CESifo Working Papers

Global Sourcing of Heterogeneous Firms: Theory and Evidence

Wilhelm Kohler
Marcel Smolka

Global Sourcing of Heterogeneous Firms: Theory and Evidence

Abstract

This paper investigates the role of firm productivity in drawing firm boundaries in global sourcing. Our analysis focuses on how productivity affects the allocation of ownership rights between the headquarter of a firm and an intermediate input supplier (vertical integration vs. outsourcing), as well as the location of intermediate input production (offshore vs. domestic). Unlike previous work, we allow for a fully flexible productivity effect with varying magnitude and sign across different industries. Our estimation strategy is motivated by the canonical economic model of sourcing due to Antràs & Helpman (2004). This model invokes the property rights theory of the firm in order to pin down firm boundaries as the outcome of an interaction between firm heterogeneity and the industry's sourcing intensity (i.e. the importance of inputs sourced from suppliers relative to headquarter inputs). We demonstrate that, at the level of the firm, the model implies a productivity effect that varies not just in magnitude (and potentially non-monotonically), but also in sign with the sourcing intensity of the industry. To estimate the effects empirically, we use Spanish firm-level data from the Encuesta sobre Estrategias Empresariales (ESEE). We find a pattern of effects whereby productivity stimulates vertical integration in industries of low sourcing intensity, but favors outsourcing in industries of high sourcing intensity. Moreover, we find that productivity boosts offshoring throughout all industries, with the effect increasing monotonically in the sourcing intensity. Our results lend strong empirical support to the property rights view of the firm in the global economy.

Wilhelm Kohler
University of Tübingen
Mohlstr. 36
Germany – 72074 Tübingen
wilhelm.kohler@uni-tuebingen.de

Marcel Smolka
Aarhus University
Fuglesangs Allé 4, Building 2632
Denmark – 8210 Aarhus V
msmolka@econ.au.dk

1 Introduction

One of the oldest and most intricate questions in economics is what determines the boundaries of firms. What motivates firms to seek control over certain parts of the value chain, beyond the degree of influence afforded by market transactions? Why do some firms seek more control than others? Why might firms aim for different degrees of control when operating in different markets? Over the past 15 years, interest in these questions was spurred by the empirical observation that the share of international trade that takes place within the firm boundaries of control has been increasing through time and, perhaps more importantly, that this share varies a lot across countries, industries and firms; see Antràs (2014a).

In this paper, we investigate the role of firm productivity in drawing firm boundaries in global sourcing. Do more productive firms deploy a different control structure over the production of their inputs? Do they seek more or less control, and why? Is the firm's productivity important also for the decision to source its inputs internationally, in the global economy, rather than in the domestic economy? And if so, what determines just how important it is? In this paper, we propose a novel approach to address these questions, and we apply this approach to firm-level data from Spain to provide new empirical answers. Our analysis focuses on how productivity affects the allocation of ownership rights between the headquarter of a firm and an intermediate input supplier (vertical integration vs. outsourcing), as well as the location of intermediate input production (offshore vs. domestic). Importantly, the estimation strategy we develop provides for a highly flexible productivity effect with varying magnitude and sign across different industries.

Our empirical strategy is motivated by two observations. First, ownership as well as offshoring decisions of heterogeneous firms are issues of considerable interest for academics and policymakers alike, as evidenced by a large body of literature; see Antràs (2014a). Secondly, assuming a uniform productivity effect, as the literature typically does, is problematic, since it conflicts with the canonical model of sourcing due to Antràs & Helpman (2004) (AH model). This model assumes a setting of incomplete contracts and relationship-specific inputs, and combines the property rights theory of the firm (Grossman & Hart, 1986; Hart & Moore, 1990) with a heterogeneous firm model of international trade.[1] In doing so, it pins down firm boundaries as the outcome of an interaction between two industry-specific parameters: the degree of firm heterogeneity (i.e. the productivity dispersion in an industry) and what we call the sourcing intensity of production (i.e. the importance of inputs sourced from suppliers relative to headquarter inputs). In this paper, we take a firm-level perspective on the AH model and demonstrate that, at the level of the firm, the model implies a productivity effect that varies not just in magnitude (and potentially non-monotonically), but also in sign with the sourcing intensity of the industry.

Our paper contributes to the literature on global sourcing in various ways. First, we develop a firm-level perspective on the AH model, and derive testable firm-level propositions about the effect of productivity on sourcing behavior. Secondly, we propose a flexibly parameterized within-industry estimator, without imposing a linear or unidirectional productivity effect. By using variation in sourcing and productivity across firms within industries, we address potential endogeneity problems arising from unobserved industry heterogeneity (such as unobservable fixed costs of sourcing). Finally, we test our firm-level propositions on an unusually rich Spanish survey data set covering both the allocation of ownership rights between headquarters and intermediate input suppliers (ownership

---

[1] Antràs (2003) introduces the property rights approach to trade; for a survey of the approach see Antràs (2014b). The AH model extends Antràs (2003) by allowing for firms to be heterogeneous in their productivity, as in Melitz (2003). An older strand of literature, dating back to Coase (1937), uses the transaction cost approach to address contract incompleteness; for a comparison in the context of input trade see Antràs (2014a).

margin) and the location of intermediate input production (offshoring margin).

We begin with a rigorous firm-level view on the AH model. The focus lies on the interaction between the firm's productivity and the industry's sourcing intensity (assumed exogenous to the firm). We first note that in industries of high sourcing intensity, where the use of sourced intermediate inputs promises higher marginal returns, outsourcing the input supply comes with lower *per-unit* production costs than vertical integration. The exact opposite holds true in industries of low sourcing intensity, where the headquarter inputs loom large in the production process. These results are not surprising, as they derive from the well-known property rights theory of the firm: ownership rights are optimally assigned to the party undertaking the more important investment (Grossman & Hart, 1986; Hart & Moore, 1990). However, the corresponding implications for the firm-level productivity effect have, in our view, not received sufficient attention in the literature on global sourcing. In particular, since the firm's productivity magnifies *any* per-unit production cost advantage, its effect at the ownership margin can clearly go either way: favoring outsourcing in industries of high sourcing intensity, and vertical integration in industries of low sourcing intensity. This theoretical ambiguity may rationalize seemingly contradictory firm-level results on the productivity effect found in the empirical literature (Defever & Toubal, 2013; Corcos et al., 2013).

We then study the productivity effect across the entire interval of possible sourcing intensities. We first demonstrate for both margins of sourcing that the effect varies in a potentially non-monotonic way when scaling up the industry's sourcing intensity. Yet, for a plausible parameter subspace of the model we find a monotonic relationship: the productivity effect is more favorable to outsourcing (and less so to vertical integration) in more sourcing-intensive industries. The reason for this relationship is intuitive: gradually increasing the sourcing intensity makes outsourcing more and more attractive in terms of per-unit production costs, as suggested by property rights theory. Hence, due to magnification, the productivity effect becomes more and more favorable to the use of outsourcing. A different pattern emerges at the offshoring margin. Our assumption of lower input prices abroad renders offshoring desirable, as firms benefit from the associated per-unit cost savings. Hence, firms with a higher productivity are more likely to choose offshoring throughout all industries irrespective of their sourcing intensity. However, we show that the strength of the productivity effect varies with the sourcing intensity of the industry, and that it may do so in non-monotonic ways.

Our theoretical results call for an empirical approach that investigates the effect of firm productivity across different industries. The approach needs to be flexible enough to capture the potentially complex variation of the productivity effect across industries. We propose a simple within-industry estimator that accommodates the productivity effect as a flexible function of the industry's sourcing intensity. Importantly, we let the data inform us about the exact shape of this function, allowing both the magnitude and the direction of the effect to vary in non-monotonic ways. Identification is based on variation in sourcing behavior and productivity across firms within industries, coupled with cross-industry variation in sourcing intensity. This strategy avoids endogeneity problems due to unobserved heterogeneity at the industry-level. An important case in point are unobservable fixed costs of sourcing (allowed to differ across sourcing strategies). These also pose a formidable empirical challenge for studies using industry-level rather than firm-level data (Yeaple, 2006; Nunn & Trefler, 2008; Bernard et al., 2010; Nunn & Trefler, 2013), as the industry equilibrium in the AH model is highly sensitive to alternative fixed cost configurations (Antràs & Helpman, 2004).

We investigate the effect of firm productivity using the Spanish Survey on Business Strategies ("Encuesta Sobre Estrategias Empresariales" — ESEE). Important for our analysis, the ESEE data set is one of the very few data sets that cover both firm-level margins of sourcing, the ownership

as well as the offshoring margin.[2] Moreover, the data document ownership and offshoring decisions separately for different sourcing locations (foreign versus domestic) and ownership structures (vertically integrated versus outsourced), respectively. This essentially allows us to provide a twofold analysis of the productivity effects at both margins of sourcing. The ESEE data are representative for the manufacturing sector in Spain, and distinguish between 20 industries based on the NACE-2009 classification. Importantly, the industries differ sharply in their production technology, covering labor-intensive activities such as textiles production as well as capital-intensive activities such as metal and chemical manufacturing. We use these differences in technology as a source of identifying variation in our empirical analysis.

Our empirical investigation demonstrates that the effect of firm productivity on firms' sourcing behavior exhibits marked differences between industries. The differences we find are consistent with the firm-level propositions that we derive from the AH model. Thus, our results lend empirical support to a property rights view of the firm in the global economy. We find that productivity stimulates vertical integration in industries of low sourcing intensity, but discourages vertical integration in industries of high sourcing intensity. The strongest effect, found towards the bottom of the distribution of sourcing intensities, implies that a doubling of productivity increases the probability of vertical integration by 10 percentage points. Overall, the effects we find are large enough to be relevant, if judged against the low incidence of vertical integration observed in our data, where just about ten percent of all firms pursue strategies of vertical integration. Moreover, our estimations reveal that, as we move along the distribution of sourcing intensities, the productivity effect becomes gradually less favorable to vertical integration (and more so to outsourcing). This monotonic adjustment is violated only for an extremely low sourcing intensity that is rarely observed in the data. Strikingly, we find almost no difference in the pattern of productivity effects across the two sourcing locations: the domestic and the foreign economy. Moreover, we find that gains in productivity push firms towards offshoring throughout all industries, but the strength of the effect varies significantly across industries. The strongest effect is found at the upper end of the distribution of sourcing intensities, where a doubling of productivity increases the likelihood of offshoring by around 15 percentage points.

The structure of our paper is as follows. In the next section we adopt a firm-level view on the AH model, and derive firm-level propositions amenable to empirical testing. In Section 3 we present the data set we use in the empirical analysis and describe salient features regarding firms' global sourcing decisions. Section 4 discusses our estimation strategy and presents the results. Section 5 concludes.

2 A firm-level view on the AH model

In this section we develop a firm-level view on the AH model and we provide a theoretical analysis of the effect of firm productivity on the sourcing behavior of firms. We first recall the assumptions of the AH model; we then present the setup for decision making at the firm-level; and we finally derive firm-level propositions about the productivity effect.

---

[2]We have used the ESEE data in previous research. In Kohler & Smolka (2011, 2012), we document that *on average across industries* vertical integration firms and offshoring firms are more productive than outsourcing firms and non-offshoring firms, respectively. In Kohler & Smolka (2014), we demonstrate that this is due to firms *self-selecting* into sourcing strategies based on their productivity. See Tomiura (2007) and Federico (2010, 2012) for related research based on, respectively, Japanese and Italian firm-level data.

2.1 Model assumptions

Firms (or headquarters) produce differentiated varieties of a final good by entering a production relationship with an input supplier. Production relies on two types of intermediates, a headquarter input provided by the firm and a manufacturing component provided by the supplier. Both intermediates are essential for the production of the final good, and both are highly customized. The headquarter input is produced domestically, while input suppliers may either be located in the domestic economy ($h = d$) or the foreign economy ($h = f$). Customization of the intermediates for a specific variety of the final good has two consequences. First, due to impossible third party verification, the two agents cannot enter an enforceable contract about the exact quality of the intermediates to be delivered. And secondly, once produced both intermediates are entirely relationship-specific and have no use outside the production relationship. As a result, the two agents bargain about sharing the revenue generated from producing and selling the final good.

Ex-post Nash bargaining is based on a certain underlying bargaining power of the headquarter, relative to the input supplier, as well as on ex-post outside options of the two agents. Under an outsourcing relationship ($j = o$), either party's outside option is normalized to zero. Under integration ($j = v$), the headquarter acquires a property right that secures part of the revenue in case bargaining breaks down. Hence, integration affords the headquarter a strictly positive outside option. Once an agreement has been reached, the final good is produced, with revenue generated on monopolistically competitive markets, and shared according to the bargaining agreement. In the preceding stage, the two agents decide about the quantities of their intermediates to produce, based on expected revenue shares as well as the marginal costs prevailing in their respective locations. And finally, in the first stage of the game, anticipating decisions in all subsequent stages, the headquarter decides whether to secure participation of a foreign or a domestic input supplier, and whether to rely on an outsourcing or an integrated production relationship.[3]

We refer to ℓ_h as the inverse unit cost of the supplier's input, with a value equal to ℓ_d if it is produced domestically, and ℓ_f if it is produced abroad. Following Antràs & Helpman (2004), we assume $\ell_d < \ell_f$, which implies a foreign cost advantage for the manufacturing component. Without loss of generality, we normalize the unit cost for the headquarter input to unity. The ex-post revenue share accruing to the headquarter is denoted by m_j, so that $1 - m_j$ accrues to the input supplier. The outside option deriving from the residual property right implies that $m_v > m_o$. Each combination of location and ownership structure of sourcing requires its own fixed cost F_{hj} which is specific to the industry that a firm belongs to. Importantly, our analysis is completely general regarding the configuration of fixed cost for the different sourcing strategies.

To nail down the headquarter's choice of a sourcing strategy $\{h, j\}$, 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 demand for the final good, and iii) a zero ex-ante outside option of the input supplier. Assumptions i) and ii) generate concavity of the revenue function. Assumption iii) ties down the participation constraint such that the return the input supplier may expect from entering the production relationship is zero.

---

[3]Antràs & Helpman (2008) introduce varying degrees of contractibility to study how contracting institutions affect the relative prevalence of vertical integration and outsourcing. Schwarz & Südekum (2014) assume a continuum of intermediates supplied by multiple suppliers, allowing for mixed ownership structures (i.e. co-existence of integration and outsourcing at the level of the firm). Antràs & Chor (2013) study ownership decisions at uniquely sequenced stages along the value-added chain of production. Bache & Laugesen (2014) enrich the revenue side of the model by adding an export activity, and they study the effects of different types of trade liberalization.

2.2 Setup for decision making

We write $\Pi(\ell_h, m_j; \zeta, \theta)$ for the headquarter's maximum operating profit, given the sourcing strategy $\{h, j\}$. In this expression, ζ denotes the industry-specific elasticity of final output with respect to the intermediate sourced from the input supplier. We call ζ the sourcing intensity of the industry. Finally, θ denotes total factor productivity which differs across firms active in the same industry.[4] The headquarter's operating profit is equal to its revenue share minus the cost of the headquarter input, plus a lump-sum transfer to the supplier that secures the supplier's participation. It is easy to show that this profit is equal to the total revenue from the production relationship minus the cost of both intermediates. Defining $\Pi(\ell_h, m_j; \zeta, \theta)$ as the *maximum* profit implies that the levels of both intermediates have been chosen optimally, given the sourcing strategy $\{h, j\}$.

The headquarter's choice of the sourcing strategy $\{h, j\}$ is then dictated by

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

Antràs & Helpman (2004) show that under the above assumptions we have

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

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

$$Z(\ell_h, m_j; \zeta) \quad := \quad Az(m_j; \zeta)C(\ell_h, m_j; \zeta), \tag{3}$$

$$z(m_j; \zeta) \quad := \quad 1 - \frac{\varepsilon - 1}{\varepsilon}\Big[m_j(1-\zeta) + (1 - m_j)\zeta\Big], \tag{4}$$

$$\text{and} \quad C(\ell_h, m_j; \zeta) \quad := \quad \Big[m_j^{1-\zeta}(\ell_h(1 - m_j))^\zeta\Big]^{\varepsilon-1}. \tag{5}$$

In these definitions, A captures the general equilibrium interrelationship between different sectors. For the exact meaning of A, see Appendix A. For the purpose of our analysis in this paper, we treat A as a constant. The term $C(\ell_h, m_j; \zeta)^{\varepsilon-1}$ may be interpreted as the *inverse* minimum unit-cost function for the final good, dual to the assumed Cobb-Douglas technology, with the prices of the two inputs inflated by $1/m_j$ and $1/(1 - m_j)$, respectively. Thus, the hold-up problem acts like an input tax, implying a lower than optimal overall input provision (and lower revenue) as well as a distorted input mix, unless $m_j = 0.5$.

The decision rule (1) requires a discrete comparison of the operating profit for $\ell_h = \ell_d, \ell_f$ and $m_j = m_o, m_v$. To describe this comparison, we introduce the following definitions:

$$\Delta_\ell\Pi(m_j; \zeta, \theta) \quad := \quad \Pi(\ell_f, m_j; \zeta, \theta) - \Pi(\ell_d, m_j; \zeta, \theta), \quad j = v, o \tag{6}$$

$$\Delta_m\Pi(\ell_h; \zeta, \theta) \quad := \quad \Pi(\ell_h, m_v; \zeta, \theta) - \Pi(\ell_h, m_o; \zeta, \theta), \quad h = d, f \tag{7}$$

$$\Delta_\ell F_j \quad := \quad F_{fj} - F_{dj}, \quad j = v, o \tag{8}$$

$$\text{and} \quad \Delta_m F_h \quad := \quad F_{hv} - F_{ho}. \quad h = d, f \tag{9}$$

Definitions analogous to (6) and (7) hold for $\Delta_\ell Z(m_j; \zeta)$ and $\Delta_m Z(\ell_h; \zeta)$. The term $\Delta_\ell\Pi(m_j; \zeta, \theta)$ gives the difference in operating profits between the two locations of sourcing, conditional on the ownership structure j. It measures the *location advantage of offshoring*, and is strictly positive due

---

[4]To avoid cluttered notation, we abstain from indexing firms and industries until we get to the point where it is necessary.

to $\ell_d < \ell_f$. In a similar way, $\Delta_m\Pi(\ell_h; \zeta, \theta)$ is the profit difference between vertical integration and outsourcing in sourcing location h. If positive, this difference indicates a *strategic advantage of integration*. If it is negative, the strategic advantage lies with outsourcing. The terms $\Delta_\ell F_j$ and $\Delta_m F_h$ measure the industry-specific fixed cost disadvantages of offshoring and integration, respectively. Either term can be positive or negative.

The fact that the sign of $\Delta_m\Pi(\ell_h; \zeta, \theta)$ is ambiguous reflects a non-monotonic relationship between the headquarter's profit and its revenue share m_j. Vertically integrating the supplier (and thus acquiring control rights in the input produced by the supplier) gives the headquarter a larger ex post share of the production revenue, $m_v > m_o$. However, anticipating a lower ex-post revenue share for itself, the supplier will bring a lower quantity of its input to the production relationship ex-ante, thereby reducing the overall production revenue. Hence, vertical integration is preferred to outsourcing in terms of operating profits only if the supplier's input is not too important for the production relationship as a whole (i.e. if the industry's sourcing intensity ζ is not too high). This is the central trade-off generated by the hold-up problem in both Antràs (2003) and Antràs & Helpman (2004).

2.3 The effect of firm productivity

The effect of firm productivity at both the ownership and the offshoring margin is found by examining the responsiveness of $\Delta_m\Pi(\ell_h; \zeta, \theta)$ and $\Delta_\ell\Pi(m_j; \zeta, \theta)$, respectively, with respect to changes in θ. Examining the *modularity properties* of these functions with respect to both θ and ζ allows us to characterize the productivity effect along the distribution of ζ:[5]

Proposition 1 (ownership margin of sourcing, conditional on location).
The effect of firm productivity on $\Delta_m\Pi(\ell_h; \zeta, \theta)$

   (a) *is heterogeneous across industries and of ambiguous sign: In industries with a low enough sourcing intensity, $\zeta < \zeta^*$, a higher productivity increases $\Delta_m\Pi(\ell_h; \zeta, \theta)$, thus favoring vertical integration, and conversely in industries with a high enough sourcing intensity, $\zeta > \zeta^*$. For the knife-edge case of $\zeta = \zeta^*$ the effect is zero.*

   (b) *The productivity effect varies monotonically with $\zeta \in \left[\underline{\zeta}, \overline{\zeta}\right]$, being more favorable to outsourcing in more sourcing-intensive industries (with $\underline{\zeta} < \zeta^* < \overline{\zeta}$). For a plausible parameter subspace of the model the interval $\left[\underline{\zeta}, \overline{\zeta}\right]$ is large.*

Proof. The productivity effect is found as $\partial\Delta_m\Pi(\ell_h; \zeta, \theta)/\partial\theta = (\varepsilon - 1)\Delta_m Z(\ell_h; \zeta)\theta^{\varepsilon-2}$. Proposition 1 and Lemma 3 in Antràs (2003) imply that the *ratio* $Z(\ell_h, m_v; \zeta)/Z(\ell_h, m_o; \zeta)$ is monotonically decreasing in ζ, with a unique threshold ζ^* implicitly defined through $Z(\ell_h, m_v; \zeta, \theta)/Z(\ell_h, m_o; \zeta, \theta) = 1$. Hence, the *difference* $\Delta_m Z(\ell_h; \zeta)$ is strictly positive for $\zeta < \zeta^*$, strictly negative for $\zeta > \zeta^*$, and equal to zero for $\zeta = \zeta^*$. This proves part (a) of the proposition. We prove part (b) of the proposition in Appendix B.1, by showing that $\Pi(\cdot)$ is submodular with respect to m and ζ for a large and plausible parameter subspace of the model. $\square$

Figure 1 illustrates Proposition 1. The solid lines depict the difference in operating profits between vertical integration and outsourcing, $\Delta_m\Pi(\ell_h; \zeta, \theta) = \Delta_m Z(\ell_h; \zeta)\theta^{\varepsilon-1}$, as a linear function of $\Theta :=$

---

[5]See Appendix B.1 for an exact definition of the modularity properties we employ in order to investigate the effect of firm productivity across different industries. We propose a "modularity view" on firm-level sourcing decisions in Kohler & Smolka (2011). Mrázová & Neary (2013) point out more generally that modularity properties lie at the heart of the selection effects discussed in modern trade literature focusing on firm heterogeneity.

$\theta^{\varepsilon-1}$. They show that the productivity effect can go either way, favoring outsourcing in sourcing-intensive industries, and vertical integration in other industries. The downward-sloping line is for an industry with a sufficiently high sourcing intensity, $\zeta_1 > \zeta^*$, where $\Delta_m Z(\ell_h; \zeta_1) < 0$ and thus the strategic advantage lies with outsourcing. An increase in productivity magnifies this advantage, as $\Theta_{i'} > \Theta_i$ implies $\Delta_m Z(\ell_h; \zeta_1)\Theta_{i'} > \Delta_m Z(\ell_h; \zeta_1)\Theta_i$. In industries with a low enough sourcing intensity, $\zeta_2 < \zeta^*$, the strategic advantage lies with integration, $\Delta_m Z(\ell_h; \zeta_2) > 0$, as depicted by the upward-sloping line. Again, this advantage is magnified by the firm's productivity: $\Delta_m Z(\ell_h; \zeta_2)\Theta_{i'} > \Delta_m Z(\ell_h; \zeta_2)\Theta_i$.

<<Figure 1 about here>>

Importantly, these relationships are independent of the fixed costs associated with outsourcing and vertical integration, as these enter (2) in an additively-separable way, and thus leave the maximum operating profit unchanged. Measuring the fixed cost disadvantage of integration $\Delta_m F_h$ on the vertical axis, and absent any idiosyncratic disturbances at the firm-level, the solid lines in the figure represent indifference lines where the strategic advantage of one or the other ownership structure is just offset by its fixed cost disadvantage.

Part (b) of Proposition 1 is a logical implication of the property rights theory of the firm, and gives us a deeper insight into the mechanics of the model. It tells us that for a large parameter subspace of the model the indifference lines in Figure 1 rotate clockwise when scaling up the sourcing intensity of the industry. This means that increases in ζ mitigate the productivity effect in industries with $\zeta < \zeta^*$, but reinforce it in industries with $\zeta > \zeta^*$. In other words, the productivity effect is more favorable to outsourcing (and less so to vertical integration) in more sourcing-intensive industries. However, the proposition also implies that this relationship does not universally apply, and that under certain parameter constellations the opposite effect may obtain. The proof of Proposition 1 in Appendix B.1 demonstrates that these constellations feature low values of both m and ε as well as high values of these two parameters.

The following proposition addresses the productivity effect at the offshoring margin of sourcing, i.e. with respect to $\Delta_\ell \Pi(m_j; \zeta, \theta)$:

Proposition 2 (offshoring margin of sourcing, conditional on ownership structure).
The effect of firm productivity on $\Delta_\ell \Pi(m_j; \zeta, \theta)$

(a) *is heterogeneous across industries but of unambiguous sign: A higher productivity increases $\Delta_\ell \Pi(m_j; \zeta, \theta)$, thus favoring offshoring irrespective of the sourcing intensity ζ.*

(b) *The productivity effect varies monotonically with $\zeta \in [0, 1]$ for $m < \tilde{m}$, being more favorable to offshoring in more sourcing-intensive industries. The effect varies non-monotonically with $\zeta \in [0, 1]$ for $m > \tilde{m}$, being more favorable to offshoring in more sourcing-intensive industries for $\zeta \in [0, \tilde{\zeta})$ and vice versa for $\zeta \in [\tilde{\zeta}, 1]$.*

Proof. Part (a) of the proposition follows straightforwardly from the definition of $\Delta_\ell \Pi(m_j; \zeta, \theta) = \Delta_\ell Z(m_j; \zeta)\theta^{\varepsilon-1}$ and the assumption that $\ell_d > \ell_f$. For part (b) we examine the modularity properties of the profit function $\Pi(\cdot)$ with respect to ℓ and ζ in Appendix B.2. $\qquad\square$

Figure 2 illustrates Proposition 2. It plots the location advantage of offshoring $\Delta_\ell \Pi(m_j; \zeta, \theta)$ as a linear function of Θ. Due to $\ell_d > \ell_f$ this function is always upward-sloping, and the productivity effect is always favorable to offshoring. While the productivity effect is again independent of the fixed costs of sourcing, its strength depends on the industry's sourcing intensity ζ, as the slope of the function

depicted in the figure is governed by ζ. Take the flatter of the two lines to start with. Its slope is equal to $\Delta_\ell Z(m_j; \zeta_2)$. Measuring the fixed cost disadvantage of offshoring $\Delta_\ell F_j$ on the vertical axis, this line is an indifference line for domestic and foreign sourcing. How does this line rotate when moving to an industry with $\zeta_1 > \zeta_2$? Part (b) of Proposition 2 tells us that the answer depends on parameter values for m. In the figure a change from ζ_2 to ζ_1 results in a steeper indifference line, so that the productivity effect is more favorable to offshoring in the more sourcing-intensive industry. Intuitively, if production is more intensive in the sourced input, for a given foreign cost advantage of this input offshoring becomes more attractive due to larger associated cost savings. Appealing as this logic may be, this is not a robust result. As indicated in the proposition, it does not obtain for other plausible parameter constellations of the model.

<<Figure 2 about here>>

3   Data

In this section we introduce the firm-level data we use in our empirical analysis. We first describe the most important features of the data set; we then explain how we measure the firm-specific productivity as well as the industry-specific sourcing intensity in our empirical analysis; and we finally provide some basic and important facts featured by our data.

3.1   Source

Our firm-level data are drawn from the "Encuesta Sobre Estrategias Empresariales" (ESEE), an annual survey of roughly 2,000 manufacturing firms in Spain. It is conducted by the "Sociedad Estatal de Participaciones Industriales", a public foundation based in Madrid.[6] To date the ESEE covers a representative panel of Spanish manufacturing firms for the years 1990-2012. Its panel structure allows us to track firms over time. The initial selection of surveyed firms (in 1990) followed a two-way sampling procedure. Questionnaires were 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 latter subset were selected through a stratified, proportional and systematic sampling (with a random seed). Later, special efforts have been made to keep the sample representative with respect to the population of reference. The survey distinguishes 20 different industries and a total of six different size groups in terms of the number of workers employed: 10-20; 21-50; 51-100; 101-200; 201-500; >500. Industries are defined according to sets of products at the NACE-2009 level.[7]

A central feature of our data is that from 2006 onwards they include information on the global sourcing activities of firms along the ownership as well as the offshoring margin of sourcing. The two relevant questions, which we deem pivotal for the quality of the data, read as follows:[8]

---

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

[7]Table C.1 in Appendix C gives a list of manufacturing industries considered in the survey. Prior to 2009, industries in ESEE data were given in terms of the older NACE-1993 classification. We have used concordance information provided by the SEPI foundation in order to account for this; see `https://www.fundacionsepi.es/esee/en/evariables/i_Cambio_clasificacion_sectorial_ESEE.asp`. More information on the survey and its sampling properties are available in English from SEPI's website at `http://www.fundacionsepi.es/esee/en/epresentacion.asp`.

[8]The survey questionnaire is distributed in Spanish and available for download at `http://www.fundacionsepi.es/esee/sp/svariables/indice.asp`.

1. *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 that these represent in the total amount of purchases of your firm in* [year].

   (a) *Spanish suppliers that belong to your group of companies or that 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]

2. *For the year* [year], *indicate whether you imported goods and services that you incorporate (transform) in the production process, and the percentage that theses imports − according to the type of supplier − represent in the total value of your imports.* [yes/no]

   (a) *From suppliers that belong to your group of companies and/or from foreign firms that 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]

We use answers to question 1.(a) to construct a dummy variable for domestic integration (abbreviated DI) that takes on the value one if the firm answers "yes", and zero if it answers "no". We proceed accordingly for domestic outsourcing (question 1.(b): DO), foreign integration (question 2.(a): FI), and foreign outsourcing (question 2.(b): FO). We then characterize each observation (i.e. firm-year combination) through a tuple of variables $\Omega = \langle DI, DO, FI, FO \rangle$. For example, we attach $\Omega = \langle 1, 0, 1, 0 \rangle$ to any firm that reports to source inputs from both a foreign and a domestic integrated supplier in a given year (but not from an independent supplier in Spain or abroad). It will sometimes prove convenient to refer to Ω as a set of tuples. An example is: $\Omega = \{\langle 1, 0, 1, 0 \rangle, \langle 0, 0, 1, 0 \rangle\}$, which we write in shorthand as $\Omega = \{\langle \cdot, 0, 1, 0 \rangle\}$.

3.2 Firm-specific productivity

A pivotal variable in our empirical analysis is a firm's productivity level θ. We employ two measures of firm productivity. In the main part of the paper we use labor productivity, defined as value added over effective work-hours. Value added is given by the real total production value plus other operating income (i.e., income from rent and leasing, industrial property, commissions, and certain services), minus the real total expenditure on intermediate inputs and external services. To check the robustness of the results obtained, we use a measure of total factor productivity (based on semi-parametric estimations of industry-specific production functions, as suggested by Olley & Pakes (1996)). Details on this estimation can be found in Appendix C.

3.3 Industry-specific sourcing intensity

The key variable at the industry level is the sourcing intensity of production ζ. This variable is not directly observed. On a very fundamental level, ζ reflects the extent to which the input suppliers are bound to bear the costs of production, and $1 - \zeta$ reflects the cost share borne by the headquarter firms. What determines the extent of cost sharing between input suppliers and headquarter firms? Antràs (2003) argues that the costs of physical capital are easier to share than the costs of labor inputs, and that headquarter firms primarily provide (or pre-finance) machinery and specialized tools and equipment, or assist their suppliers in the acquisition of capital equipment and raw materials (as

reported in Dunning (1993, 455-456)).[9] Therefore, in the logic of the AH model, the headquarter input carries more weight in the production of capital-intensive goods than in the production of labor-intensive goods. It has thus become common practice in empirical work to proxy $1 - \zeta$ by an industry-specific measure of capital intensity; see Antràs (2003), Yeaple (2006), Nunn & Trefler (2008), and Federico (2012). We proceed similarly in our empirical analysis, and use the (reversed) scale of industry-specific capital intensities to represent ζ (normalized to the unit interval $[0, 1]$). The industry-specific capital intensity is given by the "typical" capital intensity (the median value) we observe in the industry over the period from 2000 to 2012.[10] The industries Leather & Footwear and Textile & Wearing Apparel plausibly emerge with the most sourcing-intensive production. The industries Beverages and Ferrous Metals & Non-Ferrous Metals, in contrast, feature the least sourcing-intensive production.[11]

3.4 Basic facts

Figure 3 displays the evolution of firms engaged in different sourcing strategies. We define sourcing strategies in a *mutually inclusive* way, so that a firm counts for more than one sourcing strategy if it reports multiple ways of sourcing. We distinguish between domestic outsourcing ($\Omega = \{\langle \cdot, 1, \cdot, \cdot \rangle\}$), domestic integration ($\Omega = \{\langle 1, \cdot, \cdot, \cdot \rangle\}$), foreign outsourcing ($\Omega = \{\langle \cdot, \cdot, \cdot, 1 \rangle\}$), and foreign integration ($\Omega = \{\langle \cdot, \cdot, 1, \cdot \rangle\}$).[12] The figure shows pronounced differences in the fractions of firms choosing a particular sourcing strategy. It also shows that these fractions remain roughly constant over time. Domestic outsourcing is almost universally used (roughly 90% of firms), followed by foreign outsourcing (30%), domestic integration (10%), and foreign integration (5%). Thus, as far as the relative importance of sourcing strategies is concerned, we find a pattern similar to those observed for other industrialized countries such as Japan (Tomiura, 2007) and Italy (Federico, 2010, 2012).[13]

<<Figure 3 about here>>

Figure 4 shows that the use of sourcing strategies strongly depends on firm size. The figure displays the fractions of firms engaged in different sourcing strategies in 2012, the most recent year available in our data set. It does so separately for the six different firm size groups, with relative frequencies indicated by bars. As firms employing 10 to 50 workers represent about 90% of all firms in Spanish manufacturing, their use of sourcing strategies closely resembles the picture displayed in Figure 3. However, larger firms show a markedly stronger engagement in both foreign sourcing and vertical integration. In particular, strategies of foreign as well as vertical integration are each used by more than 40% of the very large firms (those with more than 500 employees). For foreign outsourcing the number is even higher, at more than 70%.

---

[9]Other references consistent with this idea and discussed in Antràs (2003) are Milgrom & Roberts (1993), Aoki (1990, 25), and Young et al. (1985).

[10]The capital intensity is defined as the real value sum of real estate, construction and equipment over the average number of workers during the year. We employ sampling weights to correct for deviations from random sampling.

[11]Figure C.1 in Appendix C shows that firms in our sample are concentrated in industries with sourcing intensities above 0.3.

[12]There is a fifth group which we call "non-sourcing" firms ($\Omega = \langle 0, 0, 0, 0 \rangle$) since they report zero volumes for input sourcing.

[13]We have also investigated the importance of different sourcing strategies in terms of the value of sourcing. First, we have computed the value shares for each sourcing strategy as firm averages. And secondly, we have computed the value shares at the industry level. In either case we find the same (ordinal) ranking of sourcing strategies as displayed in Figure 3.

Figure 5 illustrates that firms using vertically integrated production relationships are strongly concentrated in industries of low sourcing intensity (those producing beverages, certain metal products, and chemical products). Outsourcing relationships, in contrast, are considerably more important, relative to vertical integration, in sourcing-intensive industries (those producing leather & footwear or textiles). The between-industry differences we observe in the data are remarkable. In the industry Leather & Furniture, the industry with the highest sourcing intensity, the share of firms sourcing inputs intra-firm is almost zero. For Beverages, the industry with the lowest sourcing intensity, this share is 20%. Our findings for Spain resemble those found for the U.S. For instance, Figure 1 in Antràs (2003) documents that the share of related-party imports in total U.S. imports is the higher, the higher the capital intensity of the industry. Figure 6 shows that firms engaging in offshoring (whether through outsourcing or vertical integration) are not equally distributed across industries. However, differences in the prevalence of offshoring firms between industries are difficult to explain with the sourcing intensity of production: Figure 6 reveals no clear association between the two variables. For example, in the industry Chemical & Pharmaceutical Products, almost 60% of firms engage in offshoring. In the industry Mineral Products (Non-Metal Products), in contrast, the share is around 15%, although the two industries are similarly sourcing-intensive (in the vicinity of 0.4).

4   Empirical analysis

Our theoretical results call for a flexible empirical model that allows us to investigate the productivity effect along the distribution of sourcing intensities. In this section, we first describe the empirical model we use to estimate the effects at the ownership margin of sourcing, and present the results we obtain from this model; we then repeat the same exercise focusing on the offshoring margin of sourcing; and we finally describe the results of an extensive robustness analysis carried out to substantiate our findings.

4.1   The ownership margin of sourcing

Empirical model. Let firms be indexed by $i = 1, \ldots, I$, and industries by $s = 1, \ldots, S$. We use $\mathcal{I}_s$ to denote the set of firms belonging to industry s. We index time by t and define

$$\Lambda_{it,h}(\ell_h; \zeta_s, \theta_{it}) := \Delta_m \Pi(\ell_h; \zeta_s, \theta_{it}) - \Delta_m F_{st,h} + \mu_{it,h}, \quad i \in \mathcal{I}_s, \tag{10}$$

as firm i's total profit difference between vertical integration and outsourcing in location $h = f, d$ at time t, where $\mu_{it,h}$ is a composite term summarizing all effects unrelated to the economic mechanism underlying the AH model of sourcing. We assume that this term is the sum of a deterministic part, $\overline{\mu}_{it,h}$, and a stochastic part, $\tilde{\mu}_{it,h}$, the latter representing unobserved heterogeneity at the firm-level. We want to allow for the fixed cost disadvantage of vertical integration to differ across industries, and to change through time:

$$\Delta_m F_{st,h} = \Delta_m F_{s,h} + \Delta_m F_{t,h}. \tag{11}$$

We face a number of challenges when transforming (10) into an empirical model that can be estimated with our data. First, $\Lambda_{it,h}(\ell_h; \zeta_s, \theta_{it})$ is a latent variable that is not observed by the econometrician.

In the estimation, we therefore revert to a binary variable, denoted by $\text{INT}_{it,h}$, indicating the sign of $\Lambda_{it,h}(\ell_h; \zeta_s, \theta_{it})$, as revealed through the firm's observed choice of ownership structure:

$$\text{INT}_{it,h} = \begin{cases} 1 & \text{if } \Lambda_{it,h}(\ell_h; \zeta_s, \theta_{it}) \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Our data feature firms reporting *multiple* ways of sourcing, as we have seen in the previous section. Thus, we are facing fuzzy data, compared to the theoretical model where different ways of sourcing are mutually exclusive events. Our response to this challenge is to find and impose suitable sample restrictions that generate an empirical decision model for the ownership margin of sourcing as framed in the AH model. We do so separately for the two locations of sourcing, constructing two disjoint samples, one of *offshoring firms* and one of *domestically sourcing firms*. These samples are denoted by sets of tuples Ω_f and Ω_d, to be defined below.

Secondly, there is heterogeneity at the industry-level and at the firm-level that determines the fixed cost difference $\Delta_m F_{st,h}$ and feeds into $\mu_{it,h}$. Failing to account for this heterogeneity risks introducing an endogeneity bias into the estimation. We address this problem by extending the model to include an industry fixed effect $\gamma_{s,h}$, a year fixed effect $\gamma_{t,h}$, and a number of time-varying firm-specific covariates.[14] The fixed effects capture the two fixed cost components, $\Delta_m F_{s,h}$ and $\Delta_m F_{t,h}$, alongside other time-invariant industry-specific parameters (such as skill intensity, R&D intensity, or external financial dependence), and time trends common to all industries. The firm-specific variables included to control for within-industry heterogeneity in dimensions other than productivity are: capital intensity, skill intensity and export volume.[15]

Finally, we must account for the complex nature of $\Delta_m \Pi(\ell_h; \zeta_s, \theta_{it})$ and allow for a flexible productivity effect along the distribution of sourcing intensities. We do so by employing a polynomial regression framework where firm productivity is interacted with a higher-order polynomial for the industry-specific sourcing intensity. We stack the indeterminates of the polynomial into the column vector $\mathbf{\Gamma}^K(\zeta_s) = (\zeta_s^0, \zeta_s^1, \ldots, \zeta_s^K)'$. The model we estimate for the ownership margin of sourcing thus reads as follows:

$$\begin{aligned} \Pr(\text{INT}_{it,h} = 1|\cdot) &= \Pr\left(\Lambda_{it,h}(\ell_h; \zeta_s, \theta_{it}) \geq 0\right) \\ &= \Pr\left(\Delta_m \Pi(\ell_h; \zeta_s, \theta_{it}) - \Delta_m F_{st,h} + \overline{\mu}_{it,h} + \tilde{\mu}_{it,h} \geq 0\right) \\ &= \Pr\left(\boldsymbol{\lambda} \cdot \mathbf{\Gamma}^K(\zeta_s) \times \theta_{it} + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it} \geq -\tilde{\mu}_{it,h}\right) \\ &= G\left(\boldsymbol{\lambda} \cdot \mathbf{\Gamma}^K(\zeta_s) \times \theta_{it} + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}\right), \; i \in \mathcal{I}_s, \Omega_{it} \in \Omega_h, \quad (12) \end{aligned}$$

where G is a cumulative distribution function implicitly defined by the distribution of $\tilde{\mu}_{it,h}$, $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_K)$ are the parameters of interest, and $\mathbf{X}_{it} = (X_{1it}, \ldots, X_{Lit})'$ is a column vector of time-variant firm-specific control variables (along with a vector of coefficients $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_L)$ to be estimated). Notice that in this equation the industry-specific effect $\gamma_{s,h}$ absorbs the main (i.e., non-interaction) effects of the polynomial for the industry's sourcing intensity, and that Ω_h represents the sample either of offshoring or of domestically sourcing firms.

We assume that $\tilde{\mu}_{it,h}$ is uniformly distributed with $E[\tilde{\mu}_{it,h}|\cdot] = 0$, so that the polynomial regression framework in (12) simplifies to a linear probability model (LPM):[16]

$$E[\text{INT}_{it,h}|\cdot] = \Pr(\text{INT}_{it,h} = 1|\cdot) = \boldsymbol{\lambda} \cdot \mathbf{\Gamma}^K(\zeta_s) \times \theta_{it} + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, \; i \in \mathcal{I}_s, \Omega_{it} \in \Omega_h. \quad (13)$$

---

[14]See Table C.3 in Appendix C for a definition of these variables.

[15]See Bernard et al. (2012) for a survey of these dimensions of firm heterogeneity, and Corcos et al. (2013) for related evidence.

[16]For the LPM to be consistent and unbiased, we must assume that $\left[\boldsymbol{\lambda} \cdot \mathbf{\Gamma}^K(\zeta_s) \times \theta_{it} + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}\right] \in [0,1]$

The restriction to observations that satisfy $\Omega_{it} \in \Omega_h$ is crucial. Our aim is to identify the productivity effect from a sample of firms that behave differently at the ownership margin of sourcing but identically otherwise. For foreign sourcing, $h = f$, we construct a sample of offshoring firms, similar to the ones studied in Defever & Toubal (2013) and Corcos et al. (2013). In particular, we impose $\Omega_f = \{\langle \cdot, 1, \cdot, 1 \rangle\}$ and define $\text{INT}_{it,f}$ to equal one if $\Omega_{it} \in \{\langle \cdot, 1, 1, 1 \rangle\}$, and zero if $\Omega_{it} \in \{\langle \cdot, 1, 0, 1 \rangle\}$. Hence, our LPM compares offshoring firms that engage in vertical integration ($FI_{it} = 1 \wedge FO_{it} = 1$) with those that do not ($FI_{it} = 0 \wedge FO_{it} = 1$), discarding observations with $DO_{it} = 0$, but allowing for both $DI_{it} = 0$ and $DI_{it} = 1$.[17]

For $h = d$ we construct a sample of domestically sourcing firms excluding all firms engaged in some form of offshoring. We impose $\Omega_d = \{\langle \cdot, 1, 0, 0 \rangle\}$ and define $\text{INT}_{it,d}$ to equal one if $\Omega_{it} = \langle 1, 1, 0, 0 \rangle$, and zero if $\Omega_{it} = \langle 0, 1, 0, 0 \rangle$. This is a comparison of domestically sourcing firms that choose vertical integration ($DI_{it} = 1 \wedge DO_{it} = 1$) with those that do not ($DI_{it} = 0 \wedge DO_{it} = 1$). Importantly, because Ω_f and Ω_d are disjoint sets, our empirical analysis involves two self-contained and independent tests of Proposition 1.

Results. We estimate Equation (13) separately on the samples Ω_f and Ω_d, using OLS. Statistical inference is based on robust standard errors clustered by firm, i.e., we allow for arbitrary forms of both heteroskedasticity and autocorrelation. In view of Proposition 1, our primary interest lies with $\widehat{\boldsymbol{\lambda}} \cdot \boldsymbol{\Gamma}^K(\zeta)$, which is an estimate of the productivity effect $\partial \Delta_m \Pi(\ell_h; \zeta, \theta) / \partial \theta$. For all degrees of the polynomial $\boldsymbol{\Gamma}^K(\zeta)$ larger than zero, $K > 0$, this estimate is a function of the industry's sourcing intensity ζ. In stressing a productivity effect that is heterogeneous across industries, Proposition 1 clearly calls for $K > 0$. Indeed, part (b) of the proposition calls for $K > 1$, since this would allow for a non-monotonic relationship between the productivity effect and the sourcing intensity.

Figure 7 depicts the estimated polynomials using the sample of offshoring firms Ω_f. Figure 8 does the same for the sample of domestically sourcing firms Ω_d. In each case we present polynomials up to $K = 5$. The estimate for $K = 0$ gives us the *average* productivity effect across industries. A positive estimate means that the likelihood of vertical integration is increasing in productivity, which is what we find with a significance level of 5 percent for both offshoring and domestically sourcing firms. These results corroborate our earlier findings of a "productivity premium" on vertical integration over outsourcing (Kohler & Smolka, 2011, 2014).

Higher degree polynomials capture heterogeneity in the productivity effect across industries. Our

---

for all observations. We use the LPM instead of a non-linear probability model (such as the Probit model), because the interaction effect is not identified in a non-linear model with industry fixed effects. To see this, notice that in a non-linear model the interaction effect emerges as (dropping firm and industry indices):

$$\frac{\partial^2 \Pr(\text{INT} = 1|\cdot)}{\partial \theta \partial \zeta} = G'(\cdot) \times \frac{\partial \left(\boldsymbol{\lambda} \cdot \boldsymbol{\Gamma}^K(\zeta) \right)}{\partial \zeta} + G''(\cdot) \times \left(\frac{\partial \left(\boldsymbol{\lambda} \cdot \boldsymbol{\Gamma}^K(\zeta) \times \theta \right)}{\partial \zeta} + \frac{\partial \gamma}{\partial \zeta} \right) \times \boldsymbol{\lambda} \cdot \boldsymbol{\Gamma}^K(\zeta).$$

Since the industry fixed effect, $\gamma_{s,h}$ in Equation (12), absorbs the main effects of the polynomials for the industry's sourcing intensity, the derivative $\partial \gamma / \partial \zeta$ is not identified (and neither, therefore, is the interaction effect).

[17]Unlike most of the literature, our data allow us to condition on, or control for, the firm's domestic sourcing strategies. As for domestic outsourcing, we restrict the sample to observations with $DO_{it} = 1$, since this strategy is adopted by virtually all offshoring firms that we observe; see Figure 3. However, we have verified that our results do not change if we allow for $DO_{it} = 0$ and include the variable DO_{it} as a control in the regression. As for domestic integration, we allow for both $DI_{it} = 1$ and $DI_{it} = 0$ (for reasons of sample size), but we have verified that our results are not driven by the correlation between DI_{it} and FI_{it}. We have done so by including DI_{it} as an explanatory variable in the regression, and by dropping either firms with $DI_{it} = 1$ or firms with $DI_{it} = 0$ from the sample. Another feature of our data is that a large majority of observations with $FI_{it} = 1$ also exhibit $FO_{it} = 1$ (87% in 2012), which leads us to drop observations with $FO_{it} = 0$. However, our results do not change in any significant way when keeping these observations in the sample.

results attest to strong and significant heterogeneity, although the degree of heterogeneity is somewhat more pronounced, and the R^2-values consistently higher, for offshoring firms than for domestically sourcing firms. The heterogeneity follows a clear pattern: gains in productivity tend to push firms towards vertical integration in the lower range of the distribution of sourcing intensities, and towards outsourcing in the upper range of the distribution. Moreover, for an interior range of sourcing intensities (those around the knife-edge case of $\zeta^* \approx 0.7$), this relationship is monotonic, whereby a higher sourcing intensity tilts the effect of productivity away from vertical integration and towards outsourcing. Importantly, these results show up consistently across the two locations of sourcing, domestic and foreign, and hence provide strong empirical evidence for Proposition 1 and the core economic mechanism featured in Antràs & Helpman (2004).

<<Figures 7 and 8 about here>>

How are we to select among the models represented by different polynomials? Table 1 presents the results of a formal statistical procedure of model selection. To gauge the relevance of the *marginal* polynomial order introduced into the model, we test whether the estimate of λ_K is significantly different from zero in each of the models $K = 0, \ldots, 5$ (odd-numbered columns). To examine the relevance of quadratic, cubic, and higher-order polynomials, we perform a series of Wald tests (even-numbered columns). Based on this procedure, the preferred model among those depicted in Figure 7 is $K = 3$.[18] This specification supports Proposition 1, as it features a maximum for the productivity effect at $\zeta \approx 0.3$, indicating non-monotonic adjustment at the lower end of the distribution of ζ, while revealing monotonic adjustment in the larger neighborhood of $\zeta^* \approx 0.7$ where the productivity effect is muted. This means that virtually all our data fall into the part of the distribution where the productivity effect implies monotonic adjustment.[19]

Figure 8 displays the estimation results obtained for the sample of domestically sourcing firms Ω_d. The figure portrays a picture of heterogeneity across industries that is very similar to what we have found for offshoring firms above. However, the fit is not as good (judged by the R^2-values) as with foreign sourcing. Moreover, columns (3) and (4) of Table 1 reveal that the preferred polynomial is of first (instead of third) degree, with $\hat{\lambda}_1 < 0$ at the 5 percent significance level. This implies monotonicity of the productivity effect in the sourcing intensity, with productivity becoming ever more favorable to outsourcing as we move up along the distribution of sourcing intensities.

4.2 The offshoring margin of sourcing

Empirical model. We define firm i's total profit difference between foreign and domestic sourcing under ownership structure $j = v, o$ at time t as

$$\Lambda_{it,j}(m_j; \zeta_s, \theta_{it}) := \Delta_\ell \Pi(m_j; \zeta_s, \theta_{it}) - \Delta_\ell F_{st,j} + \mu_{it,j}, i \in \mathcal{I}_s, \tag{14}$$

where $\Delta_\ell F_{st,j}$ is defined in analogy to $\Delta_m F_{st,h}$ in (11). Proceeding as before we specify the following LPM for the ownership margin of sourcing:

$$E[\text{FOR}_{it,j}|\cdot] = \Pr(\text{FOR}_{it,j} = 1|\cdot) = \boldsymbol{\lambda} \cdot \boldsymbol{\Gamma}^K(\zeta_s) \times \theta_{it} + \gamma_{s,j} + \gamma_{t,j} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, i \in \mathcal{I}_s, \Omega_{it} \in \Omega_j, \tag{15}$$

---

[18]This choice appears broadly consistent with the values of the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) given in Figure 7. However, we prefer the above procedure over the AIC and BIC, because the latter contain no information about the absolute quality of the models considered.

[19]About 98 percent of the observations in our sample (and 19 out of 20 industries) feature $\zeta > 0.3$, as indicated by the vertical lines in the figures; see also the histogram in Figure C.1. The Beverages industry is thus the only industry to the left of the vertical lines (at $\zeta = 0$).

where

$$\text{FOR}_{it,j} = \begin{cases} 1 & \text{if } \Lambda_{it,j}(m_j; \zeta_s, \theta_{it}) \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

We estimate this model on two disjoint samples, using Ω_v to denote the sample of *integrating firms* and Ω_o to denote the sample of *outsourcing firms*. As for integrating firms, $j = v$, we employ $\Omega_v = \{\langle 1, 1, ., 0 \rangle\}$ and define $\text{FOR}_{it,v}$ to take on the value one if $\Omega_{it} = \langle 1, 1, 1, 0 \rangle$ and zero otherwise. We thus compare integrating firms that choose offshoring with those that do not.[20] This involves a drastic reduction in sample size (down to 701 observations and 288 firms in our benchmark estimates), so that the estimation results obtained must be interpreted with caution.[21] As for outsourcing firms, $j = o$, we impose $\Omega_o = \{\langle 0, 1, 0, . \rangle\}$ and set $\text{FOR}_{it,o} = 1$ if $\Omega_{it} = \langle 0, 1, 0, 1 \rangle$ and $\text{FOR}_{it,o} = 0$ otherwise. This involves a comparison between outsourcing firms engaged in offshoring and those that are not.

Results. Against the backdrop of Proposition 2, our primary interest lies with $\widehat{\boldsymbol{\lambda}} \cdot \boldsymbol{\Gamma}^K(\zeta_s)$, which is the estimated effect of a change in firm productivity θ on $\Delta_\ell \Pi(m_j; \zeta, \theta)$, the profit difference between offshoring and domestic sourcing. As with the ownership margin of sourcing, the proposition stresses heterogeneity of the productivity effect across industries, but unlike Proposition 1 it holds that the productivity effect is unambiguously positive, meaning that in all industries offshoring becomes more attractive as the firm becomes more productive. Moreover, for a relatively large parameter subspace, the effect is monotonically increasing in the sourcing intensity over the entire interval $\zeta \in [0, 1]$. If non-monotonic, the productivity effect is rising in ζ for low values of ζ, and falling in ζ for high values of ζ.

<<Figures 9 and 10 about here>>

Figures 9 and 10 present the estimated polynomials for the samples Ω_v and Ω_o, respectively. Again, heterogeneity of the productivity effect calls for a polynomial of degree $K > 0$, but for completeness we also present the average effect obtained by setting $K = 0$. For integrating firms, the average productivity effect is estimated with a positive sign, but not significantly different from zero, whereas for outsourcing firms it is positive at the 1 percent significance level, and quantitatively important, with a margin of 6 percentage points. For both samples we observe significant heterogeneity, judged by the estimated coefficients $\widehat{\lambda}_1$ along with the corresponding significance levels in columns (5) and (7) of Table 1, and in both cases the Wald test leads us to select the first degree polynomial as our preferred model. Thus, while non-monotonicity is a distinct theoretical possibility, our estimation suggests that it is not an empirically relevant phenomenon at the offshoring margin of sourcing. Moreover, the estimated productivity effect is positive for virtually all observations in the sample ($\zeta > 0.3$), and $\hat{\lambda}_1 > 0$ means that the pattern of heterogeneity across industries is as expected from part (b) of Proposition 1. Thus, productivity is the more favorable to offshoring, the higher the sourcing intensity, and this is true for both vertical integration and outsourcing.

4.3 Robustness analysis and extensions

Next, we investigate the robustness of our estimation results with respect to changes in the estimation sample as well as some modifications to, and extensions of, our baseline models. We describe each step of this analysis in turn and briefly summarize the main results obtained.

---

[20]As before, due to the large incidence of domestic outsourcing, we include observations with $DO_{it} = 1$ and drop those with $DO_{it} = 0$. Keeping the latter group of observations in the sample and controlling for DO_{it} in the regression does not make a significant difference for the results we obtain.

[21]Although these concerns are corroborated by standard regression model diagnostics, in the interest of completeness we report and comment on all estimation results in the text.

Richer specification of fixed costs. In the baseline models above, we assume that changes in the fixed cost differences over time occur equally in all industries, and that firms active in the same industry face the same fixed cost disadvantage of integration and offshoring, respectively. These assumptions might be considered as being too strong to be plausible. Consider for example the recent improvements in transport and communication technology that have rendered offshoring more attractive over the past decade. It is easy to imagine that some industries (e.g. those intensive in complex capital goods) have benefitted more from this development than other industries. This leads us to expect that changes in the fixed cost disadvantage of offshoring do not apply equally to all industries, but are, instead, specific to the industry. Also, it seems likely that larger firms feature a different fixed cost profile than smaller firms. We allow for these possibilities by extending our baseline models to include industry-and-year fixed effects $\gamma_{st,b}$ (rather than just industry fixed effects $\gamma_{s,b}$ and year fixed effects $\gamma_{t,b}$) as well as size group fixed effects $\gamma_{g,b}$, where g is an index for the six firm size groups and $b = h, j$. This extension gives rise to a more ambitious model where identification comes from between-firm variation in sourcing and productivity within size groups and industry-year pairs. We report the results of this extension under Model (B) in Table 2 (for convenience restricted to the case of $K = 1$). We may conclude that the core insights of our benchmark models easily survive in this extended model.

Region-specific effects. In the baseline models above, we do not distinguish between different sourcing locations other than domestic versus foreign. This is due to data limitations, as the sourcing activities are reported at the level of the firm, and not disaggregated by source country. Our firm-level analysis could thus suffer from omitted variables or aggregation bias. We address these concerns in two different ways. First, as the data include the origin of firms' imports, we augment the model for the ownership margin to include dummy variables for the different import regions reported in our data. These are the European Union; Latin America; the rest of the OECD; and the rest of the world. Such a model allows, for example, for a fixed cost disadvantage of integration that is specific to the region from which the firm is sourcing its inputs, and thus reduces the risk of omitted variables bias. Model (C) in Table 2 reveals reassuring robustness also with respect to this modification. And secondly, in order to address a possible aggregation bias, we restrict the sample to firms receiving all their imports from the European Union. This is a data-driven modification, as a sizable fraction of firms with positive imports restrict their importing activities (and thus their sourcing) to the European Union (about 38 percent of importers in the sample in 2012). Importers extending their activities beyond the EU typically do so *in addition* to importing from the EU. The estimation results (not reported) do not indicate aggregation bias.

Other sources of firm-level heterogeneity. Our model uses two sources of variation in the data to identify the parameters in $\boldsymbol{\lambda}$. The first is variation in productivity *between* firms *within* industries, and the second is variation in sourcing intensity *between* industries. This approach has the advantage that the variation in the variables of interest is large. Under the given set of assumptions, the OLS estimator is asymptotically consistent. However, if there is an omitted firm-specific variable that is correlated with the other covariates, the estimates suffer from omitted variables bias due to unobserved heterogeneity. A more satisfying approach could therefore be to exploit the *within-firm* variation in the data over time, along with the variation *between* industries. This would allow us to control for any time-invariant firm-specific variable that influences the firm's decision in favor of one or the other sourcing strategy.[22] However, because of the sample restrictions employed, there is very little within-

---

[22]One way to get rid of these variables is to within-transform the data and compute all variables relative to the firm-specific mean value over time. The so-called within-group estimator (or fixed effects estimator) is asymptotically consistent, independently of whether the (time-invariant) firm-specific variables are correlated with the other covariates or not.

firm variation in the dependent variable that could be exploited for identification purposes.[23] We must therefore follow an alternative route to reduce the risk of an omitted variables bias. In particular, we test for other sources of firm heterogeneity that are correlated with, respectively, the firm's ownership and offshoring decisions by augmenting the model to include further time-varying firm-level controls: age; employment; R&D intensity; service intensity; debt ratio; capacity utilization; and fixed effects for the type of good produced (consumption, intermediate, or undefined). A detailed description of these variables can be found in Appendix C; the results for the parameters of interest are reported under Model (D) in Table 2. Again, the outcome is a high degree of robustness.

Exclusion of foreign-owned firms. In our baseline estimations, we include firms irrespective of whether they are foreign owned or owned domestically. To the extent that firms in Spain are bound to source inputs from their foreign parental companies, the property rights model might not be an accurate representation of the firm-supplier production relationship. To see whether our results are upheld in a sample of domestically owned firms, we drop observations where foreign investors own more than 50% of the Spanish firm's joint capital. Model (E) in Table 2 yields reassuring robustness results also in this dimension.

Trimming estimator. A necessary condition for OLS to be consistent and unbiased in the LPM is that the linear predictions of the model fall into the unit interval. We have checked for violations of this condition in all our benchmark models, and found sizable fractions of the data where the linear predictions are negative. For example, approximately 11 percent of the data deliver such predictions in each of the models for the ownership margin; see Figure C.2 in Appendix C.[24] Hence, we have reason to believe that our estimates of $\boldsymbol{\lambda}$ are neither consistent nor unbiased. The results reported in Horrace & Oaxaca (2006), however, suggest that the problem can be addressed by using a "trimming" estimator which excludes the critical observations from the estimation sample. We have done so in up to three iterations and succeeded in gradually reducing the fraction of critical observations. The results reported under Models (F) to (H) in Table 2, if anything, strengthen the conclusions we have drawn from our baseline models above.

Probit model. Another way to address the drawbacks of the LPM is to estimate a non-linear Probit model. In our case, this means that the function G in Equation (12) is the cumulative distribution function of the normal distribution. Although the interaction effect between the productivity and the sourcing intensity is not identified in such a model (as we have shown earlier in footnote 16), we can still identify and consistently estimate the parameters in $\boldsymbol{\lambda}$. We have done so for $K = 1$, and our Probit estimates of λ_0 and λ_1 can serve as an indication of whether the LPM used for our baseline models suffers from severe misspecification. Overall, the results we obtain with the Probit estimator do not indicate misspecification; see Model (I) in Table 2.

Outliers with respect to θ. To see if our estimation results are driven by productivity outliers in the sample, we have re-estimated our baseline models excluding the most productive as well as the least productive firms (top and bottom 1%) in each industry-year pair from the sample. We do not find that they are; see Model (J) in Table 2.

Outliers with respect to ζ. The Beverages industry is by far the least sourcing-intensive industry in our sample ($\zeta = 0$). The industry is more than twice as capital-intensive as the second most capital-intensive industry (Ferrous Metals & Non-Ferrous Metals).[25] Excluding this industry from

---

[23]Although we use data for seven years (2006 to 2012), over that period the typical number of years in which we observe a firm is three (with gaps).

[24]The problem seems less severe for modelling the offshoring margin based on Ω_o, where less than 3 percent of the data deliver predictions outside the unit interval.

[25]The logged capital intensities are 4.94 and 4.19 for the industries Beverages and Ferrous Metals & Non-Ferrous

the sample reinforces our interpretation of substantial heterogeneity in the productivity effect across industries; see Model (K) in Table 2.

Alternative measure of productivity. In our baseline estimations, we use real value added over effective work-hours as a measure of firm productivity. Alternatively, we use total factor productivity (TFP) in our estimations. We apply the Olley & Pakes (1996) estimation algorithm, henceforth called OPA, in order to estimate industry-specific production functions. Against the benchmark of these estimates, we recover each firm's TFP level as a firm-specific, time-variant variable. More details on the TFP estimation can be found in Appendix C. The results reported under Model (L) in Table 2 reveal that using this alternative measure of productivity does not change the conclusions of our analysis.

Alternative measure of sourcing intensity. For our baseline estimations, we construct the sourcing intensity of production based on the median capital intensity of firms active in a particular industry (over the period 2000-2012). This is our preferred measure as the median statistic is robust to outliers. Alternatively, we use the mean of capital intensity, to find that the results are upheld; see Model (M) in Table 2.

Sourcing shares as dependent variables. We have so far used the discrete choice framework to model the probability of a firm to choose the one or the other of two sourcing strategies. An alternative view on the data is to look at the *quantitative* importance of the different sourcing strategies, measured by the volume of input sourcing. This information is available in our data; see the survey questions presented in Section 3.1. We exploit this variation in the data by using firm-specific relative sourcing volumes as dependent variables in the estimation (rather than dummy variables). As for the ownership margin, we use the share of intra-firm sourcing in the total volume of sourcing (specific to the foreign economy, $h = f$, or the domestic economy, $h = d$). As for the offshoring margin, we use the share of foreign sourcing in the total volume of sourcing (specific to integrated suppliers, $j = v$, or outsourced suppliers, $j = o$). Overall, the estimation results we obtain for these models are largely consistent with those from our baseline estimations; see Model (N) in Table 2.

Number of foreign affiliates. Although we cannot assign sourcing volumes reported by the firm to specific countries or transactions, we do have information on the number and location of the firm's foreign affiliates. This allows us to open up a further perspective on the ownership margin of sourcing. From our theoretical analysis we can expect more productive firms to be more inclined to vertical integration (and thus have more foreign affiliates) in industries of low sourcing intensity, and conversely in industries of high sourcing intensity. Hence, we estimate the following model by OLS:

$$\#\text{Foreign-affiliates}_{it,f} = \lambda_0 \cdot \theta_{it} + \lambda_1 \cdot (\zeta_s \times \theta_{it}) + \gamma_{s,f} + \gamma_{t,f} + \boldsymbol{\beta} \cdot \mathbf{X}_{it} + \xi_{it,f}, \ i \in \mathcal{I}_s, \Omega_{it} \in \Omega_f, \quad (16)$$

where $\#\text{Foreign-affiliates}_{it,f}$ is the number of foreign affiliates of firm i in year t, and $\xi_{it,f}$ is an error term with zero conditional mean. The results reported in Table 3 are consistent with the property rights theory of the firm.

Further industry interactions. We have so far focused on differences in the productivity effect across the distribution of sourcing intensities. However, industries exhibit important differences also in other dimensions, and the productivity effect might also be responsive to these other differences. To see to what extent the interaction effects between the firm's productivity and the industry's sourcing intensity are driven by other industry characteristics, we augment the model to include interactions between firm productivity and a number of important industry variables: skill intensity; R&D intensity; service intensity; and external financial dependence (derived from Rajan & Zingales

---

Metals, respectively.

(1998)). All of these variables, except for skill intensity, are positively correlated with the sourcing intensity; see Table C.2 in Appendix C. The results reported in Tables 4 and 5 indicate that the sourcing intensity is the only industry characteristic that consistently and significantly determines the productivity effect at both margins of sourcing. However, the results also suggest that productivity interacts with other industry characteristics. For example, in the sample of offshoring firms, the skill intensity of an industry seems to co-determine the productivity effect: in more skill-intensive industries the productivity effect is more favorable to vertical integration. This result is consistent with the empirical literature on intra-firm trade, where skill intensity and capital intensity have been used interchangeably to proxy the importance of headquarter inputs at the industry-level.

Discretized version of the model. In the baseline models above, we assume that variations in the productivity effect across industries can be adequately represented by a polynomial function of degree K. To see if this is a plausible assumption we have estimated a discretized version of the model where we interact productivity with dummy variables for quintiles of the distribution of ζ_s, the sourcing intensity of production:

$$\Pr(\text{INT}_{it,h} = 1|\cdot) = \lambda_1 \cdot \theta_{it} + \sum_{q=2}^{5} \lambda_q \cdot (d(q; \zeta_s) \times \theta_{it}) + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, \ i \in \mathcal{I}_s, \Omega_{it} \in \Omega_h, \quad (17)$$

and

$$\Pr(\text{FOR}_{it,j} = 1|\cdot) = \lambda_1 \cdot \theta_{it} + \sum_{q=2}^{5} \lambda_q \cdot (d(q; \zeta_s) \times \theta_{it}) + \gamma_{s,h} + \gamma_{t,j} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, \ i \in \mathcal{I}_s, \Omega_{it} \in \Omega_j, \quad (18)$$

where $d(q; \zeta_s)$ is a dummy variable equal to one if industry s falls into the qth quintile of the distribution of sourcing intensities, and $\lambda_1, \ldots, \lambda_5$ are the parameters of interest. In order to exploit the variation in the data where this seems most relevant, we exclude the Beverages industry from the analysis and focus on the parts of the distribution of ζ where we observe virtually all of our data, i.e., $\zeta > 0.3$. Figure 11 displays the main results from these estimations. Again, these results provide considerable empirical support for the property rights theory of the firm, as they are fully consistent with Propositions 1 and 2.

System of equations. In the baseline estimations, we use sample restrictions to generate empirical decision models that capture the basic trade-offs that firms face at both the ownership and the offshoring margin of sourcing. An alternative route is to estimate a system of four equations on the full sample where each equation has a different element of the tuple $\Omega = \langle DI, DO, FI, FO \rangle$ as the dependent variable (the firm's choice variable). Such a model can take the form of a system of seemingly unrelated regression equations (SURE) (generalizing the linear regression model; see Klette & Griliches (1962)) or a multivariate Probit model (generalizing the univariate Probit model). We have estimated both models and found a pattern of effects that corroborates the above results. A higher productivity is associated with a higher probability of vertical integration (both DI and FI) in industries of low sourcing intensity, whereas this relationship breaks down (or is even reversed) in industries of high sourcing intensity. A positive and significant productivity effect on outsourcing (both DO and FO), in contrast, is more likely the higher the sourcing intensity of production. Moreover, the estimations reveal significant and positive (residual) correlation between FI, FO, and DI.[26]

---

[26]Detailed results from the SURE and the multivariate Probit are available from the authors upon request.

5   Conclusions

We draw two main conclusions from our analysis. First, we believe our results amount to considerable empirical support of the property rights model of global sourcing (Antràs & Helpman, 2004). We have shown that in this model firms respond very differently to gains in productivity if their technologies require a differential mix of sourced inputs and headquarter inputs (the sourcing intensity of production). The model predicts distinct cross-industry patterns of firm-level responses at the ownership as well as the offshoring margin of sourcing, and the empirical results we present in this paper clearly demonstrate that these predictions find empirical support in Spanish firm-level data.

We have placed special emphasis on the ambiguity of the firm-level response at the ownership margin of sourcing, which is an important, yet overlooked theoretical implication of the property rights model. In particular, a higher productivity pushes firms towards vertical integration in those industries where headquarter inputs play a major role in production, and towards outsourcing in industries where they are less important. Moreover, we have demonstrated that the productivity effect is potentially non-monotonic in the sourcing intensity at both the ownership and the offshoring margin of sourcing. However, employing a suitably flexible estimation framework that allows us to capture such non-monotonicity (if present), our empirical results point to a monotonic adjustment. At the ownership margin, we observe that a higher sourcing intensity renders the productivity effect more favorable to outsourcing. At the offshoring margin, it makes the productivity effect work more strongly in favor of offshoring.

Secondly, our results have interesting implications for scenarios of technological change and globalization. Suppose that due to technological progress firms headquartered in industrialized countries experience a uniform increase in their productivity. Our results imply that the share of multinational firms will consequently increase in those industries where headquarter inputs loom large (capital-intensive industries), and decrease where they are of relatively minor importance (labor-intensive industries). Because multinational firms are effectively transmitting technological knowledge across borders (Guadalupe et al., 2012), this has far-reaching implications for technology diffusion and industrial structure in developing countries. In particular, what starts out as a neutral technological improvement in industrialized countries will, through its effect at the ownership margin of sourcing, arrive as a non-neutral technological improvement in developing countries, favoring input suppliers in capital-intensive industries.

As regards offshoring, our results suggest, not surprisingly, that in this scenario of technological progress firms in all industries will see offshoring become more attractive, as they are better able to exploit the foreign cost advantage associated with offshoring. However, the effects will be felt differently across industries, thus accelerating job turnover in some industries more than in others. Importantly, the cross-industry differences that we find are explained by the sourcing intensity of production, independently of the prevalence of routine and impersonal tasks (Blinder, 2009; Becker et al., 2013; Ebenstein et al., 2014).

Figures and tables

Figure 1. The ownership margin of sourcing

$\Delta_m F_h$

$\Delta_m Z(\ell_h; \zeta_2) \cdot \Theta$

$j = o$

$\zeta_2 < \zeta^* < \zeta_1$

$j = v$

Θ_i

$\Theta_{i'}$

Θ

$j = o$

$j = v$

$\Delta_m Z(\ell_h; \zeta_1) \cdot \Theta$
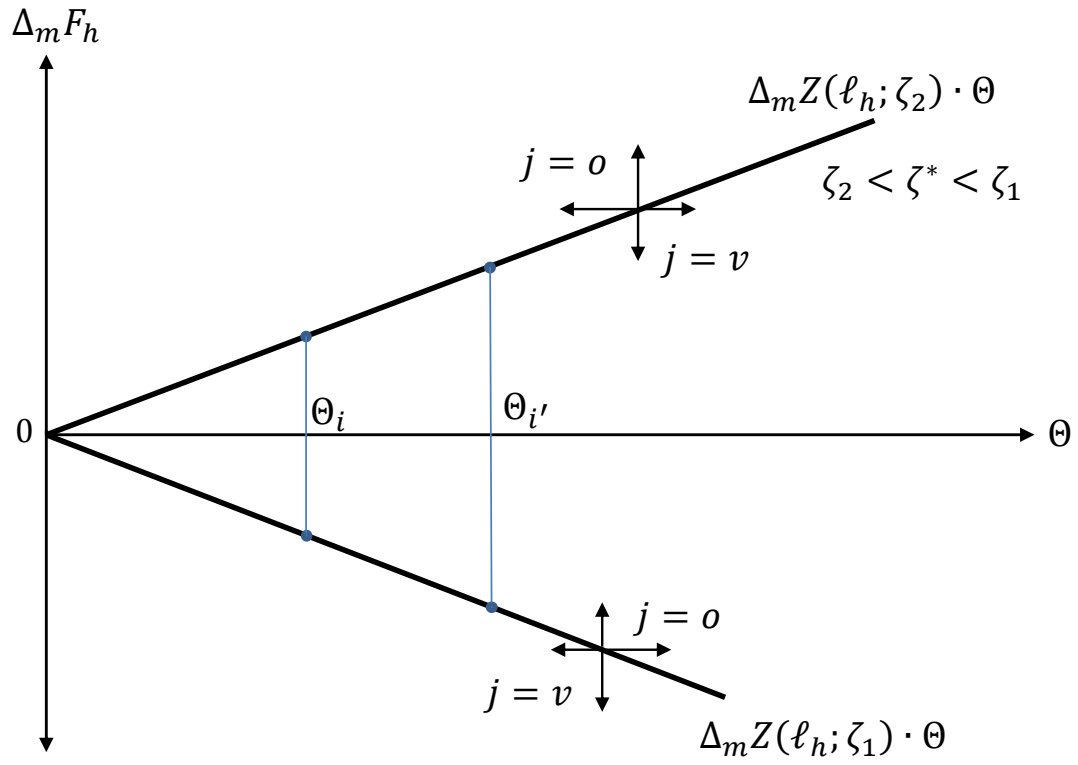
Figure 2. The offshoring margin of sourcing
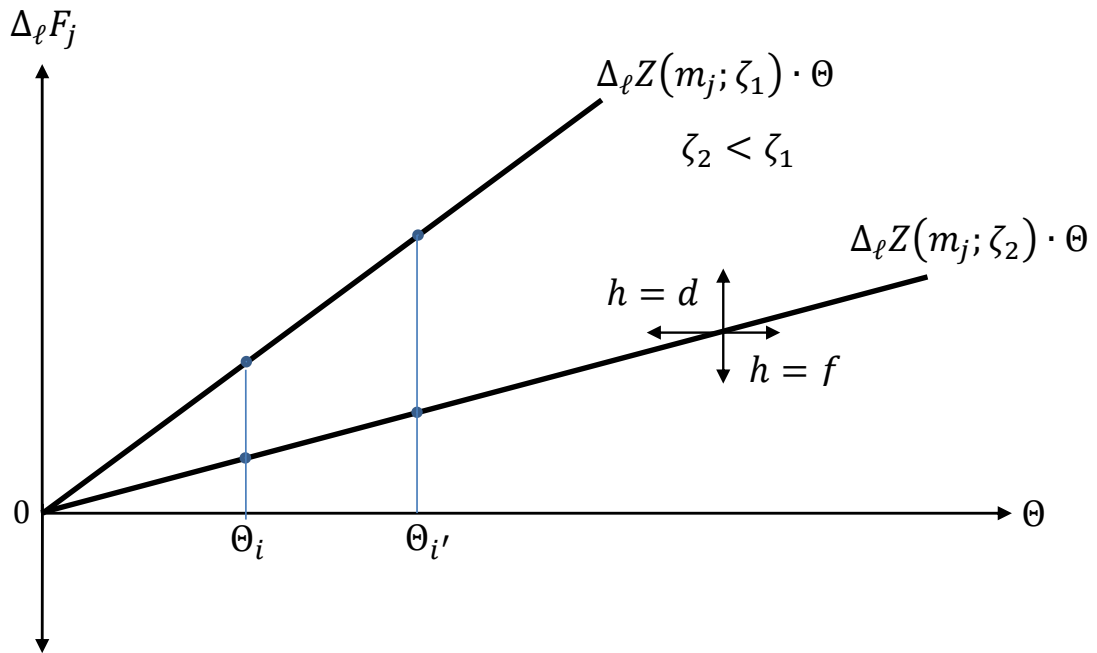
Figure 3. Sourcing strategies of Spanish firms 2006-2012[†]
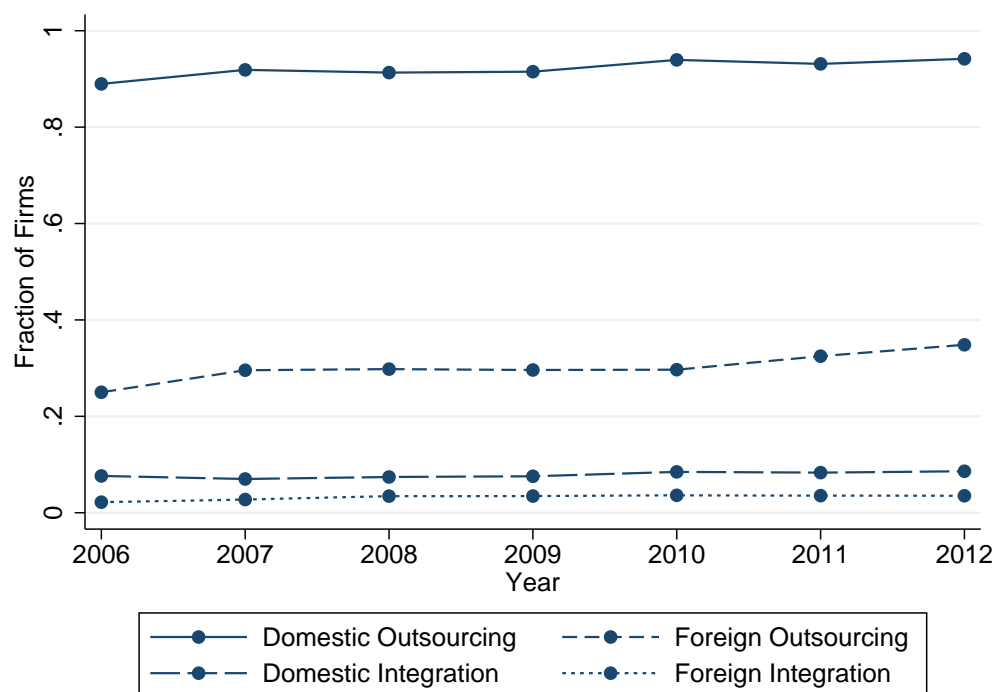
[†]Note: Sourcing strategies are mutually inclusive, i.e. domestic outsourcing includes all firms in the set $\Omega = \{\langle \cdot, 1, \cdot, \cdot \rangle\}$, and accordingly for domestic integration ($\Omega = \{\langle 1, \cdot, \cdot, \cdot \rangle\}$), foreign outsourcing ($\Omega = \{\langle \cdot, \cdot, \cdot, 1 \rangle\}$), and foreign integration ($\Omega = \{\langle \cdot, \cdot, 1, \cdot \rangle\}$). Sampling weights apply.

Figure 4. Sourcing strategies of Spanish firms depending on firm size, 2012[†]

Figure 5. Share of integrating firms in Spain by industry[†]
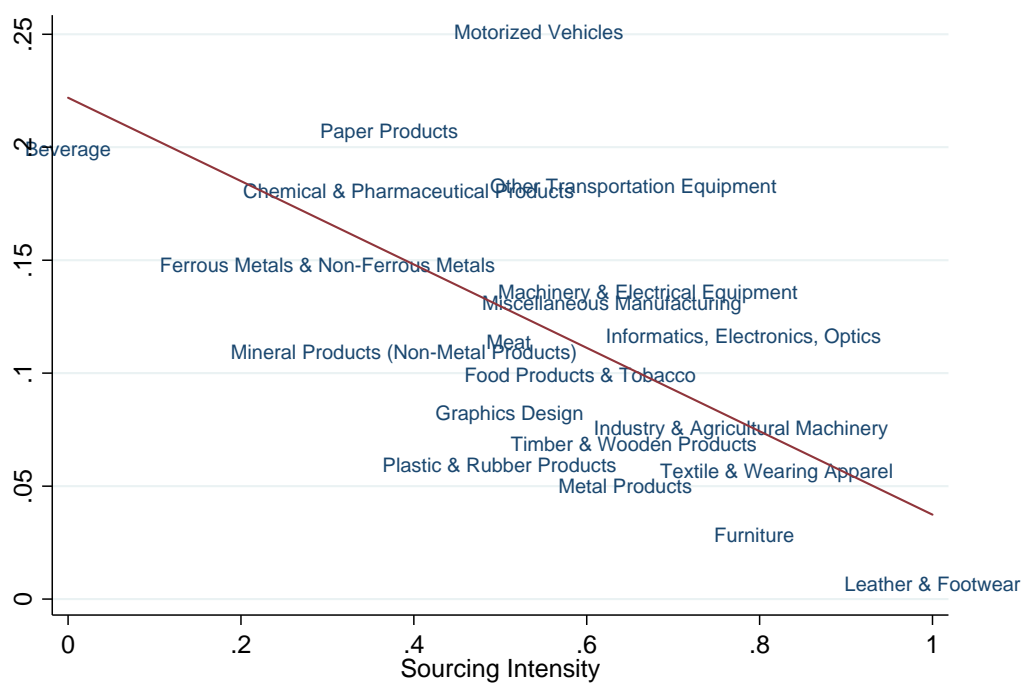
Figure 6. Share of offshoring firms in Spain by industry[†]
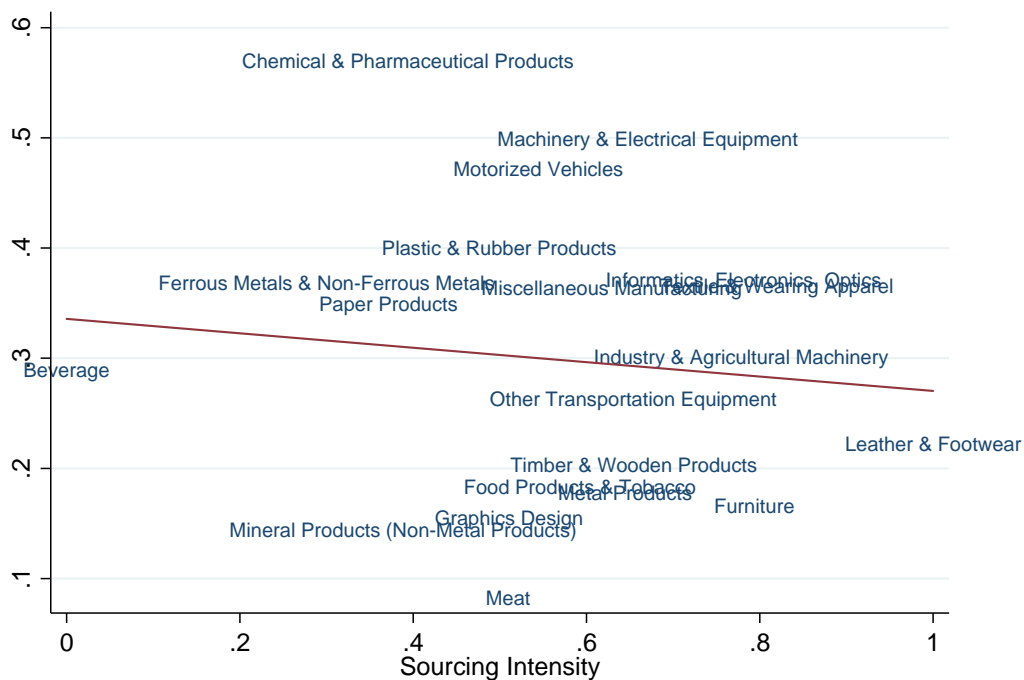
Figure 7. Productivity effect at the ownership margin (foreign economy)[†]
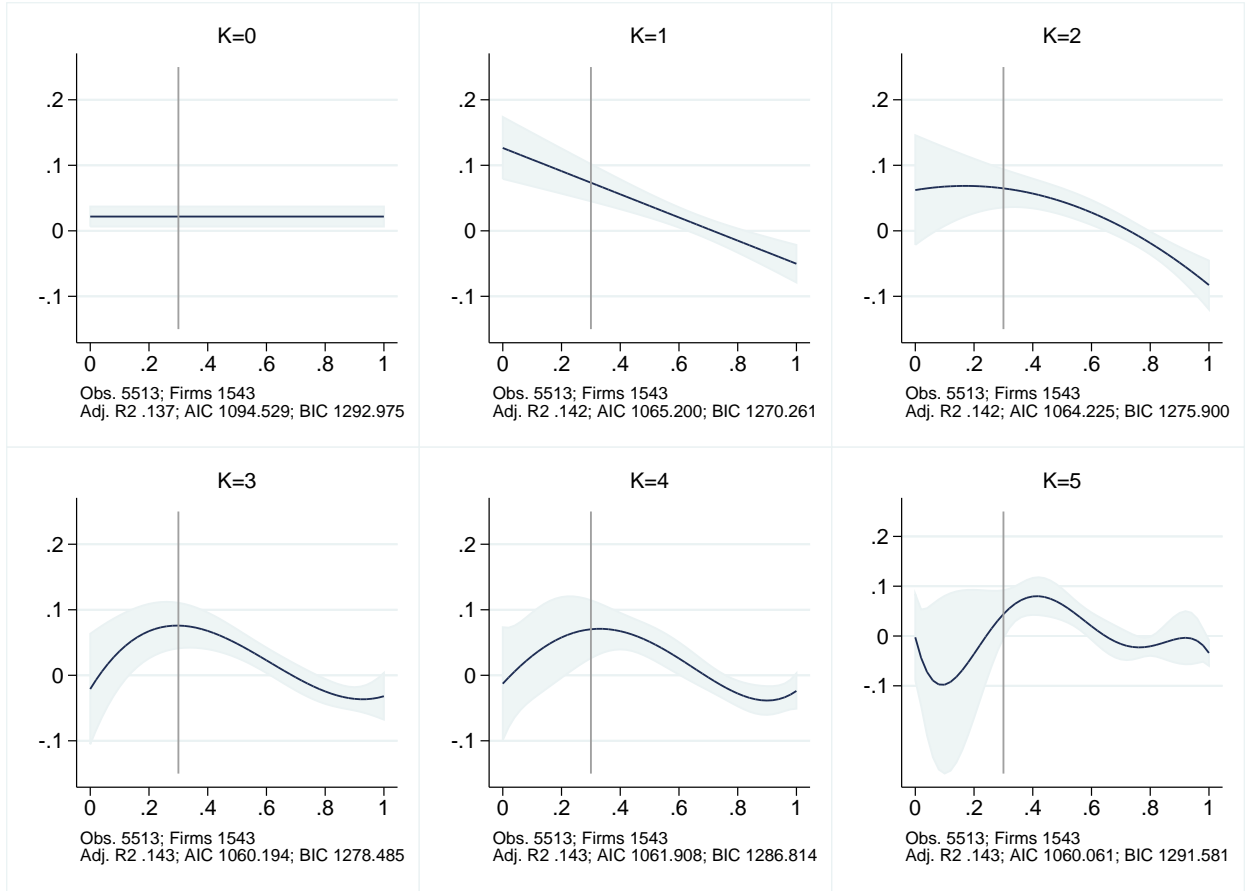
K=0
Obs. 5513; Firms 1543
Adj. R2 .137; AIC 1094.529; BIC 1292.975

K=1
Obs. 5513; Firms 1543
Adj. R2 .142; AIC 1065.200; BIC 1270.261

K=2
Obs. 5513; Firms 1543
Adj. R2 .142; AIC 1064.225; BIC 1275.900

K=3
Obs. 5513; Firms 1543
Adj. R2 .143; AIC 1060.194; BIC 1278.485

K=4
Obs. 5513; Firms 1543
Adj. R2 .143; AIC 1061.908; BIC 1286.814

K=5
Obs. 5513; Firms 1543
Adj. R2 .143; AIC 1060.061; BIC 1291.581

[†]Note: This figure shows the estimated productivity effect at the ownership margin of sourcing for $h = f$:

$$\frac{\partial \Pr(\widehat{\text{INT}_f} = 1|\cdot)}{\partial \theta} = \widehat{\boldsymbol{\lambda}} \cdot \boldsymbol{\Gamma}^K(\zeta)$$

alongside 90% confidence intervals. Estimates of $\boldsymbol{\lambda}$ are obtained from estimating Equation (12) in the LPM. Sampling weights apply.

Figure 8. Productivity effect at the ownership margin (domestic economy)[†]

K=0
Obs. 6484; Firms 1885
Adj. R2 .066; AIC -1188.98; BIC -985.66

K=1
Obs. 6484; Firms 1885
Adj. R2 .075; AIC -1248.70; BIC -1038.61

K=2
Obs. 6484; Firms 1885
Adj. R2 .075; AIC -1246.90; BIC -1030.03

K=3
Obs. 6484; Firms 1885
Adj. R2 .076; AIC -1252.59; BIC -1028.94

K=4
Obs. 6484; Firms 1885
Adj. R2 .076; AIC -1253.26; BIC -1022.84

K=5
Obs. 6484; Firms 1885
Adj. R2 .076; AIC -1253.41; BIC -1016.21

[†]Note: This figure shows the estimated productivity effect at the ownership margin of sourcing for $h = d$:

$$\frac{\partial \Pr(\widehat{\mathrm{INT}_d} = 1|\cdot)}{\partial \theta} = \widehat{\boldsymbol{\lambda}} \cdot \boldsymbol{\Gamma}^K(\varsigma)$$

alongside 90% confidence intervals. Estimates of $\boldsymbol{\lambda}$ are obtained from estimating Equation (12) in the LPM. Sampling weights apply.

Figure 9. Productivity effect at the offshoring margin (vertically integrated production)[†]

[†]Note: This figure shows the estimated productivity effect at the offshoring margin of sourcing for $j = v$:

$$\frac{\partial \Pr(\widehat{\mathrm{FOR}_v} = 1|\cdot)}{\partial \theta} = \widehat{\boldsymbol{\lambda}} \cdot \boldsymbol{\Gamma}^K(\zeta)$$

alongside 90% confidence intervals. Estimates of $\boldsymbol{\lambda}$ are obtained from estimating Equation (15) in the LPM. Sampling weights apply.

Figure 10. Productivity effect at the offshoring margin (outsourced production)[†]

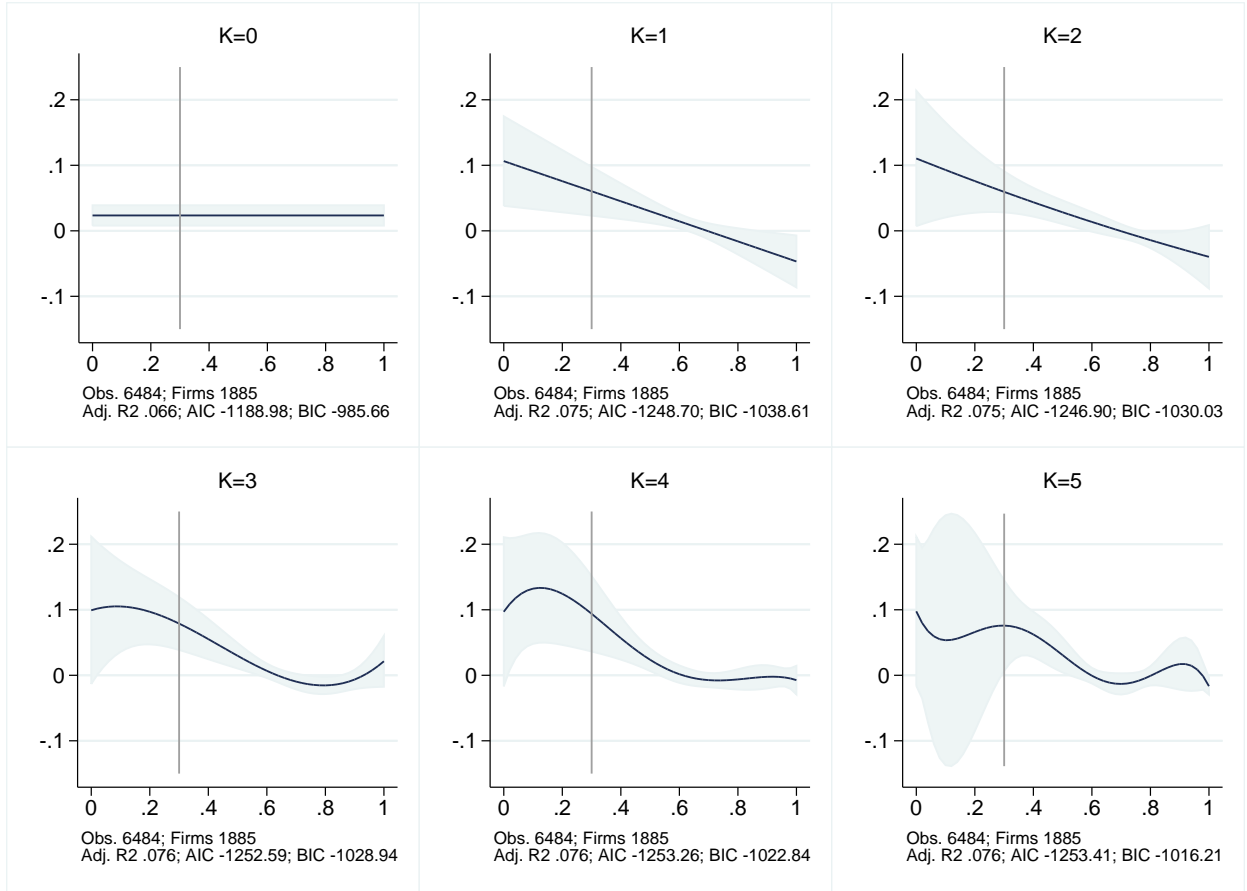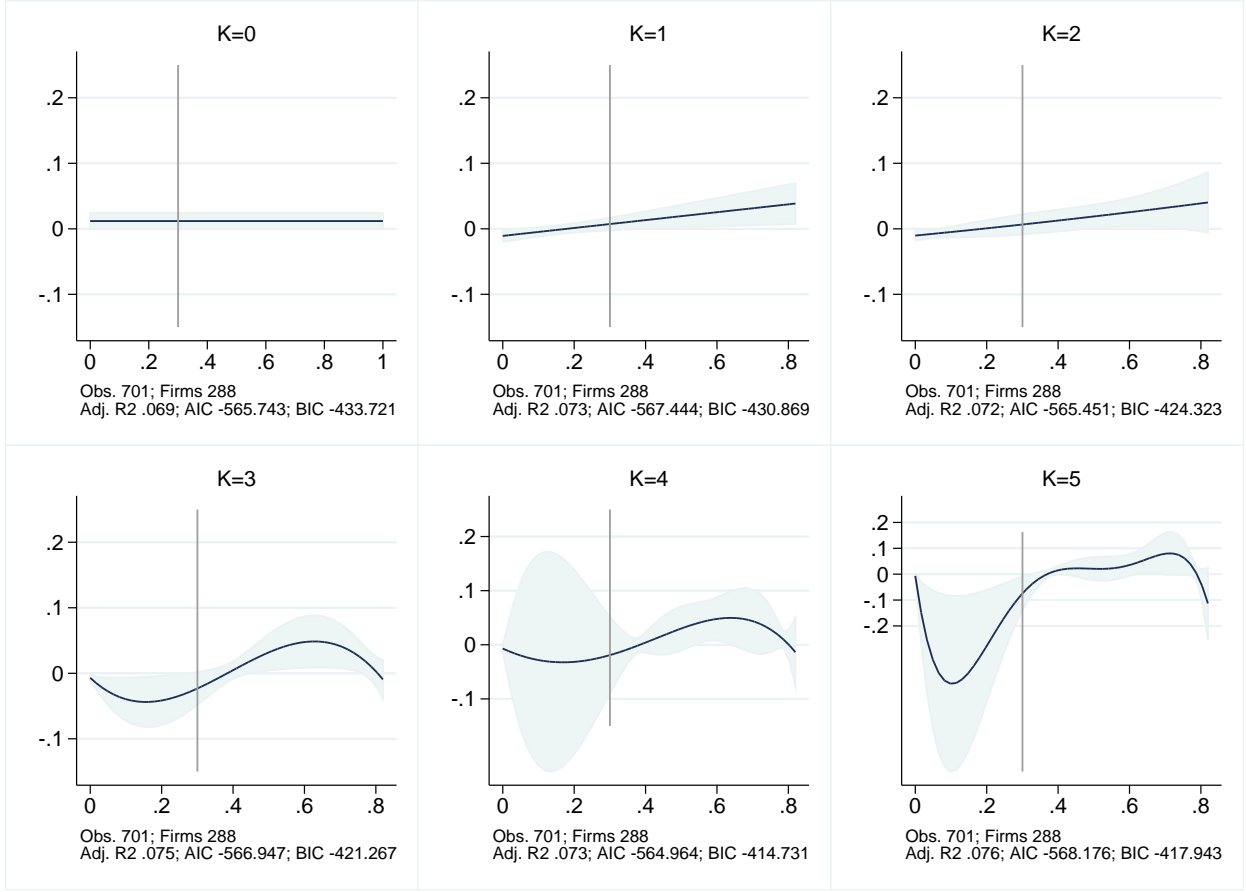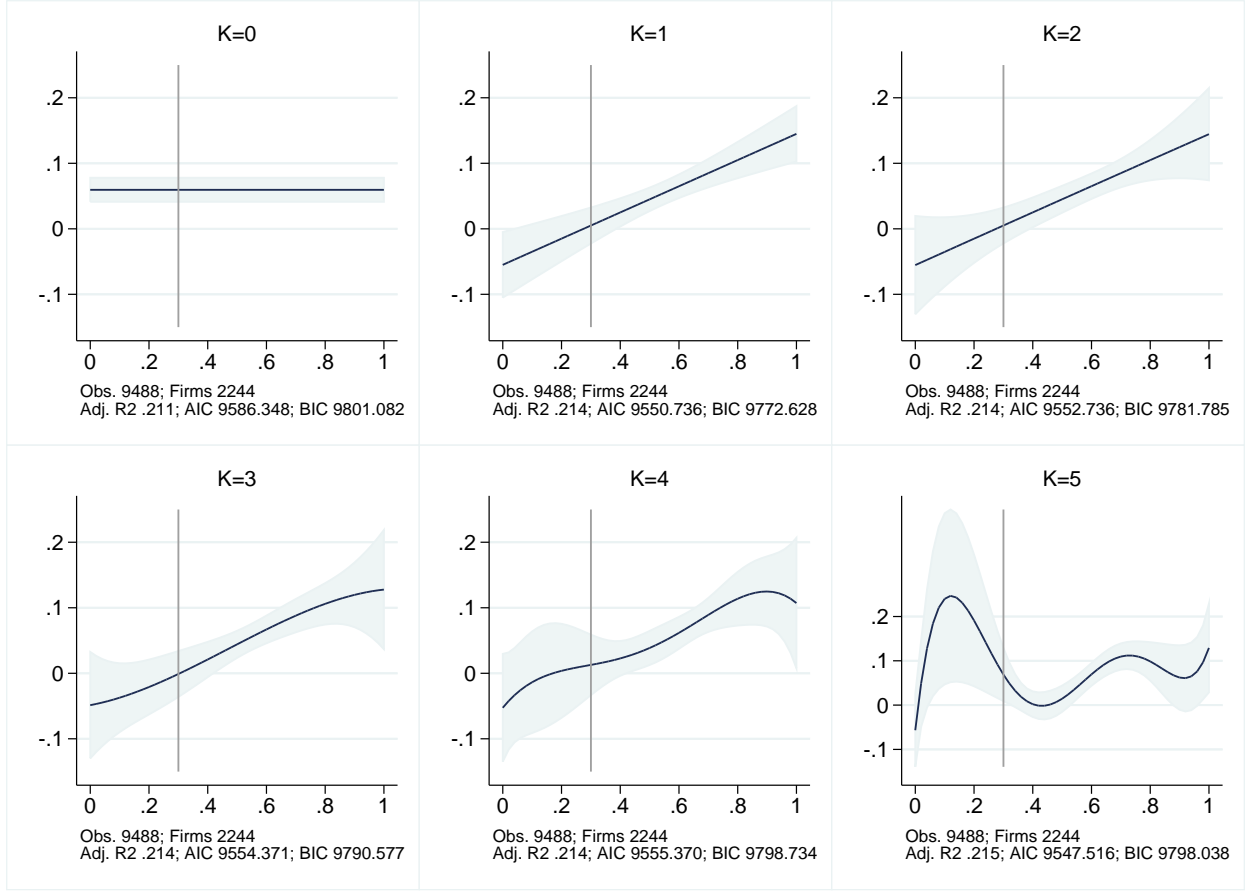[†]Note: This figure shows the estimated productivity effect at the offshoring margin of sourcing for $j = o$:

$$\frac{\partial \Pr(\widehat{\text{FOR}_o} = 1|\cdot)}{\partial \theta} = \widehat{\boldsymbol{\lambda}} \cdot \boldsymbol{\Gamma}^K(\zeta)$$

alongside 90% confidence intervals. Estimates of $\boldsymbol{\lambda}$ are obtained from estimating Equation (15) in the LPM. Sampling weights apply.

Figure 11. Productivity effect by quintiles of sourcing intensity[†]
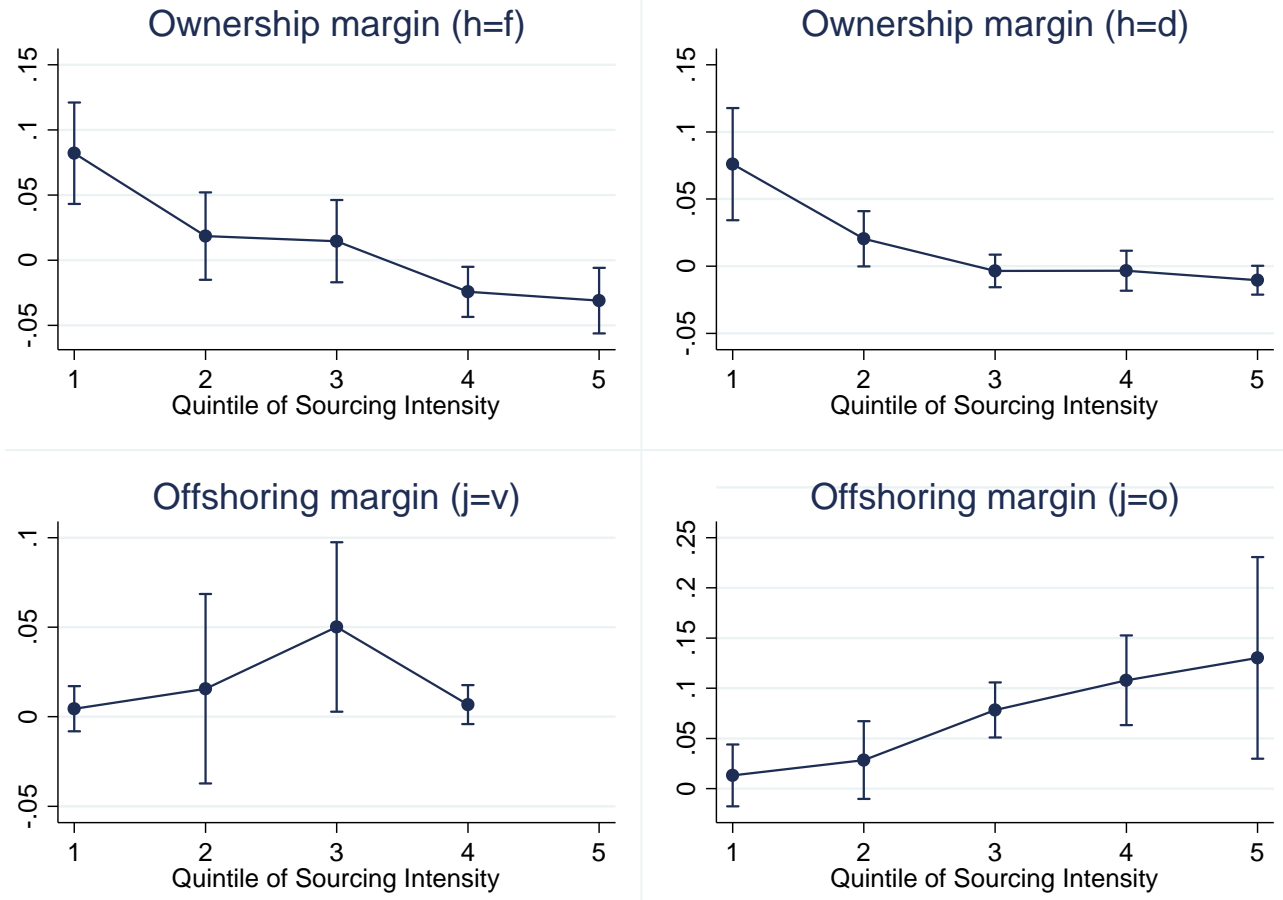
[†]Note: This figure shows the estimated productivity effect for both the ownership margin and the offshoring margin of sourcing at different quintiles of the industry's sourcing intensity (excluding the Beverages industry). The underlying models are given by

$$\Pr(\text{INT}_{it,h} = 1|\cdot) = \lambda_1 \cdot \theta_{it} + \sum_{q=2}^{5} \lambda_q \cdot (d(q; \zeta_s) \times \theta_{it}) + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, \ i \in \mathcal{I}_s, \Omega_{it} \in \Omega_h,$$

for the two subfigures at the top, and by

$$\Pr(\text{FOR}_{it,j} = 1|\cdot) = \lambda_1 \cdot \theta_{it} + \sum_{q=2}^{5} \lambda_q \cdot (d(q; \zeta_s) \times \theta_{it}) + \gamma_{s,h} + \gamma_{t,j} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, \ i \in \mathcal{I}_s, \Omega_{it} \in \Omega_j,$$

for the two subfigures at the bottom. All specifications include a firm's capital intensity, skill intensity, and export volume, along with industry fixed effects and year fixed effects. The samples employed are defined as: $\Omega_f = \{\langle \cdot, 1, \cdot, 1 \rangle\}$; $\Omega_d = \{\langle \cdot, 1, 0, 0 \rangle\}$; $\Omega_v = \{\langle 1, 1, ., 0 \rangle\}$; $\Omega_o = \{\langle 0, 1, 0, . \rangle\}$. Sampling weights apply.

Table 1. Regression analysis and model selection[†]

	Ownership margin				Offshoring margin			
	Sample: Ω_f; Dep. var.: INT_f		Sample: Ω_d; Dep. var.: INT_d		Sample: Ω_v; Dep. var.: FOR_v		Sample: Ω_o; Dep. var.: FOR_o	
	λ_K	Wald test	λ_K	Wald test	λ_K	Wald test	λ_K	Wald test
MODEL	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
K=0	0.0218**		0.0235**		0.0118		0.0595***	
	(0.00930)		(0.00945)		(0.00756)		(0.0112)	
K=1	-0.177***		-0.153**		0.0602**		0.200***	
	(0.0421)		(0.0643)		(0.0274)		(0.0520)	
K=2	-0.220	2.494	0.0282	0.0238	0.0103	0.0104	-0.00204	0.000160
	(0.139)	0.115	(0.183)	0.878	(0.100)	0.919	(0.161)	0.990
K=3	0.889**	3.953**	0.685	2.724*	-1.755*	1.688	-0.220	0.0911
	(0.401)	0.0194	(0.503)	0.0659	(0.957)	0.187	(0.519)	0.913
K=4	0.777	2.796**	-1.929	2.105*	-1.210	1.118	-1.683	0.190
	(1.679)	0.0390	(1.822)	0.0977	(12.70)	0.342	(2.645)	0.903
K=5	-18.38	1.989*	-10.67	1.669	-113.0	1.888	34.41**	1.535
	(15.17)	0.0938	(13.21)	0.154	(68.55)	0.113	(15.03)	0.189
Observations	5,513		6,484		701		9,488	
No. of firms	1,543		1,885		288		2,244	

[†]Note: This table presents regression results and test statistics for both the ownership margin and the offshoring margin of sourcing. The underlying models are given by

$$\Pr(\text{INT}_{it,h}|\cdot) = \boldsymbol{\lambda} \cdot \boldsymbol{\Gamma}^K(\zeta_s) \times \theta_{it} + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, i \in \mathcal{I}_s, \Omega_{it} \in \Omega_h,$$

in columns (1) to (4), and by

$$\Pr(\text{FOR}_{it,j}|\cdot) = \boldsymbol{\lambda} \cdot \boldsymbol{\Gamma}^K(\zeta_s) \times \theta_{it} + \gamma_{s,j} + \gamma_{t,j} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, i \in \mathcal{I}_s, \Omega_{it} \in \Omega_j,$$

in columns (5) to (8). All specifications include a firm's capital intensity, skill intensity, and export volume, along with industry fixed effects and year fixed effects. Each row represents a different model defined by the degree of the polynomial in $\boldsymbol{\Gamma}^K$ (running from $K = 0$ to $K = 5$). For each model, the coefficient reported is the estimate of λ_K (i.e., the coefficient of the highest-order indeterminate in the respective polynomial). Robust standard errors (clustered by firm) are given in parentheses. The samples employed are defined as: $\Omega_f = \{\langle \cdot, 1, \cdot, 1 \rangle\}$; $\Omega_d = \{\langle \cdot, 1, 0, 0 \rangle\}$; $\Omega_v = \{\langle 1, 1, ., 0 \rangle\}$; $\Omega_o = \{\langle 0, 1, 0, . \rangle\}$. We test the relevance of quadratic, cubic, and higher-order polynomials through Wald tests: the null hypothesis we employ is that all coefficients of indeterminates of order higher than 1 are jointly zero ($H_0 : \lambda_2 = \ldots = \lambda_K = 0$). The alternative hypothesis is that at least one of these coefficients is different from zero (H_1 : at least one $\lambda_k \neq 0$ with $k = 2, \ldots, K$). The table reports the corresponding F-statistics and two sided p-values in even-numbered columns. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

Table 2. Robustness analysis $(K = 1)^{\dagger}$

	Ownership margin				Offshoring margin			
	Sample: Ω_f; Dep. var.: INT$_f$		Sample: Ω_d; Dep. var.: INT$_d$		Sample: Ω_v; Dep. var.: FOR$_v$		Sample: Ω_o; Dep. var.: FOR$_o$	
	λ_0	λ_1	λ_0	λ_1	λ_0	λ_1	λ_0	λ_1
MODEL	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) Benchmark model	0.127***	-0.177***	0.106**	-0.153**	-0.0108*	0.0602**	-0.0550*	0.200***
	(0.0288)	(0.0421)	(0.0416)	(0.0643)	(0.00576)	(0.0274)	(0.0304)	(0.0520)
(B) Richer specification of fixed costs	0.0955***	-0.130***	0.104**	-0.154**	-0.0101	0.0539*	-0.0707**	0.220***
	(0.0273)	(0.0397)	(0.0431)	(0.0663)	(0.00733)	(0.0287)	(0.0296)	(0.0511)
(C) Region-specific effects	0.126***	-0.179***	0.109***	-0.160**				
	(0.0287)	(0.0418)	(0.0419)	(0.0648)				
(D) Additional firm-specific controls	0.0929***	-0.130***	0.109**	-0.162**	0.000721	0.0412	-0.0807***	0.250***
	(0.0275)	(0.0401)	(0.0490)	(0.0756)	(0.00944)	(0.0330)	(0.0310)	(0.0545)
(E) Exclusion of foreign-owned firms	0.0488***	-0.0639***	0.109**	-0.157**	-0.0101**	0.0372*	-0.0588*	0.209***
	(0.0166)	(0.0234)	(0.0427)	(0.0661)	(0.00447)	(0.0211)	(0.0316)	(0.0537)
(F) Trimming estimator (one iteration)	0.182***	-0.259***	0.146***	-0.210***	-0.0212**	0.132**	-0.0374	0.194***
	(0.0406)	(0.0572)	(0.0528)	(0.0790)	(0.0102)	(0.0543)	(0.0314)	(0.0546)
(G) Trimming estimator (two iterations)	0.198***	-0.286***	0.171***	-0.248***	-0.0299	0.187**	-0.0319	0.188***
	(0.0448)	(0.0628)	(0.0556)	(0.0836)	(0.0198)	(0.0915)	(0.0348)	(0.0591)
(H) Trimming estimator (three iterations)	0.171***	-0.263***	0.183***	-0.264***	-0.0522	0.267**	-0.0311	0.187***
	(0.0494)	(0.0698)	(0.0610)	(0.0917)	(0.0329)	(0.113)	(0.0349)	(0.0592)
(I) Probit estimator	0.632***	-0.921**	0.405**	-0.489	-0.475	1.469	-0.0974	0.588***
	(0.237)	(0.387)	(0.201)	(0.335)	(0.531)	(1.138)	(0.121)	(0.199)
(J) Outliers with respect to productivity	0.157***	-0.196***	0.148***	-0.212**	-0.0141**	0.0624**	-0.0532	0.220***
	(0.0385)	(0.0563)	(0.0540)	(0.0831)	(0.00695)	(0.0315)	(0.0377)	(0.0656)
(K) Outliers with respect to sourcing intensity	0.143***	-0.203***	0.119***	-0.170***	-0.0289	0.0920	-0.0541	0.197***
	(0.0324)	(0.0469)	(0.0371)	(0.0536)	(0.0294)	(0.0656)	(0.0377)	(0.0636)
(L) Alternative measure of productivity	0.290***	-0.368***	0.262***	-0.313***	2.25e-05	0.0886*	0.0433	0.146*
	(0.0513)	(0.0776)	(0.0544)	(0.0744)	(0.0137)	(0.0473)	(0.0507)	(0.0870)
(M) Alternative measure of sourcing intensity	0.127***	-0.169***	0.100**	-0.131**	-0.00833	0.0522*	-0.0352	0.155***
	(0.0289)	(0.0399)	(0.0392)	(0.0552)	(0.00608)	(0.0279)	(0.0284)	(0.0441)
(N) Sourcing shares as dependent variables	0.0541***	-0.0740***	0.0242**	-0.0311*	-0.00189	0.0253**	-0.00416	0.0340***
	(0.0130)	(0.0189)	(0.0113)	(0.0165)	(0.00223)	(0.0125)	(0.00708)	(0.0128)

†Note: This table presents the results of our robustness analysis for both the ownership margin and the offshoring margin of sourcing. Unless indicated otherwise, the underlying models are given by

$$\Pr(\text{INT}_{it,h} = 1|\cdot) = \lambda_0 \cdot \theta_{it} + \lambda_1 \cdot \theta_{it} \times \zeta_s + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, i \in \mathcal{I}_s, \Omega_{it} \in \Omega_h$$

in columns (1) to (4), and by

$$\Pr(\text{FOR}_{it,j} = 1|\cdot) = \lambda_0 \cdot \theta_{it} + \lambda_1 \cdot \theta_{it} \times \zeta_s + \gamma_{s,j} + \gamma_{t,j} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, i \in \mathcal{I}_s, \Omega_{it} \in \Omega_j$$

in columns (5) to (8). (A) Benchmark model as in Figures 7 to 10 with $K = 1$. (B) We include industry-and-year fixed effects as well as size group fixed effects. (C) We include dummy variables for import regions. (D) We include as additional time-varying firm-level controls: age; employment; R&D intensity; service intensity; debt ratio; capacity utilization; and fixed effects for the type of good produced (consumption, intermediate, or undefined). (E) We exclude all firms where foreign investors own more than 50% of the firm's joint capital. (F)-(H) We exclude all observations where the linear prediction lies outside the unit interval and re-estimate the model. (I) We use the Probit estimator. (J) We drop in each industry-year pair the top and bottom 1% of productivity. (K) We drop the Beverages industry ($\zeta = 0$). (L) We use total factor productivity à la Olley & Pakes (1996) rather than labor productivity. (M) To construct the sourcing intensity (normalized to the unit interval), we use the log of the mean capital intensity (rather than the median) observed in a particular industry over the period 2000-2012. (N) We use the share of sourcing values as dependent variables (rather than dummy variables): $\left(\frac{\text{Intra-firm sourcing}_{it,h}}{\text{Total sourcing}_{it,h}}\right)$ in columns (1) to (4); $\left(\frac{\text{Foreign sourcing}_{it,j}}{\text{Total sourcing}_{it,j}}\right)$ in columns (5) to (8). All specifications include (as control variables) a firm's capital intensity, skill intensity, and export volume, along with industry fixed effects and year fixed effects. Robust standard errors (clustered by firm) are given in parentheses. The samples employed are defined as: $\Omega_f = \{\langle\cdot, 1, \cdot, 1\rangle\}$; $\Omega_d = \{\langle\cdot, 1, 0, 0\rangle\}$; $\Omega_v = \{\langle 1, 1, \cdot, 0\rangle\}$; $\Omega_o = \{\langle 0, 1, 0, \cdot\rangle\}$. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

Table 3. Ownership margin: Number of affiliated firms by sourcing location[†]

	Sample: Offshoring firms (Ω_f)				
	Dependent variable: Number of affiliated firms				
	Total	EU	Latin America	Rest of the OECD	Rest of the World
VARIABLES	(1)	(2)	(3)	(4)	(5)
Productivity θ	0.644***	0.305***	0.112*	0.0653**	0.161**
	(0.207)	(0.116)	(0.0604)	(0.0264)	(0.0758)
Productivity θ x	-0.825***	-0.397**	-0.148*	-0.0673	-0.213**
Sourcing intensity ζ	(0.298)	(0.168)	(0.0867)	(0.0413)	(0.0951)
Observations	5,513	5,513	5,513	5,513	5,513
No. of firms	1,543	1,543	1,543	1,543	1,543
R^2	0.051	0.051	0.039	0.045	0.012

[†]Note: This table presents regression results from estimating the following equation:

$$\text{\#Foreign-affiliates}_{it,f} = \lambda_0 \cdot \theta_{it} + \lambda_1 \cdot (\zeta_s \times \theta_{it}) + \gamma_{s,f} + \gamma_{t,f} + \boldsymbol{\beta} \cdot \mathbf{X}_{it} + \xi_{it,f}, \ i \in \mathcal{I}_s, \Omega_{it} \in \Omega_f,$$

where all specifications include a firm's capital intensity, skill intensity, and export volume, along with industry fixed effects and year fixed effects. Robust standard errors (clustered by firm) are given in parentheses. The sample employed is defined as: $\Omega_f = \{\langle \cdot, 1, \cdot, 1 \rangle\}$. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

Table 4. Ownership margin: Further industry interactions[†]

	Sample: Offshoring firms (Ω_f); Dep. var.: INT_f				Sample: Domestically sourcing firms (Ω_d); Dep. var.: INT_d			
	$Y \in \{Y_{Skill}, Y_{R\&D}, Y_{Service}, Y_{Ext.Fin.Dep.}\}$		$\mathbf{Y} = (Y_{Skill}, Y_{R\&D}, Y_{Service}, Y_{Ext.Fin.Dep.})'$		$Y \in \{Y_{Skill}, Y_{R\&D}, Y_{Service}, Y_{Ext.Fin.Dep.}\}$		$\mathbf{Y} = (Y_{Skill}, Y_{R\&D}, Y_{Service}, Y_{Ext.Fin.Dep.})'$	
	λ_1	λ_Y	λ_1	λ_Y	λ_1	λ_Y	λ_1	λ_Y
INDUSTRY VARIABLE (Y)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Skill intensity	-0.116***	0.250***	-0.0817*	0.304***	-0.152**	0.00465	-0.174***	-0.0526
	(0.0378)	(0.0557)	(0.0450)	(0.0935)	(0.0610)	(0.0721)	(0.0625)	(0.115)
Observations	5,513				6,484			
No. of firms	1,543				1,885			
R^2	0.155				0.079			
R&D intensity	-0.204***	2.970***		-0.639	-0.158**	1.038		1.788
	(0.0420)	(0.926)		(1.270)	(0.0656)	(1.050)		(1.547)
Observations	5,513				6,484			
No. of firms	1,543				1,885			
R^2	0.149				0.080			
Service intensity	-0.197***	0.0146***		-0.00197	-0.158**	0.00210		0.00158
	(0.0426)	(0.00413)		(0.00468)	(0.0647)	(0.00409)		(0.00475)
Observations	5,513				6,484			
No. of firms	1,543				1,885			
R^2	0.151				0.079			
Ext. Fin. Dep.	-0.215***	0.105*		-0.0398	-0.151**	-0.00535		-0.00999
	(0.0438)	(0.0586)		(0.0726)	(0.0696)	(0.0547)		(0.0523)
Observations	5,385		5,385		6,200		6,200	
No. of firms	1,501		1,501		1,814		1,814	
R^2	0.149		0.155		0.080		0.080	

[†]Note: This table presents regression results from estimating equations of the following form:

$$\Pr(\text{INT}_{it,h} = 1 | \cdot) = \lambda_0 \cdot \theta_{it} + \lambda_1 \cdot \theta_{it} \times \zeta_s + \lambda_Y \cdot \theta_{it} \times Y_s + \gamma_{s,h} + \gamma_{t,h} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, i \in \mathcal{I}_s, \Omega_{it} \in \Omega_h,$$

where Y_s is the industry variable indicated on the left-hand side. In columns (3) & (4) and (7) & (8), we bring in all interactions at the same time ($\theta_{it} \times \mathbf{Y}_s$). Estimates of λ_0 not reported. All specifications include a firm's capital intensity, skill intensity, and export volume, along with industry fixed effects and year fixed effects. Robust standard errors (clustered by firm) are given in parentheses. The samples employed are defined as: $\Omega_f = \{\langle \cdot, 1, \cdot, 1 \rangle\}$; $\Omega_d = \{\langle \cdot, 1, 0, 0 \rangle\}$. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

Table 5. Offshoring margin: Further industry interactions†

INDUSTRY VARIABLE (Y)	Sample: Integrating firms (Ω_v); Dep. var.: FOR$_v$				Sample: Outsourcing firms (Ω_o); Dep. var.: FOR$_o$			
	Y∈{Y$_{Skill}$,Y$_{R\&D}$,Y$_{Service}$,Y$_{Ext.Fin.Dep.}$}		**Y**=(Y$_{Skill}$,Y$_{R\&D}$,Y$_{Service}$,Y$_{Ext.Fin.Dep.}$)'		Y∈{Y$_{Skill}$,Y$_{R\&D}$,Y$_{Service}$,Y$_{Ext.Fin.Dep.}$}		**Y**=(Y$_{Skill}$,Y$_{R\&D}$,Y$_{Service}$,Y$_{Ext.Fin.Dep.}$)'	
	λ_1	λ_Y	λ_1	$\boldsymbol{\lambda_Y}$	λ_1	λ_Y	λ_1	$\boldsymbol{\lambda_Y}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Skill intensity	0.0720**	0.0487	-0.217	-0.288	0.197***	-0.0148	0.103	-0.226
	(0.0319)	(0.0533)	(0.149)	(0.176)	(0.0555)	(0.0899)	(0.0791)	(0.147)
Observations	701				9,488			
No. of firms	288				2,244			
R^2	0.112				0.216			
R&D intensity	0.0441	2.329		0.432	0.189***	1.548		3.418
	(0.0288)	(1.943)		(2.658)	(0.0515)	(1.775)		(2.516)
Observations	701				9,488			
No. of firms	288				2,244			
R^2	0.114				0.217			
Service intensity	0.0292	0.00799*		0.0245*	0.186***	0.00578		0.00914
	(0.0288)	(0.00479)		(0.0129)	(0.0517)	(0.00691)		(0.00769)
Observations	701				9,488			
No. of firms	288				2,244			
R^2	0.114				0.217			
Ext. Fin. Dep.	0.0136	0.114		0.263	0.199***	-0.00815		0.0143
	(0.0435)	(0.107)		(0.181)	(0.0582)	(0.0743)		(0.0843)
Observations	634		634		9,213		9,213	
No. of firms	263		263		2,181		2,181	
R^2	0.128		0.139		0.214		0.215	

†Note: This table presents regression results from estimating equations of the following form:

$$\Pr(\text{FOR}_{it,j} = 1|\cdot) = \lambda_0 \cdot \theta_{it} + \lambda_1 \cdot \theta_{it} \times \zeta_s + \lambda_Y \cdot \theta_{it} \times Y_s + \gamma_{s,j} + \gamma_{t,j} + \boldsymbol{\beta} \cdot \mathbf{X}_{it}, i \in \mathcal{I}_s, \Omega_{it} \in \Omega_j,$$

where Y_s is the industry variable indicated on the left-hand side. In columns (3) & (4) and (7) & (8), we bring in all interactions at the same time ($\theta_{it} \times \mathbf{Y}_s$). Estimates of λ_0 not reported. All specifications include a firm's capital intensity, skill intensity, and export volume, along with industry fixed effects and year fixed effects. Robust standard errors (clustered by firm) are given in parentheses. The samples employed are defined as: $\Omega_v = \{\langle 1, 1, \cdot, 0\rangle\}$; $\Omega_o = \{\langle 0, 1, 0, \cdot\rangle\}$. *,**,*** denote significance at the 10%, 5%, 1% levels, respectively.

Appendices

A General equilibrium interrelationship

In the definition of maximum profits in Equations (1) through (5), the term A captures the general equilibrium interrelationship between sectors. In this appendix we provide more details on the exact meaning of this term in a model of the type considered in Antràs & Helpman (2004).

For easier notation, we abstain from indexing sectors. Suppose that there are quasi-linear preferences over sectoral aggregates X, composed in a CES way of differentiated varieties of final goods, such as the ones produced in the production relationship considered above. Inverse demand of the consumption aggregate is given by $P = X^{\kappa-1}$, where κ is a utility parameter satisfying $0 < \kappa < 1$. Defining $\alpha := (\varepsilon - 1)/\varepsilon$, the term A then emerges as

$$A := X^{*(\kappa-\alpha)/(1-\alpha)}\alpha^{\alpha/(1-\alpha)}, \tag{A.1}$$

where X^* is the *general equilibrium quantity* of aggregate consumption of the differentiated final goods. This is easily verified from Equations (6) and (7) in Antràs & Helpman (2004). Equilibrium sectoral expenditure may be written as $R^* := X^* P^* = X^{*\kappa}$, whence A may equivalently be written as $R^{*\varepsilon+(1-\varepsilon)/\kappa}\alpha^{\alpha/(1-\alpha)}$.

In Antràs & Helpman (2004), X^* is determined by a zero profit condition, assuming free and costless entry. Note that this type of general equilibrium closure of the model assumes $\alpha > \kappa$ (higher degree of substitution within than between sectors). This implies that maximum profits are falling in X^* (or R^*). The intuition is that a higher X^* is generated through firm entry, so that output (and thus revenue as well as profits) per firm is reduced. This guarantees that entry caused by positive profits eventually leads to a zero profit equilibrium.

B Proofs of propositions

To prove part (b) of Propositions 1 and 2, it proves convenient to introduce the notions of supermodularity and submodularity of functions.

Definition 1. (a) *The function $H(g, q)$ is called supermodular with respect to g and q, if for any two values $g_1 > g_0$ and $q_1 > q_0$ the following is true: $\Delta_g H(q_1) > \Delta_g H(q_0)$, where $\Delta_g H(q) := H(g_1, q) - H(g_0, q)$.* **(b)** *The function $H(g, q)$ is called submodular, if $-H(g, q)$ is supermodular with respect to g and q.* **(c)** *If $H(g, q)$ is twice differentiable, then it is called supermodular, if $\partial^2 H/(\partial g \partial q) > 0$, and vice versa for submodularity.*

B.1 Proof of Proposition 1, part (b)

The productivity effect is equal to $(\varepsilon - 1)\Delta_m Z(\ell_h; \zeta)\theta^{\varepsilon-2}$. This effect being more favorable to outsourcing means that $\Delta_m Z(\ell_h; \zeta)$ is falling in ζ, i.e., that the function $Z(\cdot)$ is submodular in m and ζ. Part (b) of the proposition states that submodularity holds true for values $\zeta \in [\underline{\zeta}, \overline{\zeta}]$, whereby this interval always includes ζ^*, the sourcing intensity implicitly defined by $\Delta_m Z(\ell_h; \zeta^*) = 0$.

We know from Lemma 3 in Antràs (2003) that $Z(\ell_h, m_v; \zeta)/Z(\ell_h, m_o; \zeta)$ is decreasing monotonically in ζ. We write $Z_v(\cdot, \zeta)$ for $Z(\ell_h, m_v; \zeta)$ and accordingly for $Z_o(\cdot, \zeta)$. Then

$$\frac{\partial}{\partial \zeta}\left[\frac{Z_v(\cdot, \zeta)}{Z_o(\cdot, \zeta)}\right] = \frac{1}{Z_o(\cdot, \zeta)}\left[\frac{\partial Z_v(\cdot, \zeta)}{\partial \zeta} - \frac{Z_v(\cdot, \zeta)}{Z_o(\cdot, \zeta)}\frac{\partial Z_o(\cdot, \zeta)}{\partial \zeta}\right] < 0 \tag{B.2}$$

By definition, we have $\frac{Z_v(\cdot,\zeta^*)}{Z_o(\cdot,\zeta^*)} = 1$, whence

$$\frac{\partial Z_v(\cdot,\zeta^*)}{\partial \zeta} - \frac{\partial Z_o(\cdot,\zeta^*)}{\partial \zeta} < 0. \tag{B.3}$$

The definition of $\Delta_m Z(\ell_h;\zeta)$ then implies that

$$\frac{\partial}{\partial \zeta}\Delta_m Z(\ell_h;\zeta^*) = \left[\frac{\partial Z_v(\cdot,\zeta^*)}{\partial \zeta} - \frac{\partial Z_o(\cdot,\zeta^*)}{\partial \zeta}\right] < 0. \tag{B.4}$$

This proves that in the neighborhood of ζ^* the function $\Delta_m Z(\ell_h;\zeta)$ is unambiguously falling in ζ. In other words, the interval $[\underline{\zeta},\overline{\zeta}]$ includes ζ^*.

Proposition 1(b) also states that the interval $[\underline{\zeta},\overline{\zeta}]$ is large for plausible parameter constellations. In view of definition 1, we invoke twice differentiability of the maximum profit function, which allows us to prove this by investigating $\frac{\partial^2 Z(\cdot,m)}{\partial m \partial \zeta} = \frac{\partial^2 Z(\cdot,\zeta)}{\partial \zeta \partial m}$, this latter equality following from Young's theorem. The interval $[\underline{\zeta},\overline{\zeta}]$ is defined such that for $\zeta \in [\underline{\zeta},\overline{\zeta}]$ the cross-derivative $\frac{\partial^2 Z}{\partial \zeta \partial m}$ is negative. It proves convenient to work with $\ln Z$, using the decomposition

$$\frac{\partial^2 Z}{\partial \zeta \partial m} = Z\frac{\partial^2 \ln Z}{\partial \zeta \partial m} + \frac{\partial \ln Z}{\partial \zeta}\frac{\partial Z}{\partial m}. \tag{B.5}$$

Remember that

$$Z = AzC, \tag{B.6}$$

$$\text{with} \quad z = 1 - \frac{\varepsilon-1}{\varepsilon}\left[m(1-\zeta) + (1-m)\zeta\right] \tag{B.7}$$

$$\text{and} \quad C = \left[m^{1-\zeta}\left[\ell(1-m)\right]^\zeta\right]^{\varepsilon-1}. \tag{B.8}$$

To explore the interval $[\underline{\zeta},\overline{\zeta}]$, we proceed in three steps. First, we show that $\frac{\partial^2 \ln Z}{\partial \zeta \partial m}$ is always negative. Second, we show that each of the two terms $\frac{\partial \ln Z}{\partial \zeta}$ and $\frac{\partial Z}{\partial m}$ is ambiguous, that $\frac{\partial \ln Z}{\partial \zeta}\frac{\partial Z}{\partial m}$ is ambiguous as well, and that positive values of $\frac{\partial \ln Z}{\partial \zeta}\frac{\partial Z}{\partial m}$ potentially dominate the sign of (B.5). We show that this holds true in particular for large values of m. This will prove that submodularity of $Z(\cdot)$ in m and ζ does not obtain throughout the entire parameter space of the model. The third step explores the sign of $\frac{\partial^2 Z}{\partial \zeta \partial m}$ by means of numerical simulations, in order to see how the values of $\underline{\zeta}$ and $\overline{\zeta}$ vary over the entire range of values for $m \in [0,1]$ as well as with different values of ε.

The first term on the right-hand side of (B.5) may be written as

$$\frac{\partial \ln Z}{\partial \zeta} = \frac{\partial \ln z}{\partial \zeta} + \frac{\partial \ln C}{\partial \zeta}. \tag{B.9}$$

In view of (B.6), note that the partial equilibrium nature of our analysis allows us to set $\frac{\partial \ln A}{\partial \zeta}$ equal to zero. We have

$$\frac{\partial \ln C}{\partial \zeta} = (\varepsilon-1)\frac{\partial}{\partial \zeta}\left[(1-\zeta)\ln m + \zeta\left[\ln \ell + \ln(1-m)\right]\right] \tag{B.10}$$

$$= (\varepsilon-1)\left[\ln \ell + \ln(1-m) - \ln m\right]. \tag{B.11}$$

$$\frac{\partial \ln z}{\partial \zeta} = \frac{1}{z}\frac{\partial z}{\partial \zeta} = \frac{1}{z}\left[-\frac{\varepsilon-1}{\varepsilon}(1-m-m)\right] \tag{B.12}$$

$$= \frac{\varepsilon-1}{\varepsilon}\frac{2m-1}{z}. \tag{B.13}$$

37

Bringing the terms together, we obtain

$$\frac{\partial \ln Z}{\partial \zeta} = \frac{\varepsilon - 1}{\varepsilon} \frac{2m - 1}{z} + (\varepsilon - 1) \left[\ln \ell + \ln(1 - m) - \ln m \right]. \tag{B.14}$$

$$\frac{\partial^2 \ln Z}{\partial \zeta \partial m} = \frac{\varepsilon - 1}{\varepsilon} \left[\frac{2}{z} - \frac{2m - 1}{z^2} \frac{\partial z}{\partial m} \right] + (\varepsilon - 1) \left(\frac{-1}{1 - m} - \frac{1}{m} \right). \tag{B.15}$$

We may write $\left(\frac{-1}{1-m} - \frac{1}{m} \right) = \frac{-1}{m - m^2}$ and

$$\frac{\partial z}{\partial m} = -\frac{\varepsilon - 1}{\varepsilon} (1 - \zeta - \zeta) = \frac{\varepsilon - 1}{\varepsilon} (2\zeta - 1). \tag{B.16}$$

Inserting into (B.15) we obtain

$$\frac{\partial^2 \ln Z}{\partial \zeta \partial m} = -\frac{(\varepsilon - 1)}{m - m^2} + \frac{\varepsilon - 1}{\varepsilon} \frac{1}{z^2} \left[2z - (2m - 1) \frac{\varepsilon - 1}{\varepsilon} (2\zeta - 1) \right] \tag{B.17}$$

$$= -\frac{(\varepsilon - 1)}{m - m^2} + \frac{1}{z^2} \left[2\alpha z - \alpha^2 (2m - 1)(2\zeta - 1) \right], \tag{B.18}$$

where $\alpha := \frac{\varepsilon - 1}{\varepsilon}$. Inserting for z it is straightforward to show that the bracketed term in (B.18) reduces to $\frac{2\alpha - \alpha^2}{z^2}$, so that

$$\frac{\partial^2 \ln Z}{\partial \zeta \partial m} = -\frac{(\varepsilon - 1)}{m - m^2} + \frac{2\alpha - \alpha^2}{z^2} = - \left[\frac{(\varepsilon - 1)}{m - m^2} + \frac{\alpha^2 - 2\alpha}{z^2} \right] \tag{B.19}$$

This term is strictly negative. To see this, recall that α, m, and ζ all lie strictly between zero and one. Moreover, note that the first term in the square bracket on the right-hand side (call it x) is strictly positive, and the second term (call it y) is strictly negative. For any given value of α, the values of m and ζ that minimize the denominator in y, and thus maximize the (absolute) value of the function y, are $m = 1/2$ and $\zeta = 1/2$. In turn, for any given value of α the value of m that minimizes the value of the function x is $m = 1/2$. Hence, a sufficient condition for the right-hand side of (B.19) to be negative is

$$\frac{4\alpha}{1 - \alpha} + \frac{\alpha^2 - 2\alpha}{(1 - \alpha/2)^2} > 0. \tag{B.20}$$

Straightforward manipulation of this expression yields $2\alpha - \alpha^2 > 0$, which is always true.

The next step now looks at the second term in our main decomposition (B.5), $\frac{\partial \ln Z}{\partial \zeta} \frac{\partial Z}{\partial m}$. The term $\frac{\partial \ln Z}{\partial \zeta}$ is given in (B.14) above. On account of $\ln(1 - m)$, this term converges to minus infinity as m approaches one. Conversely, it converges to infinity as m approaches zero (on account of $-\ln m$). It will also be positive for large enough values of ℓ.

The term $\frac{\partial Z}{\partial m}$ may be written as $\frac{\partial Z}{\partial m} = A \left[C \frac{\partial z}{\partial m} + z \frac{\partial C}{\partial m} \right]$, where $\frac{\partial z}{\partial m}$ is given in (B.16). We have

$$\frac{\partial C}{\partial m} = (\varepsilon - 1) \left[m^{1-\zeta} \left[\ell(1 - m) \right]^\zeta \right]^{\varepsilon - 2}$$

$$\times \left[(1 - \zeta) m^{-\zeta} \left[\ell(1 - m) \right]^\zeta - m^{1-\zeta} \zeta \left[\ell(1 - m) \right]^{\zeta - 1} \ell \right] \tag{B.21}$$

$$= (\varepsilon - 1) C \left(\frac{1 - \zeta}{m - m^2} - \frac{1}{1 - m} \right) \tag{B.22}$$

Putting things together, we have

$$\frac{\partial Z}{\partial m} = AC \left[\frac{\varepsilon - 1}{\varepsilon} (2\zeta - 1) + z(\varepsilon - 1) \left(\frac{1 - \zeta}{m - m^2} - \frac{1}{1 - m} \right) \right] \tag{B.23}$$

$$= AC(\varepsilon - 1) \left[\frac{2\zeta - 1}{\varepsilon} + \frac{z}{1 - m} \left(\frac{1 - \zeta - m}{m} \right) \right] \tag{B.24}$$

38

This term is again ambiguous. It is positive for sufficiently low values of m and negative for sufficiently high values of m. Returning to the product $\frac{\partial \ln Z}{\partial \zeta} \frac{\partial Z}{\partial m}$, we may now state that this is positive for sufficiently low as well as for sufficiently high values of m. Low values render both terms positive, whereas high values render both terms negative. Since the possible values of $\frac{\partial \ln Z}{\partial \zeta} \frac{\partial Z}{\partial m}$ include infinity, it follows that the second term in (B.5) potentially dominates the first in leading to a positive value of $\frac{\partial^2 Z}{\partial \zeta \partial m}$. Thus, submodularity with respect to m and ζ does not hold for the function $Z(\cdot)$ as a whole, i.e., for all possible parameter values.

In the final step of the proof, we now turn to a numerical simulation of $\frac{\partial^2 Z}{\partial \zeta \partial m}$ as given in (B.5) in order to demonstrate that for a plausible range of parameter values the interval $\left[\underline{\zeta}, \overline{\zeta} \right]$ is large. Figure B.1 looks at the entire parameter space for $\zeta \in [0, 1]$ and $m \in [0, 1]$, identifying subspaces of submodularity ($\frac{\partial^2 Z}{\partial \zeta \partial m} < 0$) and supermodularity ($\frac{\partial^2 Z}{\partial \zeta \partial m} > 0$) as well as the boundaries ($\frac{\partial^2 Z}{\partial \zeta \partial m} = 0$). We have conducted such simulations for five values of $\varepsilon = 2, 4, 6, 8, 10$, which cover the range of values usually considered in the literature for this elasticity. Figure B.1 reveals that violation of submodularity is more than a peripheral phenomenon. However, at least for moderate values of the elasticity of substitution in demand, the interval of submodularity, $\left[\underline{\zeta}, \overline{\zeta} \right]$ covers a large part of the unit interval. Submodularity is violated for low (high) values of m and low (high) values of ζ.

In concluding the proof of Proposition 1 (b), we must recognize that in the AH model there are but two relevant values of m, viz. m_o (for outsourcing) and m_v (for vertical integration). For the productivity effect at the ownership margin to behave as stated in the proposition, it is sufficient (though not necessary) that both values of m lie in the submodularity subspace identified in Figure B.1.

Figure B.1. Proposition 1, part (b): Modularity of $Z(\cdot)$ with respect to m and $\zeta^\dagger$
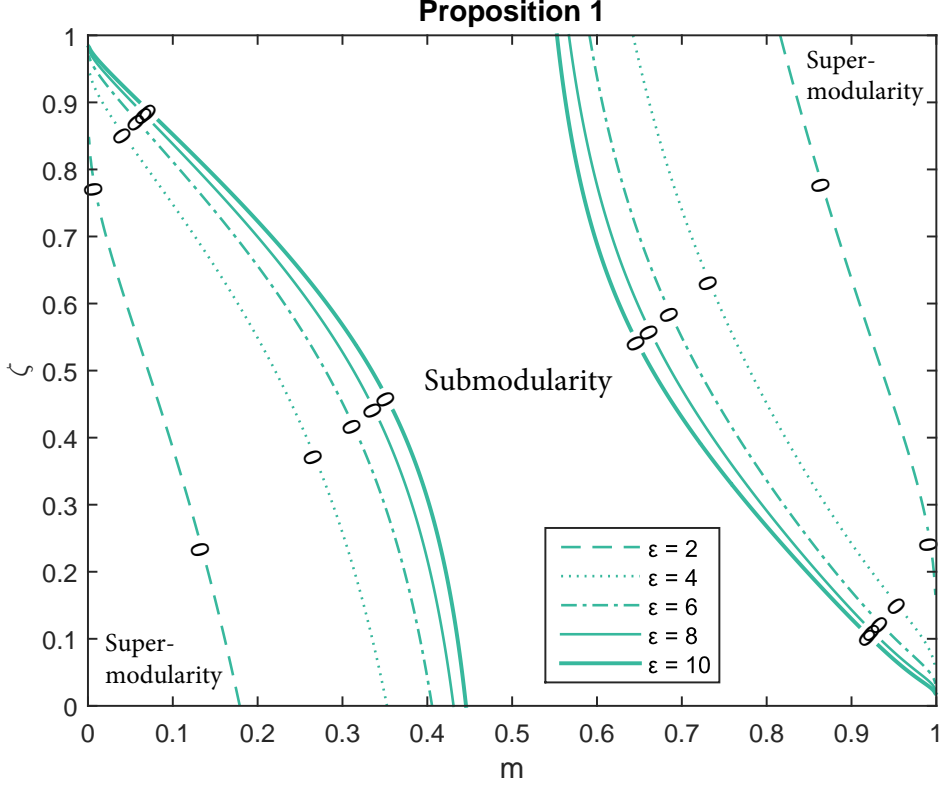
†Note: This figure displays the parameter subspaces of the model (in terms of m and ζ) for which unambiguous modularity properties obtain for $Z(\cdot)$ with respect to m and ζ. The figure assumes different values of ε and fixes $A = 1$ as well as $\ell = 1$.

B.2 Proof of Proposition 2, part (b)

The productivity effect is equal to $(\varepsilon - 1)\Delta_\ell Z(m;\zeta)\theta^{-2}$. It is more favorable to offshoring with a higher sourcing intensity, if $\Delta_\ell Z(m;\zeta)$ is rising in ζ, i.e., if the function $Z(\cdot)$ is supermodular in ℓ and ζ. Part (b) of Proposition 2 states that supermodularity does not obtain for the function as a whole, but only for a subspace of parameter values. More specifically, it states that for a sub-range of values $m \in [0, \tilde{m}]$ the function $Z(\ell, m; \zeta)$ is supermodular in ℓ and ζ, where $\tilde{\zeta}$ is a function of ℓ and ε. For values $m \in [\tilde{m}, 1]$, supermodularity obtains within an interval $\zeta \in [0, \tilde{\zeta}]$ while submodularity obtains for $\zeta \in [\tilde{\zeta}, 1]$. Ambiguous modularity implies that the productivity effect varies in a non-monotonic way across industries with different values of ζ.

Invoking (1) and twice differentiability of $Z(\cdot)$, we prove these statements by investigating $\frac{\partial^2 Z}{\partial \zeta \partial \ell}$, using the decomposition

$$\frac{\partial^2 Z}{\partial \zeta \partial \ell} = Z \frac{\partial^2 \ln Z}{\partial \zeta \partial \ell} + \frac{\partial \ln Z}{\partial \zeta} \frac{\partial Z}{\partial \ell}. \tag{B.25}$$

As with the proof of Proposition 1(b), we start by establishing that the first term in this decomposition is unambiguous and then turn to the ambiguity in the second term. From (B.14), we derive

$$\frac{\partial^2 \ln Z}{\partial \zeta \partial \ell} = \frac{\varepsilon - 1}{\ell} > 0. \tag{B.26}$$

40

Again, the partial equilibrium nature of our analysis allows us to keep A constant. The inequality follows from the assumption that $\varepsilon > 1$. Remember that ℓ is defined as the *inverse* unit cost of the manufacturing component, with the unit cost of the headquarter input normalized to one. Obviously, $\ln Z$ increases with ℓ: other things equal, a lower cost per unit of the manufacturing component increases maximum profits. Young's theorem implies $\frac{\partial^2 \ln Z}{\partial \zeta \partial \ell} = \frac{\partial^2 \ln Z}{\partial \ell \partial \zeta}$. Hence, the above inequality simply states that the effect of a lower cost of the manufacturing component on log-profits increases with the sourcing intensity ζ. The first term in (B.25) thus works in favor of supermodularity.

Turning to the second term, $\frac{\partial Z}{\partial \ell}$ may be written as

$$\frac{\partial Z}{\partial \ell} = \frac{\partial \ln Z}{\partial \ln \ell} \frac{Z}{\ell} = \frac{\partial \ln C}{\partial \ln \ell} \frac{Z}{\ell} = (\varepsilon - 1)\zeta \frac{Z}{\ell} > 0. \tag{B.27}$$

Clearly, a lower cost of the manufacturing input raises profits. However, the term $\frac{\partial \ln Z}{\partial \zeta}$ as given in (B.14) above is ambiguous, as emphasized when proving Proposition 1(b) above.

Inserting (B.26) and (B.27) as well as (B.14) into (B.25), we finally obtain

$$\frac{\partial^2 Z}{\partial \zeta \partial \ell} = \frac{(\varepsilon - 1)Z}{\ell} \left[1 + (\varepsilon - 1) \left[\zeta \frac{2m - 1}{z\varepsilon} + \ln \ell + \ln(1 - m) - \ln m \right] \right]. \tag{B.28}$$

Since $-\ln m$ converges to infinity and $\ln(1 - m)$ converges to zero as m approaches zero, there is a critical value $\tilde{m}$, such that for $m < \tilde{m}$ supermodularity ($\frac{\partial^2 Z}{\partial \zeta \partial \ell} > 0$) obtains for all values of ζ. Other things equal, $\tilde{m}$ is increasing in ℓ.

We prove the remaining part of Proposition 2(b), relating to $\tilde{\zeta}$, by means of numerical simulations. Figure B.2 looks at the entire parameter space for $\zeta \in [0, 1]$ and $m \in [0, 1]$, plotting lines satisfying $\frac{\partial^2 Z}{\partial \zeta \partial \ell} = 0$ for alternative values of ε. These lines separate the parameter space into subs-paces where supermodularity and submodularity, respectively, of $Z(\cdot)$ in ℓ and ζ obtains. Again, we have conducted the simulations for values $\varepsilon = 2, 4, 6, 8, 10$. The lines in Figure B.2 complete the proof of Proposition 2(b). While supermodularity clearly dominates the picture, submodularity is not a negligible phenomenon. It is pervasive for high values of m, particularly if the elasticity of substitution in demand is large.

In completing the proof, we must finally recognize that the AH model implies a discrete nature of m, with realizations m_o (for outsourcing) and m_v (for vertical integration). For the productivity effect to behave as stated in part (b) of Proposition 2, it is sufficient (though not necessary) that both values m_o and m_v lie in the supermodularity subspace identified in Figure B.2.

Figure B.2. Proposition 2, part (b): Modularity of $Z(\cdot)$ with respect to ℓ and $\zeta^\dagger$
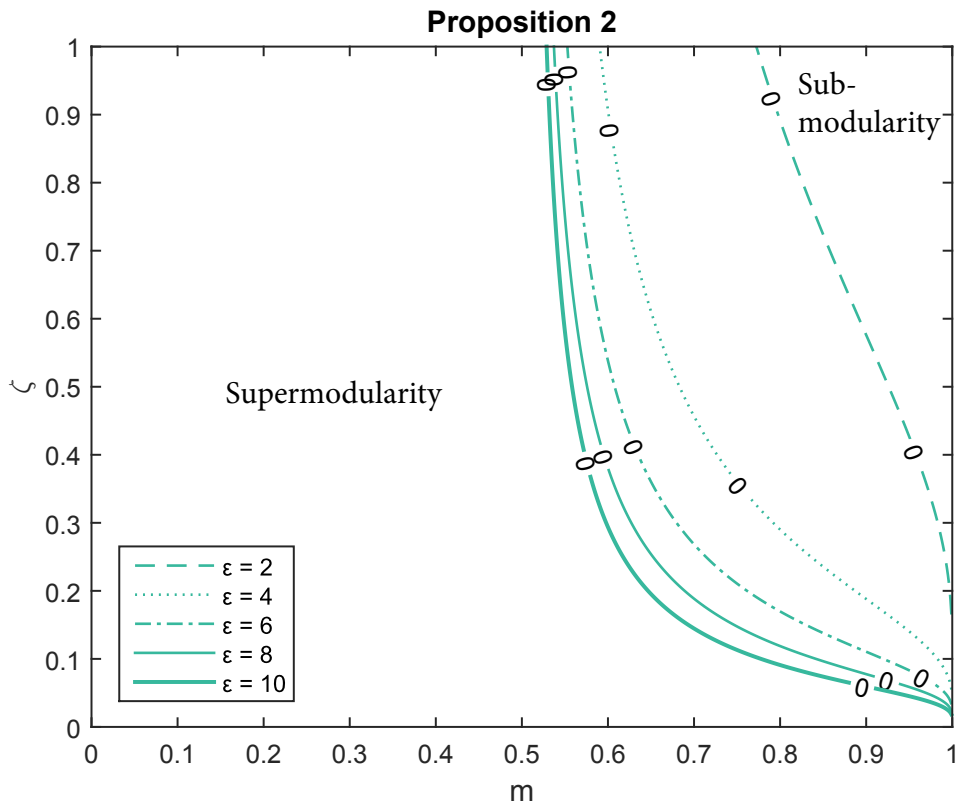
†Note: This figure displays the parameter subspaces of the model (in terms of m and ζ) for which unambiguous modularity properties obtain for $Z(\cdot)$ with respect to ℓ and ζ. The figure assumes different values of ε and fixes $A = 1$ as well as $\ell = 1$.

C Data appendix

Figure C.1. Histogram for the sourcing intensity of production[†]
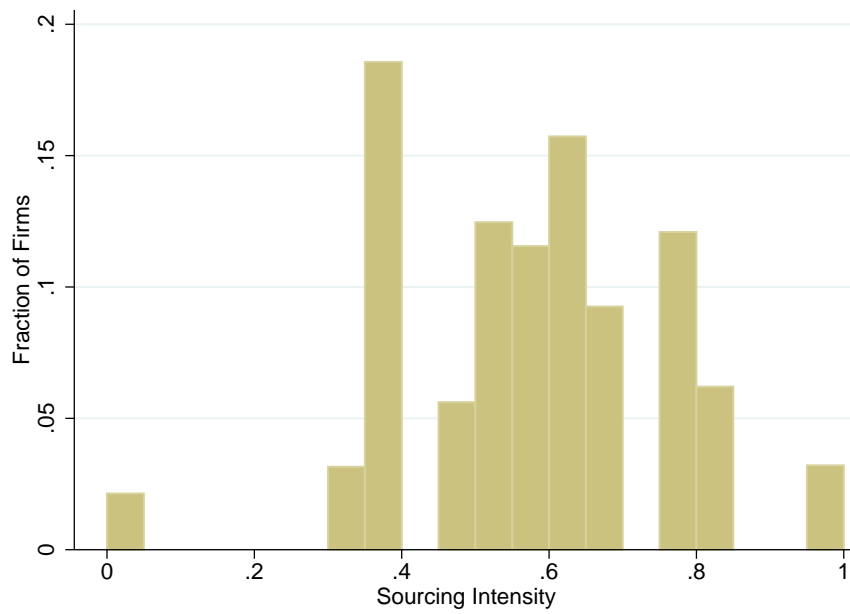
Figure C.2. Histograms for predicted probabilities $(K = 1)^{\dagger}$
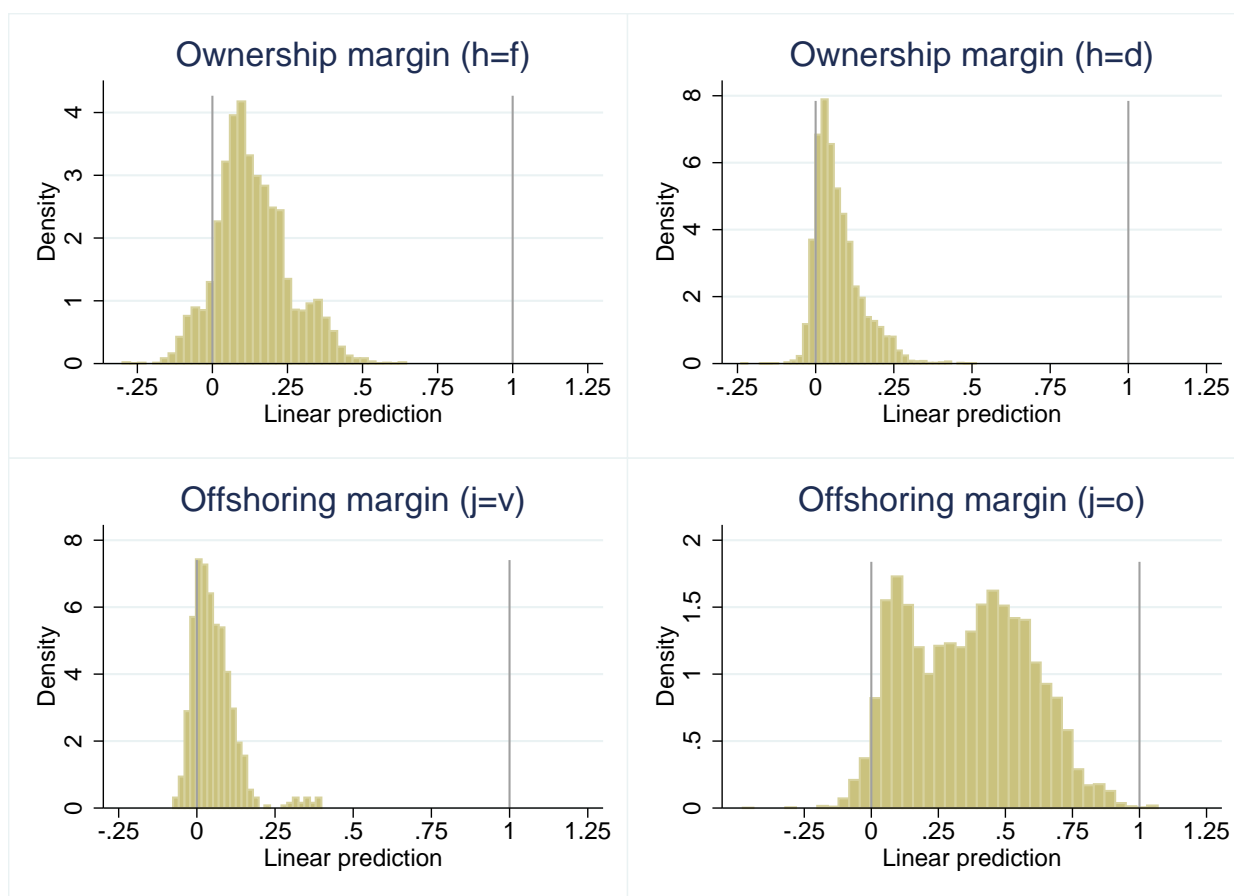
†Note: Linear predictions are from the LPM (estimated for $K = 1$). The fractions of firms outside the unit interval equal (from left to right and top to bottom): 10.6%; 11.1%; 19.8%; and 2.9%.

<div align="center">**Table C.1.** Industries in ESEE data[†]</div>

CNAE-2009 Classification	Industry
101	Meat
102-109, 120	Food Products and Tobacco
110	Beverages
131-133, 139, 141-143	Textile
151-152	Leather & Footwear
161-162	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-Ferrous 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

[†]Note: See http://www.ine.es/daco/daco42/clasificaciones/cnae09/estructura_en.pdf for individual products (or groups of products) belonging to each industry in ESEE data.

<div align="center">**Table C.2.** Correlation matrix for industry variables[†]</div>

	Sourcing intensity	Skill intensity	R&D intensity	Service intensity	Ext. fin. dep.
Sourcing intensity	1				
Skill intensity	-0.1863	1			
R&D intensity	0.2563	0.6451	1		
Service intensity	0.2376	0.6492	0.5061	1	
Ext. fin. dep.	0.3391	0.3639	0.4058	0.4141	1

[†]Note: External financial dependence not available for the Meat industry.

Table C.3. Variables and summary statistics

Variable	Description	Obs.	Mean	Std. Dev.	Min.	Max.
Firm-specific variables						
DI	Dummy variable for domestic integration	13,750	0.157	0.363	0	1
DO	Dummy variable for domestic outsourcing	13,750	0.925	0.263	0	1
FI	Dummy variable for foreign integration	13,750	0.103	0.304	0	1
FO	Dummy variable for foreign outsourcing	13,750	0.439	0.496	0	1
INT_f	Indicator for vertical integration in Ω_f	5,768	0.197	0.398	0	1
INT_d	Indicator for vertical integration in Ω_d	6,767	0.100	0.300	0	1
FOR_v	Indicator for offshoring in Ω_v	745	0.0886	0.284	0	1
FOR_o	Indicator for offshoring in Ω_o	9,889	0.384	0.486	0	1
Share of intra-firm sourcing$_f$	Share of intra-firm sourcing in Ω_f	5,767	0.0652	0.177	0	0.990
Share of intra-firm sourcing$_d$	Share of intra-firm sourcing in Ω_d	6,767	0.0330	0.131	0	0.960
Share of foreign sourcing$_v$	Share of foreign sourcing in Ω_v	732	0.0396	0.159	0	0.981
Share of foreign sourcing$_o$	Share of foreign sourcing in Ω_o	9,877	0.0686	0.152	0	0.985
# Foreign-affiliates (total)	Number of foreign affiliates	5,768	0.857	3.432	0	100
# Foreign-affiliates (EU)	Number of foreign affiliates in the EU	5,768	0.459	1.848	0	34
# Foreign-affiliates (Latin America)	Number of foreign affiliates in Latin America	5,768	0.150	0.910	0	35
# Foreign-affiliates (Rest of the OECD)	Number of foreign affiliates in the rest of the OECD	5,768	0.115	0.566	0	10
# Foreign-affiliates (Rest of the World)	Number of foreign affiliates in the rest of the world	5,768	0.133	0.986	0	47
Labor productivity (in logs)	Real value added over effective work-hours (in Euros)	13,606	3.146	0.699	-6.000	6.741
Total factor productivity (in logs)	TFP based on Olley & Pakes (1996)	11,849	1.639	0.5169	-2.470	7.289
Capital intensity (in logs)	Real value of tangible fixed assets over number of workers (in Euros)	13,705	3.872	1.157	-8.406	8.433
Skill intensity (in logs)	Graduate workers over total number of workers	13,318	0.197	0.256	0	2.760
Export volume (in logs)	Real value of exports (in Euros)	13,735	9.574	7.358	0	22.79
Age	Age of the firm in years	13,481	29.69	20.46	0	175
Employment (in logs)	Average number of workers during the year	13,748	4.123	1.429	0	9.553
R&D intensity (in logs)	Real R&D expenses over total sales	13,722	0.00759	0.0238	0	0.760
Service intensity (in logs)	Service sales over total sales	13,465	1.083	1.339	0	4.522
Debt ratio	Total debt over total assets	13,174	0.379	0.203	0	0.996
Capacity utilization	Use of production capacity (in percent)	13,729	76.34	18.02	2	100
Industry-specific variables						
Sourcing intensity	Log of the median of firm-level capital intensities over the period 2000-2012 (reverse scale; normalized to the unit interval)	20	0.577	0.222	0	1
Sourcing intensity (for robustness)	Log of the average of firm-level capital intensities over the period 2000-2012 (reverse scale; normalized to the unit interval)	20	0.618	0.248	0	1
Skill intensity	Average of firm-level skill intensities over the period 2000-2012	20	0.235	0.188	0.0523	0.734
R&D intensity	Average of firm-level R&D intensities over the period 2000-2012	20	0.00768	0.0105	0.000946	0.0399
Service intensity	Average of firm-level service intensities over the period 2000-2012	20	6.148	2.209	3.098	10.72
External financial dependence	External financial dependence derived from Rajan & Zingales (1998)	19	0.329	0.193	0.0557	0.862

Alternative measure of productivity. As an alternative measure of firm productivity, we employ total factor productivity estimated with the estimation algorithm developed by Olley & Pakes (1996). We feed the algorithm with ESEE data from 2000-2011. More specifically, we use annual information on each firm's value added, investment, capital stock, labor employment, and exit decisions. *Value added* is the sum of the total production value plus other operating income (i.e., income from rent and leasing, industrial property, commissions, and certain services), minus the total expenditure on intermediate inputs and external services, expressed in constant 2010 prices. *Investment* is the real total investment value in real estate, construction, and equipment. The *capital stock* is the real value of real estate, construction, and equipment. *Labor employment* is measured in effective work-hours. *Exit decisions* of firms allows us to distinguish between firms shutting down production and firms staying in the market, but exiting the ESEE panel.

References

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

Antràs, P. (2014a). *Global Production: Firms, Contracts, and Trade Structure*. Unpublished Manuscript (CREI Lectures on Macroeconomics 2010).

Antràs, P. (2014b). Grossman-Hart (1986) goes global: Incomplete contracts, property rights, and the international organization of production. *Journal of Law, Economics, and Organization*, *30*(1), 118–175.

Antràs, P., & Chor, D. (2013). Organizing the global value chain. *Econometrica*, *81*(6), 2127–2204.

Antràs, P., & Helpman, E. (2004). Global sourcing. *Journal of Political Economy*, *112*(3), 552–580.

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.

Aoki, M. (1990). Toward an economic model of the Japanese firm. *Journal of Economic Literature*, *28*(1), 1–27.

Bache, P. A., & Laugesen, A. (2014). Trade liberalisation and vertical integration. Economics Working Papers 2013-14, Aarhus University.

Becker, S. O., Ekholm, K., & Muendler, M.-A. (2013). Offshoring and the onshore composition of tasks and skills. *Journal of International Economics*, *90*(1), 91 – 106.

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), 444–448.

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*, *4*(1), 283–313.

Blinder, A. (2009). Offshoring: Big deal, or business as usual? In B. M. Friedman (Ed.) *Offshoring of American Jobs. What Response from U.S. Economic Policy?*, chap. 2, (pp. 19–60). Cambridge, MA: MIT Press.

Coase, R. H. (1937). The nature of the firm. *Economica*, *4*(16), 386–405.

Corcos, G., Irac, D. M., Mion, G., & Verdier, T. (2013). The determinants of intrafirm trade: Evidence from French firms. *Review of Economics and Statistics*, *95*(3), 825–838.

Defever, F., & Toubal, F. (2013). Productivity, relationship-specific inputs and the sourcing modes of multinationals. *Journal of Economic Behavior & Organization*, *94*, 345–357.

Dunning, J. H. (1993). *Multinational Enterprises and the Global Economy*. Addison Wesley Longman, Inc.

Ebenstein, A., Harrison, A., McMillan, M., & Phillips, S. (2014). Estimating the impact of trade and offshoring on American workers using the current population surveys. *Review of Economics and Statistics*, *96*(4), 581–595.

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.

Guadalupe, M., Kuzmina, O., & Thomas, C. (2012). Innovation and foreign ownership. *American Economic Review*, *102*(7), 3594–3627.

Hart, O., & Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy, University of Chicago Press*, *98*(6), 1119–1158.

Horrace, W. C., & Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, *90*(3), 321–327.

Klette, T. J., & Griliches, Z. (1962). An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. *Journal of the American Statistical Association*, *57*(298), 348–368.

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.

Kohler, W., & Smolka, M. (2014). Global sourcing and firm selection. *Economics Letters*, *124*(3), 411–415.

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

Milgrom, P., & Roberts, J. (1993). Johnson Controls Inc., Automotive Systems Group: The Georgetown, Kentucky Plant. Stanford Graduate School of Business Case S-BE-9.

Mrázová, M., & Neary, P. (2013). Selection effects with heterogeneous firms. Mimeo, 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.

Nunn, N., & Trefler, D. (2013). Incomplete contracts and the boundaries of the multinational firm. *Journal of Economic Behavior & Organization*, *94*, 330–344.

Olley, G. S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*(6), 1263–1297.

Rajan, R. G., & Zingales, L. (1998). Financial dependence and growth. *American Economic Review*, *88*(3), 559–586.

Schwarz, C., & Südekum, J. (2014). Global sourcing of complex production processes. *Journal of International Economics*, *93*(1), 123–139.

Tomiura, E. (2007). Foreign outsourcing, exporting, and FDI: A productivity comparison at the firm level. *Journal of International Economics*, *72*(1), 113–127.

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.

Young, S., Hood, N., & Hamill, J. (1985). Decision-making in foreign-owned multinational subsidiaries in the United Kingdom. ILO Multinational Enterprises Programme Working Paper No. 35, Geneva: International Labor Office.