprovision) of the headquarter input, *relative* to the supplier-input.

**Organizational form (stage 2):** In addition to the input levels $h^*$ and $m^*$, the firm's entire profit from the relationship depends on the participation fee that it may charge the input supplier, as well as on the fixed organizational cost it has to bear. Without going into further

[7] For the first stage of firm entry and a "general equilibrium" embedment of this sector, with a second numéraire sector, see Antràs & Helpman (2004).

details, we note that the participation fee must be equal to $(1-\gamma)\,R^* - w^\ell m^*$. It may be zero, positive, or negative. Hence, equilibrium firm profits are equal to

$$\begin{aligned}\Pi^* &= R^* - w^D h^* - w^\ell m^* - w^D F^{\ell\kappa} \\ &= R^*\left[1-\gamma\alpha\eta-(1-\gamma)\,\alpha\,(1-\eta)\right] - w^D F^{\ell\kappa} \qquad (3)\end{aligned}$$

where $F^{\ell\kappa}$ denotes the corresponding organizational fixed cost in units of domestic labor. The second line uses the above first-order conditions on the two input levels. It is now crucial to recognize that $R^*$ depends on the organizational choice through the above mentioned property rights that may be acquired through vertical integration. Using $c_y$ to denote the unit-cost function dual to $y(h,m)$, we may write

$$R^* = A\left[\frac{1}{\alpha\theta}c_y\left(\frac{w^D}{\gamma},\frac{w^\ell}{1-\gamma}\right)\right]^{-\alpha/(1-\alpha)} \qquad (4)$$

Notice that the elasticity of $c_y$ with respect to $w^D/\gamma$ is equal to $\eta$, and analogously for the other input price. Equilibrium firm profits then emerge as

$$\Pi^* = Z^{\ell\kappa}\theta^{\alpha/(1-\alpha)} - w^D F^{\ell\kappa} \qquad (5)$$

$$\text{where} \quad Z^{\ell\kappa} := A\frac{1-\alpha\gamma\eta-\alpha\,(1-\gamma)\,(1-\eta)}{\left[c_y\left(\frac{w^D}{\gamma},\frac{w^\ell}{1-\gamma}\right)\Big/\,\alpha\right]^{\alpha/(1-\alpha)}} \qquad (6)$$

In stage 2 of the game, the firm now chooses a sourcing strategy $\ell \in \{D,F\}$ and $k \in \{I,O\}$ that maximizes profits $\Pi^*$.

### 2.3 Sourcing hypotheses

Notice that equation (5) depicts a firm's profits from a production relationship as a linear function of $\Theta := \theta^{\alpha/(1-\alpha)}$. Remember that $0<\alpha<1$, hence $\Theta$ is unambiguously increasing in the productivity level $\theta$. The negative intercept in (5) is governed by the fixed organizational cost $F^{\ell\kappa}$ and the domestic wage rate. In turn, equation (6) reveals the slope of this line as a relatively complex function of the expected revenue share parameter $\gamma$ and the two countries' wage rates. Antràs & Helpman (2004) suggest that the organizational dimension of the sourcing strategy is best understood by first considering the effect of an unrestricted variation of the share parameter $\gamma$ on a firm's profit from the relationship with an input supplier. Of course, expected profits would be zero, if the firm's bargaining power is nil, $\gamma = 0$. Less obviously, profits would also be zero for the other extreme. Formally, letting $\gamma \to 1$ implies that $c_y(\cdot) \to \infty$, and the slope $Z^{\ell\kappa}$ becomes flat. Given the organizational cost, no firm, however productive, would expect a positive profit from such a production relationship. Intuitively, it would simply lack an essential

input, since the input supplier lacks any incentive to invest in such a relationship.

All of this suggests that there should be a level of $\gamma$ which maximizes the slope of our profit-line. Intuitively, the incentive problem of a high $\gamma$ is the more severe, the more important the supplier input for the relationship. This is measured through the elasticity $1-\eta$. Put differently, for a given sourcing location, the optimal firm-share $\gamma$ should be increasing in $\eta$, the elasticity of the output $x$ with respect to the headquarter service. Antràs & Helpman (2004) derive such a monotonic relationship for the optimal level of $\gamma$ as a function of $\eta$, whereby both $\gamma$ and $\eta$ lie between zero and 1.[8] Let us denote this relationship by $\gamma(\eta)$, with $\gamma' > 0$. It is interesting to note that the optimal value of $\gamma$ is independent of the two countries' wage rates, although the slope $Z^{\ell\kappa}$ of course does depend on the cost advantage. Both, $\gamma(\eta)$ and $Z^{\ell\kappa}$ are also independent of firm productivity $\theta$.

The key point to recognize at this stage is that the firm's revenue share is not at the firms unrestricted disposal. Hence, it cannot choose $\gamma$ at will. Available organizational options imply that with vertical integration the firm-share is equal to $\beta^{\ell I} := \delta^{\ell} + \beta\left(1-\delta^{\ell}\right)$, depending on the sourcing location, or else equal to $\beta < \beta^{\ell I}$, independent of the sourcing location. Now, for each sourcing location $\ell \in \{D, F\}$, there must be a critical value of $\tilde{\eta}$, implicitly defined through $\gamma(\tilde{\eta}) = \beta < \beta^{\ell I}$. Then, for any $\eta < \tilde{\eta}$ a firm will never choose vertical integration, whatever its productivity, and independently of the fixed organizational cost. The supplier-input is simply too important in the production relationship. Ideally, the firm would even want to incentivize the input supplier through a revenue share larger than $1-\beta$, but by assumption $\beta$ is as low as the firm's share can get.[9] Intuitively, for a value of $\eta$ which is just *marginally* above $\tilde{\eta}$, say $\tilde{\eta}+\varepsilon$, outsourcing continues to dominate vertical integration, because integration implies a *discretionary* jump from $\gamma(\tilde{\eta}+\varepsilon)$ right up to $\beta^{\ell I}$, where profits will be lower than under outsourcing. Notice that this holds true even absent all fixed organizational cost. Indeed, if we rule out such fixed cost, we realize by the same logic that the firm will choose vertical integration once $\eta$ moves sufficiently close to a value $\bar{\eta}^{\ell}$, which is implicitly determined through $\gamma(\bar{\eta}^{\ell}) = \delta^{\ell} + \beta\left(1-\delta^{\ell}\right)$. Thus, there must be a cut-off value $\eta^*$ which separates firms with different organizational choices. This leads to a first model prediction:

**Hypothesis 1 (Antràs effect without fixed cost heterogeneity)** *Suppose there are no fixed organizational cost. Then, there is a critical level $\eta^*$ for the elasticity of output $x$ with respect to the firm's headquarter investment level $h$, that needs to be surpassed for vertical integration to become an attractive organizational choice for all firms.*

[8] In their notation, the revenue-share is $\beta$, and the headquarter-service-elasticity is $\eta$.

[9] This may seem like a paradox. Why, after all, should a firm not be able to *reduce* its bargaining power? The underlying assumption here is that for some reason the firm cannot credibly commit to a lower bargaining power than $\beta$.

It might be tempting to expect vertical integration to prevail for all elasticity values $\eta > \bar{\eta}^{\ell}$. However, this is not true. While it is true that in such industries the profit line has a larger slope for vertical integration than for outsourcing, i.e., $Z^{\ell I} > Z^{\ell O}$, this does not unambiguously dictate a vertical integration strategy. The reason is that integration involves a higher fixed cost of organization. Therefore, it will only be firms with a sufficiently high productivity (leverage) beyond some threshold level $\Theta^{\ell O}$ that find this strategy worthwhile. This holds true, separately, for each of the two sourcing locations $\ell \in \{D, F\}$. We shall return to the locational question separately below.

For intermediate elasticity values that satisfy $\tilde{\eta} < \eta > \bar{\eta}^{\ell}$ the comparison of $Z^{\ell I}$ and $Z^{\ell O}$ is ambiguous. If $Z^{\ell I} > Z^{\ell O}$, which holds for $\eta$-values in an upper range of this interval, by the above logic and the assumption of $F^{\ell I} > F^{\ell O}$, it will again be the more productive firms that choose vertical integration over outsourcing. For $Z^{\ell I} < Z^{\ell O}$, in the lower range of the elasticity interval, the fixed cost ranking implies that outsourcing dominates integration, no matter how productive a firm. We may define an elasticity value $\eta^{\ell *}$ that is implicitly determined by $Z^{\ell I} = Z^{\ell O}$, and summarize in terms of the following hypothesis.

**Hypothesis 2 (Antràs effect with fixed cost heterogeneity)** *Suppose firms are restricted to a sourcing location $\ell$ and the fixed cost of organization are such that $F^{\ell I} > F^{\ell O}$. Then, provided that elasticity value $\eta$ exceeds a threshold level $\eta^{\ell *}$, there is a critical level of firm productivity $\Theta^{\ell O}$, such that firms with a productivity $\Theta > \Theta^{\ell O}$ choose a vertical integration strategy with location $\ell$ and conversely for $\Theta < \Theta^{\ell O}$. For elasticity values $\eta < \eta^{\ell *}$, all firms will choose an outsourcing strategy, no matter what their productivity. The threshold level $\eta^{\ell *}$ is larger than $\tilde{\eta}$ of hypothesis 1.*

In a nutshell, for any given sourcing location $\ell$, vertical integration is chosen by a larger set of firms, the larger the elasticity $\eta$, and for any given $\eta > \eta^{\ell *}$, it will be the more productive firms that choose vertical integration instead of outsourcing (see figure 1).

Hypothesis 2 holds separately for each possible location of input provision. As regards the sourcing location, given the ranking of fixed organizational cost, $F^{FI} > F^{FO} > F^{DI} > F^{DO}$, foreign sourcing will never be an attractive to any firm, of whatever productivity, if $c_y\left(\frac{w^D}{\gamma}, \frac{w^F}{1-\gamma}\right) > c_y\left(\frac{w^D}{\gamma}, \frac{w^D}{1-\gamma}\right)$. In this case, offshoring of any organizational mode implies a profit line with a lower intercept and a lower slope than would domestic sourcing and will, thus, never be chosen. In order to have offshoring in our picture at all, we therefore assume that the opposite inequality holds, i.e., the foreign economy offers a locational (wage cost) advantage for provision of the supplier-input. In this case $Z^{F\kappa} > Z^{D\kappa}$. Intuitively, the firm's productivity level $\Theta$ is a leverage for the cost-advantage of foreign sourcing. Indeed, the cost-advantage may be large enough and the fixed cost disadvantage of foreign sourcing small enough, so that

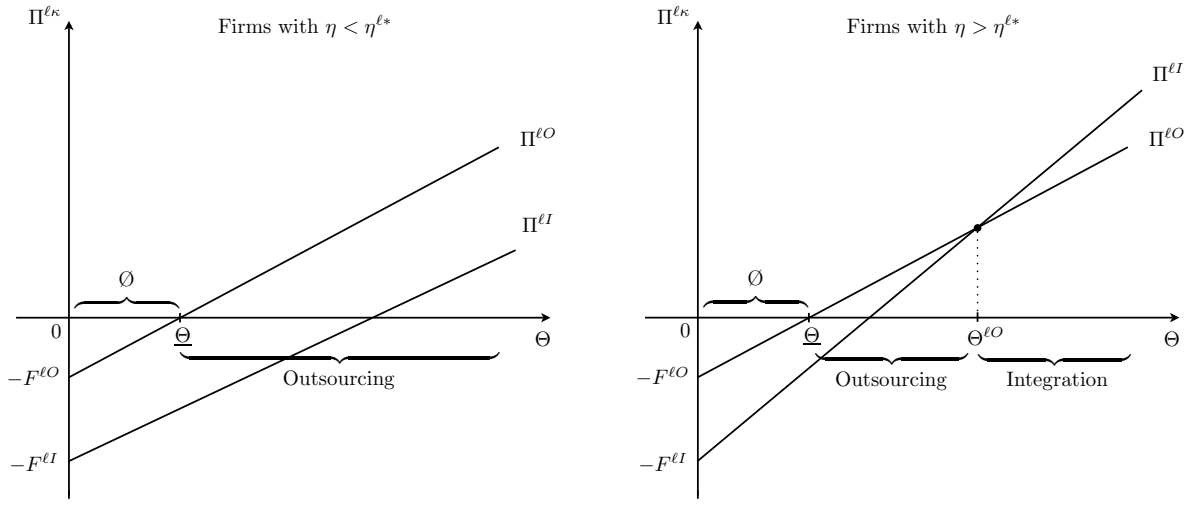


**Figure 1:** Antràs Effect with Fixed Cost Heterogeneity

the cut-off level of firm productivity $\Theta$ which yields zero-profits is lower for offshoring than for domestic sourcing. This is true if $F^{F\kappa}/Z^{F\kappa} < F^{D\kappa}/Z^{D\kappa}$. Antràs & Helpman (2004) rule this out by assumption, in order to have a richer set of equilibrium firm strategies. Although perhaps plausible, such an assumption is not crucial for the purpose of our empirical analysis, as we shall see below.

In addition to firm productivity, the locational dimension of sourcing is also determined by the headquarter-elasticity $\eta$. To see why, look at the slope term $Z^{\ell\kappa}$ in (6). Given the aforementioned assumption securing a foreign locational advantage, we always have $Z^{F\kappa} > Z^{D\kappa}$. But the slopes are determined by the value of $\eta$. As pointed out above, the elasticity of $c_y(\cdot)$ with respect to the supplier input is equal to $1-\eta$. The extent to which a given wage discrepancy $w^F < w^D$ is translated into a cost advantage of the foreign sourcing strategy, $c_y^D > c_y^F$, is increasing in $1-\eta$. Hence, a higher elasticity value $\eta$ implies a lower cost advantage. From (6) we now see that it also reduces the difference between $Z^{F\kappa}$ and $Z^{D\kappa}$. Other things equal, this increases the threshold level of firm productivity that separates firms with a domestic sourcing strategy from those engaged in offshoring, $\Theta^{D\kappa}, \kappa \in \{I, O\}$. We may then summarize the locational dimension of the sourcing strategy by the following hypothesis (see figure 2).

**Hypothesis 3 (Offshore intensity effect)** *Suppose that the fixed organizational cost satisfies the ranking $F^{F\kappa} > F^{D\kappa}$, and the domestic and foreign wage rates are such that $F^{F\kappa}/Z^{F\kappa} > F^{D\kappa}/Z^{D\kappa}$. then, for any mode of organization $\kappa \in \{I, O\}$, it will be firms with a productivity level below some critical level $\Theta^{D\kappa}$ that choose domestic input provision, while firms with a productivity exceeding this level choose offshore input provision. Moreover, for a given organizational form $\kappa$ the threshold level of productivity $\Theta^{D\kappa}$ unambiguously increases in the elasticity $\eta$.*

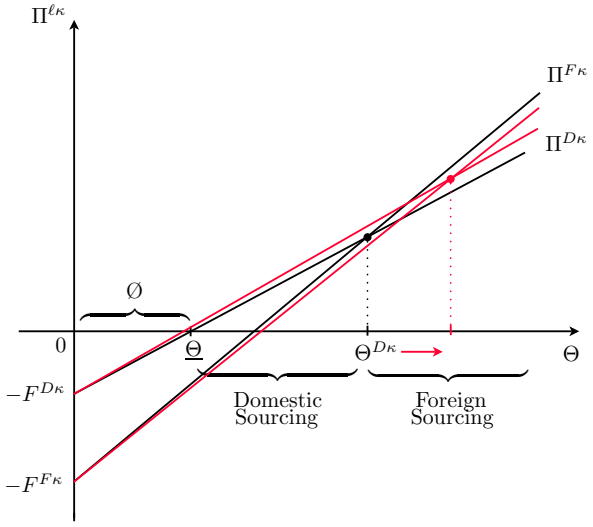


**Figure 2:** Offshore Intensity Effect

# 3 Related empirical research

Antràs (2003) and Antràs & Helpman (2004, 2008) try to explain a previously less explored part of international economic activity. Their theoretical propositions ask for empirical verification. Antràs (2003) himself rests part of his theory on the observation that the share of intra-firm imports in total U.S. imports in an industry is significantly higher, the higher the capital intensity of this industry is. As is true for nearly all related empirical contributions in the past, he limits his analysis to the case of multinational enterprises (MNEs), leaving the domestic sourcing possibility aside. Specifically, he assesses a modified version of hypothesis 1 with capital intensity as a proxy for headquarter intensity, using four years of data (1987, 1989, 1992, and 1994) for 23 industries from the Bureau of Economic Analysis (BEA) website on intra-firm and total U.S. imports. Antràs (2003) finds strong support for the hypothesis that the share of intra-firm imports in total imports in each industry is larger, the higher the industry's capital intensity. The main result is that a one percent increase in the industry's capital intensity (R&D intensity) raises the share of intra-firm imports in total imports in this industry by approximately 0.7 to 1 percent (0.4 to 0.5 percent).

Yeaple (2006) uses the same U.S. trade data from the BEA website as Antràs (2003) does for the year 1994 but he restricts his sample to U.S. MNEs with foreign subsidiaries thereby sticking closer to the theoretical set-up. In the framework of a Poisson regression he finds that the theoretical predictions in Antràs (2003) and Antràs & Helpman (2004) are highly relevant, especially for trade relations with countries for which the variable cost differences

(wage differences) compared to the U.S. are largest. In particular, a one percent increase of an industry's capital intensity raises the share of intra-firm imports in total imports in that industry by on average 1.15 (0.79) percentage points if the sample is restricted to the least developed countries (emerging countries) as trading partners, but it has no significant effect if only North-North trade is being considered. Higher productivity dispersions in a given industry seem to follow exactly the same pattern while the logic is reversed for an industry's R&D intensity: in North-North trade, higher R&D intensities contribute significantly more to higher shares of intra-firm imports in total imports than is the case for North-South trade. Yeaple (2006) suggests that "the least developed countries do not produce on average a set of goods that require 'knowledge inputs' from U.S. MNEs but that do require physical capital inputs."

There are several recent empirical studies with a huge body of mainly descriptive statistics analyses by Tomiura (2005, 2007a, 2007b, 2007c) dealing with production strategies of Japanese manufacturing firms. Without going into detail, the overwhelming evidence is that only a very small fraction of Japanese firms engages in offshore outsourcing, and the minority thereof in vertical FDI. All offshoring firms show outstanding characteristics regarding capital intensity, skill intensity, and productivity compared to purely domestic firms. Also, firms with foreign affiliates tend to be more capital-abundant and more productive than firms that source internationally through independent suppliers only. Altogether, the evidence, which is based on an extraordinarily huge number of firms, is to a very large extent in line with the predictions in Antràs (2003) and Antràs & Helpman (2004).

Nunn & Trefler (2008) carry out an important empirical work on the propositions of the whole class of models. On the basis of intra-firm and total trade data at the 6-digit Harmonized System (HS6) level from the U.S. Census Bureau for the years 2000 and 2005, they perform a regression very similar to the one in Antràs (2003) but with a higher level of disaggregation. The findings suggest that both an industry's capital and skill intensity favor vertical FDI relative to offshore outsourcing. In a second step productivity heterogeneity is accounted for by including interaction terms between the capital (and skill) intensity and an industry-specific measure of productivity dispersion. The results in Nunn & Trefler (2008) indicate that the share of intra-firm imports in total imports is higher in industries with higher degrees of productivity dispersion, given that the industry's capital (skill) intensity is sufficiently high. For industries with a comparably low capital (skill) intensity, the effect is much less pronounced. The authors conclude that this is "dramatic confirmation" of the productivity-dependent sorting pattern of firms put forward by Antràs & Helpman (2004, 2008). Third, Nunn & Trefler (2008) consider the possibility of partially incomplete contracting. According to Antràs & Helpman (2008) improvements in the enforceability of contracts that affect the supplier's intermediate input favor vertical FDI relative to offshore outsourcing. Nunn & Trefler (2008) rely on changes in the contractibility of foreign intermediate inputs which provokes — in the terminology of

Nunn & Trefler (2008) — a 'surprise effect' and a 'standard effect' which makes offshoring profit-maximizing for more medium-productivity firms thus increasing the relative number of foreign-outsourcing firms. The data point towards dominance of the 'surprise effect' over the 'standard effect' in headquarter-intensive industries, just as Antràs & Helpman (2008) implies.

Nunn & Trefler (2008) run all regressions with a restricted and an unrestricted sample, restricted meaning that only those countries are kept in the analysis for which at least two-thirds of intra-firm U.S. imports are from a foreign affiliate of a U.S. parent company. The restricted sample excludes important countries such as China, Japan, Germany, United Kingdom, South Korea, Taiwan, Venezuela, and France, and according to the U.S. Census Bureau website all of them are among the top ten U.S. trading partners with regard to imports and account for nearly 40 percent of total U.S. imports in 2005.[10] Considerably, for all those countries except for China the imports by foreign-controlled U.S. affiliates from their foreign parent company account for at least two-thirds (France) of all intra-firm imports from the respective country.[11] In the majority of cases the percentage rate is much higher. For example, more than 84 percent of total U.S. intra-firm imports from Germany can be attributed to a situation in which a German parent company exports to her controlled affiliate in the U.S. These considerations underline that intra-firm imports from foreign parent companies are a major issue here, and it is beyond the scope of the models in Antràs (2003) and Antràs & Helpman (2004, 2008). Interestingly, however, the results in Nunn & Trefler (2008) show now qualitative difference between the estimations with the restricted and the unrestricted sample.

Defever & Toubal (2007) are the first to use very detailed French transaction and firm-level data from the 1999 "International Intra-group exchanges" survey within the framework of a binomial logit regression. They find convincing evidence for the theory-induced productivity-dependent sorting pattern (hypothesis 2) of firms, even though in the opposite direction.[12] Descriptive and inductive data exploration suggests that, in the case of France, fixed costs under integration are lower than under outsourcing, so that the model actually predicts the reversal in the sourcing strategies. The data allow Defever & Toubal (2007) to depart from earlier empirical works in that they may capture firm-specific component intensity $(1 - \eta)$ by the share of inputs from suppliers in a firm's output instead of capital or skill intensity. Even after controlling for industry- and product-specific fixed effects as well as other determinants of the binary decision of whether to engage in vertical FDI or offshore outsourcing, a higher component intensity (that is, a lower headquarter intensity) in the production process of French

[10] See `http://www.census.gov/foreign-trade/statistics/highlights/top/top0512.html`.

[11] The percentages are from Zeile (2003) and belong to the year 1997.

[12] Interestingly, however, Corcos et al. (2008) build on a larger French dataset of which the Defever & Toubal (2007) data are a subset only and find oppositional results.

firms increases the probability of importing goods at arm's length (hypothesis 1). In particular, if the marginal effects are evaluated at the sample means, going from the lowest to the highest component intensity augments the probability of outsourcing by roughly 30 percentage points. Other covariates such as the firm's capital and skill intensity are statistically significant and have the expected signs as well. The theoretical prediction about the interaction of productivity and component intensity is tested by including an according multiplicative term in the regression. The mean interaction effect is found to be positive and highly significant. Defever & Toubal (2007) also reveal that contractual improvements that overproportionally affect the suppliers' intermediate inputs may constitute a relevant issue; see Antràs & Helpman (2008).

Bernard et al. (2007) combine detailed transaction-level U.S. trade data for 1997 with industry and country characteristics. In brief, there is evidence that intra-firm trade shares on product level, on country level, and for product-country-combinations are larger the higher the capital and skill intensities of the corresponding industry and the exporting country are. Most importantly, the authors conclude that special attention has to be paid to interaction effects between industry and country characteristics. For example, contrary to Yeaple (2006), the hypothesis 1-effect is found to be largest for capital-abundant exporting countries. Higher productivity dispersion in a given industry favors vertical FDI relative to offshore outsourcing for the imported products that belong to this industry. The authors also evaluate the role of a product's contractibility in the decision of vertically integrating the foreign supplier or of trading at arm's length. Higher contractibility of a product — measured as the degree of intermediation used to import the product — is found to be negatively related to intra-firm trade shares of the respective product category. This result is at odds with Antràs & Helpman (2008).

# 4 Data description and exploration

Our empirical analysis is based on the Spanish "Survey on Business Strategies" (ESEE)[13] from the Fundación SEPI[14]. The Fundación SEPI surveys approximately 2,000 firms with ten or more employees from all over Spain on a yearly basis since 1990 and asks for specific strategic business issues as well as for data on income statement and balance sheet statistics. The coverage of firms is representative for the Spanish manufacturing industry as a whole, and special care has been taken to control for the time dynamics in the panel which result from firms exiting and entering markets.[15]

[13] "Encuesta sobre estrategias empresariales".

[14] "Sociedad Estatal de Participaciones Industriales".

[15] See `http://www.funep.es/esee/en/einfo_que_es.asp` for details.

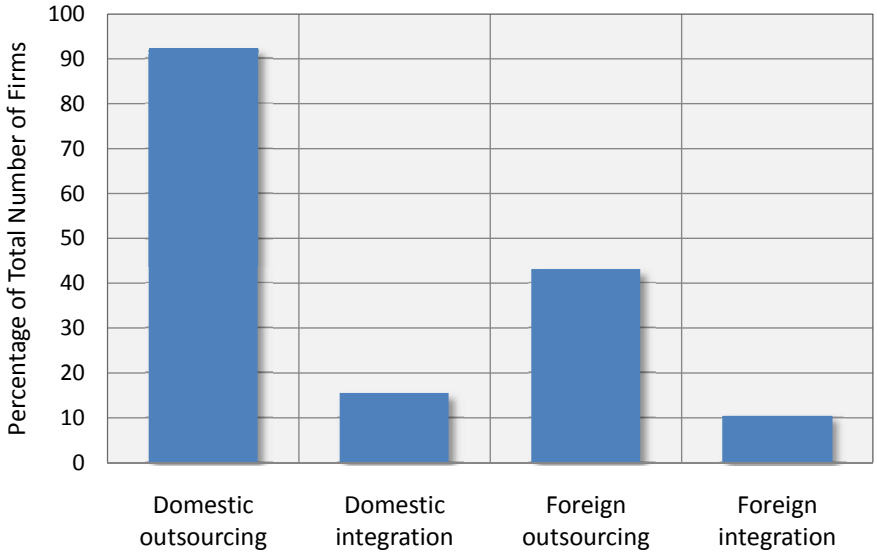


**Figure 3:** Use of Sourcing Modes by Spanish Manufacturing Firms in 2007

There are three primary advantages associated with the use of this data compared to prior empirical research. First, the data are observed at the level of the firm, not the industry. This allows us to map the micro-founded theory directly onto the data without relying on strong additional assumptions. Second, we can derive firm-specific measures for the extent to which intermediate inputs have been acquired through each of the four sourcing strategies (foreign integration, foreign outsourcing, domestic integration, and domestic outsourcing). Intermediate inputs — classified as such by the reporting firm — are goods that have been acquired from a *different legal entity* (either related or unrelated to the firm) and "incorporated and transformed in the production process" which perfectly fits into the modelling setup.[16] Clearly, the simultaneous and coherent coverage of all four sourcing modes generates a much more comprehensive picture of the firms' global sourcing decisions than was the case with earlier studies.[17] Third, all firms in the sample are categorized by the degree of participation of other companies in a firm's joint capital. By restricting the sample accordingly, we can construct a subset of firms which is in line with the underlying modelling assumption that the supplier never has a controlling stake in the headquarter's joint capital.

Upon examining the data by means of descriptive statistics we observe three facts that stand out. First, in 2007 nearly half of firms are active in foreign sourcing in the sense that they do

[16] We again emphasize that the modelling framework does not envisage a situation in which there is a command hierarchy in the strict sense of the word between firm and input producer.

[17] In principle, our data set covers the years 2000-2007 but we can distinguish the four strategic options for intermediate input sourcing only for the two most recent years. Nonetheless, we take the whole panel into account where its use provides us with additional insights or more accurate estimates (for example, we rely on the whole panel data set when it comes to total factor productivity estimation).

engage either in vertical FDI or in offshore outsourcing, or both. Noticeably, the equivalent percentage number for Japan's manufacturing firms is five percent; see Tomiura, 2007a. The high share of international firms in Spain reflects the high degree of economic integration in Europe. Indeed, more than 70 percent of the total value of imported goods have a member country of the European Union as origin, but only five percent come from Latin America.[18] Second, integration as opposed to outsourcing of intermediate input suppliers is reserved for only a minority of firms. 309 firms (15 percent) procure intermediate inputs from related parties in Spain, 209 firms (10 percent) acquire them from related parties abroad. By contrast, nearly all firms outsource part of their production to independent suppliers in the domestic economy, and more than 40 percent do so in the foreign economy (see figure 3). Third, close to 50 percent of all firms in the sample engage in two or more sourcing modes simultaneously. Such complex or mixed strategies are irrelevant from the Antràs (2003) and Antràs & Helpman (2004) perspective because, theoretically, each firm is reliant on one specific manufacturing component only, and she will choose one sourcing mode for the acquisition of this intermediate input. But, from an empirical point of view, the distinction is highly relevant and points towards the need for further theoretical modelling.

Table 1 reports variations in the intra-firm share of intermediate input acquisition in 2007, using 3-digit CNAE-93 industries. Of the total value of intermediate input sourcing, roughly one third has been acquired from abroad. Of the value of all components that have been imported by Spanish manufacturing firms, more than 40 percent can be attributed to imports from a related party. In 2006 the corresponding number is one third which is similar to the finding in Antràs (2003). For the value of all intermediates that have been acquired domestically, the related-party share happens to be significantly lower with about 20 percent. From a theoretical point of view, the difference in the shares could be attributed to higher degrees of contractual frictions for international transactions, if the lack of contractual enforcement affects both headquarter services and manufacturing components; see Antràs & Helpman (2008). We observe substantial cross-industry differences in the intra-firm share of intermediate input imports in total imports. In line with earlier empirical evidence, the intra-firm share of intermediate input imports is relatively high for industries producing highly differentiated, high-technology goods like "General & Electric Machinery", "Motorized Vehicles", and "Plastic & Rubber Products" but also for capital-abundant "Chemical Products" which are to a very large extent homogeneous goods. Conversely, the share is low or even zero for the "Textile", "Leather & Footwear", and "Timber & Wooden Products" industries. What is maybe surprising is that there is no obvious link between the industry-specific related-party shares for internationally and domestically sourced inputs. Indeed, in some industries related-party imports have a considerable weight in total

[18] These numbers correspond to imports in 2006.

imports but related-party transactions in the home country are of only marginal importance (or vice versa). For example, the "Plastic & Rubber Products" industry acquires 70 percent of its total component imports through a related party, but it buys less than seven percent of its domestically sourced goods from a related party.

**Table 1:** Acquisition of Intermediate Inputs by 3-Digit CNAE-93 Industry, 2007

| | | Foreign Sourcing | | Domestic Sourcing | |
|---|---|---|---|---|---|
| CNAE-93 | Industry | Total Value | Related-Party Share | Total Value | Related-Party Share |
| 311-316, 321-323 | General & Electric Machinery | 2,650 | 0.728 | 2,350 | 0.092 |
| 251-252 | Plastic & Rubber Products | 1,280 | 0.693 | 1,470 | 0.064 |
| 152-158, 160 | Food Products & Tobacco | 2,000 | 0.670 | 5,290 | 0.073 |
| 261-268 | Mineral Products (Non-Metal Products) | 606 | 0.482 | 3,150 | 0.282 |
| 341-343 | Motorized Vehicles | 10,200 | 0.468 | 15,300 | 0.218 |
| 241-247 | Chemical Products | 2,250 | 0.384 | 7,270 | 0.250 |
| 281-287 | Metal Products | 435 | 0.349 | 2,590 | 0.106 |
| 291-297 | Industry & Agricultural Machinery | 537 | 0.298 | 1,900 | 0.101 |
| 171-177, 181-183 | Textile | 315 | 0.266 | 519 | 0.116 |
| 300, 331-335 | Office Machinery & Data Processing | 369 | 0.227 | 598 | 0.036 |
| 211-212 | Pulp & Paper products | 435 | 0.204 | 1,760 | 0.345 |
| 361 | Furniture Industry | 206 | 0.079 | 677 | 0.098 |
| 351-355 | Other Transportation Equipment | 1,060 | 0.075 | 2,120 | 0.021 |
| 362-366, 371-372 | Miscellaneous Manufacturing | 86 | 0.068 | 162 | 0.003 |
| 271-275 | Ferrous Metals & Non-Ferrous Metals | 3,130 | 0.061 | 6,110 | 0.365 |
| 201-205 | Timber & Wooden Products | 41 | 0.037 | 563 | 0.047 |
| 151 | Meat | 36 | 0.025 | 3,010 | 0.286 |
| 221-223 | Publishing & Graphics Design | 64 | 0.000 | 774 | 0.193 |
| 159 | Beverage | 369 | 0.000 | 972 | 0.141 |
| 191-193 | Leather & Footwear | 16 | 0.000 | 98 | 0.011 |
| | Total | 26,100 | 0.420 | 56,700 | 0.201 |

* Values are in millions of euros. Only intermediate inputs are considered. Industries are ranked according to their related-party import share of intermediate inputs from highest to lowest.

Before turning to the exposition of the empirical strategy and the estimation results, we shortly present a description of the variables that we use in our application. Table 2 provides detailed information about explained and explanatory variables. In the light of the theoretical hypotheses we concentrate on firm characteristics in order to contribute to a better understanding of the sourcing behaviour of firms. Specifically, we seek to explain firm-specific variations in the share of both intra-firm imports of intermediate goods in total imports of intermediate goods ($FORINT$) and domestic intra-firm procurement of intermediate inputs in total procurement of domestically produced inputs ($DOMINT$). In a methodologically equivalent setting, we use the share of intermediate input imports in total intermediate input acquisition as the dependent variable ($OFFSHORE$). Thus, we are able to evaluate the relevance of the modelling framework with respect to both a firm's ownership structure of production as well as a firm's locational choice of sourcing; see hypotheses 1 to 3. In order to capture the importance of headquarter services for final-good production, $\eta$, we build on the Antràs (2003) premise that headquarter

firms primarily intervene in capital investment decisions rather than in personnel (that is labor-related) directives of their suppliers. The purpose of this paper is not to explicitly evaluate the validity of this premise, although some of our results might be interpreted as evidence for or against it. For each firm we work out the capital-labor ratio by scaling tangible fixed assets by the average number of employees ($KAPINT$). Furthermore, we calculate a measure for a firm's skill intensity by dividing the number of graduated employees by the number of production workers ($SKILLINT$). For robustness checks we also control for a firm's size measured as the average number of workers over the year ($SIZE_L$) and the amount of sales ($SIZE_S$). Moreover, we include as right-hand-side variables a firm's export status and age where this is meaningful ($EXPORT$ and $AGE$). Note that all these firm variables refer to a specific legal entity and not to a corporate group comprising a number of economically interdependent entities. Again, this is important since we can be sure that all variables truly reflect the headquarter characteristics that we are interested in.

Throughout our analysis, we use four different measures of firm productivity $\theta$ in order to evaluate the sensitivity of our results in terms of different productivity specifications. More specifically, we compute labor productivity ($PROD_L$) as valued added over the hours effectively worked and total factor productivity through sectoral production function estimates in standard ordinary least squares (OLS) regressions, in a fixed-effects model, and in an Olley & Pakes (1996) framework ($PROD_{OLS}$, $PROD_{FE}$, and $PROD_{OP}$, respectively).[19] The big advantage of the production function estimates with the Olley & Pakes (1996) method is that potential endogeneity problems related to unobserved productivity shocks and firms' market exit decisions are convincingly accounted for. All regression results are to a very large extent robust to using different productivity variables with one exception: results with total factor productivity estimated by simple OLS ($PROD_{OLS}$) differ substantially, but it is plausibly argued that this is the poorest productivity estimate in our analysis. Hence, in what follows we only discuss and present regression tables with total factor productivity estimated along the lines of the Olley & Pakes (1996) methodology.[20]

There is another somewhat subtle, nonetheless crucial, point to mention with respect to the data. In each year there is a small fraction of about six percent of all firms that do not purchase any intermediate inputs neither from a related nor from an independent party. Strictly speaking, such firms manufacture and combine all inputs necessary for final good production within the legal entity's very command structure, which rules out any kind of "hold-up" problem. From a theoretical point of view, with this kind of production organization the firm will choose optimal,

[19] See appendix for details.

[20] All other regression tables are available from the authors upon request.

**Table 2:** Data Description[1]

| Variable | Abbreviation | Description |
|---|---|---|
| *Sourcing Mode Variables* | | |
| Foreign integration share | $FORINT$ | Share of related-party imports of intermediate inputs in total imports of intermediate inputs. |
| Domestic integration share | $DOMINT$ | Share of domestic related-party sourcing of intermediate inputs in total domestic sourcing of intermediate inputs. |
| Offshore intensity | $OFFSHORE$ | Share of intermediate input imports in total intermediate input acquisition. |
| Foreign integration | $FI$ | Dummy variable which takes on the value one if the firm acquires at least some inputs through foreign integration and zero otherwise. |
| Foreign outsourcing | $FO$ | Dummy variable which takes on the value one if the firm acquires at least some inputs through foreign outsourcing (but not through foreign integration) and zero otherwise. |
| Domestic integration | $DI$ | Dummy variable which takes on the value one if the firm acquires at least some inputs through domestic integration (but not through foreign integration or foreign outsourcing) and zero otherwise. |
| Domestic outsourcing | $DO$ | Dummy variable which takes on the value one if the firm acquires at least some inputs through domestic outsourcing (but not through foreign integration, foreign outsourcing, or domestic integration) and zero otherwise. |
| *Firm Characteristics* | | |
| Capital intensity | $KAPINT$ | Net fixed assets over total year-average personnel. Expressed in thousands of euros per employee. |
| Skill intensity | $SKILLINT$ | Number of graduated employees per production worker (rounded to one digit after the decimal point) as of December 31st. |
| Labor productivity | $PROD_L$ | Labor productivity. See appendix for details. |
| Total factor productivity (a) | $PROD_{OLS}$ | Total factor productivity estimated with OLS. See appendix for details. |
| Total factor productivity (b) | $PROD_{FE}$ | Total factor productivity estimated by fixed-effects. See appendix for details. |
| Total factor productivity (c) | $PROD_{OP}$ | Total factor productivity estimated with Olley & Pakes (1996) methodology. See appendix for details. |
| Size (Personnel) | $SIZE_L$ | Approximation of the average number of workers over the year. It is calculated as the sum of full-time regular personnel, 1/2 of part-time regular personnel (both items as of December 31st), and the average number of eventual workers. |
| Size (Sales) | $SIZE_S$ | Total firm sales over the year. |
| Export status | $EXPORT$ | Dummy variable for a firm's export status. Takes on the value one if the firm exports goods to a foreign market and zero otherwise. |
| Age | $AGE$ | Firm age in years. |

[1] All data are firm-specific. All sourcing mode variables as well as the skill intensity variable are available for the years 2006 and 2007 while the productivity, capital intensity, size, export status, and age variables are available for the period 2000-2007. Intermediate inputs — classified as such by the reporting firm — are goods that have been "incorporated and transformed in the production process". Variables *KAPINT*, *SKILLINT*, *PROD*, and *SIZE* are in natural logs. Currency is euro.

that is, efficient investment levels in all inputs, but it will never be able to leverage factor price differences across countries. It is left for future work to take up this issue and incorporate it into a convincing modelling framework. For our purposes here it is sufficient to note that we have to exclude such firms whenever we compute share variables *FORINT*, *DOMINT*, and *OFFSHORE*.

# 5 Econometric modelling and estimation results

## 5.1 The Antràs effect and the offshore intensity effect

We start our econometric analysis by sticking very close to the cross-industry regressions of Nunn & Trefler (2008) and others who use the share of related-party imports in total imports as dependent variable. However, we extend the existing literature by using firm-level data, adapting the same exercise to the domestic sourcing decision and investigating the determinants of the locational choice of intermediate input production. In fact, we estimate three equations separately with *FORINT*, *DOMINT*, and *OFFSHORE*, respectively, as dependent variables. One econometric concern about this approach in our application is that there is a considerable amount of observations for which the dependent variables are zero since a majority of firms do not source from related parties (or from abroad). As is well-known, the application of simple OLS regression methods — be it on all observations or on positive observations only — would result in biased estimates. We tackle this problem by estimating different versions of the TOBIT model (Tobin, 1958) to account for the non-linearity in the regression. One advantage of this model is that it makes use of potentially valueable information by not excluding the limiting values. Thereby, it allows to compute both changes in the probability of being above the limit and changes in the values of the endogenous variable if it is already above the limit (McDonald & Moffitt, 1980). In particular, we apply the following panel Tobit model.

$$y_{it}^* = \mathbf{x}_{it}\boldsymbol{\lambda} + u_{it}, \qquad i = 1, ..., N, \quad t = 1, ..., T, \tag{7}$$

$$y_{it} = \max(0, y_{it}^*), \tag{8}$$

$$u_{it}|\mathbf{x}_{it} \sim \text{Normal}(0, \sigma^2),$$

where $y_{it}^*$ is the unobserved latent variable for firm $i$ at time $t$, $y_{it}$ is the observed dependent variable, $\mathbf{x}_{it} = (1, x_{it2}, ..., x_{itk})$ is a vector of explanatory variables, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_k)'$ is a vector of the intercept and the parameters to estimate, and $u_{it}$ is the error term.

First, we simply pool the data and apply Tobit as if it were one long cross section thus ignoring the time dimension. Here, $\mathbf{x}_{it}$ does not include lagged dependent or independent variables. Strictly speaking, however, we do not need to assume the data on $\mathbf{x}_{it}$ to be strictly exogenous, that is, we do not need $Cov[u_{it}, \mathbf{x}_{js}] = 0$. Thus, it is straightforward to exploit this property

of the pooled TOBIT analysis in future work. Besides, the error term may feature arbitrary serial correlation; the pooled Tobit estimator is still consistent and asymptotically normal if the model is correctly specified. Second, we control for unobserved heterogeneity by estimating a random-effects Tobit model. Suppose

$$u_{it} = c_i + \xi_{it}, \tag{9}$$
$$\xi_{it}|\mathbf{x}_i, c_i \sim \text{Normal}(0, \sigma_\xi^2),$$

where $c_i$ is a time-invariant firm characteristic that is unobserved by the econometrician, and $\mathbf{x}_i$ contains $\mathbf{x}_{it}$ for all $t$. Then, we have structural serial correlation in $u_{it}$ and, in contrast to the pooled Tobit estimator, we have to impose strict exogeneity on $\mathbf{x}_{it}$ conditional on $c_i$. Moreover, we assume that $\mathbf{x}_i$ and $c_i$ are independent and that $c_i$ follows a normal distribution. Third, we generalize the random-effects Tobit model by allowing for varying parameter vectors across time. In fact, the second model can be seen as a special case of the $T$ equation multivariate Tobit model in which the parameter vectors are time-invariant. Since we observe the dependent variables at two points in time only we estimate

$$y_{it}^* = \mathbf{x}_{it}\boldsymbol{\lambda}_t + u_{it}, \qquad i = 1, ..., N, \quad t = 1, 2, \tag{10}$$
$$y_{it} = \max(0, y_{it}^*), \tag{11}$$
$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \sim N\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right), \tag{12}$$

where $\rho$ is the coefficient capturing correlation between the two equations. Technically, and in contrast to the previous random-effects TOBIT model, we can estimate this bivariate TOBIT model with a balanced panel only. For robustness checks we also constrain this last model to have constant parameter vectors across the two time periods, which is simply the random-effects model with a balanced panel.

In all Tobit models we include as independent variables the factor intensities (*KAPINT*, *SKILLINT*), total factor productivity ($PROD_{OP}$), the age of the firm (*AGE*), and industry and year dummies to control for probable sector- and year-specific fixed effects. In the case where we use the offshore intensity of production organization as dependent variable (*OFFSHORE*) we additionally control for a firm's export status (*EXPORT*) in order to see whether this changes any of our results. Since the estimated coefficients in corner solution models such as ours are, in themselves, not informative about the economic magnitude of the independent variables' effects, we make use of the fact that the expected value of $y$ conditional on $\mathbf{x}$ is equal to the probability of being above zero conditional on $\mathbf{x}$, $\Pr[y > 0|\mathbf{x}]$, multiplied by the expected value of $y$ conditional

on being above the limit and on $\mathbf{x}$ (see McDonald & Moffitt, 1980), that is,

$$E[y|\mathbf{x}] = \Pr[y > 0|\mathbf{x}]E[y|y > 0, \mathbf{x}]. \tag{13}$$

Now, it is clear that a marginal change in one of the explanatory variables affects two factors, first, the probability of being above zero, and second, the expected value of $y$ conditional on being above the limit and on $\mathbf{x}$. Taking the derivative of (13) with respect to $x$ yields

$$\frac{\partial E[y|\mathbf{x}]}{\partial x} = \Pr[y > 0|\mathbf{x}]\frac{\partial E[y|y > 0, \mathbf{x}]}{\partial x} + E[y|y > 0, \mathbf{x}]\frac{\partial \Pr[y > 0|\mathbf{x}]}{\partial x}. \tag{14}$$

In tables 3 to 5 we always report the marginal effects of changes in the explanatory variables on the probability of observing a positive outcome. We run all regressions on two different samples: the first sample is unrestricted in all respects and the second sample excludes firms that are no true headquarter firms in the sense that another domestic or foreign firm has a controlling stake in the firm's joint capital. In fact, the set-up of the theoretical model does not envisage situations in which affiliated firms acquire intermediate inputs from their headquarter firms. Unfortunately, a serious problem of our maximum-likelihood models with the restricted sample is that estimations rarely converge, especially when using *FIRELINT* as the dependent variable. In all feasible cases, however, the results obtained with the two different samples do not show any significant differences, at least not in qualitative terms. In the following we therefore report all results for the unrestricted sample only.[21]

Let us first turn to the two separate cases in which we use the intra-firm shares of intermediate input sourcing as explained variables (*FORINT* and *DOMINT*). We directly want to test hypotheses 1 and 2 (Antràs effect) for both the domestic and the foreign economy. In light of hypothesis 1 we expect positive coefficients of the capital intensity variable since vertical integration requires sufficiently high capital intensities. Relatedly, hypothesis 2 suggests that in each sourcing market only the more productive firms should be active in integration so that we expect a positive coefficient for $PROD_{OP}$ in both equations. Columns (1) to (5) in tables 3 and 4 present the results for the foreign and the domestic sourcing case, respectively. We find positive and significant coefficients for capital intensity, skill intensity and total factor productivity in nearly all specifications. A unit-increase in the capital intensity variable, for example, increases the probability of acquiring components from a related party by two to five percentage points in the foreign sourcing case and by four to six percentage points in the domestic sourcing case. This is convincing evidence for the existence of the Antràs effect. Similarly, if an average firm faces a unit-increase in the total factor productivity variable, her probability of purchasing inputs from

[21] All regression results obtained from estimates with the restricted sample are available from the others upon request.

a vertically integrated supplier rises by approximately two percentage points, both in the foreign and the domestic economy. These results strongly confirm and extend prior empirical findings and are in line with hypotheses 1 and 2.

**Table 3:** Intra-firm Share of Foreign Intermediate Input Sourcing (TOBIT Model)[1]

| | Pooled | Random-Effects | Bivariate Constrained | Bivariate Unconstrained | |
|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4)-(2006) | (5)-(2007) |
| *KAPINT* | 0.045*** | 0.029*** | 0.030*** | 0.022* | 0.039*** |
| | (0.013) | (0.008) | (0.011) | (0.012) | (0.013) |
| *SKILLINT* | 0.127*** | 0.091*** | 0.101** | 0.106** | 0.125*** |
| | (0.046) | (0.029) | (0.042) | (0.043) | (0.048) |
| $PROD_{OP}$ | 0.066*** | 0.023*** | 0.021* | 0.031** | 0.011 |
| | (0.022) | (0.009) | (0.011) | (0.014) | (0.015) |
| *AGE* | 0.001** | 0.001*** | 0.001* | 0.001* | 0.001* |
| | (0.001) | (0.000) | (0.001) | (0.001) | (0.001) |
| Constant | -1.726*** | | | | |
| | (0.250) | | | | |
| $\rho$ | | 0.931*** | 0.923*** | 0.936*** | |
| | | (0.010) | (0.015) | (0.012) | |
| Industry Fixed Effects | Yes | Yes | Yes | Yes | |
| Year Fixed Effects | Yes | Yes | Implicit | Implicit | |
| Observations | 1482 | 1482 | 1118 | 1118 | |
| Pseudo $R^2$ | 0.11 | | | | |
| Log (Pseudo-)Likelihood | -765.724 | -561.084 | -407.395 | -392.155 | |
| F(.,.) / Wald $\chi^2$ | 6.55 | 95.52 | 517.77 | 2937.38 | |
| $Prob > F / Prob > \chi^2$ | 0.000 | 0.000 | 0.000 | 0.000 | |

[1] The table gives the marginal effects on the probability of being above zero evaluated at the sample means for the intra-firm share of foreign intermediate input acquisition (*FORINT*) as a function of the explanatory variables *KAPINT*, *SKILLINT*, $PROD_{OP}$, and *AGE* (see data description in table 2). The other corresponding marginal effects always have the same sign and significance. The sample does not exclude firms with domestic or foreign parent companies. Robust standard errors in columns (1) and (2) are clustered by firm and given in parentheses. *,**,*** indicate significance at the 10%, 5%, 1% levels, respectively.

We now turn to the case where we regress the offshore intensity (*OFFSHORE*) on our set of explanatory variables. Taking the model seriously and following the whole empirical literature in interpreting capital intensity as headquarter intensity, we expect to find a negative sign for the capital intensity variable; see hypothesis 3. The reason is that foreign sourcing should be a less probable production strategy if the supplier-input — and therefore the respective unit-cost advantage associated with offshoring — is of relatively low importance. At the same time, however, hypothesis 3 suggests that only the most productive firms source components abroad: only they should be able to shoulder the higher fixed costs and thus leverage the presumed unit cost-advantage involved with foreign sourcing. The results that we present in table 5 both confirm and contradict important parts of such a view on firms' sourcing activities. First, it is true that high-productivity firms are more likely to offshore intermediate input production. To be more precise, a unit-increase in the total factor productivity variable of an average firm boosts the probability of foreign sourcing by approximately four percentage points. Second, we see that

capital-intensive firms feature significantly higher probabilities of sourcing supplier-inputs from abroad. Quantitatively, this effect is even more pronounced than in the previous regressions: a unit-increase in the capital intensity variable of an average firm boosts the probability of offshoring by eight to ten percentage points. This is clearly at odds with the model. Importantly, this result is not driven by a firm's export status since we already control for this. It seems as if here the headquarter proxy is working through a different channel from the one derived in Antràs (2003) and Antràs & Helpman (2004).

**Table 4:** Intra-firm Share of Domestic Intermediate Input Sourcing (TOBIT Model)[1]

| | Pooled | Random-Effects | Bivariate Constrained | Bivariate Unconstrained | |
|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4)-(2006) | (5)-(2007) |
| *KAPINT* | 0.058*** | 0.048*** | 0.053*** | 0.044*** | 0.045*** |
| | (0.010) | (0.007) | (0.009) | (0.009) | (0.007) |
| *SKILLINT* | 0.076** | 0.078*** | 0.079*** | 0.088*** | 0.061** |
| | (0.030) | (0.025) | (0.031) | (0.028) | (0.030) |
| $PROD_{OP}$ | 0.032** | 0.010 | 0.019** | 0.011 | 0.021** |
| | (0.013) | (0.045) | (0.009) | (0.008) | (0.010) |
| *AGE* | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | (0.000) | (0.001) | (0.000) | (0.000) | (0.000) |
| Constant | -1.127*** | | | | |
| | (0.122) | | | | |
| $\rho$ | | 0.890*** | 0.895*** | 0.901*** | |
| | | (0.011) | (0.014) | (0.013) | |
| Industry Fixed Effects | Yes | Yes | Yes | Yes | |
| Year Fixed Effects | Yes | Yes | Implicit | Implicit | |
| Observations | 3282 | 3282 | 2878 | 2878 | |
| Pseudo $R^2$ | 0.11 | | | | |
| Log (Pseudo-)Likelihood | -1306.356 | -960.526 | -831.133 | -801.957 | |
| F(.,.) / Wald $\chi^2$ | 8.38 | 147.30 | 173.05 | 700.74 | |
| $Prob > F / Prob > \chi^2$ | 0.000 | 0.000 | 0.000 | 0.000 | |

[1] The table gives the marginal effects on the probability of being above zero evaluated at the sample means for the intra-firm share of domestic intermediate input acquisition (*DOMINT*) as a function of the explanatory variables *KAPINT*, *SKILLINT*, $PROD_{OP}$, and *AGE* (see data description in table 2). The other corresponding marginal effects always have the same sign and significance. The sample does not exclude firms with domestic or foreign parent companies. Robust standard errors in columns (1) and (2) are clustered by firm and given in parentheses. *,**,*** indicate significance at the 10%, 5%, 1% levels, respectively.

A fundamental aspect of hypothesis 2 is the interplay between a firm's capital intensity and her productivity level. A drawback of our analysis so far is that we are only looking at the direct marginal effects of changes in an average firm's capital intensity and productivity. But, in non-linear models such as the TOBIT model, all marginal effects (and their respective standard errors) depend on the values of all explanatory variables. To make the potential heterogeneity across observations visible, we compute for each observation in the sample its marginal effects on the probability of observing a non-zero outcome. Standard errors are obtained with the

delta method.[22] Then, we can nicely explore the hypothesized interaction between a firm's productivity level and her capital intensity by plotting the marginal effects of $PROD_{OP}$ and their corresponding standard errors against the firms' capital intensities $KAPINT$. In light of hypothesis 2 we expect the marginal effects to grow with the capital intensity for the two cases in which we use the intra-firm shares of intermediate input sourcing as dependent variables. More precisely, there should be in each location a cut-off headquarter intensity $\eta^{*\ell}$ above which the marginal effects of productivity on the probability of sourcing from a related party are positive and significantly different from zero while they should essentially be zero below this cut-off. In the case in which we use the offshore intensity as dependent variable, we have no firm view on what kind of pattern we expect in light of the fact that the marginal effect of capital intensity for an average firm is positive and significantly different from zero in itself.

Figure 4, in which each dot represents one firm at a specific point in time, exactly reveals the described patterns.[23] First, marginal effects of total factor productivity on the probability of sourcing from a vertically integrated supplier vary with a firm's capital intensity in a systematic way. In fact, firms with the highest capital intensities in the sample feature the highest quantitative effects of changes in the productivity variable. Accordingly, for firms with the lowest capital intensities such effects approach zero, both in the foreign and the domestic sourcing case. Second, looking at the z-statistics that we obtain from dividing the marginal effects by their standard errors, we see that for sufficiently small capital intensities the productivity effects are not significantly different from zero (z-statistics below 1.96). Third, there are important differences between these two cases and the offshoring-intensity case: although the magnitude of the productivity effects again seems to be positively correlated with capital intensity, its variance is relatively small and the marginal effects are practically always significantly different from zero.

The problem with these graphs is that we are looking at real firms, each of which has unique characteristics not only in terms of different capital intensities. Indeed, if we are interested in the pure interaction between productivity and capital intensity we have to hold other firm characteristics constant when changing the capital intensity variable. That is what we do in figure 5. We evaluate all marginal effects of productivity on the probability of observing a positive outcome at the sample means of all explanatory variables except for a firm's capital intensity. Note that these are now the effects for fictional firms with different capital intensities but average (and completely identical) characteristics in all other respects. We find, in point of

[22] For details, see Wooldridge (2003).

[23] We only show the figures for pooled TOBIT estimates in order to save on space. The results of the other TOBIT models are, however, congruent. Figures are available from the authors upon request.

**Table 5:** Offshore Intensity of Intermediate Input Sourcing (TOBIT Model)[1]

| | Pooled | | Random-Effects | | Bivariate Constrained | | Bivariate Unconstrained | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7)-(2006) | (8)-(2007) | (9)-(2006) | (10)-(2007) |
| *KAPINT* | 0.097*** | 0.067*** | 0.102*** | 0.078*** | 0.101*** | 0.078*** | 0.101*** | 0.106*** | 0.079*** | 0.082*** |
| | (0.011) | (0.011) | (0.010) | (0.010) | (0.012) | (0.012) | (0.012) | (0.013) | (0.013) | (0.014) |
| *SKILLINT* | 0.163*** | 0.141*** | 0.181*** | 0.155*** | 0.222*** | 0.194*** | 0.258*** | 0.199*** | 0.233*** | 0.173*** |
| | (0.050) | (0.046) | (0.044) | (0.044) | (0.058) | (0.053) | (0.060) | (0.061) | (0.057) | (0.057) |
| $PROD_{OP}$ | 0.084*** | 0.063*** | 0.044*** | 0.038*** | 0.034*** | 0.029** | 0.033** | 0.034** | 0.028* | 0.029* |
| | (0.017) | (0.016) | (0.012) | (0.013) | (0.013) | (0.014) | (0.016) | (0.016) | (0.016) | (0.016) |
| *AGE* | 0.002*** | 0.001** | 0.002*** | 0.001** | 0.002*** | 0.001** | 0.001** | 0.002*** | 0.001 | 0.001** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| *EXPORT* | | 0.287*** | | 0.240*** | | 0.237*** | | | 0.226*** | 0.257*** |
| | | (0.026) | | (0.020) | | (0.028) | | | (0.030) | (0.031) |
| Constant | -2.195*** | -1.857*** | | | | | | | | |
| | (0.178) | (0.181) | | | | | | | | |
| $\rho$ | | | 0.819*** | 0.796*** | 0.812*** | 0.789*** | 0.817*** | | 0.795*** | |
| | | | (0.012) | (0.013) | (0.019) | (0.021) | (0.018) | | (0.020) | |
| Industry Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | | Yes | |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Implicit | Implicit | Implicit | | Implicit | |
| Observations | 3356 | 3356 | 3356 | 3356 | 2958 | 2958 | 2958 | | 2958 | |
| Pseudo $R^2$ | 0.16 | 0.22 | | | | | | | | |
| Log (Pseudo-)Likelihood | -1665.322 | -1543.356 | -1199.282 | -1129.238 | -970.012 | -913.803 | -956.906 | | -900.636 | |
| F(.,.) / Wald $\chi^2$ | 15.53 | 20.87 | 380.83 | 503.06 | 286.81 | 387.58 | 310.38 | | 414.90 | |
| $Prob > F / Prob > \chi^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | | 0.000 | |

[1] The table gives the marginal effects on the probability of being above zero evaluated at the sample means for the offshore intensity of total intermediate input acquisition (*OFFSHORE*) as a function of the explanatory variables *KAPINT*, *SKILLINT*, $PROD_{OP}$, *AGE*, and *EXPORT* (see data description in table 2). The other corresponding marginal effects always have the same sign and significance. The sample does not exclude firms with domestic or foreign parent companies. Robust standard errors in columns (1) and (2) are clustered by firm and given in parentheses. *,**,*** indicate significance at the 10%, 5%, 1% levels, respectively.

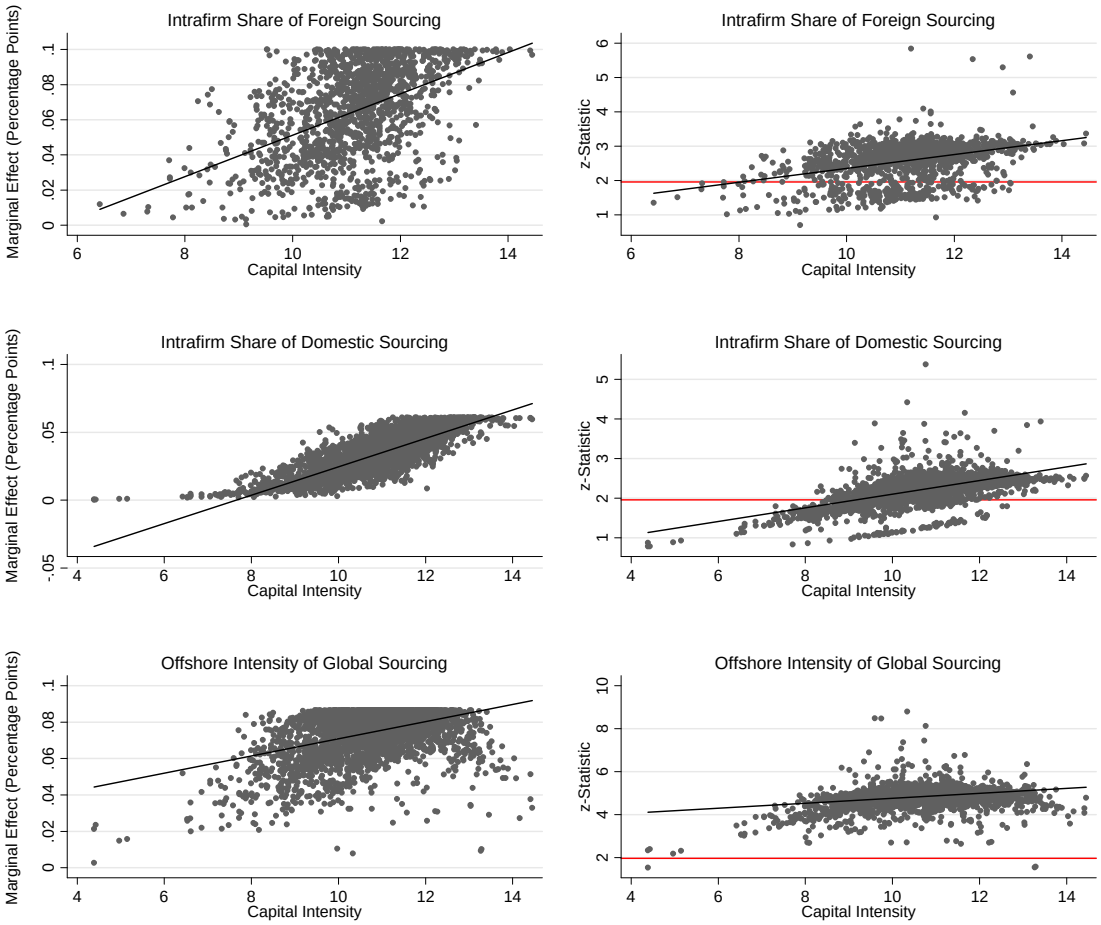


**Figure 4:** Marginal Effects of Productivity and z-Statistics After Pooled Tobit

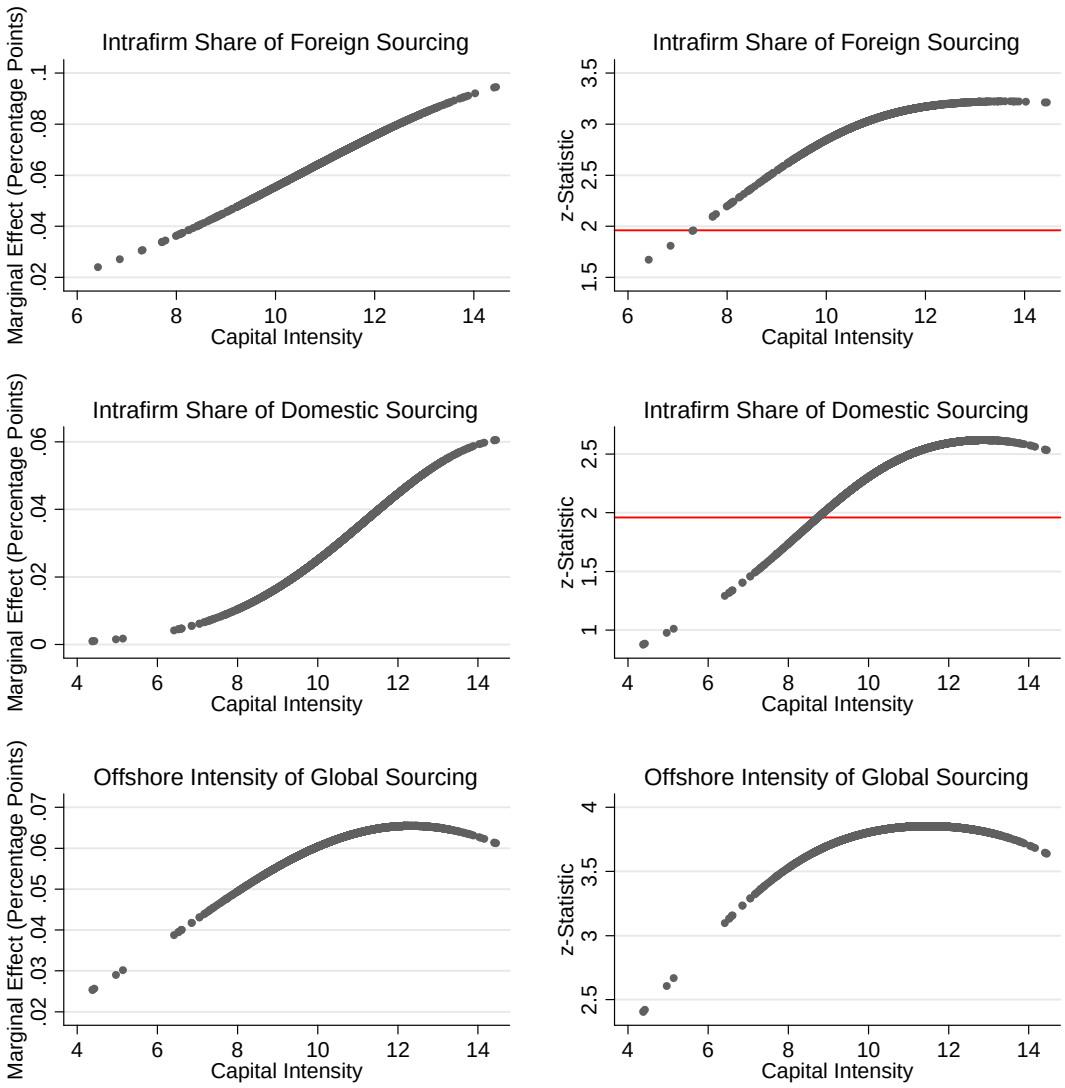


**Figure 5:** Marginal Effects of Productivity and z-Statistics After Pooled Tobit; $PROD_{OP}$ held at Mean.

fact, dramatic confirmation for the Antràs effect with fixed cost heterogeneity: the productivity effects on the probability of acquiring inputs from a related party are larger for higher capital intensities, and such effects disappear for sufficiently low capital intensities. Any productivity increase of an average firm with a capital-labor ratio of $\ln(K/SIZE_L)$ smaller than 7.5 (8.5) will not lead to higher probabilities of engaging in foreign (domestic) integration. This result suggests that the foreign cut-off headquarter intensity $\eta^{F*}$ is smaller than the domestic counterpart, $\eta^{D*}$, which means that the fixed cost ratio of integration over outsourcing is comparably smaller abroad than at home. We repeat the same exercise twice, evaluating the marginal effects at the means of all explanatory variables but once at the minimum and once at the maximum of the productivity variable (see appendix). The Antràs effect with fixed cost heterogeneity holds through.

What is to be learnt from this analysis? A crucial result is that coefficients on a factor intensity variables reflect confounding factors. The point is the following: if we interpret capital intensity as headquarter intensity in the sense of $\eta$, then our estimations cast serious doubts on the validity of an important aspect of the theoretical model, namely the offshore intensity effect. However, if we abstain from this interpretation, then the logical consequence is that much more research is needed to evaluate the role that capital intensity plays (from a theoretical point of view) in the global sourcing decisions of firms. In our view, a firm's capital intensity reflects the effects of at least two factors. The first factor involves the headquarter-intensity channel as convincingly rationalized in Antràs (2003) and Antràs & Helpman (2004). As stated above, we find impressive evidence that the Antràs effect with fixed cost heterogeneity is indeed at work in the empirical model. The second factor, however, must have to do with factor price differences across countries. In our theoretical section we have slightly reformulated the Antràs & Helpman (2004) model in more general terms as regards the functional form of final-good production. Without going into details, we thus point towards the unit-cost function $c_y$ through which the factor-intensity variable could plausibly affect the offshoring intensity.

### 5.2 Productivity Ranking of Firms

In order to establish a comprehensive productivity ranking of firms in a unified econometric framework we adopt the Bernard & Jensen (1999) methodology which has seen extensive use in estimating productivity and size premia of exporting relative to non-exporting firms. Specifically, we estimate such premia for firms that pursue different sourcing strategies, relative to a baseline category. For this purpose we regress a firm's total factor productivity ($PROD_{OP}$) and size in terms of sales ($SIZE_S$) in separate specifications on mutually exclusive sourcing dummies ($FI$ for foreign integration, $FO$ for foreign outsourcing, $DI$ for domestic integration, and $DO$ for domestic outsourcing). If a firm is active in two or more sourcing modes simultaneously,

she is assigned to the category which is supposed to be associated with the highest fixed costs. Note that we completely ignore the Antràs effect in this approach. In particular, we apply the following panel model.

$$\begin{aligned} y_{it} &= \mathbf{x}_{it}\boldsymbol{\lambda} + u_{it}, \qquad i = 1, ..., N, \quad t = 1, ..., T, \\ u_{it} &= c_i + \xi_{it}, \end{aligned} \tag{15}$$

where $y_{it}$ is the observed dependent variable for firm $i$ at time $t$, $\mathbf{x}_{it} = (1, x_{it2}, ..., x_{itk})$ is a vector of explanatory variables, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_k)'$ is a vector of the intercept and the parameters to estimate, and $u_{it}$ is the composite error term including an unobserved effect $c_i$ and an idiosyncratic error $\xi_{it}$.

To get the most out of the data, we implement three common estimation methods. First, we simply pool the data and apply ordinary least squares (OLS) as if it were one long cross section thus ignoring the time dimension. This estimator provides us with consistent estimates as long as there is only arbitrary correlation between the composite error term and the explanatory variables in each period $t$. To incorporate serial correlation within panel units through the unobserved effect we always report robust standard errors clustered by firm. Second, we average out all variables across the two time periods and apply OLS where we again disregard the time dimension in the data. This estimator is called between-estimator in the literature. The model reads as

$$\begin{aligned} \bar{y}_i &= \bar{\mathbf{x}}_i\boldsymbol{\lambda} + \bar{u}_i, \qquad i = 1, ..., N, \\ \bar{u}_i &= c_i + \bar{\xi}_i, \end{aligned} \tag{16}$$

where the bars indicate mean values over time. For this model to deliver consistent estimates the data on $\mathbf{x}_{it}$ has to be strictly exogenous conditional on $c_i$, and $\mathbf{x}_{it}$ has to be independent from $c_i$. Third, we explicitly treat the data as a panel and control for unobserved heterogeneity across firms by estimating a random-effects model. Under the same assumptions as before, this is the model with the fully efficient estimator.

As additional control variables we include a firm's age ($AGE$), export status ($EXPORT$), and size in terms of employees ($SIZE_L$), as well es industry and year dummies. As in the previous approach, we perform all regressions with an unrestricted and a restricted sample. Again, in contrast to the former, the latter excludes firms with foreign or domestic parental companies. Furthermore, we distinguish between a sample with non-sourcing firms excluded and non-sourcing firms included. In the former case, we estimate all productivity and size premia relative to the domestic outsourcing category. In the latter case, the non-sourcing mode serves as the baseline category.

Hypothesis 3 suggests a productivity-dependent sorting pattern according to which only the most productive firms source from a vertically integrated supplier abroad and only the least productive firms outsource intermediate input production in the domestic economy. Moreover, foreign-outsourcing firms are meant to feature relatively higher productivity levels than domestic-integration firms. Consequently, we expect the following coefficient ranking for the three (four) dummy variables (from highest to lowest): *FI*, *FO*, *DI*(, and *DO*). Looking at the results for the total factor productivity premia for the restricted sample in columns (1) to (6) in tables 6 and 7, we exactly find the hypothesized ranking of coefficients in the overwhelming majority of cases. First, the foreign-integration dummy has always the largest coefficient of all sourcing dummies (even when we control for size), and it is almost always significantly different from zero. Additionally, it is significantly larger than the coefficients of the other sourcing dummies when size is not controlled for. Strikingly, foreign-integration firms are roughly 50 percent more productive than domestic-outsourcing firms when we do not control for size.[24] Second, the domestic-outsourcing category is always associated with the lowest productivity level: in table 6 all coefficients of all other sourcing dummies are positive and mostly significant, in table 7 the domestic-outsourcing dummy has always the lowest coefficient of all sourcing dummies. We see that foreign-outsourcing firms and domestic-integration firms are somewhat above and somewhat below than 20 percent more productive than domestic outsourcing-firms, respectively.[25] Third, we cannot establish a statistically significant difference between foreign-outsourcing firms and domestic-integration firms, although the point estimates of the *FO* and *DI* dummies suggest that the former are more productive than the latter. Note that the productivity ranking of firms that grows out of the theoretical model is not as clear-cut as it might seem at first sight. It only holds for a situation where both cut-off headquarter intensities have been surpassed. Definitely, domestic-integration firms and foreign-outsourcing firms are partly located in the same productivity range once you compare high headquarter-intensity firms ($\eta > \eta^{\ell *}$) with medium headquarter-intensity firms ($\eta^{\ell *} > \eta > \eta^{l *}, \ell \neq l$). Conclusively, the minor hint towards this kind of ambiguity that we find in the data can by no means be seen as evidence against the model.

In the text we only report the results for the restricted sample but we find at least two interesting differences between the estimates with the unrestricted and the restricted samples. First, the coefficient ranking slightly changes once we include non-headquarter firms in the sample.[26] Indeed, what we observe is a reversal in the ranking of foreign-outsourcing firms and domestic-integration firms in terms of productivity, and this reversal is at least marginally

[24]The premia are calculated from the coefficients in table 6 as follows: $PREMIUM^{\ell\kappa} = 100 \times (\exp(\lambda^{\ell\kappa}) - 1)$.

[25]Again, we look at table 6 at the columns where we do not control for size.

[26]See appendix for estimation results with the unrestricted sample.

**Table 6:** Size and Productivity Premia According to Different Sourcing Modes I (Restricted Sample)[1]

| | $PROD_{OP}$ | | | | | | $SIZE_S$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pooled OLS | | Between Estimator | | Random-Effects | | Pooled OLS | | Between Estimator | | Random-Effects | |
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *FI* | 0.432*** | 0.213* | 0.553*** | 0.298** | 0.347*** | 0.176* | 2.634*** | 0.490*** | 2.963*** | 0.590*** | 0.288** | 0.116 |
| | (0.120) | (0.117) | (0.152) | (0.149) | (0.098) | (0.091) | (0.307) | (0.087) | (0.333) | (0.167) | (0.128) | (0.088) |
| *FO* | 0.204*** | 0.134*** | 0.237*** | 0.157*** | 0.150*** | 0.098*** | 0.875*** | 0.302*** | 0.979*** | 0.344*** | 0.088*** | 0.078*** |
| | (0.033) | (0.031) | (0.038) | (0.037) | (0.032) | (0.031) | (0.080) | (0.038) | (0.087) | (0.044) | (0.021) | (0.019) |
| *DI* | 0.213** | 0.059 | 0.202** | 0.019 | 0.132* | 0.021 | 1.461*** | 0.254** | 1.756*** | 0.328*** | 0.156*** | 0.079 |
| | (0.085) | (0.090) | (0.091) | (0.090) | (0.070) | (0.070) | (0.177) | (0.118) | (0.205) | (0.102) | (0.048) | (0.051) |
| *AGE* | -0.002* | -0.004*** | -0.002** | -0.004*** | -0.001 | -0.003*** | 0.018*** | 0.003*** | 0.017*** | 0.003*** | 0.028*** | 0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) | (0.003) | (0.001) |
| *EXPORT* | 0.184*** | 0.094*** | 0.191*** | 0.097*** | 0.179*** | 0.097*** | 1.018*** | 0.283*** | 1.067*** | 0.290*** | 0.228*** | 0.165*** |
| | (0.032) | (0.033) | (0.034) | (0.035) | (0.030) | (0.030) | (0.075) | (0.040) | (0.080) | (0.041) | (0.042) | (0.034) |
| $SIZE_L$ | | 0.137*** | | 0.137*** | | 0.138*** | | 1.090*** | | 1.081*** | | 1.083*** |
| | | (0.015) | | (0.015) | | (0.015) | | (0.017) | | (0.017) | | (0.017) |
| Constant | 9.751*** | 9.306*** | 9.712*** | 9.274*** | 9.752*** | 9.299*** | 14.990*** | 11.460*** | 14.93*** | 11.48*** | 15.510*** | 11.590*** |
| | (0.076) | (0.086) | (0.087) | (0.096) | (0.076) | (0.086) | (0.195) | (0.120) | (0.206) | (0.114) | (0.247) | (0.122) |
| Industry Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | No | No | Yes | Yes | Yes | Yes | No | No | Yes | Yes |
| H0: $FI = FO$ | 0.05 | 0.49 | 0.04 | 0.34 | 0.04 | 0.38 | 0.00 | 0.02 | 0.00 | 0.14 | 0.12 | 0.66 |
| H0: $FI = DI$ | 0.13 | 0.27 | 0.04 | 0.10 | 0.07 | 0.16 | 0.00 | 0.09 | 0.00 | 0.16 | 0.33 | 0.71 |
| H0: $FO = DI$ | 0.92 | 0.38 | 0.71 | 0.13 | 0.79 | 0.27 | 0.00 | 0.68 | 0.00 | 0.88 | 0.15 | 0.98 |
| Observations | 2278 | 2278 | 2278 | 2278 | 2278 | 2278 | 2487 | 2487 | 2487 | 2487 | 2487 | 2487 |
| $R^2$ | 0.18 | 0.23 | 0.21 | 0.26 | | | 0.43 | 0.87 | 0.46 | 0.87 | 0.46 | 0.87 |

[1] The table gives estimation results obtained with the Bernard & Jensen (1999) methodology. Each column represents a separate regression where the dependent variable at the top of each column is a function of dummy variables for foreign integration (*FI*), foreign outsourcing (*FO*), domestic integration (*DI*), and a firm's age (*AGE*) and export status (*EXPORT*). In columns with even numbers we additionally control for firm size ($SIZE_L$). The sourcing modes are mutually exclusive. If a firm is active in two or more sourcing modes simultaneously, she is assigned to the category which is supposed to be associated with the highest fixed costs. The sample excludes firms with domestic or foreign parent companies as well as non-sourcing firms. The lower part of the table gives p-values of tests for equality of coefficients. Robust standard errors are clustered by firm (except for the between estimator) and given in parentheses. The size indicators are in natural logs. *,**,*** indicate significance at the 10%, 5%, 1% levels, respectively.

**Table 7:** Size and Productivity Premia According to Different Sourcing Modes II (Restricted Sample)[1]

| | $PROD_{OP}$ | | | | | | $SIZE_S$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pooled OLS | | Between Estimator | | Random-Effects | | Pooled OLS | | Between Estimator | | Random-Effects | |
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *FI* | 0.415*** | 0.183 | 0.540*** | 0.274* | 0.326*** | 0.149 | 2.708*** | 0.487*** | 3.048*** | 0.588*** | 0.304** | 0.116 |
| | (0.124) | (0.122) | (0.161) | (0.158) | (0.101) | (0.095) | (0.318) | (0.096) | (0.359) | (0.180) | (0.136) | (0.092) |
| *FO* | 0.188*** | 0.103** | 0.228*** | 0.133* | 0.131*** | 0.070* | 0.959*** | 0.303*** | 1.080*** | 0.346*** | 0.096** | 0.074*** |
| | (0.051) | (0.048) | (0.072) | (0.071) | (0.043) | (0.042) | (0.120) | (0.058) | (0.169) | (0.083) | (0.043) | (0.028) |
| *DI* | 0.198** | 0.030 | 0.193* | -0.004 | 0.113 | -0.006 | 1.547*** | 0.253** | 1.862*** | 0.328*** | 0.159*** | 0.074 |
| | (0.093) | (0.093) | (0.109) | (0.107) | (0.076) | (0.076) | (0.199) | (0.126) | (0.251) | (0.125) | (0.057) | (0.053) |
| *DO* | -0.012 | -0.027 | -0.006 | -0.021 | -0.016 | -0.024 | 0.092 | 0.004 | 0.102 | 0.003 | 0.002 | 0.001 |
| | (0.047) | (0.045) | (0.069) | (0.067) | (0.039) | (0.038) | (0.108) | (0.052) | (0.162) | (0.079) | (0.035) | (0.025) |
| *AGE* | -0.002* | -0.004*** | -0.002** | -0.004*** | -0.001 | -0.003*** | 0.018*** | 0.003*** | 0.017*** | 0.003*** | 0.028*** | 0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) | (0.003) | (0.001) |
| *EXPORT* | 0.195*** | 0.103*** | 0.198*** | 0.104*** | 0.182*** | 0.100*** | 1.034*** | 0.290*** | 1.077*** | 0.297*** | 0.217*** | 0.156*** |
| | (0.031) | (0.032) | (0.033) | (0.034) | (0.028) | (0.029) | (0.072) | (0.038) | (0.078) | (0.040) | (0.039) | (0.032) |
| $SIZE_L$ | | 0.137*** | | 0.137*** | | 0.139*** | | 1.091*** | | 1.083*** | | 1.089*** |
| | | (0.015) | | (0.014) | | (0.015) | | (0.017) | | (0.017) | | (0.017) |
| Constant | 9.765*** | 9.336*** | 9.730*** | 9.308*** | 9.777*** | 9.328*** | 14.860** | 11.420*** | 14.82*** | 11.45*** | 15.490*** | 11.550*** |
| | (0.086) | (0.092) | (0.106) | (0.111) | (0.082) | (0.090) | (0.210) | (0.124) | (0.250) | (0.133) | (0.239) | (0.121) |
| Industry Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | No | No | Yes | Yes | Yes | Yes | No | No | Yes | Yes |
| H0: $FI = FO$ | 0.05 | 0.49 | 0.04 | 0.34 | 0.04 | 0.37 | 0.00 | 0.02 | 0.00 | 0.14 | 0.11 | 0.64 |
| H0: $FI = DI$ | 0.13 | 0.27 | 0.04 | 0.09 | 0.06 | 0.15 | 0.00 | 0.09 | 0.00 | 0.16 | 0.29 | 0.68 |
| H0: $FI = DO$ | 0.00 | 0.07 | 0.00 | 0.04 | 0.00 | 0.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.20 |
| H0: $FO = DI$ | 0.91 | 0.39 | 0.71 | 0.13 | 0.80 | 0.27 | 0.00 | 0.67 | 0.00 | 0.86 | 0.16 | 1.00 |
| H0: $FO = DO$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| H0: $DI = DO$ | 0.01 | 0.51 | 0.03 | 0.85 | 0.06 | 0.79 | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.14 |
| Observations | 2426 | 2426 | 2426 | 2426 | 2426 | 2426 | 2645 | 2645 | 2645 | 2645 | 2645 | 2645 |
| $R^2$ | 0.18 | 0.23 | 0.21 | 0.26 | | | 0.44 | 0.87 | 0.46 | 0.87 | | |

[1] The table gives estimation results obtained with the Bernard & Jensen (1999) methodology. Each column represents a separate regression where the dependent variable at the top of each column is a function of dummy variables for foreign integration ($FI$), foreign outsourcing ($FO$), domestic integration ($DI$), domestic outsourcing ($DO$), and a firm's age ($AGE$) and export status ($EXPORT$). In columns with even numbers we additionally control for firm size ($SIZE_L$). The sourcing modes are mutually exclusive. If a firm is active in two or more sourcing modes simultaneously, she is assigned to the category which is supposed to be associated with the highest fixed costs. The sample excludes firms with domestic or foreign parent companies but includes non-sourcing firms. The lower part of the table gives p-values of tests for equality of coefficients. Robust standard errors are clustered by firm (except for the between estimator) and given in parentheses. The size indicators are in natural logs. *,**,*** indicate significance at the 10%, 5%, 1% levels, respectively.

significant in several specifications. Second, all point estimates of the sourcing coefficients are smaller with the unrestricted sample than with the restricted sample, with one exception: the *DI* dummy features a larger coefficient with the unrestricted sample when non-sourcing firms are excluded. Generally speaking, this points towards the idea that true headquarter firms have to overcome higher productivity hurdles relative to the domestic outsourcing category than non-headquarter firms whose sourcing decisions might not always be subject to their own command structures.

We only comment very briefly on the size premia that we also investigate in the same econometric framework. The reason is that the model primarily argues on the grounds of total factor productivity differences across firms. In any case, we again find substantial heterogeneity across firms in terms of size, and this heterogeneity exhibits a structural pattern depending on the firms' sourcing strategies. Firms sourcing from a vertically integrated supplier in the foreign economy are by far the largest and firms engaging in domestic outsourcing are by far the smallest ones in the sample. In contrast to what we observed before, the coefficients on the domestic-integration dummy are larger than those on the foreign-outsourcing dummy with few exceptions. Often, these differences are statistically significant. We conclude that there is indeed an important structural difference in the productivity and size premia ranking according to different sourcing modes.

# 6 Final remarks

Which firms acquire intermediate inputs abroad, and which at home? And why do they organize the ownership structure of input production in different ways? A new strand of the international economics literature — pioneered by Antràs (2003) — approaches these previously less explored questions bringing together insights from the firm and contracting theory with recent advancements in the study of international trade. In our paper we test empirically the most relevant theoretical propositions that are associated with the modelling framwork in Antràs (2003) and Antràs & Helpman (2004). We rest our empirical analysis on a novel Spanish firm-level data set thereby extending existing empirical literature considerably.

Our findings suggest that Antràs (2003) and Antràs & Helpman (2004) indeed elucidate important channels through which a firm's global sourcing strategy is being affected. In particular, we find the Antràs effect with fixed cost heterogeneity to be strikingly consistent with the data. Only sufficiently capital-intensive firms may find it optimal to integrate the input-supplier into the firm boundaries of control, depending on their total factor productivity levels. A productivity-increase of such firms significantly boosts their probability of engaging in vertical integration. On the other hand, low capital-intensity firms never find it optimal to acquire

property rights in the supplier-input irrespective of their productivity levels. Considerably, as suggested by the model, this interaction between headquarter intensity and productivity is shaping production organization both at home and abroad in an almost identical manner. Furthermore, our data impressively reflect the productivity ranking of firms engaging in different sourcing strategies as put forward by Antràs & Helpman (2004). We see that firms acquiring intermediate inputs from a vertically integrated supplier abroad are on average more productive than all other firms in the sample, and domestic-outsourcing firms are the least-productive ones. There is also evidence that foreign-outsourcing firms are slightly more productive than domestic-integration firms.

Despite this overwhelming evidence in favor of the modelling framework, there is one result in our analysis that we find especially worthwhile considering. Our capital-intensity variable, the common proxy for headquarter intensity, seems to operate not only in the spirit of Antràs (2003) and Antràs & Helpman (2004) but also as a means of capturing and reflecting comparative advantage. Indeed, the more capital-intensive is a firm the larger is her offshore intensity. It is left for future work to convincingly incorporate this effect into the existing modelling set-up. Our research suggests a bunch of further starting points for future theoretical as well as empirical work in this area. What is the rationale behind completely integrating intermediate input production into a firm's *legal entity* (non-sourcing firms)? How do we explain the existence of firms engaging in various sourcing modes simultaneously in the data? And, is there a way of empirically investigating the effects of varying degrees of contractibility in both inputs in a more explicit way than was the case with earlier studies? Besides what we state in the introduction as regards trade policy, and without going into detail, we are convinced that these questions should and will be put on the future research agenda.

# Appendix

This appendix gives information on how we compute the different measures of productivity at the level of the firm. Where necessary, we calculate real values in terms of 2007 prices using sectoral price indexes provided by the national statistics institute (INE) in Spain.

**Labor productivity** $PROD_L$. Computed as value added divided by the approximation of hours effectively worked. The units of the hourly productivity may be interpreted as euros per hour. This variable is calculated only for companies with non-negative values added. Value added is defined as the sum of sales, changes in the stock of inventory, and other operating income minus the sum of purchases and the costs of external services.

**Total factor productivity** $PROD_{OLS}$. Computed as the residual of sectoral production function estimates according to the following regression equation (estimated by standard OLS).

$$y_t = \lambda_1 k_t + \lambda_2 l_t + \lambda_3 a_t + u_t, \tag{17}$$

where $y_t$ is log-value added in year $t = 2000, ..., 2007$, $k_t$ is the log value of the capital stock net of amortization and depreciation, $l_t$ is the log value of the average number of employees per year, $a_t$ is the age of the firm (in years), and $u_t$ is the error term. The capital stock excludes all immaterial and financial assets. The average number of employees over the year is the sum of full-time regular employees, 1/2 of part-time regular employees (both items as of December 31st), and the average number of eventual workers.

**Total factor productivity** $PROD_{FE}$. Computed as the residual of sectoral estimates of equation (17) with a fixed-effects model in order to account for unobserved time-invariant firm heterogeneity. The basic idea is as follows. If $u_t = \Omega + \eta_t$ where $\eta_t$ is an error term with usual properties and $\Omega$ is a time-invariant productivity differential that is being observed by the firm (but not the econometrician), then there may be systematic correlation between $u_t$ and $k_t$ and $l_t$ because the firm will incorporate this additional information in the decision of how many inputs to use (*simultaneity bias*). Thus, the OLS estimates are biased. With the fixed-effects specification we eliminate $\Omega$ and get consistent estimates.

**Total factor productivity** $PROD_{OP}$. Computed as the residual of estimates of equation (17) with the Olley & Pakes (1996) methodology. Suppose $\Omega$ changes over time such that $u_t = \Omega_t + \eta_t$. Then, the fixed-effects model yields biased estimates. We account for this endogeneity issue by applying a consistent three-step estimation procedure proposed by Olley & Pakes (1996) where capital investment $I_t$ is meant to proxy for the unobserved productivity shock $\Omega_t$. In doing so, we automatically address the *selection bias* problem that results from inefficient firms exiting the market and constitutes another potential source of inconsistency. For details regarding the

estimation algorithm see Olley & Pakes (1996).

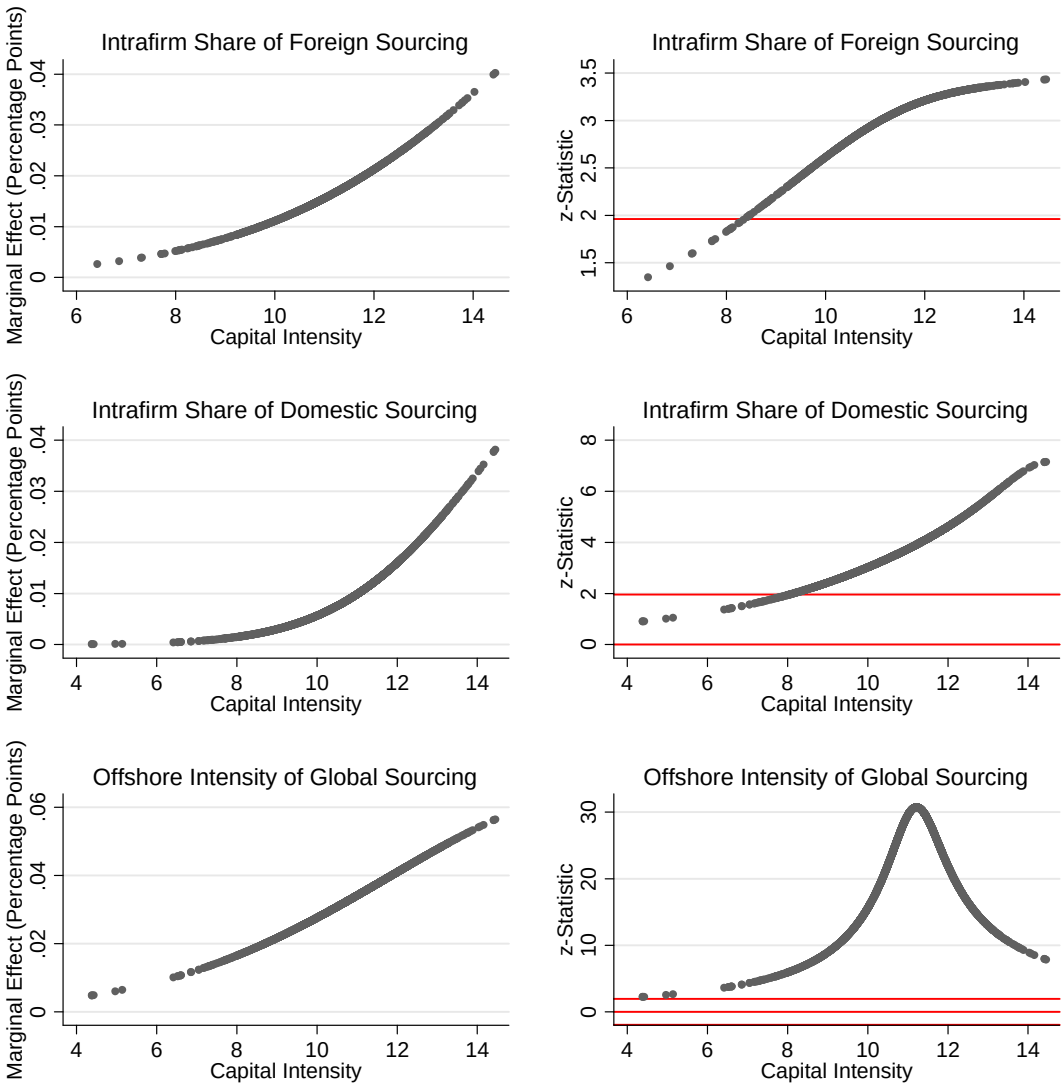


**Figure A1:** Marginal Effects of Productivity and z-Statistics After Pooled Tobit; $PROD_{OP}$ held at Minimum.

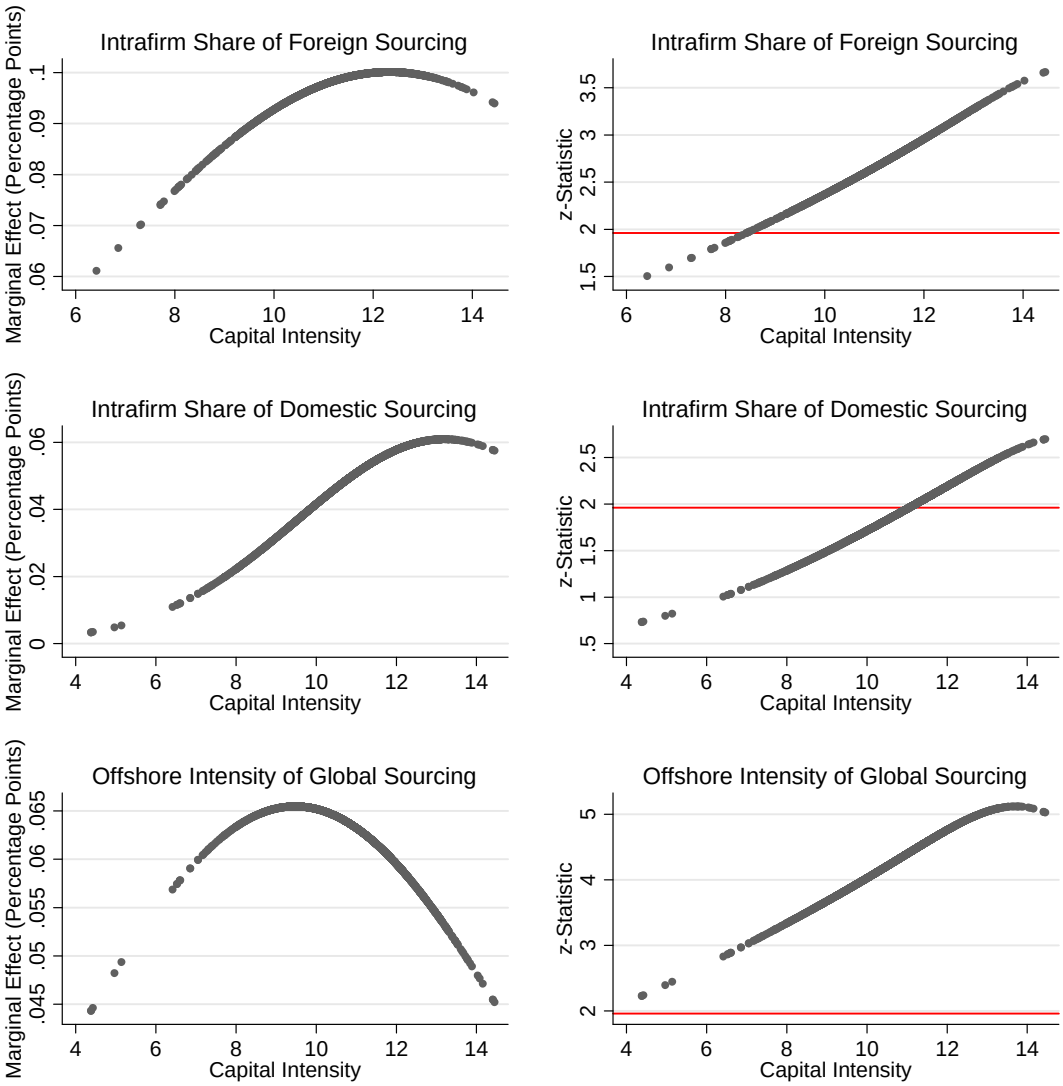


**Figure A2:** Marginal Effects of Productivity and z-Statistics After Pooled Tobit; $PROD_{OP}$ held at Maximum.

**Table A1:** Size and Productivity Premia According to Different Sourcing Modes I (Unrestricted Sample)[1]

| | $PROD_{OP}$ | | | | $SIZE_S$ | | | |
|---|---|---|---|---|---|---|---|---|
| | Pooled OLS | | Random-Effects | | Pooled OLS | | Random-Effects | |
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *FI* | 0.383*** | 0.160*** | 0.300*** | 0.123*** | 2.244*** | 0.425*** | 0.224*** | 0.111*** |
| | (0.045) | (0.047) | (0.041) | (0.043) | (0.112) | (0.054) | (0.044) | (0.033) |
| *FO* | 0.192*** | 0.111*** | 0.143*** | 0.084*** | 0.935*** | 0.279*** | 0.097*** | 0.095*** |
| | (0.030) | (0.028) | (0.027) | (0.026) | (0.073) | (0.033) | (0.021) | (0.018) |
| *DI* | 0.277*** | 0.127** | 0.216*** | 0.099** | 1.513*** | 0.310*** | 0.141*** | 0.102*** |
| | (0.051) | (0.050) | (0.044) | (0.043) | (0.136) | (0.067) | (0.051) | (0.032) |
| *AGE* | -0.002** | -0.004*** | -0.001* | -0.004*** | 0.017*** | 0.002*** | 0.028*** | 0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) |
| *EXPORT* | 0.212*** | 0.091*** | 0.207*** | 0.092*** | 1.220*** | 0.263*** | 0.316*** | 0.160*** |
| | (0.028) | (0.029) | (0.026) | (0.026) | (0.072) | (0.036) | (0.045) | (0.030) |
| $SIZE_L$ | | 0.144*** | | 0.145*** | | 1.127*** | | 1.126*** |
| | | (0.011) | | (0.011) | | (0.013) | | (0.014) |
| Constant | 9.792*** | 9.324*** | 9.814*** | 9.331*** | 14.970*** | 11.310*** | 15.780*** | 11.450*** |
| | (0.073) | (0.081) | (0.074) | (0.083) | (0.159) | (0.105) | (0.210) | (0.115) |
| Industry Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| H0: $FI = FO$ | 0.000 | 0.267 | 0.000 | 0.344 | 0.000 | 0.003 | 0.003 | 0.631 |
| H0: $FI = DI$ | 0.074 | 0.573 | 0.113 | 0.652 | 0.000 | 0.125 | 0.116 | 0.813 |
| H0: $FO = DI$ | 0.100 | 0.741 | 0.095 | 0.711 | 0.000 | 0.644 | 0.310 | 0.822 |
| Observations | 3393 | 3393 | 3393 | 3393 | 3737 | 3737 | 3737 | 3737 |
| $R^2$ | 0.215 | 0.270 | | | 0.505 | 0.898 | | |

[1] The table gives estimation results obtained with the Bernard & Jensen (1999) methodology. Each column represents a separate regression where the dependent variable at the top of each column is a function of dummy variables for foreign integration ($FI$), foreign outsourcing ($FO$), domestic integration ($DI$), and a firm's age ($AGE$) and export status ($EXPORT$). In columns (2), (4), (6), and (8) we additionally control for firm size ($SIZE_L$). The sourcing modes are mutually exclusive. If a firm is active in two or more sourcing modes simultaneously, she is assigned to the category which is supposed to be associated with the highest fixed costs. The sample includes firms with domestic or foreign parent companies but excludes non-sourcing firms. The lower part of the table gives p-values of tests for equality of coefficients. Robust standard errors are clustered by firm and given in parentheses. The size indicators are in natural logs. *,**,*** indicate significance at the 10%, 5%, 1% levels, respectively.

**Table A2:** Size and Productivity Premia According to Different Sourcing Modes II (Unrestricted Sample)[1]

| | $PROD_{OP}$ | | | | $SIZE_S$ | | | |
|---|---|---|---|---|---|---|---|---|
| | Pooled OLS | | Random-Effects | | Pooled OLS | | Random-Effects | |
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *FI* | 0.349*** | 0.126** | 0.278*** | 0.104** | 2.144*** | 0.371*** | 0.225*** | 0.090** |
| | (0.056) | (0.056) | (0.049) | (0.049) | (0.161) | (0.068) | (0.052) | (0.038) |
| *FO* | 0.161*** | 0.078* | 0.123*** | 0.063* | 0.861*** | 0.235*** | 0.104*** | 0.074*** |
| | (0.045) | (0.043) | (0.038) | (0.036) | (0.136) | (0.054) | (0.038) | (0.027) |
| *DI* | 0.247*** | 0.095 | 0.191*** | 0.075 | 1.440*** | 0.266*** | 0.135** | 0.082** |
| | (0.061) | (0.058) | (0.051) | (0.049) | (0.178) | (0.080) | (0.054) | (0.036) |
| *DO* | -0.025 | -0.029 | -0.015 | -0.018 | -0.053 | -0.036 | 0.005 | -0.012 |
| | (0.043) | (0.040) | (0.034) | (0.033) | (0.128) | (0.051) | (0.032) | (0.024) |
| *AGE* | -0.002** | -0.004*** | -0.001* | -0.004*** | 0.018*** | 0.002** | 0.028*** | 0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) |
| *EXPORT* | 0.224*** | 0.102*** | 0.210*** | 0.096*** | 1.244*** | 0.275*** | 0.297*** | 0.154*** |
| | (0.027) | (0.028) | (0.025) | (0.026) | (0.070) | (0.035) | (0.042) | (0.028) |
| $SIZE_L$ | | 0.144*** | | 0.146*** | | 1.131*** | | 1.130*** |
| | | (0.011) | | (0.011) | | (0.013) | | (0.013) |
| Constant | 9.814*** | 9.353*** | 9.832*** | 9.350*** | 14.960*** | 11.310*** | 15.750*** | 11.430*** |
| | (0.081) | (0.086) | (0.079) | (0.086) | (0.195) | (0.112) | (0.208) | (0.116) |
| Industry Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| H0: $FI = FO$ | 0.000 | 0.270 | 0.000 | 0.323 | 0.000 | 0.005 | 0.004 | 0.633 |
| H0: $FI = DI$ | 0.087 | 0.597 | 0.096 | 0.578 | 0.000 | 0.159 | 0.078 | 0.826 |
| H0: $FI = DO$ | 0.000 | 0.001 | 0.000 | 0.003 | 0.000 | 0.000 | 0.000 | 0.002 |
| H0: $FO = DI$ | 0.096 | 0.721 | 0.113 | 0.773 | 0.000 | 0.642 | 0.458 | 0.802 |
| H0: $FO = DO$ | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| H0: $DI = DO$ | 0.000 | 0.012 | 0.000 | 0.026 | 0.000 | 0.000 | 0.007 | 0.002 |
| Observations | 3579 | 3579 | 3579 | 3579 | 3937 | 3937 | 3937 | 3937 |
| $R^2$ | 0.217 | 0.273 | | | 0.507 | 0.900 | | |

[1] The table gives estimation results obtained with the Bernard & Jensen (1999) methodology. Each column represents a separate regression where the dependent variable at the top of each column is a function of dummy variables for foreign integration ($FI$), foreign outsourcing ($FO$), domestic integration ($DI$), domestic outsourcing ($DO$), and a firm's age ($AGE$) and export status ($EXPORT$). In columns (2), (4), (6), and (8) we additionally control for firm size ($SIZE_L$). The sourcing modes are mutually exclusive. If a firm is active in two or more sourcing modes simultaneously, she is assigned to the category which is supposed to be associated with the highest fixed costs. The sample includes firms with domestic or foreign parent companies as well as non-sourcing firms. The lower part of the table gives p-values of tests for equality of coefficients. Robust standard errors are clustered by firm and given in parentheses. The size indicators are in natural logs. *,**,*** indicate significance at the 10%, 5%, 1% levels, respectively.