Global Sourcing: Towards an Empirical Test of the Hold-up Model*

Wilhelm Kohler†
University of Tübingen
CESifo and GEP

Marcel Smolka‡
University of Tübingen

February 2013

Abstract

We use Spanish firm-level data to test the hold-up model of global sourcing proposed by Antràs & Helpman (2004). We propose a novel representation of the model which guides us in bringing the theory to the data. We estimate a discrete choice model of firms' sourcing behavior, separately for the location choice and the ownership choice of sourcing. We find that a firm's productivity interacts with an industry's headquarter intensity in governing both dimensions of sourcing in the way suggested by the Antràs & Helpman (2004) model. In addition, estimated "ex ante sourcing premia" of firms lend brought support to the idea that firms *self-select* into sourcing activities.

JEL-Classification: F14, F23, L22, L23
Keywords: global sourcing, incomplete contracts, firm productivity, firm-level data.

---

*We would like to thank Pol Antràs, Fabrice Defever, Jörn Kleinert, Hong Ma and Jagadeesh Sivadasan for helpful discussions. Christian Glebe and Marc-Manuel Sindlinger have provided excellent research assistance. Financial support from the Volkswagen Foundation under the project "Europe's Global Linkages and the Impact of the Financial Crisis: Policies for Sustainable Trade, Capital Flows, and Migration" is gratefully acknowledged.

†University of Tübingen, Nauklerstrasse 47, 72074 Tübingen, Germany
Phone: +49 (0) 7071 2976016, wilhelm.kohler@uni-tuebingen.de

‡University of Tübingen, Nauklerstrasse 47, 72074 Tübingen, Germany
Phone: +49 (0) 7071 2978183, marcel.smolka@uni-tuebingen.de

1 Introduction

In this paper, we propose a new empirical strategy for testing the canonical model of global sourcing introduced by Antràs & Helpman (2004). This model explains the *organizational form* of sourcing (vertical integration vs. outsourcing) as an attempt to minimize the inefficiency that derives from a hold-up problem in the provision of inputs, while the *location* of input provision (domestic vs. foreign) is explained by conventional cost considerations. A central prediction of the model is that firms select themselves into different sourcing strategies, based on their productivity levels and on how these strategies compare in terms of fixed cost. However, with two dimensions of sourcing there are 4 different strategies and the detailed selection pattern predicted by the model is highly sensitive to the assumed ranking of these strategies in terms of fixed cost. This poses a formidable challenge for empirical testing. Indeed, the model still waits for a proper empirical test, mainly because these fixed costs are essentially unobservable. In this paper, we develop a modeling strategy that does away with any need to observe fixed cost rankings, thus facilitating such a test. We apply our strategy to a data set on Spanish firms, finding strong support of the Antràs-Helpman (AH) model.

Well over half of world trade is trade in intermediate products. In the year 2000, the share of intermediates in non-energy inputs for manufacturing production in 12 main OECD countries was well above 25 percent. In the year 2005, on average across OECD and emerging market economies, as much as a quarter of the manufacturing export value reflects the cost of imported intermediate inputs, up from 20 percent in 1995. This indicates that firms are sourcing their inputs on a global scale. Firms also use different organizational forms for their sourcing. During the past decade , the share of related party trade in US manufacturing imports has remained stable at around 50 percent. For for services it has increased from around 19 in 2002 to 25 percent in 2009.[1]

Modern trade theory emphasizes that buyers and sellers of intermediate inputs enter a contractual relationship which is fundamentally different from a transaction in final goods. Input specifications often do not lend themselves to third-party verification, which rules out enforceable contracts. At the same time, product differentiation implies that such inputs have little, if any, use outside the specific production relationship. As a result, trade in intermediate inputs will often be plagued by a hold-up problem. The canonical model of global sourcing under such hold-up problems is due to Antràs (2003) who draws on the property rights theory of the firm as developed by Grossman & Hart (1986).[2] In this framework, differentiated final goods are produced using two types of specialized inputs, a headquarter service and a material input, or manufacturing component. Both inputs are essential for production. The headquarter service stands for all inputs that the headquarter of a firm is able to provide for itself, whereas for the material input the headquarter must draw on a separate input supplier. The model assumes that both inputs are not contractible and have no use outside the production relationship for a specific final good. This gives rise to a hold-up problem: The prospect of ex post Nash bargaining about the surplus derived from the production relationship reduces both agents' incentives to invest into input provision. As a result, the potential gains from this production relationship will not be fully exploited. Against this backdrop, the headquarter decides about the organizational form of input provision as well as the location of material input production. Specifically, by choosing vertical integration, it may secure a residual property right over the supplier's input. This enhances the headquarter's position in ex post bargaining, relative to an arm's length relationship (outsourcing) with the input supplier, but at the same time it reduces the incentive for

[1]See Miroudot et al. (2009) and Lanz & Miroudot (2011).

[2]The model is enriched by Melitz (2003)-style firm heterogeneity in Antràs & Helpman (2004) and extended to a variable degree of contractual incompleteness in Antràs & Helpman (2008). See Antràs (2012) for a survey.

the input supplier to invest into the production relationship. Vertical integration is a superior form of input provision only if the former effect dominates the latter.

How well does this model describe the observed global sourcing behavior of firms? An important prediction of the model is that firms will select themselves into different organizational forms and locations of sourcing, depending on their levels of productivity. Importantly, however, the detailed pattern of this sorting depends on the significance of the headquarter service for the production relationship in question. In Antràs & Helpman (2004) this is measured by the elasticity of the final output with respect to headquarter service, also called the headquarter intensity. A central tenet of the AH model is that, other things equal, the incentive for the headquarter to choose vertical integration is the larger, the larger the headquarter intensity of the production relationship. In the refined version of the model developed in Antràs & Helpman (2008), this same mechanism implies, somewhat counter-intuitively, that the outsourcing incentive becomes more powerful, if the range of material inputs that defy contractibility increases.

There is a nascent literature that tries to bring the AH model to the data. Papers using industry-level data and reporting evidence in support of the model are Yeaple (2006), Nunn & Trefler (2008), and Bernard et al. (2010). Corcos et al. (2012) and Defever & Toubal (2011) investigate determinants of the choice between vertical FDI and offshore outsourcing, using data for a cross-section of French firms. Federico (2012) reports evidence on the global sourcing activities of Italian firms. In Kohler & Smolka (2011) we draw on Spanish firm level data to provide evidence that supports the general notion of productivity-based sorting into different locations and organizational forms of sourcing. In Kohler & Smolka (2012) we use non-parametric techniques and regression analysis inspired by the AH-model in order to describe the relationship between sourcing behavior and firm-level productivity.

None of these studies, however, provides a rigorous empirical test of the AH model. Such a test is made difficult by the model's complex and casuistic set of predictions about firms' sourcing behavior. Each element of the 2×2 "global sourcing matrix" is characterized, not just by a distinct marginal cost for the material input, but also by a distinct fixed cost of operating the respective sourcing regime. Existing formulations of the AH model therefore suggest that a test of the model requires independent observations on both the location advantage for material input production and the ranking of different sourcing modes in terms of the associated fixed cost. Barring any reliable information on this fixed cost ranking, existing empirical studies therefore fall short of a proper test of the AH model.

In this paper, we propose a reformulation of the AH model that lends itself to an empirical strategy for testing the model which does away with the need to take any stance on, or observe, the fixed cost ranking of alternative sourcing regimes. Our reformulation frames the interaction between the productivity of a firm and the headquarter intensity of a production relationship in terms of *supermodularity* and *submodularity*, respectively, of the firm's maximum profit function (Mrázowá & Neary, 2011). In very simple terms, supermodularity implies that any advantage of one sourcing strategy over some other is reinforced by a higher productivity level of the firm, and conversely for submodularity. Analogously, supermodularity (or submodularity) may also obtain with respect to the headquarter intensity of the production relationship. Applying these concepts to the organizational form and the location of input sourcing, we derive propositions on how the productivity level interacts with the headquarter intensity of a firm in determining the selection of firms into different sourcing strategies.

Based on these propositions, our empirical strategy then comes in two steps. First, we estimate so-called "ex ante mode premia" on vertical integration. Such an "ex ante premium" emerges if, in a cross section of firms starting out with a non-vertical-integration strategy, superior first period performance (e.g., high levels of productivity) explains the *subsequent switch* to a vertical integration strategy. We estimate such premia separately for domestic and foreign sourcing. This is a theory-consistent way of exploiting the time variation in our panel data set, provided that there is some degree of sluggishness firms' choice of sourcing strategies.[3] Arguably, it is a better way of identifying selection patterns than contemporaneous correlations.[4]

We find significant "ex ante mode premia", which indicates that productivity-based self selection of firms into different sourcing strategies does take place. However, this does not amount to a test of the model. The model predicts that sorting is driven, not just by firms' productivity levels, but also by the extent to which production requires non-contractible inputs (the inverse of the headquarter intensity). Moreover, it predicts that the direction of sorting is importantly influenced by the fixed cost ranking of the sourcing strategies. The second step of our empirical analysis turns to a full test of the model by explicitly addressing these concerns. Based on our reformulation of the AH model in section 2, we estimate linear probability models for the choice of vertical integration (as opposed to outsourcing) and for foreign (as opposed to domestic) sourcing. Conditional on the location of sourcing (domestic or offshore), the AH model predicts a higher likelihood of vertical integration with a higher productivity level of the firm. More importantly, this relationship is reinforced with an increasing headquarter intensity. This latter effect is known as the Antràs (2003) effect in the literature. We can show that in our linear probability model both of these effects obtain, irrespective of any fixed cost ranking of the sourcing strategies involved. In our estimation, the effect of this ranking is absorbed by an industry-specific fixed effect.

In a similar vein, we estimate a linear probability model for the location of sourcing, conditional on the organizational form (vertical integration of outsourcing). Our theoretical results in section 2 imply that, given an assumed cost advantage for foreign sourcing, the likelihood of a firm selecting to source offshore increases with its productivity level. Moreover, it suggests that this relationship is reinforced by a low headquarter intensity of production. To the best of our knowledge, this second effect has so far gone unnoticed in the literature. Again, we can show that both of these effects hold for any fixed cost ranking, which is again absorbed by an industry-specific fixed effect. A crucial, but plausible assumption of our procedure is that the headquarter intensity is an industry-specific variable.

We implement this strategy using Spanish firm level data. Our data set has several noteworthy features. First, it allows us to address both of the dimensions of sourcing highlighted by the AH model (organizational form and location) in a unified setting. Secondly, it allows us to focus on cases where the unit of observation indeed decides about the sourcing strategy, rather than itself being the object of a different party's sourcing strategy. Third, and most importantly for the present context, the data feature a panel structure that allows us to address selection in a novel, dynamic way. Using this data set to estimate our linear probability model, and performing statistical tests on the coefficients obtained in line with the aforementioned hypotheses, we cannot reject the AH model.

The remainder of this paper is structured as follows. Section 2 presents our novel representation of the AH model and derives the key propositions on sourcing strategies. In section 3, we briefly introduce the Spanish firm-level survey data source and present a set of informative descriptive statistics on the

[3]Several empirical studies using firm-level data find evidence for self-selection (e.g. Fariñas et al. (2010); Wagner (2011)), but none of them looks at all four sourcing modes at the same time. Perhaps more importantly, none of them looks at selection in the true sense of the word, i.e., as firms *switching between modes of sourcing over time*. Federico (2010) does have information on all four sourcing modes, but his data are available for a single cross-section only.

[4]Compared to aspects of global sourcing, the relationship between export activities and firm performance is extensively researched. This literature has gained momentum ever since the pioneering contribution by Bernard & Jensen (1995) who document that exporting firms are larger, pay higher wages, and are more productive than their domestic peers in the same industry. See Bernard et al. (2012) for a recent summary of key differences between trading and non-trading firms and how these can be reconciled with theory.

pattern of sourcing observed in this data set. Section 4 investigates the self-selection hypothesis by estimating ex ante sourcing premia. Section 5 derives a probabilistic choice model of sourcing, and it presents the results from a linear probability estimation and a statistical test of the associated AH parameter hypotheses. The final section concludes the paper.

2 The hold-up model of global sourcing

In this section, we reformulate the hold-up model of global sourcing as developed by Antràs & Helpman (2004), henceforth AH model, in order to derive propositions on how heterogeneous firms select themselves into different sourcing strategies. A sourcing strategy is defined as a unique combination of location and organizational form of obtaining necessary inputs. Our reformulation uses the concept of supermodularity as proposed by Mrázowá & Neary (2011), and it aims at deriving propositions that may be tested with available data.

The underlying assumptions of the AH model are indicated in the introduction. A set of firms produce differentiated final goods using two customized intermediate inputs, both of which are essential. One of the two inputs is produced domestically by the headquarter of the firm (called the "headquarter service"), while the other (called the "manufacturing component") needs to be provided by a separate input supplier which may be located in the domestic or the foreign economy. The model aims to explain the organizational form (ownership structure) of input provision as well as the location of the input supplier that the headquarter of each firm chooses for its differentiated final good. We shall use indices $h = d, f$ and $j = v, o$ to denote the two available locations (domestic and foreign) and ownership structures (vertical integration and outsourcing) of input provision. The *location* of manufacturing component production determines the cost of this component. It proves useful to introduce ℓ as the inverse of this cost, with a value equal to ℓ_d if the input is produced domestically, and ℓ_f if it is produced in the foreign economy. Following Antràs & Helpman (2004), and without loss of generality, we assume $\ell_d < \ell_f$, meaning a foreign cost advantage for the manufacturing component.

The *organizational form* of input provision determines the bargaining position in the final stage of the game when the two parties must bargain over revenue sharing. Such bargaining is necessary, because both inputs are essential and completely specific to the production relationship, and because enforceable contracts cannot be written, for reasons indicated in the introduction. The ex post revenue share accruing to the headquarter is denoted by m, so that $1 - m$ accrues to the input supplier. Importantly, m determines the incentive that the two agents face when investing in input provision in the second stage of the game. Invoking property rights theory as pioneered by Grossman & Hart (1986), the AH model distinguishes between two organizational forms: *vertical integration* (v), where the headquarter obtains ownership rights in the manufacturing component, and *outsourcing* (o) which involves no such ownership. Ownership secures an improved outside option for the headquarter when bargaining in the final stage, whence $m_v > m_o$.[5] Each combination of location and organizational form (ownership structure) of sourcing requires its own (industry-specific) distinct fixed cost F_{hj}.

[5]Antràs & Helpman (2004) assume that with outsourcing both parties have a zero ex post outside option at the final stage of bargaining, but this is immaterial; what matters is $m_v > m_o$. The AH model has been generalized to more complex setups. For instance, Antràs & Helpman (2008) allow for a given continuum of material inputs which falls into two ranges of contractible and non-contractible tasks, respectively. Schwarz & Suedekum (2012) similarly allow for a continuum of intermediate inputs, modeled on an interval of variable length which is endogenously determined by the final goods producer. Although the entire continuum of intermediates is assumed non-contractible, it falls into two subranges where the final goods producer relies on outsourcing and vertical integration, respectively, vis à vis multiple input providers. For the purpose of this paper, the simpler version of the model originally proposed by Antràs & Helpman (2004) proves most useful. See Antràs (2012) for a survey.

However, the model does not take a stance on relative fixed-cost advantages of different sourcing strategies.

In stage one of the game, the headquarter's choice of a sourcing strategy (h, j) maximizes the excess of its revenue share over the cost of the two necessary inputs, subject to a participation constraint for the input supplier, and net of the fixed cost F_{hj}. In order to nail down this choice, Antràs & Helpman (2004) make three further assumptions: i) A Cobb-Douglas technology for final goods production, ii) a uniform and constant perceived price elasticity of final goods demand, and iii) a zero ex ante outside option of the input supplier. We write $\Pi(\ell, m; \eta, \theta)$ for the headquarter's *maximum* operating profit, where η denotes the elasticity of final output with respect to the headquarter service, usually referred to as the *headquarter intensity*, and θ denotes the firm's total factor productivity, respectively. The headquarter's sourcing strategy is then dictated by

$$\max_{h,j}\{\Pi(\ell_h, m_j; \eta, \theta) - F_{hj}\}. \tag{1}$$

Antràs & Helpman (2004) show that under the above assumptions the maximum operating profit for the headquarter emerges as

$$\Pi(\ell_h, m_j; \eta, \theta) = Z(\ell_h, m_j; \eta)\theta^{\varepsilon-1}, \tag{2}$$

where $\varepsilon \equiv 1/(1-\alpha) > 1$ denotes the perceived price elasticity of demand for final goods (in absolute value), and where

$$Z(\ell_h, m_j; \eta) \;:=\; A[1 - \alpha m_j \eta - \alpha(1 - m_j)(1 - \eta)]C(\ell_h, m_j; \eta), \tag{3}$$

$$C(\ell_h, m_j; \eta) \;:=\; [m_j^{\eta}(\ell_h(1 - m_j))^{1-\eta}]^{\varepsilon-1}, \tag{4}$$

In these definitions, A captures the size of the market,[6] while the term $C^{1/(1-\varepsilon)}$ denotes the minimum unit cost function dual to the Cobb-Douglas technology for the final goods production. Notice, however, that in this term the costs of the two inputs are "inflated" by the hold-up problem. Thus, normalizing the unit cost of producing the headquarter service to 1, the entire headquarter cost enters the Cobb-Douglas minimum unit cost function in (4) with a value equal to $1/m_j$, and accordingly for the manufacturing component with a value of $(1/\ell_h)/(1 - m_j)$.[7]

Obviously, the decision rule in (1) requires a discrete comparison of the operating profit for $\ell_h = \ell_d, \ell_f$ and $m_j = m_o, m_v$. Our principal focus in this paper lies on the equilibrium sorting pattern across firms differing in productivity levels θ, and across industries differing in the headquarter intensity of production η. We follow Mrázowá & Neary (2011) in describing the sorting mechanism in terms of *supermodularity*. Supermodularity captures a complementary relationship between the effects of two different variables on the maximum profit level as given in (2).

Definition 1. *(a) A function $H(g, q)$ is supermodular with respect to g and q, if and only if the following relationship holds:*

$$\Delta_g H(q_1) > \Delta_g H(q_0), \text{ for } \textbf{any } q_1 > q_0, \tag{5}$$

*where $\Delta_g H(q) := H(g_1, q) - H(g_0, q)$ for **any** $g_1 > g_0$. The function $H(g, q)$ is submodular, if $\Delta_g H(q_1) < \Delta_g H(q_0)$ for any $g_1 > g_0$. (b) If H is twice differentiable, then supermodularity between g and q is equivalent to $H_{gq} > 0$.*

[6] More specifically, if B is a market-size parameter, the term A may be written as $B[(\varepsilon - 1)/\varepsilon]^{\varepsilon-1}$.

[7] Notice also that according (1) to the sourcing strategy is chosen so as to maximize the excess of revenue of the cost of *both* inputs. This result derives from the assumption that input suppliers have a zero ex ante outside option, so that the headquarter firm absorbs the supplier's ex post profit by means of a suitable participation fee, which reduces the value of the production relationship to the input supplier down to zero.

Obviously, if H is supermodular with respect to g and q, then $-H$ is submodular with respect to g and q.

In what follows, we apply this definition to $\Delta_\ell \Pi(m_j;\eta,\theta) := \Pi(\ell_f, m_j;\eta,\theta) - \Pi(\ell_d, m_j;\eta,\theta)$ which we shall call the *location advantage* of offshoring.[8] Remember that we have assumed $\ell_f > \ell_d$, which implies $\Delta_\ell \Pi(m_j;\eta,\theta) > 0$, i.e., a foreign cost advantage for the manufacturing component. In a similar vein, we refer to $\Delta_m \Pi(\ell_h;\eta,\theta) := \Pi(\ell_h, m_v;\eta,\theta) - \Pi(\ell_h, m_o;\eta,\theta)$ as the *incentive advantage* of vertical integration over outsourcing. As emphasized above, the AH model invokes the theory of property rights in assuming that $m_v > m_o$. However, a key feature of the model is that the sign of $\Delta_m \Pi(\ell_h;\eta,\theta)$ remains ambiguous. This is because the term $Z(\ell_h, m_j;\eta)$, while monotonic in ℓ, is non-monotonic in m. It will prove convenient to apply the above definitions of supermodularity also to the function Z as given in (3), introducing $\Delta_\ell Z(m_j;\eta) := Z(\ell_f, m_j;\eta) - Z(\ell_d, m_j;\eta)$ as well as $\Delta_m Z(\ell_h;\eta) := Z(\ell_h, m_v;\eta) - Z(\ell_h, m_o;\eta)$.

2.1 The location of sourcing

Checking for supermodularity of the functions Π and Z now helps us to pin down the possible selection patterns of firms that are generated by the model under different parameter constellations of η and θ. We first look at the location choice of sourcing.

Lemma 1. *(a) Given $l_f > l_d$, the function Π is supermodular with respect to the inverse unit cost of the manufacturing component ℓ and the firm's productivity θ. (b) For sufficiently high values of m_j, the function Z is submodular with respect to ℓ and the industry's headquarter intensity η.*

Proof. **(a)** Supermodularity requires that $\partial[\Delta_\ell \Pi(m_j;\eta,\theta)]/\partial\theta > 0$. We compute $\partial[\Delta_\ell \Pi(m_j;\eta,\theta)]/\partial\theta = \Delta_\ell Z(m_j;\eta)(\varepsilon-1)\theta^{\varepsilon-2}$. We have assumed $\ell_f > \ell_d$, which implies that $\Delta_\ell Z(m_j;\eta) > 0$. Recalling that $\varepsilon > 1$ completes the proof. **(b)** Submodularity of Z requires that $\partial[\Delta_\ell Z(m_j;\eta)]/\partial\eta < 0$. We rely on intuition for the proof. $\Delta_\ell Z(m_j;\eta)$ measures the profit advantage that derives from producing the manufacturing component abroad, given $\ell_f > \ell_d$ which means that the manufacturing component is cheaper to produce abroad than at home. This advantage is equal to $A[1-\alpha m_j\eta - \alpha(1-m_j)(1-\eta)]$ times the cost advantage in terms of the cost function $C(\cdot)$ introduced in (4) above. In percentage terms this cost advantage is equal to $1-\eta$ times the percentage difference in the manufacturing cost between home and foreign. Hence, given ℓ_h and ℓ_f, this is falling in η. This is the direct cost effect of a higher η. However, a change in η also has strategic implications that derive from the hold-up problem captured by the bracketed term in 3. The marginal derivative of that term with respect to η is $\alpha(1-2m_j)$, which is negative if $m_j > 1/2$. Given the aforementioned direct cost effect, this is certainly sufficient for submodularity as in part (b) of the lemma, but the critical lower bound of m_j is lower than 1/2. □

Part (a) of the lemma is straightforward. So is the direct cost behind part (b), which implies that a low value of η and thus a high elasticity of output with respect to the component input in effect leverages the cost advantage. But from the headquarter's perspective this cost advantage is watered down by a low ex post revenue share, hence the direct cost effect only dominates for a sufficiently high value of m_j. From the first part of Lemma 1, the assumed location advantage of offshoring is magnified by a higher productivity. We depict this magnification effect in figure 1, plotting the fixed cost disadvantage of offshoring, $\Delta_\ell F_j := F_{fj} - F_{dj}$, on the vertical axis and a measure of the firm's

[8] A negative value of $\Delta_\ell \Pi(m_j;\eta,\theta)$ implies a location *dis*advantage of offshoring, or, equivalently, a location advantage of domestic sourcing.

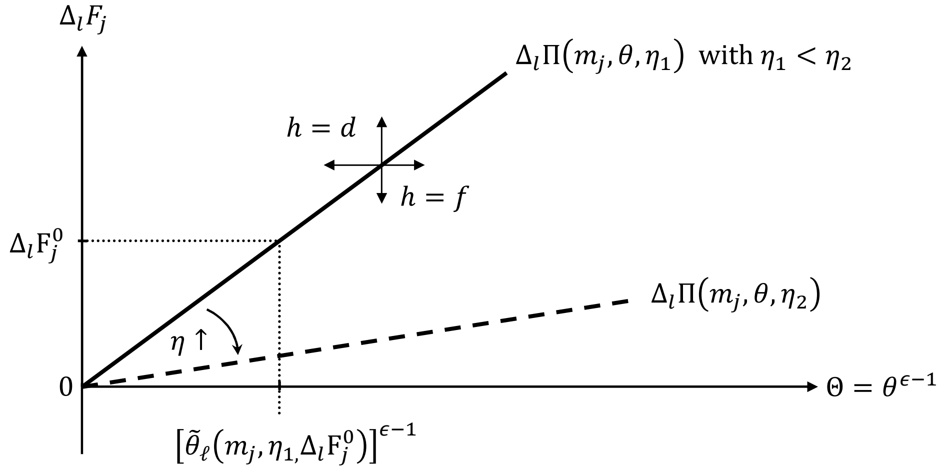

Figure 1. The location advantage of offshoring

productivity level, $\Theta := \theta^{\varepsilon-1}$, on the horizontal axis. The solid straight line is a demarcation line that separates value combinations for $\Delta_\ell F_j$ and Θ where the decision rule (1) leads the headquarter to choose offshoring (to the south-east) and domestic sourcing (to the north-west), respectively. The line holds for a specific headquarter intensity η_1. Given supermodularity of the maximum profit function with respect to ℓ and θ, this demarcation line is rising in Θ, with a slope given by $\Delta_\ell Z(m_j;\eta_1)$. Part (b) of the above lemma implies that, for a sufficiently low value of m_j this line rotates clockwise with an increase in the headquarter intensity. The reason for this is that the location choice applies to the manufacturing component provided by the supplier, not to the headquarter's service input. Hence, the foreign unit cost advantage is more "valuable" for low than for high values of the headquarter intensity.

What does this figure tell us about firms selecting themselves into different locations of sourcing. Obviously, no selection arises if $\Delta_\ell F_j < 0$, which implies a fixed cost advantage of offshoring. Given supermodularity as established in Lemma lemma:modularity-ell there is a unique industry-specific critical productivity level $\tilde{\theta}_\ell(m_j\eta_1, \Delta_\ell F_j)$ which solves $\Delta_\ell \Pi(m_j\eta_1, \theta) = \Delta_\ell F_j$. If positive, the critical value of θ serves as a selection criterion for the location choice with respect to a firm's productivity level. Firms with productivities above $\tilde{\theta}_\ell(m_j;\eta_1)$ choose offshoring as their optimal sourcing strategy, while firms with productivities below this threshold go for domestic sourcing. We summarize this sorting result in the following proposition.

Proposition 1. *(a) Any foreign location advantage for the manufacturing component is magnified by the firm's productivity level. (b) This magnification effect is the stronger, the lower the headquarter intensity of production. (c) Given that $l_f > l_d$, if there is selection with respect to the sourcing location, then firms with a productivity level above (below) a certain threshold level will choose offshoring (domestic sourcing). (d) This threshold level is increasing in the headquarter intensity of production. (e) Statements (a) through (d) apply separately for both types of ownership structure, vertical integration and outsourcing.*

Proof. Follows from Lemma 1. □

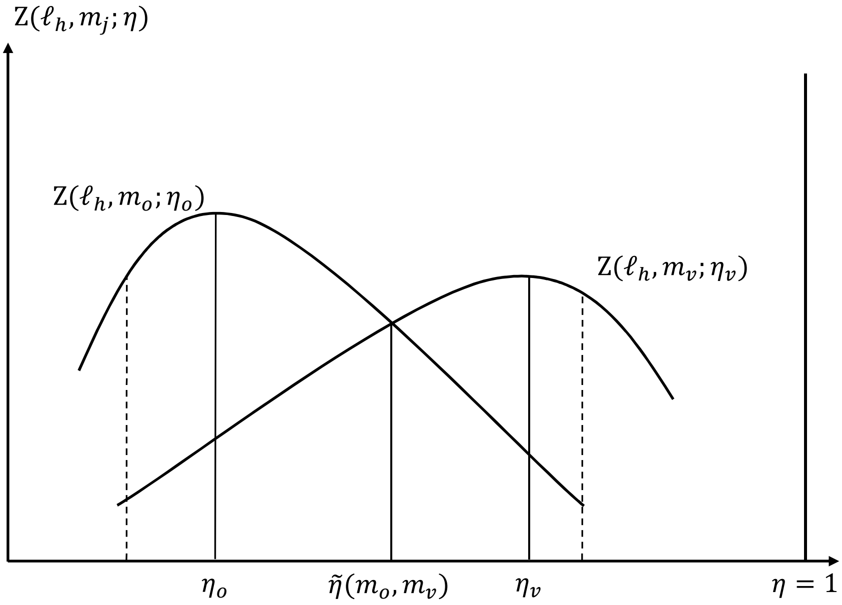

Figure 2. Threshold value of η according to lemma 2.

2.2 The organizational form of sourcing

The organizational form of sourcing determines the bargaining position of the headquarter, once the two inputs have been produced. Vertical integration affords the headquarter an enhanced outside option when having to negotiate about ex post revenue sharing. It therefore implies a higher ex post revenue share accruing to the headquarter, $m_v > m_o$.[9] Thus, vertical integration enhances the headquarter's incentive at the expense of the component supplier's incentive in stage two when the two agents invest in input production.

Intuitively, maximizing the revenue from the production relationship requires that the headquarter's (supplier's) incentive should be aligned with the importance of the headquarter service (the manufacturing component) in production of the final good, measured by the headquarter intensity η. Antràs & Helpman (2004) show that for any given value of η there exists a unique value $m^*(\eta)$ that maximizes the value of Z appearing in the maximum profit function given in (2) and defined in (3) above. This relationship features $m^{*\prime} > 0$, because a larger η makes incentivizing the headquarter more important, relative to incentivizing the input supplier.

However, m cannot be chosen at will. The institutional environment constrains the choice to two specific values for the ex post headquarter share of revenue, m_v and m_o, with $m_v > m_o$. This is of special relevance, because the function $Z(\ell_h, m; \eta)$ and thus the maximum profit function Π as defined in (2) is non-monotonic in m. This is implied by the relationship $m^*(\eta)$. The incentive structure of this complex situation may be characterized succinctly by the following lemma.

Lemma 2. *(a) Given the "ownership parameters"* (m_o, m_v)*, there is a unique threshold level* $\tilde{\eta}(m_o, m_v)$ *of the headquarter intensity of production, such that the maximum profit function* Π *as defined in (2) is supermodular with respect to* m *and the firm's productivity,* θ*, if* $\eta > \tilde{\eta}(m_o, m_v)$*, while submodu-*

[9]This invokes the theory of property rights as pioneered by Grossman & Hart (1986); for more details see Antràs & Helpman (2004).

larity obtains for $\eta < \tilde{\eta}(m_o, m_v)$*. (b) The function* Z *is supermodular with respect to* $m_v > m_o$ *and* η *for values of* η *such that* $\eta_{min} < \eta < \eta_{max}$*, where* η_{min} *is defined through* $\partial\Pi(\ell_h, m_o; \eta, \theta)/\partial\eta = \partial\Pi(\ell_h, m_v; \eta, \theta)/\partial\eta$ *for* $\eta < \tilde{\eta}(m_o, m_v)$ *and* η_{max} *is defined by this same equality for* $\eta > \tilde{\eta}(m_o, m_v)$*.*

Proof. The proof is illustrated figure 2 and by figure 3, which reproduces $m^*(\eta)$ from Antràs & Helpman (2004) and adds Z-lines for three distinct values of η.[10] We first proof that a unique value of $\tilde{\eta}(m_o, m_v)$ exists for any pair of "ownership parameters" (m_o, m_v). Define $\eta_o = m^{*-1}(m_o)$ and analogously for η_v. We have $\eta_o < \eta_v$, since $m_o < m_v$ and $m^*(\eta)$ is monotonically increasing. Then, by definition of $m^*(\eta)$ as maximizing profits Π, we have $\Pi(\ell_h, m_o; \eta_o, \theta) > \Pi(\ell_h, m_v; \eta_o, \theta)$ and $\Pi(\ell_h, m_v; \eta_v, \theta) > \Pi(\ell_h, m_o; \eta_v, \theta)$. Specifically, we have $\Pi(\ell_h, m_o; \eta, \theta) < \Pi(\ell_h, m_o; \eta_o, \theta)$ for all $\eta > \eta_o$ and $\Pi(\ell_h, m_v; \eta, \theta) < \Pi(\ell_h, m_v; \eta_v, \theta)$ for all $\eta < \eta_v$. From this it follows that there is a unique value $\tilde{\eta}(m_o, m_v)$ such that $\Pi[\ell_h, m_o; \tilde{\eta}(m_o, m_v), \theta] = \Pi[\ell_h, m_v; \tilde{\eta}(m_o, m_v), \theta]$; see figure 2. From figure 2 and the upper half of figure 3 it is clear that $\Delta_m Z(\ell_h; \eta) := Z(\ell_h, m_v; \eta) - Z(\ell_h, m_o; \eta)$ is increasing as we move from η_1 to $\eta_2 > \eta_1$. This is indicated by the dashed arrows. The same holds true for a change from any $\eta < \eta_2$, whereby we must note that $\Delta_m Z(\ell_h; \eta_2) = 0$. By the same token, $\Delta_m Z(\ell_h; \eta)$ is falling as we move from any $\eta > \eta_2$, such as η_3, down to $\eta_2 < \eta_3$. This is highlighted this by dotted arrows. It is obvious from these figures that for η values satisfying $\eta_o < \eta < \eta_v$ $\partial[\Delta_m Z(\ell_h; \eta)]/\partial\eta > 0$. The same holds true for values just marginally outside this interval, but a little reflection reveals that there exist values η_{min} and η_{max}, that define a larger range for which supermodularity of Z as in part **(b)** of the lemma. Part **(a)** of the lemma follows from the fact that the aforementioned threshold value $\tilde{\eta}(m_v, m_o)$ is unique. □

As with the location choice of sourcing, we now apply Lemma 2, in order to identify the selection pattern of firms with respect to the organizational form of sourcing. It results from an interplay between m and η. Figure 4 highlights possible selection patterns. The upward-sloping solid line identifies combinations of firm productivity Θ and the fixed cost disadvantage of vertical integration $\Delta_m F_h := F_{hv} - F_{ho}$ for which the headquarter firm is indifferent between vertical integration and outsourcing, given a headquarter intensity equal to $\eta_2 > \tilde{\eta}(m_o, m_v)$. This is the case of supermodularity of Π with respect to m and θ. Due to a relatively high headquarter intensity and $m_v > m_o$, vertical integration is the preferred incentive structure. The slope of this line, equal to $\Delta_m Z(\ell_h; \eta_2)$ measures the extent of the incentive advantage of vertical integration. However, if there is a fixed cost *dis*advantage of vertical integration, it will not be attractive for all firms. Vertical integration will in fact be chosen only for parameter combinations to the right and below the solid line in the top half of the figure. From the second part of Lemma 2, we know that the demarcation rotates clock-wise as the headquarter intensity of production η increases. It is a horizontal line for $\eta = \tilde{\eta}(m_o, m_v)$.

For each headquarter intensity, we may thus again identify a unique industry-specific critical productivity level $\tilde{\theta}_m(\ell_h, \eta, \Delta_m F_h)$ solving $\Delta_m \Pi(\ell_h, \eta, \theta) = \Delta_m F_h$. In headquarter-intensive industries with $\eta > \tilde{\eta}$, firms with productivities larger than $\tilde{\theta}_m(\ell_h, \eta, \Delta_m F_h)$ choose vertical integration as their optimal sourcing strategy, while firms with productivities below this threshold go for domestic sourcing. The opposite holds true for the case of component-intensive industries with $\eta < \tilde{\eta}$.

[10] Given that either input is essential, Z takes on a value of zero for the two extreme values of $m = 0$ and $m = 1$. The slope of any such Z-line approaches infinity as m converges to zero; and it converges to minus infinity as m converges to 1. This follows from the assumed Cobb-Douglas technology which implies that the marginal productivity of any input approaches infinity as the quantity of this input converges to zero (Inada condition). In the neighborhood of $m = 0$ the slope of the Z-line is lower for a higher value of η. The intuition is that the effect of disincentivizing the component supplier (through a higher m) weighs less heavily if the headquarter intensity is larger. The same reasoning applies, mutatis mutandis, for the slope of the Z line in the neighborhood of $m = 1$.

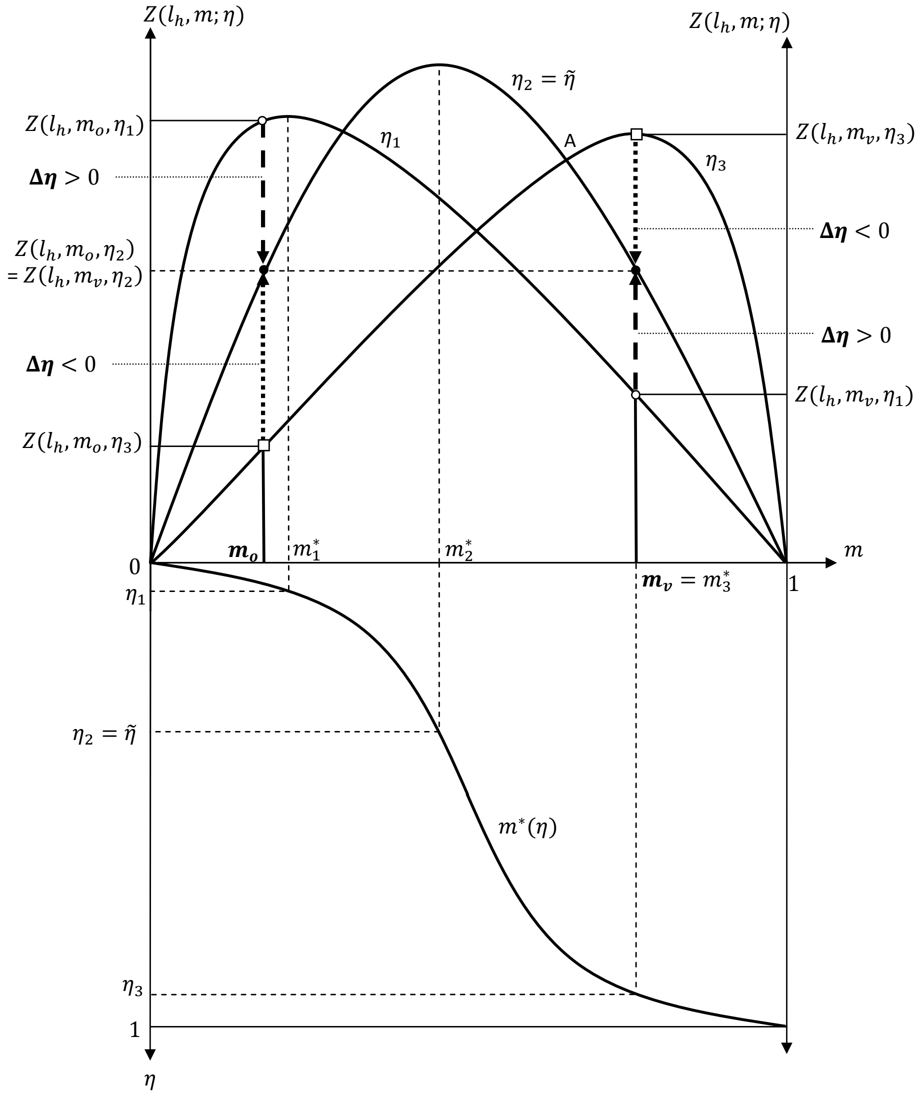

Figure 3. Incentive advantage of vertical integration and headquarter intensity

Proposition 2. *(a) For given ownership parameters* (m_o, m_v)*, if there is selection with respect to the ownership structure of sourcing, then there is a threshold value of the headquarter-intensity* $\tilde{\eta}(m_o, m_v)$*, such that in industries with* $\eta > \tilde{\eta}(m_o, m_v)$ *high-productivity firms choose vertical integration and low-productivity firms choose outsourcing. The opposite holds true with respect to selection in component-intensive sectors with* $\eta < \tilde{\eta}(m_o, m_v)$*. (b) The higher the headquarter intensity, the stronger the magnification effect of a firm's productivity level on the incentive advantage of vertical integration. (c) The opposite holds true for the magnification effect of a firm's productivity on the incentive advantage of outsourcing. (d) All statements apply separately to domestic and foreign sourcing.*

Proof. Follows from Lemma 2. □

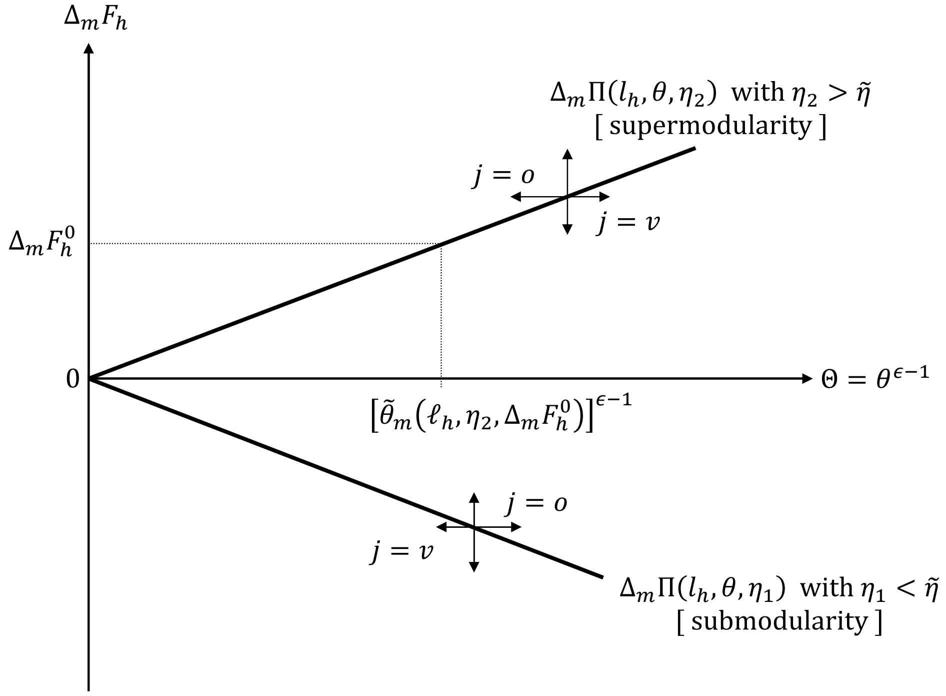

Figure 4. The incentive advantage of integration and outsourcing

3 Micro-level survey data on sourcing and productivity in the panel

In this section, we provide a short description of the firm-level survey data that we employ in our empirical analysis, and we briefly describe our performance measure at the level of the firm.[11] The data source is "Encuesta Sobre Estrategias Empresariales" (ESEE), an annual survey of roughly 2,000 manufacturing firms in Spain.[12] It is arranged by "Sociedad Estatal de Participaciones Industriales", a Madrid-based public foundation.[13] To date the ESEE covers a representative panel of Spanish manufacturing firms for the years 1990-2009. Its panel structure allows us to track manufacturing firms and industries in Spain over a relatively long time span. The selection of surveyed firms follows a two-way sampling scheme. Survey questionnaires are sent out to all firms employing more than 200 workers and to a subset of firms employing between 10 and 200 workers. Firms in this subset are randomly drawn from each of a total of 80 independent strata which are defined in advance for combinations of industries and size groups. The survey distinguishes 20 different industries, each comprising a number of products at NACE-2009 level, and four different size groups for firms employing between 10 and 200 individuals.[14] Given the sampling properties of the survey, we are able to emulate a random sample at the level of individual industries and the manufacturing sector at

[11]More detailed information on the data along with a stylized picture of sourcing heterogeneity among Spanish manufacturing firms can be found in Kohler & Smolka (2011).

[12]The English translation is "Survey on Business Strategies."

[13]Detailed information on the foundation's history and activities are available at `http://www.funep.es`.

[14]Table B.1 in the appendix gives the full list of manufacturing industries considered in the survey. Before the year 2009, firms have been classified into industries according to the older NACE-1993 classification. We have used relevant concordance information provided by the SEPI foundation where necessary to correct for this change in industrial classification.

large for each year by weighting each observation by the inverse of its probability of being sampled.[15]

A unique advantage of the Spanish firm-level data, compared to others used in the literature on firm selection, is that, starting in 2006, they allow us to compute the values of intermediate inputs acquired through each of four sourcing channels featured by the hold-up model described above. As will become evident below, these channels nicely fit into a firm's four-way sourcing problem of choosing among foreign integration (FI), foreign outsourcing (FO), domestic integration (DI), and domestic outsourcing (DO), as suggested by Antràs & Helpman (2004, 2008). The categorization of firms regarding their sourcing strategies relies on the following questions in the survey questionnaire.[16]

- *Of the total amount of purchases of goods and services that you incorporate (transform) in the production process, indicate – according to the type of supplier – the percentage which these represent in the total amount of purchases of your firm in* [year].
  - (a) *Spanish suppliers which belong to your group of companies or which participate in your firm's joint capital.* [yes/no] / [if yes, then percentage rate]
  - (b) *Other suppliers located in Spain.* [yes/no]/[if yes, then percentage rate]
- *For the year* [year]*, indicate whether you imported goods and services that you incorporate (transform) in the production process, and the percentage which theses imports – according to the type of supplier – represent in the total value of your imports.* [yes/no]
  - (a) *From suppliers which belong to your group of companies and/or from foreign firms which participate in your firm's joint capital.* [yes/no]/[if yes, then percentage rate]
  - (b) *From other foreign firms.* [yes/no]/[if yes, then percentage rate]

Answers to these questions can be combined with available information on total purchases and imports in order to compute for each firm in each of the years 2006, 2007, 2008, and 2009 the value of intermediate inputs acquired from a related party and, separately, from an outside supplier, both at home and abroad. Notice that, given the framing of the above questions, values of intermediate inputs reported to be obtained from related parties may refer to subsidiary firms, parental firms, or both. This distinction carries some relevance, especially because surveyed firms could be bound to confer the control over their sourcing decisions to a related party, provided this party has a controlling stake in the reporting firms' joint capital. Fortunately, we can control for the participation rate of other companies in the firm's joint capital, since that this information is also available in our data.

A pivotal variable in our regression analysis is a firm's relative productivity level in a given industry. It is well-known by now that firm-level estimates of total factor productivity (TFP) are often plagued by biases originating in endogenous selection into markets, simultaneous choice of input factors, omitted firms' input and output prices, and endogenous product mixes. As is standard in the literature, we apply the Olley & Pakes (1996) estimation algorithm in order to estimate sector-specific production functions. From these estimates we can recover each firm's total factor productivity level as a firm-specific, time-variant characteristic, computed as deviations from estimated industry-level productivities. The Olley & Pakes (1996) estimation algorithm takes care of the selection bias and the simultaneity bias.

---

[15]More information on the survey and its sampling properties are available in English from SEPI's website at `http://www.funep.es/esee/en/einfo_que_es.asp`.

[16]The survey questionnaire is distributed in Spanish and, separately for each survey year, downloadable at `http://www.funep.es/esee/sp/svariables/indice.asp`.

We employ the ESEE firm-level data from 2000-2009 in applying the Olley & Pakes algorithm, using annual information on each firm's real output, real investment, real capital stock, real purchases, labor employment, and exit decisions. *Real output* is the total production value plus other operating income, expressed in terms of prices of the year 2000. We deflate production values and other types of operating revenue by combining firm-level information on goods price variations from the ESEE with an industry-level price index from the INE for years with missing data. Using goods price variations at the level of individual firms is important to avoid estimation biases stemming from firm-specific mark-up pricing, firm-specific demand shocks, or firm-specific market access; see Klette & Griliches (1996) and De Loecker (2007). *Real investment* is the total investment value in real estate, constructions, and equipment, deflated with an industry-price index from the Spanish National Statistics Institute (INE). The *real capital stock* is the reported value of real estate, constructions, and equipment, net of depreciation, and deflated with an industry-price index from the INE. Similarly for *real purchases* which are defined as the total expenditures on intermediate inputs and external services. *Labor employment* is measured by effectively worked hours, which reduces the possibility of measurement bias, relative to standard measures used in the literature. *Exit decisions* of firms are also reported in the data, which is important in order to distinguish between firms shutting down production and firms staying in the market but exiting the ESEE panel.

4 Heterogeneity in firms' global sourcing activities

As we have argued in the introduction, little is known about the sourcing behavior of individual firms, especially when it comes to time-variant information in the panel. Innumerable papers have been published discussing performance gaps between exporting and non-exporting firms, some dealing with differences between offshoring and non-offshoring firms. This section follows a descriptive approach to generate new insights into the pervasive heterogeneity in firms' global sourcing activities. The panel dimension in our data set allows us to document recent trends in the global sourcing behavior of firms and investigate the prevalence of firms switching in and out of global sourcing modes over time. Furthermore, we show that certain types of sourcing strongly correlate with a number of important firm-specific variables such as sales, capital intensity, and productivity, to name just a few.

Table 1 reports the prevalence of certain sourcing activities in the Spanish data for two different years, 2006 and 2009, separately for two types of firms: small firms with at most 200 employees and large firms with more than 200 employees. We define sourcing activities in a mutually inclusive way, such that a single firm may count for more than one sourcing group due to multiple ways of sourcing. The first four rows from above distinguish among four sourcing activities: domestic sourcing, foreign sourcing, outsourcing, and integration. There is a fifth group of firms which we refer to as non-sourcing. These firms report zero volumes for input sourcing under the relevant question asked in the survey questionnaire. We find that in terms of location choice domestic sourcing activities are the most common in the data. More than 90% of firms source inputs domestically, independently of the year and the firm size group. When it comes to the ownership structure of input sourcing, the same statement applies to outsourcing activities. By contrast, foreign sourcing activities are reserved for a significantly lower share of firms. One third of small firms and two thirds of large firms pursue foreign sourcing. As to the choice of ownership structure of input production, less than one quarter of small firms use an integrated sourcing mode, while the corresponding number for large firms is about one half.

Another important finding in table 1 is that the use of all mentioned sourcing activities has seen an increase from the year 2006 to 2009. This increase applies to both small firms and large firms to a similar extent. For example, the share of firms going for an integrated sourcing activity has gone up by 10%-points for both groups of firm size. For foreign sourcing activities, the corresponding share of firms has risen by 4.9%-points in case of small firms and by 3.7%-points for large firms. The changes in the shares of domestically sourcing and outsourcing firms are less pronounced, but these are almost universally applied activities anyway. We interpret the trends in the numbers as an indication for the oft-cited growing fragmentation and diversification of production.

In table 1, we also look at each of the four combinations of the two sourcing dimensions (choice of sourcing location and choice of ownership structure), distinguishing among domestic outsourcing, domestic integration, foreign outsourcing, and foreign integration. We see that domestic outsourcing is by far the most widely used sourcing activity (> 90% of firms), with foreign outsourcing ranking second (one third of small firms and two thirds of large firms). Domestic integration is a less common phenomenon. Yet, its use has experienced the sharpest increase over the years among small firms, starting out from below 10% in 2006 and surpassing a level of 20% in 2009. Foreign integration is a very rare sourcing activity, at least for small firms. In fact, for both integrated sourcing modes (domestic and offshore), the fractions of large firms are significantly larger than those of small firms. The use of all four sourcing modes is subject to a clear upward trend. This is also reflected in the average number of sourcing modes used by firms. From 2006 to 2009, this number has risen from 1.3 to 1.5 for small firms and from 2.1 to 2.3 for large firms.

<<Table 1 about here>>

Table 2 looks at firms adjusting their sourcing strategy over time. We compute the numbers of switching moves for certain sourcing activities, relative to the maximum number of switching moves between 2006 and 2009. We do so separately for small and large firms, and for switching moves *into* a certain activity (columns (I) and (III)) and those *out* of a certain activity (columns (II) and (IV)). By analogy to table 1, the first four rows distinguish among domestic sourcing, foreign sourcing, outsourcing, and integration. This is followed by the combinations domestic outsourcing, domestic integration, foreign outsourcing, and foreign integration. We find that switching moves are, in general, neither pervasive nor irrelevant. Most of them are found for foreign sourcing activities, and in particular foreign outsourcing. For example, over the period considered, in close to 13% of all possibilities, small firms have changed their foreign outsourcing status from one period to another. More specifically, in 6.7% of all possibilities, this change has taken the form of switching into the foreign outsourcing activity. In 5.9% of all possibilities, it has been a move out of the foreign outsourcing activity. Slightly larger numbers are obtained for large firms. Fewer switching moves are found for integrated sourcing activities of small firms, and especially for those abroad. In 1.6% of all occasions, small firms report a move into the foreign integration activity or out of it. For large firms, the data report a significantly higher incidence of this type of changes in the sourcing status over time. The fraction of firms changing at least one of their sourcing activities once is relatively high, amounting to one third in case of small firms and one half in case of large firms.

<<Table 2 about here>>

Table 3 reports estimates of what we call *sourcing premia* with respect to various firm characteristics. With the term sourcing premia we refer to mean differences in a number of relevant firm-specific variables (denoted by Y), depending on *where* (domestic *versus* foreign economy) and *how* (integrated *versus* outsourced production) firms organize their intermediate input production. The numbers in table 3 are estimated coefficients of a firm-and-time-specific *sourcing group* identifier, SOURCING_{it}, in pooled regressions (2006-2009) of the following form:

$$\begin{aligned}\ln(Y_{it}) &= b_0 + b_1\text{SOURCING}_{it} + b_2\text{INDUSTRY} + b_3\text{YEAR} + b_4\text{INDUSTRY} \times \text{YEAR} \\ &\quad + e_{it} \end{aligned} \tag{6}$$

or

$$\begin{aligned}\ln(Y_{it}) &= b_0 + b_1\text{SOURCING}_{it} + b_2\text{INDUSTRY} + b_3\text{YEAR} + b_4\text{INDUSTRY} \times \text{YEAR} \\ &\quad + b_5\text{EXPORT}_{it} + e_{it}, \end{aligned} \tag{7}$$

depending on whether or not we condition on a firm's export status. The subscripts i and t index firms and years, respectively, INDUSTRY is a vector of industry dummies (see table B.1 in the appendix for a list of industries), YEAR is a vector of year dummies, EXPORT_{it} takes on the value one if firm i exports in year t and zero otherwise, and e_{it} is an independent error term with zero mean. The dummy variable SOURCING_{it} indicates membership of firm i at time t in one of two (pre-defined) sourcing groups. As we explain shortly, the definition of a sourcing group depends on the exact sourcing premium we are interested in.

Estimation of sourcing premia is somewhat more involved than estimation of the well-known exporter premia, for two reasons. First, we distinguish among four sourcing modes instead of two markets (foreign *versus* domestic). Second, firms engaging in multiple-mode sourcing are the rule rather than the exception; see Kohler & Smolka (2011) for details. Hence, we need to put some structure on the analysis. The left panel in table 3 looks at differences between outsourcing firms ($\text{SOURCING}_{it} = 0$) and firms opting for integration, *in addition* to outsourcing ($\text{SOURCING}_{it} = 1$).[17] It does so separately for foreign sourcing (upper-left corner) and domestic sourcing (lower-left corner)[18] and for small firms and large firms. The right panel proceeds analogously for firms restricting themselves to domestic sources ($\text{SOURCING}_{it} = 0$) and those sourcing inputs from abroad, *in addition* to sourcing domestically ($\text{SOURCING}_{it} = 1$).[19] We distinguish between intra-firm acquisition (upper-right corner) and arm's length acquisition (lower-right corner).

The estimated sourcing premia are informative as to the extent of firm heterogeneity in sourcing activities in relation to some key performance measures of firms (total sales, total employment, total equity, labor productivity, total factor productivity) as well as variables of factor intensity (capital intensity, skill intensity).[20] Given that we include a comprehensive set of industry dummies, year dummies, and industry-and-year dummies in the regressions, these premia are purged from any time-invariant as well as time-variant cross-industry differences. The general picture we obtain from this exercise is that virtually all firm characteristics employed vary systematically with a firm's sourcing group membership, and they do so in a pronounced way. Such variations follow certain patterns to be discussed below and cannot be explained by firms' exporting activities alone.

<<Table 3 about here>>

We start out by looking at differences between integrating and outsourcing firms; see columns (I) to (IV). The robustly significant positive sourcing premia indicate that firms which operate an

[17] Note that this definition excludes "pure" integration firms, which are very rarely observed in the data.

[18] Offshoring firms are excluded for the case of domestic sourcing.

[19] This definition excludes "pure" foreign sourcing firms, which are, again, rarely observed.

[20] For a definition of these variables, see table B.2 in the appendix.

integrated input production structure are larger (in terms of sales, employment, and equity) and more productive (in terms of labor productivity and total factor productivity) than those which do not. With few exceptions, these results hold true for both sourcing locations (abroad and at home), both groups of firm size (> 200 or ≤ 200 employees) and irrespective of a firm's export status. Similar differences are found with respect to firms' capital and skill intensity, but here the pattern appears to be less stable than that for the above-mentioned size and productivity measures.

We find particularly large differences in the size of integrated firms. Their sourcing premia in terms of total sales range between 80% and 100% for small firms and between 35% and 50% for large firms. In general, the estimated size premia of integration relative to outsourcing are larger for offshoring than for non-offshoring firms. Conditioning on export status, in turn, makes a surprisingly little difference. Robustly significant sourcing premia in the vicinity of 20%-25% are found for small firms' labor productivity and total factor productivity, while those for large firms are smaller in magnitude ($\approx 10\%$) in case of foreign sourcing or insignificant in case of domestic sourcing.

Next, we turn to the differences between firms sourcing inputs offshore and those sourcing inputs domestically; see columns (V) to (VIII). First and foremost, there is a significant size premium for offshoring firms relative to non-offshoring firms, whether the input production structure is integrated or outsourced. In general, this size premium is not homogeneous across the two firm size groups or the two ownership structures of input production, but it survives if we condition on a firm's export status. It is largest for small firms purchasing inputs via arm's length transactions, fluctuating between 60% and 120%, depending on the precise size measure. A further result we obtain from estimating sourcing premia is that firms outsourcing abroad are more skill intensive, more capital intensive, and more productive than their domestic peers, at least if we look at the group of small firms. For the group of firms sourcing intra-firm abroad, the picture is somewhat less clear but points into a similar direction.

The bottom line is that both dimensions of firms' sourcing decisions (choice of ownership structure and choice of sourcing location) are significantly correlated with important performance measures of firms. Two findings stand out. First, firms operating an integrated input production structure perform better than those which do not. This holds true for both sourcing locations. Second, firms obtaining inputs from offshore perform better than those which do not. This holds true for both ownership structures. Our integrated view on the two sourcing dimensions allows us to form expectations on the self-selection behavior of firms and the estimation of *ex-ante* sourcing premia, as we will argue below.

5 Self-selection into global sourcing modes

The correlations between sourcing strategies and firm performance found in the previous section suggest some selection behavior of firms. They can be thought of as being broadly consistent with theory, if we are willing to assume a positive incentive advantage of vertical integration (offshoring), coupled with higher fixed costs for vertical integration (offshoring) relative to outsourcing (domestic sourcing). However, the correlations do not lend themselves to ready conclusions about the direction of causality. This section takes the analysis a step further into this direction. More precisely, we ask whether the sourcing premia reported above derive, at least partly, from *ex-ante* performance differentials of firms making a choice for the one or the other sourcing mode *ex-post*. Identification comes from switching behavior of firms belonging to the same sourcing group initially but changing their sourcing group status over time.

Taking our theoretical considerations in section 2 and our estimation in section 4 into account, we expect sufficiently large changes in productivity to induce firms to change their sourcing group status over time. This section takes a closer look at the values of a variety of relevant firm characteristics and subsequent sourcing behavior of firms. We ask whether differences in initial firm characteristics can explain why some firms belonging to the group of outsourcing firms switch into the integration mode while others do not. We take a separate look at this question for each sourcing location (abroad and at home). Along the same lines, we then ask whether such differences can also explain the move from the group of domestically sourcing firms into the group of offshoring firms. Again, we do so separately for each ownership structure (integration and outsourcing). We call these initial differences in firm characteristics *ex-ante* sourcing premia. In case of significant *ex-ante* sourcing premia, we interpret the evidence as supporting the self-selection hypothesis. The Antràs & Helpman (2004) model suggests that firms endogenously decide upon their sourcing strategies, once their (exogenously given) productivities are revealed. This chain of reasoning requires firms to *self-select* into the one or the other location and organization of sourcing, which is a hypothesis that can be brought to our data. As we have argued in the introduction, the literature so far lacks an empirical investigation of the relevance of such self-selection behavior, at least as far as both the location and the organization dimension of sourcing are concerned.

We start out by looking at the integration-*versus*-outsourcing decision. Since it is generally not quite clear how fast firms can be expected to adjust their sourcing behavior over time, we test for ex-ante differentials in the *levels* of certain firm-specific attributes and, separately, in their *growth rates* in the year before a firm acquires inputs through intra-firm input acquisition *for the first time.* In the case of foreign sourcing, we thus restrict our estimation sample to offshoring firms which *(i)* never rely on foreign integration in the four consecutive years 2006-2009 or *(ii)* which *switch* to foreign integration over the period 2006-2009. In the case of domestic sourcing, the estimation sample excludes all offshoring firms and restricts the sample further to firms which *(i)* never rely on domestic integration in the four consecutive years 2006-2009 or *(ii)* which *switch* to domestic integration over the period 2006-2009. For our purposes, we set up two empirical models, the first of which reads as:

$$\ln(Y_{it}) = b_0 + b_1 \text{SOURCING}_{it+1} + b_2 \text{INDUSTRY} + b_3 \text{YEAR} + b_4 \text{INDUSTRY} \times \text{YEAR} + e_{it}, \quad (8)$$

where all variables except for the sourcing indicator are defined as in the previous section. Additional specifications of this model also include a firm's export status, EXPORT_{it}. The sourcing dummy variable is equal to one in case firm i switches to foreign integration at time $t+1$ for the first time, and zero otherwise. The estimated parameter β_1 can be used to compute the *ex-ante* sourcing premium (in percent) of integrating firms relative to non-integrating firms, one year ahead of switching from non-integration status to integration status. The second model is given by:

$$\ln\left(\frac{Y_{it+1}}{Y_{it}}\right) = b_0 + b_1 \text{SOURCING}_{it+1} + b_2 \text{INDUSTRY} + b_3 \text{YEAR} + b_4 \text{INDUSTRY} \times \text{YEAR} + e_{it}, \quad (9)$$

where other specifications again condition on a firm's export status, EXPORT_{it}. In light of the sourcing premia found in the previous section, we expect a positive estimated parameter β_1 in both models. We set up two completely analogous models for the estimation of *ex-ante* sourcing premia of offshoring firms relative to domestically sourcing firms, separately for the two ownership structures of production. The estimation samples are restricted accordingly. Again, we expect positive estimates of β_1, because offshoring firms have been revealed superior relative to non-offshoring firms by our estimates reported in table 3.

For the ownership decision of sourcing, the regression results for equations (8) and (9), estimated via ordinary least squares, are shown in columns (I) to (IV) of table 4. The upper panel gives the results for the foreign ownership choice, while the lower panel those for the domestic ownership structure. We find significant and positive *ex-ante* sourcing premia in the levels of various measures of firm size and firm productivity. Importantly, these are relevant for both sourcing locations and of about the same magnitude as the sourcing premia reported for small firms in table 3.[21] Conditioning on a firm's export status does not alter the conclusions either. This is clear-cut evidence in favor of the idea that firms acquiring inputs through an integrated sourcing mode for the first time are *ex-ante* significantly more productive and larger than firms which do not. This finding supports the notion of firms self-selecting into an integrated production structure, although the regression analysis does not permit the numbers to be interpreted in a strict causality sense. On the other hand, we see that the *ex-ante* growth rates of firm characteristics are not significantly correlated with ownership decisions of sourcing.

Columns (V) to (VIII) of table 4 report the estimation results for the location decision of sourcing. The upper panel gives estimated sourcing premia for offshoring firms acquiring inputs intra-firm relative to non-offshoring firms engaging in intra-firm input acquisition.[22] The lower panel does the same for arm's length input acquisition. In case of intra-firm sourcing, the results point towards a significant *ex-ante* size premium (in levels) of offshoring *versus* non-offshoring firms, but there is no significant *ex-ante* difference in productivity or other firm characteristics. As with the integration-*versus*-outsourcing decision, there are no *ex-ante* sourcing premia in the growth rates of size, productivity, or capital intensity, either. A more clear-cut picture emerges regarding the differences between offshoring and non-offshoring firms in the case of arm's length input sourcing. Here, we find robustly significant and positive *ex-ante* sourcing premia (in levels) in all firm dimensions, a firm's skill intensity being the only exception. These premia survive if we condition on a firm's export status. This can again be interpreted as striking evidence in favor of the self-selection hypothesis. Yet, there is again no robust evidence for significant differences in the growth rates of firm characteristics in the run-up to the switching decision.

<<Table 4 about here>>

6 An empirical model of global sourcing

In spite of using industry fixed effects, our estimation of (*ex-ante*) sourcing premia above treats the entire manufacturing sector in Spain as a single industry, ignoring cross-industry variation in a number of dimensions. Most importantly, the headquarter intensity of production, η, is deeply rooted in an industry's production technology and fundamentally shapes the industry equilibrium in the AH model. In fact, as we have argued in the theory section, this parameter lies at the heart of the AH model and can be used to empirically discriminate this model from other models of sourcing; see also Kohler & Smolka (2011). In what follows, we estimate a simple linear model of firms' sourcing behavior derived from theory, separately for the ownership choice and the location choice of intermediate input production. The model highlights the interaction between the productivity of the firm and the headquarter intensity of the production process in shaping both choices. First, a higher headquarter intensity increases a firm's probability of sourcing inputs through an integrated production structure, whether abroad or at home. This effect should be larger for firms with a higher productivity. At

[21]In the estimation of *ex-ante* sourcing premia, we consider all firms simultaneously and weight each observation by the inverse of the probability of being sampled.

[22]Here the sample excludes foreign outsourcing firms but includes domestic outsourcing firms due to the high incidence of this latter strategy in our data.

the sectoral level, and under additional assumptions on the distribution of productivity, it is often referred to as the Antràs-effect in the literature. Second, a higher headquarter intensity decreases a firm's probability of offshoring, whether through integration or outsourcing. Again, this negative effect should be larger for firms with a higher productivity. It is entirely novel to the literature and requires mode-specific data on both domestic and foreign sourcing.

6.1 The ownership choice of intermediate input production

Consider the subset of either offshoring or domestic firms. A firm i acquires its inputs through intra-firm sourcing at time t, $V_{it}^k = 1$, if maximum operating profits net of fixed organizational costs are larger under integration than under outsourcing, $\Pi(l_h, m_v; \eta, \theta) - F_{hv} > \Pi(l_h, m_o; \eta, \theta) - F_{ho}$. This can be expressed as:

$$V_{it}^h = \begin{cases} 1 & \text{if } \Delta_m Z(l_h; \eta)\Theta_{it} - \Delta_m F_h > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

The term $\Delta_m Z(l_h; \eta)$ captures the incentive advantage (or *dis*advantage) of integration. Suppose $\Delta_m Z(l_h; \eta) > 0$, meaning that there is an incentive advantage of integration which becomes magnified through a firm's productivity, Θ_{it}. In this case, we should observe *at least* some firms sourcing their inputs through an integrated production structure, acknowledging linearity in Θ_{it} and abstracting from prohibitively large fixed costs. All firms should go for vertical integration, if $\Delta_m F_h < 0$, i.e., if outsourcing entails higher fixed costs than integration. However, if $\Delta_m F_h > 0$, we should observe a productivity-dependent sorting pattern of firms according to which high-productivity firms select into the integration mode and low-productivity firms into outsourcing. Carefully speaking, this equilibrium structure is the one most visible in the data.

Given that the term $\Delta_m Z(l_h; \eta)$ is monotonically increasing in the headquarter intensity of production η, the incentive advantage of integration is an increasing function of η. Given the multiplicative relationship between $\Delta_m Z(l_h; \eta)$ and Θ_{it}, there is a positive interaction between the headquarter intensity of production and a firm's productivity level. We explore this interaction in a non-structural linear probability model (LPM)[23], controlling for the fixed cost difference between integration and outsourcing, $\Delta_m F_h$. In industries with a relatively high headquarter intensity ($\eta > \tilde{\eta}$), an increase in a firm's productivity should come with a higher probability of choosing vertical integration. In industries with a low headquarter intensity ($\eta' < \eta$), this effect should be smaller. For the knife-edge case, $\eta' = \tilde{\eta}$, it should be zero and for $\eta' < \tilde{\eta}$, it should be negative. The LPM takes the the following form:

$$V_{it}^h = \beta_0 \theta_{it-1} + \beta_1 \theta_{it-1} \times \eta_s + \gamma_{st} + \beta \mathbf{X}_{it-1} + \gamma_i + \mu_{it}, \tag{11}$$

where θ_{it-1} is firm i's productivity at time $t-1$, η_s is the (time-invariant) headquarter intensity of industry s, γ_{st} is an industry-and-time fixed effect, $\mathbf{X}_{it-1}$ is a vector of time-varying firm characteristics[24], γ_i is a firm fixed effect, μ_{it} is a random error term with zero mean, and the βs are parameters to estimate. The dependent variable, V_{it}^h, takes on the value one if firm i acquires inputs from an integrated supplier in period t, and zero otherwise. Firm productivity is estimated by the Olley & Pakes (1996) algorithm. The headquarter intensity of production is approximated by an industry's capital

[23]We revert to the simple LPM instead of more advanced non-linear estimation techniques for two reasons. First, if industry fixed effects are controlled for, the interaction effect between the firm's productivity and the industry's headquarter intensity is not identified in a Probit model. Second, the LPM allows us to straightforwardly control for firm fixed effects by within-transforming the data.

[24]The vector includes a firm's age, capital intensity, skill intensity, export status, and a four-valued ordered variable indicating the largest proportion of other firms' capital in the reporting firm's joint capital (0%,0-25%,25-50%,>50%).

intensity, computed as the weighted mean over all firms active in a particular industry. Similar proxy variables are widely used in the literature, given that the headquarter intensity is not directly observed (Antràs, 2003). The industry-and-time fixed effects capture all time-variant as well as time-invariant shocks affecting the ownership choice of sourcing, whether they are specific to industries or not. Most importantly, they control for the difference in fixed costs between integration and outsourcing, as well as the main effect of the headquarter intensity of production. The explanatory variables are given in one-year lags to avoid simultaneity bias.

In light of our empirical results obtained so far, we conjecture that most industries feature high headquarter intensities above the cut-off value $\tilde{\eta}$. Hence, we expect a positive productivity effect in regressions including productivity without the interaction term between productivity and the headquarter intensity. This expectation applies to both sourcing locations (foreign and domestic). The more interesting expectation concerns the interaction effect. A significantly negative or insignificant estimate of β_0 plus a significantly positive estimate of β_1 would be strong support for the AH model. It would imply that a high productivity favors integration if and only if the headquarter intensity is sufficiently high. This logic is turned upside down in industries with sufficiently low headquarter intensities. Again, we want to stress that a key advantage of our data relative to existing literature is that we can bring this proposition to the data independently for both sourcing locations.

We first estimate the model in equation (11) by pooling the data and applying OLS, assuming that the firm fixed effect, γ_i, is uncorrelated with the other covariates. We do so separately for the set of offshoring firms (ownership choice abroad) and the set of non-offshoring firms (ownership choice at home).[25] Standard errors are clustered at the level of the firm. As is well-known, under the given set of assumptions, the OLS estimator is asymptotically consistent but not efficient. Hence, we also estimate a random effects (RE) model which is, under the same set of assumptions, the efficient estimator. Both estimators, OLS and RE, use the between-variation in the data for parameter identification, i.e., they rely on differences in the ownership structure and productivity *across firms*. This cross-firm approach has the advantage that the variation in the variables of interest is large. However, it runs into risk of introducing a potential omitted variables bias due to unobserved heterogeneity. This would happen, if the firm fixed effect, γ_i, is not random but instead correlated with the other covariates. A common way of getting rid of the firm fixed effect is to within-transform the data and compute all variables relative to the firm-specific mean value over time. Since the firm fixed effect is constant over time, it is eliminated by the within-transformation. The so-called fixed effects (FE) estimator is asymptotically consistent, whether the firm fixed effect is correlated with the other covariates or not. If it is not, however, the RE estimator is more efficient than the FE estimator. Identification of the parameters β_0 and β_1 in the FE model comes from the within-variation in the data, i.e., from the variation of ownership structure and productivity *over time*. A problem in our application is that the short time span of our data (four minus one periods) implies relatively low variation in the variables of interest over time. Hence, the FE model may suffer from a lack of sufficient identifying variation. For the sake of completeness, however, we also report the results obtained from FE estimation.

Table 5 presents all estimation results of the model in equation (11). The upper panel gives the results for the ownership choice of intermediate input production in the foreign economy and the lower panel those for the ownership choice in the domestic economy. For each of the two panels, columns (I) to (VI) report the OLS estimates, columns (V) to (VIII) the RE estimates, and columns (IX) to (XII) the FE estimates. For each estimator, we first estimate the model without the interaction term and then with the interaction term included (each version with and without the vector of firm-level

[25]This structure is dictated by the prevalence of multi-mode sourcing strategies; see table 1. The set of offshoring firms includes firms sourcing domestically, but the set of non-offshoring firms does not include firms sourcing abroad.

controls).[26] As to the models without the interaction effect, we find a robustly significant and positive effect of firm productivity on the probability of sourcing through an integrated production structure. Our results apply to both sourcing locations and are qualitatively independent of the inclusion of firm-level controls. The only exception is the FE model for the domestic location, where the estimated coefficient is positive but insignificant.

<<Table 5 about here>>

In all specifications employed, the effect of productivity on the probability of sourcing inputs intra-firm is somewhat larger for offshoring firms than it is for domestic firms, ranging between 10%-points and 25%-points for a 1%-increase in total factor productivity. This is a quantitatively huge effect, which confirms our expectations derived from the robustly significant and positive sourcing premia obtained for integrated firms relative to outsourcing firms in previous steps of our analysis. It is also in line with our belief that a significant share of firms operate in headquarter-intensive industries. It is important to understand, though, that this positive productivity effect is independent of the relative fixed costs of the two organizational forms of sourcing. More precisely, it is perfectly consistent with a fixed cost advantage of integration as well as a fixed cost *dis*advantage of integration. The reason is that the fixed cost difference is an additive term in inequality (10). It is thus controlled for in the estimation through industry-and-time effects, even if this difference changes over time.

When it comes to the interaction effect between firm productivity and sectoral headquarter intensity, we find a surprisingly clear and consistent picture. Across both models and all specifications, estimates of the parameter β_0 (the main effect of productivity) have a negative sign and those of β_1 (the interaction effect) a positive sign. In most cases, we can reject the null hypothesis that the parameters are equal to zero. Thus, our estimates point towards a significant positive interaction effect between firm productivity and an industry's headquarter intensity, as suggested by theory.[27] For sufficiently headquarter-intensive industries, the effect of productivity on the probability of intra-firm sourcing is positive, and vice versa for sufficiently component-intensive industries. This Antràs-effect is economically significant and found independently for both domestic sourcing as well as foreign sourcing.

6.2 The location choice of intermediate input production

Consider now the subset of either integrating or outsourcing firms. A firm i acquires its inputs from abroad at time t, $A_{it}^{j} = 1$, if maximum operating profits net of fixed organizational costs are larger abroad than at home, $\Pi(l_f, m_j; \eta, \theta) - F_{fj} > \Pi(l_d, m_j; \eta, \theta) - F_{dj}$. This can be expressed as:

$$A_{it}^{j} = \begin{cases} 1 & \text{if } \Delta_l Z(m_j; \eta)\Theta_{it} - \Delta_l F_j > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

The logic behind the location choice of intermediate input production is similar to that of the ownership choice. If positive, the term $\Delta_l Z(m_j; \eta)$ measures the location advantage of offshoring. High-productivity firms get the most out of this advantage, due to the multiplicative relationship between

[26]To save on space, parameters of firm-level controls are not reported.

[27]One exception is the FE model for the foreign sourcing location, but the point estimates of the parameters in the FE model are relatively close to those in the RE model. We cannot apply the familiar Hausman test in order to provide a formal test whether or not the differences are significant, because we weight each firm observation by the inverse of the probability of being sampled.

$\Delta_l Z(m_j;\eta)$ and Θ_{it}. Hence, as with the previous model, we should observe *at least* some firms sourcing their inputs offshore, if the fixed cost of offshoring are not too high. If, moreover, $\Delta_l F_j < 0$, all firms should opt for the foreign source, given that there is a fixed cost advantage of offshoring, coupled with the location advantage. The high share of purely domestic firms in the data, however, reject this type of parameter constellation right away. If, in turn, $\Delta_l F_j > 0$, we should again observe a productivity-dependent sorting pattern of firms with high-productivity firms sourcing abroad and low-productivity firms sourcing domestically.

In what follows, we want to focus on the effect of the headquarter intensity η on the firm's location choice of sourcing. We know from above that the term $\Delta_l Z(m_j;\eta)$ is a decreasing function of the headquarter intensity of production. In words, this means that the location advantage of offshoring is less important in headquarter-intensive industries than it is in component-intensive industries. Headquarter-intensive industries rely more heavily on headquarter services than on inputs provided by a (domestic or foreign) supplier, relative to component-intensive industries. The location advantage of offshoring, however, is irrelevant for the provision of headquarter services, which, by definition, are provided domestically by the headquarter firm itself. Hence, the more headquarter-intensive the production technology, the less important is the variable cost advantage of offshoring for the production of the final good. This logic gives rise to a negative interaction between the headquarter intensity of production and a firm's productivity level, because the effect of (exogenous) changes in η on the location choice of input production is magnified by Θ_{it}. This interaction is explored in the following LPM:

$$A_{it}^j = \beta_0\theta_{it-1} + \beta_1\theta_{it-1}\times\eta_s + \gamma_{st} + \beta\mathbf{X}_{it-1} + \gamma_i + \mu_{it}, \tag{13}$$

where A_{it}^j is equal to one for offshore input acquisition of firm i at time t, and zero otherwise. All other variables are defined as in the previous subsection. We may again speculate to expect a positive productivity effect in regressions in which the interaction term between productivity and the headquarter intensity is not included. This expectation derives from the large and positive sourcing premia of offshoring firms relative to non-offshoring firms estimated above. It applies to both ownership structures of input acquisition (integration and outsourcing). The interaction term, on the other hand, is expected to feature a negative estimated coefficient, based on the above-mentioned argument. To the best of our knowledge, this prediction has so far not been confronted with the data, although it is an integral part of the full-fledged version of the AH model. Arguably, an empirical test of this prediction permits a formidable way of checking the inner consistency of the model to an extent that goes well beyond the existing empirical literature.

Table 6 reports the estimation results of the model in equation (13). The table is structured in a way which is completely analogous to that of the previous subsection. The upper panel gives the results for an integrated production structure, where all firms outsourcing abroad are excluded from the estimation sample.[28] The lower panel gives the results for an outsourced production structure, where all firms sourcing intra-firm are excluded from the estimation sample. Without the interaction term, the OLS and the RE model both report a positive and significant productivity effect on the probability of offshoring for both ownership structures. This is perfectly in line with our expectations. Maybe more importantly, it is consistent with the results reported in previous empirical papers on offshoring, even though these do not estimate the productivity effect separately for the two ownership structures, as we do here. In fact, we find that the productivity effect is way more pronounced for the subset of outsourcing firms, with point estimates for the coefficient of the productivity variable being as much as ten times as large as those for the subset of integrating firms. Certain doubts on the

[28] Restricting the sample to pure integration firms would reduce the sample size dramatically due to the large incidence of domestic outsourcing.

robustness of the positive productivity effect arise, however, because the fixed effects model reports all insignificant effects, irrespective of the precise specification. This holds true also for the model including the interaction term between firm productivity and the headquarter intensity of production. That said, we should stress here again that the FE model makes high demand on the data, and that one explanation for the poor performance of the FE model is the lack of sufficient variation in our data over time.

Our exposition of the AH model guides us to expect a differential effect of productivity on a firm's probability of offshoring, depending on the industry's headquarter intensity of production. This expectation is strongly confirmed by the estimation results from the OLS and the RE model. In all specifications employed, the main effect of productivity has a positive sign and the interaction effect a negative sign. In the full model with firm-level controls, these effects are significantly different from zero at reasonable levels of confidence. This pattern is independent of the ownership structure of intermediate input production and can be interpreted as further evidence in favor of the AH model.

<<Table 6 about here>>

7 Conclusion

We have argued that the existing empirical literature on global sourcing strategies by heterogeneous firms fails to provide a strict test of the hold-up model developed by Antràs & Helpman (2004). Drawing on recent work by Mrázowá & Neary (2011), we reframe this model in terms of the fundamental properties of supermodularity and submodularity, respectively, between the key characteristics of a sourcing strategy relating to the location and the ownership structure of input provision and a firm's productivity level. This sheds additional light on the fundamental mechanisms at work in driving patterns of selection among heterogeneous firms into different sourcing strategies.

We then turn to an empirical analysis drawing on a unique firm level data set for Spanish manufacturing industries. The Spanish data allows us to identify sourcing activities that correspond very closely to the theoretical model by Antràs & Helpman (2004). We first describe key stylized facts of this data set and then exploit its panel structure to identify true selection effects through econometric estimation of ex ante sourcing premia. This is the first novel contribution of this paper to the existing literature. Our findings corroborate the idea that firms *self-select* into global sourcing activities, a fundamental presumption in the hold-up model of global sourcing.

The second part of our empirical analysis then turns to a discrete choice model of sourcing. Based on our novel representation of the AH model, we translate the theory into a linear probability model explaining observed global sourcing activities by the interaction of the firm's productivity level with the industry's headquarter intensity. This part of our empirical analysis provides empirical support for the hold-up model of global sourcing which goes well beyond what has been obtained in the existing empirical literature.

Figures and tables

Table 1. Prevalence of sourcing activities and changes over time, 2006/2009

	Small firms			*Large firms*		
	2006	*2009*	*Δ 06/09*	*2006*	*2009*	*Δ 06/09*
	(I)	(II)	(III)	(IV)	(V)	(VI)
Domestic sourcing	91.2	93.5	+2.3	92.1	95.1	+3.0
Foreign sourcing	31.6	36.5	+4.9	66.4	70.1	+3.7
Outsourcing	91.5	94.4	+2.9	93.4	95.8	+2.4
Integration	11.0	22.8	+11.8	48.4	57.8	+9.4
Non-sourcing	7.5	5.0	-2.5	3.0	2.1	-0.9
Domestic outsourcing	90.2	92.8	+2.6	88.9	93.2	+4.3
Domestic integration	8.8	20.1	+11.3	35.0	40.0	+5.0
Foreign outsourcing	30.6	35.9	+5.3	62.0	66.3	+4.3
Foreign integration	3.2	4.4	+1.2	23.8	29.0	+5.2
Average number of sourcing modes	1.328	1.532	+0.204	2.096	2.284	+0.188

Columns (I), (II), (IV), and (V) give the percentages of firms in the various sourcing categories. All percentages are of the total number of firms in the respective size category. Columns (III) and (VI) give the percentage point changes from 2006 to 2009. A single firm may show up in more than one sourcing category due to multiple ways of sourcing. Large firms employ more than 200 workers. The last row gives the average number of sourcing modes per firm.

Table 2. Switching moves and fraction of switching firms, 2006-2009

	Switching moves			
	Small firms		*Large firms*	
	In	*Out*	*In*	*Out*
	(I)	(II)	(III)	(IV)
Domestic sourcing	3.0	2.8	2.8	2.1
Foreign sourcing	6.5	5.8	7.5	6.3
Outsourcing	3.1	2.5	2.3	1.6
Integration	1.9	1.7	6.4	4.1
Domestic outsourcing	3.4	3.0	4.0	2.9
Domestic integration	1.4	1.4	4.7	4.2
Foreign outsourcing	6.7	5.9	8.2	6.7
Foreign integration	0.8	0.8	5.6	3.4
Fraction of switching firms (2006-2009)	32.3		47.5	

This table gives for each sourcing category the fraction of sourcing decisions deviating from those in the previous period. All percentages are of the total number of possible changes in sourcing status over the period 2006-2009. "In" refers to a firm's decision to make use of a sourcing channel which it had not used in the previous year. "Out" refers to a firm's decision not to make use of a sourcing channel which had used in the previous year. Large firms employ more than 200 workers.
The last row gives the share of sampled firms changing at least one of their sourcing channels at least once in the period from 2006 to 2009.

Table 3. Firm heterogeneity in location and ownership decisions of input acquisition (pooled 2006-2009)

	Integration versus outsourcing					*Offshore versus domestic sourcing*			
	Unconditional		*Conditional on export status*			*Unconditional*		*Conditional on export status*	
	Small firms	*Large firms*	*Small firms*	*Large firms*		*Small firms*	*Large firms*	*Small firms*	*Large firms*
	(I)	(II)	(III)	(IV)		(V)	(VI)	(VII)	(VIII)
Abroad					*Intra-firm*				
Total sales	1.063	0.445	0.900	0.431	*Total sales*	0.492	0.434	0.248	0.409
Total employment	0.669	0.245	0.546	0.230	*Total employment*	0.335	0.318	0.203	0.287
Total equity	1.032	0.379	0.831	0.354	*Total equity*	0.603	0.642	0.372	0.572
Skill intensity	0.074	0.094	0.074	0.092	*Skill intensity*	0.106	0.095	0.099	0.104
Capital intensity	0.382	(0.050)	0.305	(0.038)	*Capital intensity*	(0.033)	(-0.067)	(-0.079)	(-0.081)
Labor productivity	0.201	0.114	0.172	0.112	*Labor productivity*	(0.012)	0.130	(-0.020)	0.133
TFP	0.238	0.106	0.210	0.107	*TFP*	(0.100)	(0.062)	0.081	0.085
Observations	[1,593;2,015]	[1,289;1,418]	[1,593;2,015]	[1,289;1,418]	*Observations*	[481;673]	[716;797]	[481;673]	[716;797]
At home					*At arm's length*				
Total sales	0.867	0.363	0.797	0.360	*Total sales*	1.077	0.145	0.760	0.132
Total employment	0.431	0.165	0.400	0.167	*Total employment*	0.589	(0.042)	0.404	(0.024)
Total equity	0.819	0.370	0.751	0.388	*Total equity*	1.139	0.192	0.767	0.161
Skill intensity	0.028	(0.021)	0.024	(0.010)	*Skill intensity*	0.063	(0.006)	0.047	(0.008)
Capital intensity	0.278	(0.069)	0.246	(0.108)	*Capital intensity*	0.467	0.068	0.299	(0.048)
Labor productivity	0.264	(-0.007)	0.248	(-0.016)	*Labor productivity*	0.245	(0.043)	0.167	(0.044)
TFP	0.235	(0.052)	0.222	(0.041)	*TFP*	0.198	(0.012)	0.141	(0.017)
Observations	[1,873;3,433]	[517;593]	[1,873;3,433]	[517;593]	*Observations*	[3,440;5,405]	[1,789;1,991]	[1,789;1,991]	[1,789;1,991]

Columns (I), (II), (V), and (VI) give estimated coefficients of a sourcing dummy in regressions of the form:
$\ln(Y_{it}) = b_0 + b_1\text{SOURCING}_{it} + b_2\text{INDUSTRY} + b_3\text{YEAR} + b_4\text{INDUSTRY} \times \text{YEAR} + e_{it}$;
columns (III), (IV), (VII), and (VIII) give estimated coefficients of a sourcing dummy in regressions of the form:
$\ln(Y_{it}) = b_0 + b_1\text{SOURCING}_{it} + b_2\text{INDUSTRY} + b_3\text{YEAR} + b_4\text{INDUSTRY} \times \text{YEAR} + b_5\text{EXPORT}_{it} + e_{it}$,
where i indexes firms and t years. The vectors INDUSTRY and YEAR collect comprehensive industry and year dummy variables, respectively. EXPORT_{it} is an indicator for a firm's export status. All estimated coefficients except for those given in parentheses are significantly different from zero at the 10% level. Large firms employ more than 200 workers. The panel in the upper-left corner looks at differences between firms sourcing through foreign outsourcing (baseline category) and those opting for foreign integration in addition to foreign outsourcing. Accordingly for the panel in the lower-left corner for domestic outsourcing and domestic integration. The underlying sample excludes all offshoring firms. The panel in the upper-right corner looks at differences between firms sourcing through domestic integration (baseline category) and those opting for foreign integration in addition to domestic integration. Accordingly for the panel in the lower-right corner for domestic outsourcing and foreign outsourcing.

Table 4. Ex-ante sourcing premia in levels and growth rates, 2006-2009

	Integration versus outsourcing					*Offshore versus domestic sourcing*			
	Unconditional		*Conditional on export status*			*Unconditional*		*Conditional on export status*	
	Levels	*Growth rate*	*Levels*	*Growth rate*		*Levels*	*Growth rate*	*Levels*	*Growth rate*
	(I)	(II)	(III)	(IV)		(V)	(VI)	(VII)	(VIII)
<u>Abroad</u>					<u>*Intra-firm*</u>				
Total sales	1.128	(-0.063)	1.010	(-0.061)	*Total sales*	0.949	(0.104)	0.833	(0.103)
Total employment	0.829	(-0.025)	0.734	(-0.024)	*Total employment*	0.641	(0.118)	(0.569)	(0.115)
Total equity	1.028	(-0.038)	0.891	(-0.040)	*Total equity*	1.035	(-0.022)	(0.952)	(-0.023)
Skill intensity	(0.034)		(0.034)		*Skill intensity*	(0.106)		(0.107)	
Capital intensity	0.249	(0.001)	(0.200)	(0.001)	*Capital intensity*	(0.014)	(-0.024)	(-0.028)	(-0.020)
Labor productivity	(0.201)	(-0.073)	(0.182)	(-0.071)	*Labor productivity*	(0.269)	-0.273	(0.266)	-0.273
TFP	0.217	(-0.010)	0.201	(-0.011)	*TFP*	(0.090)	(0.041)	(0.074)	(0.043)
Switching firms	115	115	115	115	*Switching firms*	42	42	42	42
Observations	[1,578;1,607]	[1,547;1,607]	[1,578;1,607]	[1,547;1,607]	*Observations*	[556;580]	[551;580]	[556;580]	[551;580]
<u>*At home*</u>					<u>*At arm's length*</u>				
Total sales	0,891	0.091	0,723	0.091	*Total sales*	0.670	0.048	0.425	0.048
Total employment	0,515	(0.019)	0,425	(0.016)	*Total employment*	0.331	(0.024)	0.195	(0.020)
Total equity	1,101	(-0.062)	0.930	(-0.062)	*Total equity*	0.671	(-0.039)	0.370	(-0.040)
Skill intensity	(0.072)		(0.061)		*Skill intensity*	(0.006)		(-0.009)	
Capital intensity	(0.415)	(0.136)	(0.310)	(0.141)	*Capital intensity*	0.342	(0.001)	0.193	(0.008)
Labor productivity	0.374	(0.022)	0,332	(0.024)	*Labor productivity*	0.157	(-0.026)	0.094	(-0.020)
TFP	0.209	(0.005)	0,182	(0.007)	*TFP*	0.115	(-0.001)	0.076	(0.002)
Switching firms	36	36	36	36	*Switching firms*	251	251	251	251
Observations	[2,034;2,106]	[1,994;2,106]	[2,034;2,106]	[1,994;2,106]	*Observations*	[2,279;2,354]	[2,239;2,354]	[2,279;2,354]	[2,239;2,354]

Columns (I) and (V) give estimated coefficients of a sourcing dummy in regressions of the form
$\ln(Y_{it}) = b_0 + b_1\text{SOURCING}_{it+1} + b_2\text{INDUSTRY} + b_3\text{YEAR} + b_4\text{INDUSTRY} \times \text{YEAR} + e_{it}$,
and accordingly for columns (III) and (VII) with an additional export dummy variable, EXPORT_{it}.
Columns (II) and (VI) give estimated coefficients of a sourcing dummy in regressions of the form
$\ln(Y_{it+1}) - \ln(Y_{it}) = b_0 + b_1\text{SOURCING}_{it+1} + b_2\text{INDUSTRY} + b_3\text{YEAR} + b_4\text{INDUSTRY} \times \text{YEAR} + e_{it}$,
and accordingly for columns (IV) and (VIII) with an additional export dummy variable, EXPORT_{it}.
The indexes i and t refer to firms and years, respectively. The vectors INDUSTRY and YEAR collect comprehensive industry and year dummy variables. All estimated coefficients except for those given in parentheses are significantly different from zero at the 10% level. See the text for a precise description of the estimation sample and the definition of the sourcing dummy for each of the four panels.

Table 5. Linear probability model for the ownership choice of intermediate input production, 2006-2009

Sourcing location: Foreign economy
Estimation sample: Offshoring firms
Dependent variable: Indicator variable for integrated input production in the foreign economy

	OLS				*Random effects*				*Fixed Effects*			
VARIABLES	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Total factor productivity	0.223***	0.104***	-0.464**	-0.415*	0.151***	0.106***	-0.283	-0.234	0.105**	0.105**	-0.203	-0.167
	(0.030)	(0.031)	(0.218)	(0.215)	(0.033)	(0.031)	(0.214)	(0.199)	(0.052)	(0.052)	(0.300)	(0.293)
Total factor productivity ×			0.196***	0.149**			0.126**	0.098*			0.091	0.080
Headquarter intensity			(0.064)	(0.062)			(0.063)	(0.059)			(0.091)	(0.088)
Number of observations	2,459	2,404	2,459	2,404	2,459	2,404	2,459	2,404	2,459	2,458	2,459	2,458
Firm-level controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Industry-and-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.128	0.270	0.136	0.275	.	.	.	.	0.041	0.043	0.042	0.043

Sourcing location: Domestic economy
Estimation sample: Non-offshoring firms
Dependent variable: Indicator variable for integrated input production in the domestic economy

	OLS				*Random effects*				*Fixed Effects*			
VARIABLES	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Total factor productivity	0.126***	0.061***	-0.517***	-0.391**	0.086***	0.043***	-0.353**	-0.243*	0.028	0.029	-0.280	-0.288*
	(0.023)	-0.023	(0.180)	(0.193)	(0.017)	(0.016)	(0.139)	(0.129)	(0.019)	(0.021)	(0.172)	(0.173)
Total factor productivity ×			0.187***	0.132**			0.128***	0.084**			0.090*	0.092*
Headquarter intensity			(0.056)	(0.060)			(0.043)	(0.040)			(0.054)	(0.054)
Number of observations	2,672	2,612	2,672	2,612	2,672	2,612	2,672	2,612	2,672	2,672	2,672	2,672
Firm-level controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Industry-and-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.100	0.217	0.114	0.224	.	.	.	.	0.049	0.050	0.053	0.054

This table gives estimates of the model in equation (11). *Total factor productivity* is time-variant and firm-specific and estimated with the Olley & Pakes (1996) estimation algorithm. *Headquarter intensity* is time-invariant and industry-specific and gives the industry mean of firm-level capital intensity. Standard errors are given in parentheses. For the OLS estimator and the FE estimator, these are clustered by firm. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

Table 6. Linear probability model for the location choice of intermediate input production, 2006-2009

Ownership structure: Integrated production
Estimation sample: Non-outsourcing firms
Dependent variable: Indicator variable for foreign sourcing through intra-firm trade

	OLS				*Random effects*				*Fixed Effects*			
VARIABLES	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Total factor productivity	0.051***	0.023*	0.170*	0.166*	0.024***	0.011**	0.070**	0.072**	0.002	0.002	0.001	0.001
	(0.016)	(0.013)	(0.090)	(0.092)	(0.007)	(0.005)	(0.033)	(0.033)	(0.002)	(0.002)	(0.006)	(0.006)
Total factor productivity × *Headquarter intensity*			-0.035	-0.042*			-0.013	-0.018**			0.000	0.000
			(0.022)	(0.024)			(0.008)	(0.009)			(0.002)	(0.002)
Number of observations	2756	2696	2756	2696	2756	2708	2756	2696	2756	2756	2756	2756
Firm-level controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Industry-and-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.054	0.128	0.056	0.131	.	.	.	.	0.087	0.005	0.005	0.005

Ownership structure: Outsourced production
Estimation sample: Non-integrated firms
Dependent variable: Indicator variable for the location of outsourced input production

	OLS				*Random effects*				*Fixed Effects*			
VARIABLES	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Total factor productivity	0.294***	0.159***	1.074***	0.893***	0.231***	0.143***	0.735***	0.685***	0.059	0.052	-0.186	-0.161
	(0.029)	(0.029)	(0.230)	(0.220)	(0.026)	(0.026)	(0.207)	(0.207)	(0.049)	(0.049)	(0.370)	(0.374)
Total factor productivity × *Headquarter intensity*			-0.230***	-0.216***			-0.148**	-0.159***			0.072	0.062
			(0.067)	(0.064)			(0.061)	(0.061)			(0.110)	(0.111)
Number of observations	3876	3809	3876	3809	3876	3809	3876	3809	3876	3875	3876	3875
Firm-level controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Industry-and-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.134	0.215	0.140	0.220	.	.	.	.	0.028	0.029	0.028	0.029

This table gives estimates of the model in equation (13). *Total factor productivity* is time-variant and firm-specific and estimated with the Olley & Pakes (1996) estimation algorithm. *Headquarter intensity* is time-invariant and industry-specific and gives the industry mean of firm-level capital intensity. The estimation sample in the upper panel excludes firms outsourcing abroad but not firms outsourcing domestically. Restricting the sample to pure integration firms would reduce the sample size dramatically due to the large incidence of domestic outsourcing. Standard errors are given in parentheses. For the OLS estimator and the FE estimator, these are clustered by firm. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

B Data appendix

Table B.1. List of sampled manufacturing industries

CNAE-09 Classification	Industry
101	Meat
102-109, 120	Food Products and Tobacco
110	Beverages
131-133, 139, 141-143	Textile
151-152	Leather & Footwear
261-262	Timber & Wooden Products
171-172	Paper Products
181-182	Graphics Design
201-206, 211-212	Chemical & Pharmaceutical Products
221-222	Plastic & Rubber Products
231-237, 239	Mineral Products (Non-Metal Products)
241-245	Ferrous Metals & Non-Ferrour Metals
251-257, 259	Metal Products
281-284, 289	Industry & Agricultural Machinery
261-268	Informatics, Electronics, Optics
271-275, 279	General & Electric Machinery
291-293	Motorized Vehicles
301-304, 309	Other Transportation Equipment
310	Furniture Industry
321-325, 329	Miscellaneous Manufacturing

Table B.2. Definition of firm-specific variables

Variable	Definition
Total sales	Log of total sales expressed in euros
Total employment	Log of the average number of workers during the year
Total equity	Log of total equity capital expressed in Euros
Skill intensity	Number of graduate workers (university and three-year degrees) over the total number of workers as of December 31st
Capital intensity	Ratio of capital assets to the average number of workers during the year (expressed in thousands of Euros per worker)
Labor productivity	Log of real value added over effectively worked hours
Total factor productivity	Log of total factor productivity obtained from production function estimates à la Olley & Pakes (1996)

References

Antràs, P. (2003). Firms, contracts, and trade structure. *The Quarterly Journal of Economics*, *118*(4), 1375–1418.

Antràs, P., & Helpman, E. (2004). Global sourcing. *Journal of Political Economy*, *112*(3), 552–80.

Antràs, P., & Helpman, E. (2008). Contractual frictions and global sourcing. In E. Helpman, D. Marin, & T. Verdier (Eds.) *The Organization of Firms in a Global Economy*, chap. 1, (pp. 9–54). Harvard University Press: Cambridge, MA.

Bernard, A. B., & Jensen, J. B. (1995). Exporters, jobs, and wages in US manufacturing: 1976-87. *Brookings Papers on Economic Activity: Microeconomics*, *1995*(1995), 67–112.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2010). Intrafirm trade and product contractibility. *American Economic Review: Papers & Proceedings*, *100*(2), 444–48.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics,* forthcoming.

Corcos, G., Irac, D. M., Mion, G., & Verdier, T. (2012). The determinants of intrafirm trade: Evidence from French firms. CEP Discussion Paper No 1119.

De Loecker, J. (2007). Do exports generate higher productivity? Evidence from Slovenia. *Journal of International Economics*, *73*(1), 69–98.

Defever, F., & Toubal, F. (2011). Productivity, relationship-specific inputs and the sourcing modes of multinationals. CEPR Discussion Paper No 8656.

Fariñas, J. C., López, A., & Martín-Marcos, A. (2010). Foreign sourcing and productivity: Evidence at the firm level. *The World Economy*, *33*(3), 482–506.

Federico, S. (2010). Outsourcing versus integration at home or abroad. *Empirica*, *37*(1), 47–63.

Federico, S. (2012). Headquarter intensity and the choice between outsourcing versus integration at home or abroad. *Industrial and Corporate Change*, *21*(1), 1–22.

Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, *94*(4), 691–719.

Klette, T. J., & Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, *11*(4), 343–61.

Kohler, W., & Smolka, M. (2011). Sourcing premia with incomplete contracts: Theory and evidence. *B.E. Journal of Economic Analysis and Policy*, *11*(1), 1–37.

Kohler, W., & Smolka, M. (2012). Global sourcing: Evidence from Spanish firm-level data. In R. M. Stern (Ed.) *Quantitative Analysis of Newly Evolving Patterns of International Trade*, chap. 4, (pp. 139–189). World Scientific Studies in International Economics.

Lanz, R., & Miroudot, S. (2011). Intra-firm trade: Patterns, determinants and policy implications. OECD Trade Policy Working Papers 114, OECD.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*(6), 1695–1725.

Miroudot, S., Lanz, R., & Ragoussis, A. (2009). Trade in intermediate goods and services. OECD Trade Policy Working Papers 93, OECD.

Mrázowá, M., & Neary, P. (2011). Selection Effects with Heterogeneous Firms. Discussion Paper 588, University of Oxford, Department of Economics.

Nunn, N., & Trefler, D. (2008). The boundaries of the multinational firm: An empirical analysis. In E. Helpman, D. Marin, & T. Verdier (Eds.) *The Organization of Firms in a Global Economy*, chap. 2, (pp. 55–83). Harvard University Press: Cambridge, MA.

Olley, G. S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64* (6), 1263–97.

Schwarz, C., & Suedekum, J. (2012). Global sourcing of complex production processes. Mimeo 93.

Wagner, J. (2011). Offshoring and firm performance: Self-selection, effects on performance, or both? *Review of World Economics*, *147* (2), 217–247.

Yeaple, S. R. (2006). Offshoring, foreign direct investment, and the structure of U.S. trade. *Journal of the European Economic Association*, *4* (2-3), 602–611.