DISCUSSION PAPER SERIES

No. 7530

THE DETERMINANTS OF INTRA-FIRM TRADE

Gregory Corcos, Delphine M. Irac,
Giordano Mion and Thierry Verdier

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7530.asp**

ISSN 0265-8003

THE DETERMINANTS OF INTRA-FIRM TRADE

Gregory Corcos, Norwegian School of Economics
Delphine M. Irac, Banque de France
Giordano Mion, London School of Economics, NBB and CEPR
Thierry Verdier, Paris School of Economics and CEPR

Discussion Paper No. 7530
November 2009

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7530

November 2009

ABSTRACT

The Determinants of Intra-Firm Trade

How successful is the theory of the firm in explaining intra-firm trade? To answer this question we exploit a unique dataset of 1,141,393 French import transactions, spanning across firm, countries and products in 1999, and reporting whether a transaction is intra-firm. Overall, we find support for the main predictions of the partial equilibrium property-rights approach and further deliver facts that can be useful for further theoretical development. We document substantial within-industry heterogeneity while providing evidence of the importance of the firm dimension of sourcing choices as well as of the key distinction between the extensive and intensive margins.

JEL Classification: F12, F19 and F23
Keywords: extensive margin, firm heterogeneity, incomplete contracts, intensive margin, internationalization strategies, intra-firm trade, outsourcing and quality of institutions

Gregory Corcos
NHH
Department of Economics
Helleveien, 30
N-5045 Bergen
NORWAY

Email: gregory.corcos@nhh.no

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=167813

Delphine M. Irac
Banque de France
46-1405 DAMEP
31 rue Croix des Petits Champs
75049 Paris Cedex 01
FRANCE

Email:
Delphine.irac@banque-france.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=166319

Giordano Mion
Department of Geography
London School of Economics
Houghton Street
London WC2A 2AE
UK

Email: G.Mion@lse.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157937

Thierry Verdier
Paris-Jourdan Sciences Economiques
48 Boulevard Jourdan
75014 Paris
FRANCE

Email: verdier@pse.ens.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=113317

Submitted 28 October 2009

First Draft: 26 January 2008

We thank Pol Antràs, Andrew Bernard, Thierry Mayer, Dalia Marin, Jim Markusen, Steve Redding, Sébastien Roux, and Alessandro Sembenelli as well as seminar participants at ESEM 2009, NOITS 2009 (Copenhagen), ERWIT 2008 (Appenzell), the CESifo Venice Summer Institute 2009 ("Heterogeneous Firms in International Trade"), the Ludwig-Maximilian University of Munich and the workshop "The International Firm: Access to Finance and Organizational Modes" (Milan) for helpful comments and suggestions. Financial support from FIRB, Università degli Studi di Milano, and Centro Studi Luca d'Agliano is gratefully acknowledged. The usual disclaimer applies.

1 Introduction

About a third of current international trade transactions occurs within a multinational firm, while an additional third involves at least one multinational as one party to the transaction. This centrality of multinational firms in the process of globalization is important in many respects. In public debates, this pattern of cross border production networks and FDI flows has generated widespread concerns about the relocation of production facilities to low wage developing countries. So-called 'offshoring' and foreign subcontracting activities of intermediate inputs have stimulated intense discussions and have often been associated to dramatic wage and job consequences for workers in the OECD countries, or deteriorating working conditions in industrialized and developing countries. The current pattern of multinationalization of firms has also attracted much attention among economists. In particular, substantial research efforts has been devoted to the central question of why are some international transactions carried out within a firm, while others are carried out at arms' length on markets.

A first well established literature considered explanations based on the importance of intangible assets such as knowledge and reputation.[1] Recently, new research has extended this literature by taking on an explicit contract-theoretical approach of multinationals. This approach has two key features. It emphasizes the costs associated with writing contracts for specialized inputs and it highlights the importance of the allocation of residual rights of control in such incomplete contracting setting. A core objective of this rapidly growing literature is then to discuss the determinants of intra-firm trade transactions and understand their implications for the functioning of the global economy[2]. One of the key conclusions is that in a context with relation-specific investments between two parties, ownership and control of assets should be assigned to the party whose investment contributes most to the relationship. For instance, when investments in 'headquarters' services (those provided by the sourcing firm) contribute relatively more to the creation of value added than other intermediate inputs, the allocation of property rights should go to the

[1] See the surveys in Markusen (1995) and Barba Navaretti and Venables (2004).

[2] See for instance McLaren (2000), Antràs (2003), Grossman and Helpman (2002, 2003, 2004, 2005), Antràs and Helpman (2004; 2008), Marin and Verdier (2003, 2008). Good surveys of the literature are found in Helpman (2006), Spencer (2007) and Antras and Rossi-Hansberg (2008). Some of the most illustrative recent work along this line of research is published in Helpman, Marin and Verdier (2008).

final good producer. Conversely, when production is intensive in downstream 'intermediate input' related investments, it is optimal to assign control rights to the intermediate good supplier. In the former (latter) case, transactions are predicted to occur inside (outside) the boundaries of the firm.

Along with the rapidly expanding theoretical literature, a limited empirical literature has started to develop. Exclusively focused on US data, a series of papers has discussed the empirical determinants of the share of US imports that is intra-firm. (Antràs 2003, Yeaple 2006, Nunn and Trefler 2008, Bernard et al. 2008). These studies broadly find support for the predictions by Antràs (2003) and Antràs and Helpman (2004; 2008) by relating aggregate industry and/or product level intra-firm trade shares to the intensity of downstream production in intermediate inputs. While useful and important first steps, these analyzes face the limitation to be confined at an aggregate level, whereas the theory explicitly emphasizes the importance of the firm as the key unit of analysis to understand intra-firm trade.

The main purpose of this paper is to 'focus the lens *on the firm level*' in order to provide evidence about *firm, country, and product* attributes that determine the optimal international sourcing mode. In particular, by taking advantage of a dataset documenting imports of manufactured goods by French firms in 1999, we empirically assess the predictions made in the recent theoretical literature on multinationals' organizational choices. Our paper is related to the above cited empirical literature as well as to Carluccio and Fally (2009) and Defever and Toubal (2009).

We start by confirming existing evidence in support of the partial equilibrium property-rights approach. We first examine the effect of firm characteristics on sourcing choices, controlling for unobserved product and country heterogeneity, and then focus on product and country characteristics, controlling for unobserved firm heterogeneity. The explanatory power of the underlying theory and proxy used does suggest that further theoretical and empirical analysis is needed. However, when compared to a 'benchmark' model explaining firm export status on standard firm-level variables such as size and productivity we find that our model is no less satisfying.

We provide four main contributions with respect to the existing literature.

i) First, the paper establishes the importance of the intensive margin (value of a transaction for a given sourcing mode) versus the extensive (likelihood to engage in an intra-firm transaction)

margin to understand firm-level sourcing patterns. For example, our analysis at the firm level reveals the striking result that firms are more likely to engage in intra-firm trade with labor-abundant countries, contrary to the prediction in Antràs (2003). However, at the industry or product level, intra-firm import shares do indeed vary positively with the capital abundance of the origin country, as in the above cited papers. Our analysis reveals that sectoral composition plays a minor role and that it is crucial to investigate the intensive margin to reconcile such results.

ii) Second, we document a fact of general interest that is particularly relevant for theories of intra-firm trade: even at a finely defined sectoral level one observes a degree of heterogeneity in factor intensity across firms comparable to firm productivity. Magnitudes suggest that technology heterogeneity at the firm level is not only related to measurement error and/or to the scale of production (as is usually assumed in the recent trade theory with firm heterogeneity), but also to factor intensities and cost shares across firms. In our framework, this new source of heterogeneity is important to note as factor intensities have clear consequences for the allocation of control rights and incentives inside the firm. Beyond that, variation in factor intensity across firms is likely to generate differences in factor price effects across firms within the same sector. This could contribute to explain wage inequality between workers with similar observable individual characteristics but allocated to different firms within the same sector. More generally, the observation of significant within-sector heterogeneity in factor intensity may also have implications for the debate on the distributive consequences of trade and technology in the global economy. In particular, it suggests that one should pay more attention to the firm level consequences of technological biases on wages rather than rely exclusively to the standard 'factor-biased' or 'sector-biased' dimensions of technological change.

iii) Third, the paper finds that intra-firm imports are more likely to come from countries with strong judicial institutions, as measured by the World Bank's Rule of Law index. This result conforms with property-rights theory, especially as we find that the effect is stronger for highly productive and skill-intensive firms. In these theories better enforcement reduces the need to offer high-powered incentives (ownership) to a foreign supplier.

iv) Fourth, the paper provides robust empirical evidence that complex goods and inputs are more likely to be produced within firm boundaries. Our results complement Carluccio and Fally

(2009) who find higher shares of intra-firm imports of complex inputs from countries with a lower level of financial development.

The plan of the paper is the following. In section 2 we discuss the basic theoretical background of the literature and its testable predictions. Section 3 presents the description of our dataset and the definitions and interpretations of our variables. In section 4, we discuss the empirical results. In Section 5 we try to reconcile our firm-level evidence with existing evidence at the product and industry levels, looking at both the intensive and the extensive margins of international sourcing. Finally section 6 concludes and suggests avenues for future research.

2 Theoretical Background

In this section we review three theoretical models of the determinants of internalization decisions: Antràs (2003), Antràs and Helpman (2004; 2008). All three models explain internalization decisions using property-rights theory (Grossman and Hart, 1986, Hart and Moore, 1990).

A first building block of these models is a partial equilibrium model of organizational choice which we briefly sketch here. Consider a final-good producer who needs to obtain a specific intermediate input from a supplier. Production of the input requires a non-contractible and relationship-specific investment by each party. Asset specificity and contract incompleteness create a two-sided holdup problem, implying under-investment by both parties.

Property rights over a productive asset provide its owner with residual rights of control. They create an outside option in bargaining over ex post surplus, and therefore greater incentives to invest ex ante.

It is assumed that utility is costlessly transferrable between parties ex ante and that ex post bargaining is efficient. Therefore the organizational form that maximizes expected joint surplus is chosen. The central result is that the higher the intensity in an investment, the more ownership should be given to the party responsible for this investment.

2.1 Internalization and Final Production Technology

In Antràs (2003) the two parties involved in production are a headquarter firm and a foreign supplier. The headquarter firm provides capital investment while the supplier provides labor. The maintained assumption in that model is that only labor-intensive production can be outsourced, mirroring stylised facts on US multinationals' internal cost-sharing practices. Antràs (2003) shows that efficiency in labor investments is higher under outsourcing. Therefore his first (partial-equilibrium) prediction is that intra-firm imports are more likely in capital-intensive industries/products, all else equal.[3]

The models in Antràs and Helpman (2004; 2008) have the same qualitative predictions, albeit with a different interpretation of the factor used intensively by the downstream firm (called 'headquarter services'). This intensity is denoted by $\eta \in [0, 1]$, a notation we will follow in the empirical section.

One innovation of these models is intra-industry TFP differences. Due to the CES specification the previous result is unchanged unless fixed costs vary across organizational forms. However, under the additional assumption that fixed costs are greater under integration that outsourcing, two predictions emerge. First, *within a given industry, firms engaging in intra-firm trade are more productive than those that outsource.* Second, *intra-firm trade is more likely to occur, the higher the contribution of the final good producer in the value-added chain (η)* (Antràs and Helpman, 2004). Note that the model only allows for inter-industry differences in η.

2.2 Internalization and Host-country Factor Abundance

Antràs (2003) embeds the model sketched above in a 2x2x2 general equilibrium framework. Assuming free entry, identical and homothetic preferences, and that immobile endowments are in the Factor Price Equalization set, he shows that *the share of intrafirm imports increases in the country's capital/labor ratio.* In his Helpman-Krugman model the capital-intensive industry (where all trade is intra-firm) has a greater number of varieties in capital-abundant countries than the labor-intensive industry (where all trade is at arms' length).

[3] See Antràs (2003), equation 8 p. 1390. As he points out, the absence of factor prices in the ratio of integration to outsourcing profits is an artefact of the Cobb-Douglas production function.

The two factors that Antràs (2003) considers are labor and capital. Empirically, he finds that the share of US intrafirm imports increases with capital abundance in the origin country, even when controlling for human capital abundance in that country. However, it is unclear whether his theoretical prediction generalizes to a model with more industries or more factors.

2.3 Internalization and the Extent of Contract Incompleteness

Antràs and Helpman (2008) extend their 2004 model to allow for partial contractibility of production tasks. There they relate organizational choice in offshore operations to a country's contracting environment. [4] They consider a composite component m and composite headquarter services h. Both can be decomposed into a continuum of tasks of mass one, some of which are non-contractible. The extent of contract incompleteness is captured by the range of non-contractible tasks in both activities, denoted by $(\mu_h^c, 1]$ and $(\mu_m^c, 1]$, where c refers to the country. The model includes the Antràs and Helpman (2004) model as a special case where $\mu_h = \mu_m = 0$.

Consider a change in the contractible content of component production tasks (which we refer to as 'contractibility of the input'), all else equal. For sectors or firms using components intensively, that were fully outsourcing their input production, this does not change anything. However, the model offers a different prediction for headquarter-intensive firms or sectors.

Holding headquarter services contractibility constant, an improvement in input contractibility abroad has two effects[5] *on headquarter-intensive firms:*

- *the most productive domestic producers switch to offshore outsourcing ('Standard Effect');*
- *the most productive firms resorting to offshore outsourcing insource from foreign affiliates ('Surprise Effect'). This is because the need to provide incentives to component producers is now lower.*

In sum, improved contract enforcement in the origin country favors international sourcing but does not clearly favor one sourcing mode. As explained by Antràs and Helpman (2008),

[4] An important motivation is the finding in Nunn (2007) that cross-country differences in contracting institutions matter as much as cross-country differences in human capital to explain the variance of trade flows.

[5] Nunn and Trefler (2008) term these two effects the 'Standard Effect' and the 'Surprise Effect', respectively.

"the relative prevalence of alternative organizational forms depends not only on cross-country differences in contractibility, but also on the degree to which contractual institutions are biased toward inputs controlled by the final-good producer or other suppliers."

Note that in their model with fully non-contractible investments, Antràs and Helpman (2004) found an unambiguously positive effect of tighter contract enforcement in the origin country on the share of intra-firm imports. When one allows judicial systems to make some tasks or inputs contractible, the relative contractibility of these inputs matters to organizational choice as much as their relative contribution to total output. It is therefore desirable to capture the contractibility of different tasks in empirical applications.

3 Data Sources and Variables Used

3.1 Firm-level Imports Data by Country of Origin, Product and Sourcing mode

We build a unique firm level dataset of French import flows in 1999 by merging two different data sources.

The first database, named EIIG (Échanges Internationaux Intra-Groupe), is a survey conducted in 1999 by SESSI (Service des Études Statistiques Industrielles, French Ministry of Industry). The survey was addressed to all French firms trading more than 1 million euros, owned by manufacturing groups that control at least fifty percent the equity capital of an foreign affiliate. The response rate was 53%. However, respondent firms represent 82% of total exports and imports of French multinationals.

The survey provides a detailed country of origin breakdown of French firms' yearly imports at product level (either CPA96 or HS4 4digit) and their sourcing modes - through independent suppliers and/or affiliates - for the year 1999. Intra-firm trade is defined as trade with an affiliate controlled by a single French entity with at least fifty percent of its equity capital.

Aggregating firms' imports flows by origin country, product classification (CPA96 4digit) and sourcing mode we obtain 76,364 firm-product-country triples corresponding to 4,193 importers.

31.28% of our observations correspond to intra-firm trade and the rest is outsourcing.[6] This data has been used by Defever and Toubal (2009) to test the Antràs and Helpman (2004) model.

However, given the fact that firms in the EIIG survey firms have been selected on the basis of having substantial ownership and commercial links with foreign firms, the sample is clearly biased towards intra-firm trade. Almost by definition, each firm in the EIIG database has at least one intra-firm trade flow. Indeed, while the SESSI estimates that around 36% of the total value of manufacturing imports is intra-firm (Guannel and Plateau, 2003), in the EIIG data the corresponding value is much higher (55.4%). This raises a serious selection bias issue. In order to solve this problem we make use of another data source.

The second database, coming from French Customs and used by Eaton et al. (2004) among others, is the universe of yearly import and export flows operated by French firms coming from custom declarations.[7] Aggregating firms' imports in 1999 by country of origin and product (CPA96 4digit) we obtain 1,252,462 observations referring to 126,953 firms.

In what follows, we assume that *all trade flows that are reported in the French Customs dataset but not in the EIIG occur with a third party.* Furthermore, *we refer to an individual record of our final imports database as a 'transaction'.* After combining the EIIG with the 1999 French Customs data, we further eliminate flows who report France as the origin country of imports (basically trade with overseas French territories). The final import flows dataset covers 1,141,393 firm-country-product-sourcing mode combinations, corresponding to 126,926 importers, 232 destinations, and 274 products.

3.2 Data on Firm Characteristics

The primary data source for firm's characteristics is the EAE databases (Enquête Annuelle Entreprise) provided by both SESSI and SCEES (Service Central des Enquêtes et des Études Statistiques, French Ministry of Agriculture). The database provides detailed balance sheet information on all French firms with at least 20 employees and we focus in our analysis on firms whose primary

[6] See Appendix A for details.

[7] For trade outside the EU15, there is no minimal amount for data to be recorded. Within the EU, only trade whose total annual amount for a given country-product couple exceeds 250,000 euros per year should be registered. In practice however, many trade flows below this threshold are still registered. Indeed, the total value of imports in the database represents 99% of French aggregate imports in 1999 as reported by EUROSTAT.

activity is in the manufacturing sector (NACE rev1 D category). Firms in the EAE database represent 9.8% of the total number of French manufacturing firms but 87.2% of production in 1999 as reported by EUROSTAT.

3.3 Description of the Variables Used in the Empirical Analysis

We index firms by i, imported products by p, the final product produced by a firm by f, and countries by c. Our set of covariates aim to capture determinants of the internalization decision at the level of the firm, country, imported input and final product.

3.3.1 Dependent Variable

Our dependent variable $y_{i,p,c}$ is a dummy that takes value 1 (intra-firm) or 0 (outsourcing) depending on the nature of the recorded total annual import flow at the firm level of product p from country c in the year 1999. In the case a French firm i import product p from country c from a foreign affiliate then $y_{i,p,c} = 1$, while if the transactions occurs with a third party $y_{i,p,c} = 0$.

We use binary data for two reasons. First, only a few product-country-firm triples involve both intra-firm and arms' length imports.[8] Second, the theoretical predictions we consider apply to this binary choice rather than to the value of import transactions under each mode.

3.3.2 Firm Level Variables

TFP_i stands for Total Factor Productivity of firm i. We estimate TFP as the residual (plus the constant) of a log-linearized three-factor Cobb-Douglas production function with labor, capital and material inputs. We use the value-added based Levinsohn and Petrin (2003) estimator (henceforth LP).[9]

k_i is the log of the ratio between the capital stock and employment of firm i and we use it as a measure of capital intensity. $\eta_i^{hq} \in [0, 1]$ is the ratio of value added over total sales of a firm

[8] We keep most of this 'mixed' transactions information by recording as intra-firm or outsourcing a transaction for which at least 80% of the total value occurs with one of the two sourcing modes. As for neglected transactions, they would just provide us with 0.4% more observations. See Appendix A for details

[9] Details on the estimation procedure are provided in Appendix A. We have also experimented with the Olley and Pakes (1996) estimator (henceforth OP), obtaining qualitatively identical results. We prefer LP because the OP procedure obliges us to throw away quite a few firms, due to non-positive values of investments.

i. We consider it as a proxy of the relative importance of the final production stage in the value added chain. Finally η_i^{sk} is the log of the ratio between total wage expenses and employment of firm i. This variable is meant to capture the average skills of workers of firm i with the underlying hypothesis being that more skilled workers are paid higher salaries. Indeed, when aggregating η_i^{sk} across two digit NACE rev1 industries, we obtain a high correlation (0.67) with the share of the workforce having at least a secondary education. Furthermore, we will show later on that while the correlation between η_i^{sk} and productivity is positive, it is too low to claim that η_i^{sk} reflects firm productivity only.

In some robustness specifications we add firm size ($size_i$) as a control. Note that theory has little to say about its direct impact on the sourcing mode[10]. We have experimented with both the log of sales and the log of employment of firm i obtaining virtually identical results. In what follows, we report results based on log sales.

3.3.3 Imported Products Variables

One advantage of our data is that we observe both the imported product category and the industry of the importer, unlike Levchenko (2007), Nunn (2007), and Nunn and Trefler (2008). This is important when looking at how imported product characteristics affect sourcing modes. In our data a firm typically imports several products that vary substantially in their degree of sophistication. While previous studies attributed a synthetic measure of contract complexity at the importing industry level[11], we capture that heterogeneity by measuring the contract complexity of each imported input.

To do so we follow the idea of Nunn (2007) that products that are sold on an organized exchange or reference priced are less likely to involve relationship-specific investments. We also use the Rauch (1999) product classification.[12] Denoting by $R_j^{neither}$ a dummy variable that takes

[10]For example, although in Antràs and Helpman (2004; 2008) firms engaging in intra-firm trade are expected to be larger, this is entirely due to their productivity being higher than that of outsourcing firms. Size *per se* has no additional effects.

[11]Levchenko (2007) builds a Herfindahl index measuring the degree of variety of inputs needed in the production process. Nunn (2007) and Nunn and Trefler (2008) also reconstruct average contract intensity in the importing industry. They identify inputs that require relationship-specific investments as proxied by the Rauch (1999) classification, and build an index for each industry using US input-output matrix coefficients as weights.

[12]In particular we use his 'Liberal' classification. Results are virtually unchanged if we use the 'Conservative' one.

value 1 if the HS6 product j is neither sold on an organized exchange nor reference priced, and by $\theta_{p,j}$ the share of the HS6 product j in the French imports of CPA96 4digit product p in 1999 we have:[13]

$$\mu_p = 1 - (\sum_j \theta_{p,j} R_j^{neither})$$

Finally, we consider the capital (k_p) and skill intensity (h_p) of the imported product p that we construct using French technology.[14] We introduce such variables because in Antràs (2003) factor intensities of the imported product play a key role.

3.3.4 Final Product Variables

As explained we also measure the contractibility of a final good f, in addition to the contractibility of an imported input, again using the Rauch classification. Defining $R_j^{neither}$ as dummy variable that takes value 1 if the PRODCOM2002 8 digit product j is neither sold on an organized exchange nor reference priced, and by $\theta_{f,j}$ the share of the PRODCOM2002 8 digit product j in the French production of CPA96 4digit product f in 1999 we have:[15]

$$\mu_f = 1 - (\sum_j \theta_{f,j} R_j^{neither})$$

In Table I we give summary statistics on both input and final product contractibility measures. Two things are worth noting. First, looking at transaction patterns, the correlation between μ_p and μ_f is positive (.28) and significant. This means that firm producing final complex goods import complex inputs. Second, μ_f is highly correlated with the Nunn (2007) measure of complexity in the whole production process. Comparing them across the 29 ISIC rev2 3 digit sectors (the only classification for which data are comparable) the correlation is .78.

[13] Summary statistics for this variable are given in Table I below. See Appendix A for more details.
[14] See Appendix A for details.
[15] See Appendix A for details.

3.3.5 Origin Country Variables

As to the exporting country c characteristics, k_c and h_c stand (respectively) for the capital and skill abundance of country c, while Q_c is a measure of the quality of institutions (Rule of law index from Kaufmann et al. [2003]). Tax_c stands for the corporate rate tax in country c, while $OECD_c$ is a dummy indicating membership to the OECD in 1999.

$Same - leg - orig_c$ is a dummy indicating whether country c has a French civil law system, while $Distw_c$ is the log of distance of country c to France. Finally, $Colony_c$ is dummy indicating whether country c is a former French colony and $Language_c$ is a dummy indicating whether French is spoken in country c.

Since these are rather standard variables we refer to Appendix A for further details.

4 The Intra-firm vs. Outsourcing Decision: Firm-, Country-, and Product-Level Determinants

The population of interest in our analysis consists of importing firms. Our data covers them exhaustively.[16] We use a probit model to estimate the impact of the various determinants of sourcing mode, and denote by $y_{i,p,c}$ the binary response dependent variable that takes value 1 if the transaction is intra-firm and 0 if it corresponds to outsourcing.[17] Our goal is to study the firm-level likelihood to undertake intra-firm trade vs outsourcing (what we define as the 'extensive margin'). In our analysis we use the information constructed matching the EIIG and Custom databases for the year 1999, i.e *1,141,393 firm-country-product-sourcing mode combinations, corresponding to 126,926 importers, 232 destinations and 274 products.* The main results are reported in Sections 4.1 and 4.2 while Appendix B contains supplementary findings based on the interactions of some variables of interest.

Throughout the analysis we will provide tests of the empirical predictions of certain models (Results) as well as some valuable facts concerning intra-firm and outsourcing (Facts). Before

[16] The fact that we deal with firms engaged in either international intra-firm or outsourcing activities only (thus neglecting firms that have transactions with French affiliates or source inputs within France) is not an issue because the theoretical predictions we test concern precisely this set of firms.

[17] Results are robust to the alternative logit specification.

going into estimations we report a first fact:

- *Fact 1: Intra-firm import transactions are rare but typically involve large amounts.*

Indeed in our data only 2% of transactions are intra-firm but they correspond to 24.9% of total imports' value.[18] Even when restricting the analysis to multinationals' transactions from the EIIG database, we find that 31.3% are intra-firm and account for 55.4% of imports' value. Figure I shows the kernel-smoothed distribution of log transactions' value (in euros) for both intra-firm and outsourcing. As showed by the Figure, the distribution of intra-firm transactions values is somehow right-shifted with respect to outsourcing. The shape of the two distributions is rather similar with the upper bounds of the two supports being rather close (21.39 for intra-firm and 21.82 for outsourcing).

There are certainly many possible interpretations of Fact 1. One possibility, in line with Antràs and Helpman (2004; 2008), is that, due to higher fixed costs entailed by intra firm activities compared to outsourcing, higher volumes are necessary in order to break even.

4.1 Firm-Specific Determinants of Intra-Firm Trade

In this Section we explore the firm-level determinants of the intra-firm vs outsourcing choice. To this purpose we merge our import flows data with the firm level information coming from the EAE database. We obtain a smaller dataset of 247,528 firm-country-product-sourcing mode combinations corresponding to 16,383 importers, 202 countries, and 272 products. These observations still represent more than 60% of French imports from the manufacturing sector.[19] Focusing on firm-level determinants, we will use sector, country and product dummies to control for (potentially endogenous) unobservable characteristics in the 3 dimensions.

Our choice of regressors is influenced by the property-rights approach to multinational firm boundaries. As mentioned in Section 2, Antràs and Helpman (2004) predict that the most productive firms within an industry engage in intra-firm trade.

[18] Zeile (1997) reports for the US a 42.7% share of imports' value accounted by intra-firm trade. However, such a figure is not directly comparable to France because the definition of a foreign affiliate in the EIIG is much more demanding (fifty percent or more of the equity capital) than the US definition (ten percent). In addition the EIIG lacks coverage of around twenty percent of multinationals imports due to non-respondents.

[19] In this smaller dataset 5.4% of transactions are intra-firm, representing 34% of the value of imports.

Second, we also consider the result by Antràs (2003) that intra-firm trade is more prevalent in capital-intensive industries in the US. However, the same theoretical mechanism should work within a sector. Provided there is substantial variation in capital intensity within narrowly defined sectors, we investigate whether firm-level capital intensity k_i can affect the intra-firm decision.

Finally, another key variable in Antràs (2003) and Antràs and Helpman (2004; 2008) is the intensity in the input provided by the Northern firm (denoted by η in Section 2). The latter two models predict that no intra-firm should take place in in low η (component-intensive) industries. This prediction is not valid in our data. Intra-firm trade and outsourcing coexist in virtually all NACE rev1 4 digit industry level (roughly 250 industries). While we cannot exclude that all industries are above the model's capital-intensity threshold, an appealing alternative explanation is that there is within-industry heterogeneity in η. We use two measures of headquarter intensity: the ratio of value added produced in the North over sales (as suggested by Defever and Toubal, 2009)[20], η_i^{hq}, and the log of the ratio of wage expenses over employment, η_i^{sk}.

We start by providing some descriptive statistics on our firm-level regressors. One contribution of this paper is to show that the correct unit of analysis for capital intensity as well for η is the firm and not the product. Although it has been overlooked in theoretical models, firms are not only highly heterogeneous in their productivity but also in their capital, skill and input choices.

Table II provides summary statistics for firm-level total factor productivity (TFP_i), capital intensity (k_i), headquarter intensity (η_i^{hq}), and skill intensity (η_i^{sk}) on our whole data. Keeping in mind that TFP_i, k_i and η_i^{sk} are constructed using logs (so are unit of measurement independent) and that η_i^{hq} varies between 0 and 1, one can see from standard deviations that there is a lot of heterogeneity across firms in the whole manufacturing sector with TFP and capital intensity showing the highest variability. Correlations between TFP and the other three firm-level variables are either negligible or relatively small as in the case of η_i^{sk}. Furthermore, correlations between k_i, η_i^{hq} and η_i^{sk} are also pretty low.

We measure within-industry heterogeneity in these variables by computing deviations to industry averages. Define $\Delta_s k_i$ the value of k_i minus its NES114 industry classification average

[20] In a world of complete contracts and competitive markets η would simply be the cost share of services provided in the North. Under incomplete contracts the link between factor intensity and cost shares is less evident.

(56 manufacturing industries), and apply the same notation to other variables. We report the standard deviations and correlations of these *within-sector* variables in Table III. As one can notice, the standard deviation of productivity within sectors is around half of the corresponding value on the whole manufacturing. This means that half of the standard deviation of firm-level productivity is due to differences across sectors. However, the variability of k_i, η_i^{hq}, and η_i^{sk} is only slightly reduced when accounting for differences across industries. *Even in narrowly defined industries, the variance of capital, headquarter and skill intensity at the firm level is still very high. In particular capital intensity ($\Delta_s k_i$) displays more heterogeneity than TFP. The same qualitative pattern emerges if we narrow our definition of industry to go down at the NACE rev1 4 digit level.*

Table III shows a weak correlation between $\Delta_s TFP_i$ and the other variables while cross-correlations between $\Delta_s k_i$, $\Delta_s \eta_i^{hq}$ and $\Delta_s \eta_i^{sk}$ are never very high. Furthermore, principal components analysis reveals that the first 3 components do not exceed 85% of the total variation. *All this suggests that the 4 variables provide different pieces of information on the technology of the final good producer.*[21] We must acknowledge two qualifications on our finding. First, part of this heterogeneity is likely to come from measurement error. Second, our assumption of a Cobb-Douglas production function may not be valid, implying that firms having different capital, skill and headquarter intensities could be using the same technology. However, given the magnitude of standard deviations and correlations, as well as the principal components analysis result, it is difficult to believe that all heterogeneity is attributable to errors or differences in scale across producers using the same non-CES technology. We can thus state that:

- *Fact 2: the correct unit of analysis for k and η is the firm and not the industry.*

In order to test the relevance of these firm-level variables we estimate the following probit model:

$$y^*_{i,p,c} = \alpha + \beta_1 TFP_i + \beta_2 k_i + \beta_3 \eta_i^{hq} + \beta_4 \eta_i^{sk} + \mathbf{D}_{NES} + \mathbf{D}_p + \mathbf{D}_c + \varepsilon_{i,p,c} \tag{1}$$

[21] A caveat on TFP estimation is that even the most sophisticated techniques assume identical input shares within an industry. Our finding therefore casts doubt on the reliability of TFP estimation in general. This issue has been raised previously in the firm productivity literature (e.g. Griliches and Mairesse, 1998), and there is no clear consensus on how to proceed. This is in fact more of an issue with the definition of TFP rather than an econometric challenge. In unreported regressions we use a more conservative measure of productivity (value added per worker), and obtain the same qualitative results.

where $\mathbf{D}_{NES}$, $\mathbf{D}_p$ and $\mathbf{D}_c$ stands, respectively, for sector, product and country dummies and the observable binary variable for intra-firm is $y_{i,p,c} = 1$ if $y^*_{i,p,c} \geq 0$. Table IV reports estimations of different specifications of 1. Even though this is rather demanding in terms of inference, we have clustered standard errors at the firm level to account for the possible correlation in the sourcing choice of the same firm across products and origin countries.

From the first to the fourth column we estimate, and report marginal effects of, each of the 4 firm-specific variables separately while in the fifth column we estimate them altogether. In the sixth column we include firm size as an additional covariate and show that results are robust. In unreported estimations, available upon request, we have also tried time-lags of the firm variables in order to control for simultaneity bias obtaining the same qualitative pattern.

Explanatory variables have overall positive and significant coefficient suggesting that:

- *Result 1: Intra-firm trade is more likely, the higher is firm productivity.* This finding is in line with the theoretical predictions of both Antràs and Helpman (2004) and Antràs and Helpman (2008). This is a new empirical result with respect to other firm-level studies on imports, such as Defever and Toubal (2009), as explained below, or Tomiura (2007) who does not identify affiliated party trade in his detailed firm-level analysis of Japanese imports.

- *Result 2: Capital, headquarter and skill intensity all favor intra-firm trade. However, they need to be evaluated at the level of the firm.* This new empirical finding is in line with the residual property rights literature predictions but further suggests that heterogeneity in capital, headquarter and skill intensity needs to be accounted for.[22]

A special comment is in order. Defever and Toubal (2009) find an opposite sign for TFP on a subset of our data (the EIIG survey). However, they do find a similar sign when they control for firm heterogeneity in the relative magnitude of outsourcing and integration costs.[23] A more

[22] In unreported estimations we have used capital, headquarter and skill intensity computed at the NACE 3 digit industry level, as alternative explanatory variables in 1. Results show similar signs for industry intensities, but firm-level intensities have a systematically higher explanatory power.

[23] Antràs and Helpman (2004) simply assume greater fixed costs of integration. However most replies in the EIIG survey on the relative 'organizational costs' of outsourcing and integration point at the opposite conclusion. Defever and Toubal (2009) show that the sign of the TFP coefficient switches with the answer to that question in a way that is consistent with the theory. Note that they look only at multinationals, in which case the cost of setting up an affiliate is already sunk, and the Antràs and Helpman (2004) assumption is less likely to hold.

general issue with the Defever and Toubal (2009) study is sample selection. All firms in the EIIG survey have foreign affiliates with nearly no firm declaring outsourcing transactions only. These firms typically engage in both intra-firm and outsourcing transactions, depending on the product and country. Since each firm has a unique TFP measure, identification of the TFP coefficient does not come from comparing firms that do vs firms that do not engage in intra-firm, but rather from the share of intra-firm transaction within a firm. For this reason we find our sample more appropriate to test the theory.

As a final remark we note that, as indicated by Pseudo R^2s, there is a sizeable amount of unexplained heterogeneity that needs to be investigated further, from both a theoretical and an empirical perspective. Nevertheless, the model we fit to the data is at least as good in terms of explanatory power than a simple cross-section model that predicts export status based on standard firm-level characteristics. There is indeed a large body of empirical and theoretical literature pointing out that productivity and size are key to understand firm export status[24] with the widespread opinion being that this is a rather 'satisfying' model. When using a probit model to predict export status among firm in the EAE database based on TFP_i, $size_i$, k_i, η_i^{hq}, η_i^{sk}, as well as industry dummies, we get a Pseudo R^2s of 0.1547. The comparison of this 'benchmark' model with the intra-firm trade closest relative (column 6) in Table IV suggests the the underlying theory we build on is no less satisfying.[25]

4.2 Country and Product Determinants of Intra-Firm Trade

In this Section we explore the country and product specific determinants of the intra-firm vs outsourcing choice in the light of the property-rights approach literature with incomplete contracts. We exploit the full import flow dataset (1,141,393 firm-country-product triples) using various methods to control for firm characteristics.

Antràs (2003) predicts that intra-firm trade is more likely with capital abundant countries. Another determinant suggested by Antràs and Helpman (2004) is the quality of institutions in

[24] See Bernand and Jensen (2004) and Melitz (2003) among others.

[25] When adding lagged firm export status, as in Bernard and Jensen (2004), in our benchmark model the Pseudo R^2s jumps to 0.5134. However, such dynamic version of the Melitz-style selection model is not directly comparable to our cross-sectional model of intra-firm trade.

the origin country (South). To test these theories we use skill abundance (h_c), capital abundance (k_c) and the quality of judicial institutions (Q_c) as regressors.

Antràs (2003) also predicts that capital- and skill-intensive products are more likely to be produced in-house. These effects are captured by our variables k_p and h_p. Antràs and Helpman (2008) point out that the degree of contractibility of imported intermediates and final products matters for the choice of optimal sourcing mode. We therefore include our measures of product contractibility μ_p and μ_f.

We also control for country-specific variables that are not explicitly dealt with in our theoretical background, but that can affect the optimal sourcing mode. We denote such variables as $\mathbf{CC_c}$ First, one may argue that imports from OECD countries are of a different nature. Second, the practice of transfer pricing makes it reasonable to consider the issue of taxes. We thus include both an OECD dummy and the corporate tax rate prevailing in the sourcing country as additional regressors. Furthermore, we also use gravity related variables like the log of distance of country c to France ($Distw_c$) as well as past colonial status ($Colony_c$), common language ($Language_c$) and common legal origin ($Same - leg - orig_c$) indicators.[26]

In order to test the relevance of these country and product level variables we start by estimating a simple pooled probit model:

$$y^*_{i,p,c} = \alpha + \beta_1\ k_c + \beta_2\ h_c + \beta_3\ \mu_p + \beta_4\ \mu_f + \beta_5\ Q_c + \beta_6\ k_p + \beta_7\ h_p + \mathbf{CC_c b} + \varepsilon_{i,p,c} \qquad (2)$$

Results are presented in Columns (1) and (2) of Table V with standard errors being clustered at the firm level. In particular in Column (1) we do not consider the variable μ_f that is essentially available for manufacturing firms only, sharply decreasing the number of observations.[27] Column

[26] We do not include GDP per capita for two reasons. First, it is highly correlated with the capital/labor ratio, the human capital/labor ratio as well as with the quality of institutions. Second, although wages can affect the sourcing choice (e.g. in Antràs and Helpman, 2004), GDP per capita is at best a poor proxy for labor costs. Wages and productivity vary across countries and what we would really need is a productivity-deflated measure of wages in country c (we leave this exercise for future work).

[27] Our contractibility measure is restricted to products classified by Rauch (1999), essentially manufacturing, agriculture and mining goods. The same issue applies to a lesser extent to the contractibility of imported products.

(2) presents results for the subset of firms for which information on μ_f is available.

In order to control for firm characteristics we also estimate a conditional fixed effects logit model. In our data many firms import different products from many countries under different sourcing modes. These observations provide the source of identification of our conditional fixed effect logit model.

$$y^*_{i,p,c} = \alpha + \beta_1\ k_c + \beta_2\ h_c + \beta_3\ \mu_p + \beta_4\ Q_c + \beta_5\ k_p + \beta_6\ h_p + \mathbf{CC_c b} + fi + \varepsilon_{i,p,c} \tag{3}$$

where fi is a firm-specific fixed effect potentially correlated with explanatory variables. Results are presented in Column (3) of Table V.

The fixed effects logit specification is very general as it allows us to control for unobserved firm-level characteristics. However, the main drawback is that identification relies only on firms that are engaged in both intra-firm and outsourcing activities in different countries and/or products. This reduces a lot the actual number of observations used in estimations (see at the bottom of Table V) and raises sample selection issues. Another drawback is that we cannot identify the impact of the contractibility of the final good μ_f, as it is firm-specific.

Therefore we also estimate the probit model of Equation (2) on the smaller dataset of 247,528 firm-country-product triples for which firm-level information is available, using additional controls:

$$y^*_{i,p,c} = \alpha + \beta_1\ k_c + \beta_2\ h_c + \beta_3\ \mu_p + \beta_4\ \mu_f + \beta_5\ Q_c + \beta_6\ k_p + \beta_7\ h_p + \mathbf{CC_c b1} + \mathbf{FC_i b2} + \mathbf{D}_{NES} + \varepsilon_{i,p,c} \tag{4}$$

where $\mathbf{FC_i}$ denotes a vector of observable firms characteristics (firm productivity as well as capital, headquarter and skill intensity) and $\mathbf{D}_{NES}$ denotes a full set of industry dummies. Estimation results are shown in column (4) of Table V. To save space we do not report coefficients of covariates

We thank Sébastien Roux for providing us with data on the NACE code of the whole population of French firms. Such information allow us to match μ_f also to manufacturing firms not included in the EAE database.

$\mathbf{FC_i}$ and industry dummies.

Looking across the different sets of estimates in Table V reveals that, with very few exceptions, the sign and significance of coefficients depicts a pretty clear and coherent picture. In particular, we can state the following results:

- *Result 3: Intra-firm trade is more likely with capital scarce countries.* This original finding is at odds with Antràs (2003). This result is robust to considering either k_c as the only explanatory variable or k_c and a subset of the other covariates. In particular the result still holds if we do not include the OECD dummy and/or h_c.

- *Result 4: Intra-firm trade is more likely with countries having good judicial institutions.*

Result 3 stands in sharp contrast with Antràs (2003), who predicts that the share of intra-firm trade should increase with k_c. To what extent do our empirical findings invalidate this prediction? A first caveat is that his result is derived under the very strong assumptions of capital immobility and factor price equalization. However, in an unreported robustness check, we interact capital abundance with OECD membership (a proxy for a common diversification cone) and find a non-significant coefficient. This suggests that the prediction by Antràs fails to hold even in the favorable setting of OECD origin countries.

Second, Result 3 might be driven by legal constraints to sourcing choices, such as foreign ownership restraints (as applied to China in 1999) and/or by origin countries' level of financial development. It is indeed reasonable to believe that FDI (leading to intra-firm trade) can alleviate financial constraints in countries characterized by a low financial development. However, Table IX shows that, controlling for the level of financial development of country c ($Fin - Dev_c$) and excluding China from the data, does not change our findings.[28]

Last but not least, Result 3 regards the firm-level likelihood to undertake intra-firm trade vs outsourcing . This is not directly comparable to empirical evidence on US data that point to the contrary using the share of intra-firm trade in total imports.[29] We devote Section 5 to bridging the gap between our and the above-mentioned results.

[28]The level of financial development is measured by the amount of credit from banks and other financial institutions to the private sector as a share of GDP. The variable is drawn from Beck (2002).

[29]See Antràs (2003), Yeaple (2006), Nunn and Trefler (2008), and Bernard et al. (2008)

Result 4 states that the better is a country's judicial system (high Q_c), the less likely firms are to engage in arms' length relationships. This is in line with the predictions of the Antràs and Helpman (2004) model. As discussed in Section 2, in the more general Antràs and Helpman (2008) model product contractibility in the origin country has two opposite effects: a Standard Effect and a Surprise Effect. The Standard Effect of increasing contractibility points to more arms' length relationships while the Surprise Effect, i.e. a weaker need to provide the supplier with high-powered incentives, goes in the direction of more intra-firm trade. We interpret the positive coefficient of Q_c as evidence that the Surprise Effect dominates.[30] This result is robust to controlling for both imported and final good contractibility as well as firm heterogeneity.

To the extent that stronger legal protection reduces costs of agents' interactions outside the firm, Result 4 challenges the transaction-costs theory of the multinational firm boundaries (e.g. McLaren [2000] and Grossman and Helpman [2002]). Unlike models where property rights provide incentives to imported good suppliers, transaction-cost models cannot rationalize Result 4.

Moving to product characteristics, we report a consistent pattern on the role of intermediate and final product contractibility:

- *Fact 3: The production of complex intermediate and final goods (low μ_p and μ_f) is more likely to occur within the firm boundaries.*

This original finding is not a theoretical prediction of any residual property rights approach model. Nevertheless Fact 3 is qualitatively in line with the transaction-cost approach. To the extent that μ_p and μ_f are negatively correlated with asset specificity, transaction-cost theory predicts that products with low μ's are more likely to be processed within firm boundaries.[31] In

[30] This result at the firm-level confirms the findings in Nunn and Trefler (2008) at the product level. Nunn and Trefler (2008) use an interaction term, that could resemble $\mu_p * Q_c$, to test for the relative importance of the Surprise vs. Standard Effect. However, we do not believe this is a good strategy. Antràs and Helpman (2008) consider an improvement of intermediate input contractibility in the the South, while keeping the level of contractibility of the same input in the North constant. These comparative statics apply to countries, not products, since the theory is essentially silent on the issue of optimal sourcing in the case of many inputs. Empirically, contractibility in the South is a function $f(\cdot)$ of μ_p and Q_c. Therefore we need to estimate the partial derivative of $f(\cdot)$ with respect to Q_c. In the simple log-linear specification we adopt, Q_c and μ_p are separate regressors. The partial derivative of $f(\cdot)$ corresponds to the coefficient multiplying Q_c. In unreported regressions we introduce an additional interaction term, $\mu_p * Q_c$, and find that it is not significant.

[31] Our variables proxy for asset specificity, which creates appropriable quasi-rents. The superiority of ownership in preventing costly haggling over appropriable quasi-rents originates from transaction-cost theory, starting with Williamson (1971).

addition, complex inputs are also likely to embody costly R&D efforts and require physical and legal protection to prevent imitation. Firm boundaries may represent a safe place to put these valuable intangible assets.

As far as skills are concerned, neither skill intensity (h_p) nor skill abundance (h_c) have a clear effect. Coefficients take either sign and are not significant in some cases.

Finally, as for the the case of firm level determinants of intra-firm trade, the model fit suggests that a sizeable amount of heterogeneity is still unexplained by the variables suggested by theory. At the same time, the Pseudo R^2s of the specification in column (4), which is possibly the empirical model of intra-firm trade we estimated so far that is most comparable with our benchmark export status model, reveals that the underlying theory is at least as successful in predicting when trade occurs within firm-boundaries than the productivity/size self-selection mechanisms for export status.

5 Reconciling Firm- and Industry-Level Evidence: the Intensive and Extensive Margins

A number of previous studies have analyzed sourcing modes at the sector or product level: Antràs (2003), Yeaple (2006), Nunn and Trefler (2008), Bernard et al. (2008). By using firm-level data we close the gap between the theory, which is essentially at the firm level, and the empirics. Some of our results, in particular Result 3, challenge previous findings. The aim of this section is to show that the distinction between the extensive and intensive margins of international sourcing is crucial to obtain a coherent picture.

5.1 France is Not Different From the US!

What can explain the difference between our findings and those of the above cited authors? We begin by ruling out structural differences in the pattern of intra-firm trade in France and the US. These may indeed be caused by differences in the specialization of the two economies, or differences in the definition of an affiliate (over fifty percent of the equity in France, over ten percent in the US). We therefore start by replicating the US results with our French data.

Table VI reproduces the cross-industry and cross-country regressions of Antràs (2003) for France. The dependent variable $Share_{cs}$ is the share of intra-firm imports at the country-sector level. We work with the NACE rev1 4-digit industry classification. k_s, η_s^{hq} and η_s^{sk} are sector averages of the corresponding firm-level variables. $Population_c$ is the log of country c's population in 1999, taken from the IMF World Economic Outlook database.

Our results confirm the two key findings in Antràs (2003). We find that intra-firm trade is more likely, the higher capital intensity at the sector level and capital abundance at the country level. This second finding contrasts with our result at the firm level in Table V.

Bernard et al. (2008) look at intra-firm shares at the country-product level ($Share_{pc}$). They point out that at a high level of disaggregation the variable $Share_{pc}$ has many zeros. This suggests that the participation rate in intra-firm trade plays a crucial role, as in Helpman et al. (2008). A selection procedure is thus needed to account for both observed and unobserved determinants of the decision to have foreign affiliates. Bernard et al. (2008) use a Heckman two-stage procedure to account for selection. In particular they run a first-step probit model on the variable $\widetilde{Share}_{pc} = 1$ if $Share_{pc} > 0$.

In Table VII we replicate their estimation strategy. Our excluded variables are $Colony_c$, $Same - leg - orig_c$ and $Population_c$. We find broadly similar results. First, we confirm the positive coefficient of k_c at the product-country level. Second, we also find that the quality of institutions (Q_c) has a positive effect in the selection equation but a negative effect on the the second-stage equation. In other words, the quality of institutions positively affects the likelihood that a given product-country pair will have intra-firm trade, while decreasing the value of intra-firm relative to outsourcing imports.

5.2 Determinants of Sourcing Modes at the Intensive and Extensive Margins

In Section 4 we showed that firm heterogeneity is crucial to understand the discrete choice between intra-firm trade and outsourcing, which we refer to as the extensive margin, and we documented how country and product attributes affect such choice. We now examine how firm-, product- and country-level characteristics affect the value of a transaction for a given sourcing mode ('intensive margin'). This has two advantages. First one possible reason, with the alternative being compo-

sition effects across industries, of why the industry share of intra-firm trade increases with capital abundance, but not the likelihood of an intra-firm transaction, may be that the two margins react differently. Second, we hope to offer new insights to theory since it offers little guidance on the intensive margin.

To meet these two objectives we use a Heckman selection procedure to properly identify the determinants of the intensive margin. We first run a selection equation, based on our preferred specification described in Table V (Column 4). We then run two separate second-stage regressions, one for intra-firm import values and one for outsourcing import values. To save space results of both regressions are reported in Table VIII. [32]

Our exclusion restriction is based on a firm's multinational status in 1994.[33] The persistency of multinational status, that we find in the data, suggests the presence of substantial sunk costs of creating a foreign affiliate. It is our contention that, conditional on the set of other firm variables, past multinational status (very much like past export status in Roberts and Tybout [1997] or Bernard and Jensen [2004]) conveys information on a firm's incentives to currently engage in intra-firm imports without *directly* affecting their value.[34]

In Columns 3 and 4 of Table VIII k_c has a negative and significant coefficient at the outsourcing intensive margin, and a positive and significant coefficient at the intra-firm intensive margin. While capital abundance of a country decreases the likelihood to engage in intra-firm trade, it also increases the value of intra-firm imports relative to outsourcing imports. In light of the positive coefficient in Table VI, we conclude that the effect of capital abundance at the intensive margin dominate when aggregating at the industry level. Note that this result holds within a sector (industry dummies are included), so that composition effects are already controlled for.

How can we interpret this result? First, recall that the models we consider offer little guidance on the *value* of intra-firm and outsourcing transactions. Second, fixed and variable costs may

[32]Because we focus on the intensive margin we now include firm-country-product observations with mixed sourcing modes. We break down each of these observations into an intra-firm and an outsourcing transaction, while attributing the corresponding import values.

[33]We use data from the LIFI ('Liaisons Financières') database. The LIFI is collected by the French national Statistical Office (INSEE) and provides information on ownership relationships across firms that have a legal entity in France.

[34]In our dataset the correlation between multinational status in 1994 and 1999 is 0.38. The correlation between between multinational status in 1994 and y_{ipc} is 0.25.

have different factor intensities, which is assumed away in Antràs (2003). However the negative coefficient at the extensive margin would require that the fixed costs of integration are less capital-intensive than those of outsourcing, relative to the variable cost intensities. This seems rather implausible. Finally, going beyond the property-rights framework, we can offer a conjecture that relates capital abundance to upstream market thickness. Suppose that the upfront entry costs of suppliers cannot be transferred to foreign investors. The greater capital abundance, the thicker the upstream market. Matches with buyers become more valuable, which increases the relative profits of outsourcing (as in McLaren, 2000). Since entry is more costly to finance in capital-scarce countries, the externality does not play as strongly there, rationalizing a negative coefficient for k_c at the extensive margin.

Another interesting finding of the Heckman estimations is that the coefficient of Q_c is positive at the extensive margin but negative at the intensive margin for both intra-firm and outsourcing import values, with the latter coefficient seemingly greater than the former. This is reminiscent of the result in Bernard et al. (2008) as well as in our Table VII based on product-country intra-firm shares. One simple explanation is that legal systems matter more for the fixed costs of integration while they matter more for the variable costs of outsourcing.

Finally the contractibility of the imported product μ_p has a negative effect on both the extensive and intensive margins of intra-firm trade while positively affecting outsourcing import values. Not surprisingly we thus find in Table VII that product contractibility decreases the share of intra-firm trade at the product-country level. This result is consistent with Bernard et al. (2008), although our measure of contractibility is less disaggregated.

6 Conclusion

We have built a unique dataset of 1,141,393 French import transactions in 1999 (corresponding to 126,926 importers, 232 destinations and 274 products) where we can identify whether a transaction is intra-firm or at arms' length. We have conducted a detailed examination of firm-, country- and product-level determinants of intra-firm trade while providing evidence of the key distinction between the extensive and intensive margins of international sourcing modes.

Our first finding is that many dimensions of firm heterogeneity matter for the internalization choice. Firms differ in their productivity, but also in their capital-, skill-, and headquarter-intensity, even within narrowly defined industries. Highly productive, capital-, skill- and headquarter-intensive firms are more likely to engage in intra-firm trade. This result is in line with the basic mechanisms of the residual property rights approach due to Antràs (2003), Antràs and Helpman (2004), and Antràs and Helpman (2008). Importantly, these many dimensions of importer heterogeneity may matter in our understanding of the relationship between trade and income inequality.

Organization theorists may not be surprised that the firm, not the industry, is the appropriate level of disaggregation to explain internalization decisions. After all, there are several non-mutually exclusive theories of the firm, and their importance can vary from firm to firm. Moreover, these theories are highly abstract and may require some firm-specific tailoring in empirical applications (as pointed out by Gibbons, 2005). But, interestingly, our findings reveal some strongly significant empirical regularities rather than idiosyncrasies only.

A second finding is that a firm is more likely to engage in intra-firm imports from *capital-scarce* countries. This holds even when controlling for observable and unobservable firm characteristics. This surprising result goes against the Antràs (2003) prediction as well as consistent evidence on *industry- and product-level* US imports. However, we find that the share of French intra-firm imports is positively correlated with capital abundance at the industry-country and product-country levels, just as in US studies. Our two-stage regression analysis reveals that capital abundance has a positive impact on the value of intra-firm transactions relative to outsourcing imports, and that effect dominates in industry- and product-level regressions.

Interpreting this surprising result is not straightforward. Existing theory does not offer predictions on the value of intra-firm relative to outsourcing transactions (as acknowledged in Antràs, 2003). We may however offer a conjecture. Suppose that independent suppliers cannot fully transfer the capital costs of entry to foreign investors. In that case capital-rich countries are more likely to benefit from 'thick-market externalities' among independent suppliers, for example through the alleviation of ex-post hold-up problems, as in McLaren (2000), or search frictions, as in Grossman and Helpman (2005). That way capital abundance could increase the relative

profitability of outsourcing.

Obviously our data are not rich enough to test this conjecture rigorously. Nevertheless, our results do suggest that the supporting evidence for Antràs (2003) may be less compelling than previously thought, when the extensive and intensive margins of intra-firm imports are examined separately. This confirms the importance of firm-level evidence.

Third, we find that intra-firm imports are more likely to come from countries with strong judicial institutions. This result conforms with property-rights theory, especially as we find that the effect is strongest for highly productive, capital-, skill-, headquarter-intensive firms. In these theories better enforcement reduces the need to offer high-powered incentives (ownership) to a foreign supplier. In contrast, it is hard to make sense of this result using transaction-cost theory.

This result also relates with empirical findings that the quality of law enforcement affects trade flows (Anderson and Marcouiller 2002, Nunn 2007, Ranjan and Tobias 2007). Whether the trade-reducing effects of judicial uncertainty are stronger within firms than between firms (as our Table VIII suggests) is an interesting topic for future research.

Fourth, we find some robust empirical evidence that complex goods and inputs are more likely to be produced within firm boundaries. This is consistent with the property-rights model by Carluccio and Fally (2009), where the desirability of transferring ownership to suppliers of complex products is limited by the latter's financial constraints. Our finding, however, is also consistent with a dissipation of intangible assets explanation. Complex inputs embody costly R&D efforts or the use of other intangible assets, which are likely to more effectively protected against imitation within firm boundaries.

Overall, we believe that our findings offer a useful test of various internalization theories. With the important caveat that we do not observe domestic outsourcing, they also shed additional light on the more general question of the make-or-buy choice.

A Data Appendix

A.1 The EIIG database

Intra-firm trade is defined in the EIIG database as trade with an affiliate controlled by a single French entity with at least fifty percent of its equity capital. The SESSI defines two types of trade with independent suppliers: 1) formal contractual relationships that refer to alliances, franchising, joint-ventures, and licensing agreements; 2) 'informal' relationships that involve less stringent contractual links. We consider both types of trade with independent suppliers as outsourcing. In the data 20,952 out of the 81,217 import flows (25.80%) are 'pure' intra-firm (in the sense that 100% of imports of product p from country c come from a foreign affiliate), 50,021 (61.59%) are 'pure' outsourcing,[35] and 10,244 (12.61%) are 'mixed transactions'. For a detailed description of the EIIG database see Guannel and Plateau (2003).

A.2 TFP estimations

The primary EAE database consists of an unbalanced panel of 28,587 firms over 3 years (1998 to 2000) for a total of 74,120 observations. Observations with negative values of value added, production, capital stock and material inputs are eliminated. Outliers, identified as observations falling outside the 1st and 99th percentile of the distributions of value added per worker and capital stock per worker, are also not considered for TFP estimation. This leaves us with TFP information on 22,928 firms for the core year 1999. TFP estimation has been carried out separately for each of the 56 NES114 industries in manufacturing.

Total Factor Productivity of firm i is the residual (plus the constant) of a log-linearized three-factor Cobb-Douglas production function, with labor, capital and material inputs as production inputs. We use the value-added based Levinsohn and Petrin (2003) estimator. Labor is the full time equivalent average number of employees in a given year. Material inputs are calculated as input purchases minus inventory variation. Deflators for value added and material inputs are obtained from the national accounts system of the French Statistical Office (INSEE) at the NACE

[35] In particular 48,603 are pure informal third party imports, 1,093 are pure formal third party imports and 325 are mixed formal and informal third party imports.

rev1 two digit level.

A.3 Construction of contractibility variables

The basic data needed to construct contractibility measures comes from Rauch (1999) and are organized on the basis of the SITC rev2 4 digit (975 products for which information is available). Our import data are at the CPA96 4digit classification (490 products). However, the Rauch classification covers almost exclusively manufacturing and agricultural goods. Restricting ourselves to such goods leaves us with 247 CPA96 4digit products.

In order to aggregate the Rauch data to construct a measure of contractibility for imported goods, we have first established a correspondence between HS6 and SITC rev2 4 digit and a correspondence between HS6 and CPA96 4digit.[36] We have then used import trade data in 1999 for France at the HS6 level (provided by EUROSTAT) as weights to aggregate the original SITC rev2 4 digit information to the CPA96 4digit.

As to final product contractibility, we have first used a correspondence table form the PRODCOM2002 8 digit classification to the HS6 provided by EUROSTAT. Then, exploiting the previously build HS6 to SITC and HS6 to CPA correspondence tables, we have used production data in 1999 for France at the PRODCOM2002 8 digit classification level (provided by EUROSTAT) as weights to aggregate the original SITC rev2 4 digit information to the CPA96 4digit.

A.4 Other imported product variables

In order to build k_p and h_p, we start by using a correspondence table between the industry classification NACE rev1 4digit (available in our EAE firm dataset) and the product classification CPA96 4digit. We then compute the average capital intensity (log of capital/labor ratio) and skill intensity (log of total wage expenses/number of full time equivalent workers) of French firms associated to a given CPA96 4digit product.

[36] Correspondence tables have been obtained using RAMON data available from EUROSTAT at the website: http://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL

A.4.1 Origin Country Variables

k_c and h_c are (respectively) the log of the capital/labor and human capital/labor ratios provided by Hall and Jones (1999).

Q_c is the "Rule of Law" index from Kaufmann, Kraay, and Mastruzzi (2003). This is a weighted average of a number of variables that measure individuals' perceptions of the effectiveness and predictability of the judiciary and the enforcement of contracts in each country between 1997 and 1998. Tax_c is the top corporate tax rate prevailing in a given country in 1999 taken from the World Tax Database (University of Michigan). $Same - leg - orig_c$ is borrowed from Djankov et al. (2003).

The last set of variables ($Distw_c$, $Colony_c$, and $Language_c$) comes from CEPII (Centre d'Etude Prospectives et d'Informations Internationales). $Distw_c$ is the log of distance of country c to France. Distance is calculated by aggregating regional distances at the country level, using regional populations as weights. Further details can be found in Head and Mayer (2002).

B Interaction Between Firm Heterogeneity and Country/Product Characteristics on Intra-Firm Trade

In Subsection 4.1 we have explored the role of firm heterogeneity in explaining the offshore sourcing mode, while in Subsection 4.2 we have looked at the impact of some country and product characteristics. We push here the analysis of heterogeneity further by looking at at interactions between firm and product/country variables.

This risk of such an exercise is to run into a taxonomy of stylized facts that would not be very valuable for the reader. However there are at least two interesting cases to analyze.

First, Antràs and Helpman (2008) show that the Standard Effect comes from the sub-population of relatively low productivity firms. By contrast, the opposite Surprise Effect comes from high productivity firms. Thanks to our firm-level data we can identify the tension between the Standard and the Surprise Effect by looking at the interaction between Q_c and productivity (as well as capital, skill and headquarter intensity).

In particular, for each NES114 industry, we have computed the 25th, 50th, and 75th percentile of the distributions of TFP_i, k_i, η_i^{hq}, and η_i^{sk}. We have then created two dummies for each variable to identify observations below and above the median and attributed to the former (letter) the corresponding value of the 25th (75th) percentile. Finally we have build a cross product between each of the two dummies and Q_c and re-estimated equation (4), with errors clustered at the firm level, while adding these interaction effects. We have performed 4 different estimations, one for each of the 4 interacting variables. A chi-test is then run to detect significantly (at 5%) different coefficients.

Results of such estimations, available upon request, lead to the following result:

- *Result 5: The 'surprise' effect is significantly stronger for more productive and skill intensive firms.*

This result echoes the heterogeneous impact of Q_c in Antràs and Helpman (2008). The evidence on capital and headquarter intensity goes in the same direction of productivity and skill intensity but it is not robust to firm clustering.

The second intriguing question is related to contractibility. We have seen that firms prefer to produce complex inputs and final goods within the firm boundaries. However, it would be interesting for future theoretical work to know whether firms with certain observable characteristics are more likely to do so. For each NES114 industry, we again interact dummies (constructed using the median) taking the value of the the 25th and 75th percentile of the distributions of TFP_i, k_i, η_i^{hq}, and η_i^{sk} with respectively μ_p and μ_f. Again, standard errors are clustered by firm and a chi-test is used to detect significantly (at 5%) different coefficients:

- *Fact 4: Complex intermediate goods are more likely to be involved in an intra-firm transaction the higher (lower) is the capital (headquarter) intensity of a firm. No significant pattern arises with respect to final goods complexity.*

References

[1] Anderson, James E. and Douglas Marcouiller, "Trade, Insecurity, and Home Bias: An Empirical Investigation," *The Review of Economics and Statistics 84:2* (2002), 342-352

[2] Antràs, Pol, "Firms, Contracts and Trade Structure," *Quarterly Journal of Economics*, 118 (2003), 1375-1418.

[3] Antràs, Pol, and Elhanan Helpman, "Global Sourcing," *Journal of Political Economy*, 112 (2004), 552-580.

[4] Antràs, Pol, and Elhanan Helpman, "Contractual Frictions and Global Sourcing," forthcoming in Helpman, E., D. Marin, and T. Verdier, *The Organization of Firms in a Global Economy*, Harvard University Press, 2008.

[5] Antràs, Pol, and Esteban Rossi-Hansberg, "Organizations and Trade," *NBER Working Paper* 14262, 2008

[6] Barba Navaretti, Giorgio, and Anthony J. Venables, *Multinational Firms in the World Economy* (Princeton: Princeton University Press, 2004).

[7] Beck, Thorsten "Financial Development and International Trade. Is There a Link?" *Journal of International Economics*, 57 (2002), p.107-131.

[8] Bernard, Andrew B. and J. Bradford Jensen, "Why Some Firms Export," *The Review of Economics and Statistics*, 86(2) (2004), 447-464.

[9] Bernard, Andrew B., J. Bradford Jensen, Stephen Redding and Peter K. Schott, "Intra-Firm Trade and Product Contractibility," mimeo, 2008.

[10] Carluccio, Juan F. and Thibault Fally, "Global Sourcing Under Imperfect Capital Markets, "mimeo Paris School of Economics, 2009.

[11] Defever, Fabrice, and Farid Toubal, "Productivity and the Sourcing Modes of International Firms: Evidence from French Firm-Level Data,"mimeo, 2009. A previous substantially different version circulated as a CEP Discussion Paper No 842 (2007).

[12] Djankov, Simeon, Rafael La Porta, Florencio López-de-Silanes, and Andrei Shleifer, "Courts", *Quarterly Journal of Economics*, 118 (2003), 453-517.

[13] Eaton, Jonathan, Samuel Kortum, and Francis Kramarz "Dissecting Trade: Firms, Industries, and Export Destinations, " *American Economic Review Papers and Proceedings*, 94 (2004), 150-154.

[14] Feinberg, Susan and Michael Keane, "Accounting for the Growth of MNC-based Trade using a Structural Model of U.S. MNCs,"*American Economic Review*, 96(5) (2006), 1515-1558

[15] Griliches, Z. and Mairesse, J., "Production functions: the search for identification."In: Strm, S., Editor, *Econometrics and Economic Theory in the 20th Century* (Cambridge: Cambridge Univ. Press, 1998), 169203.

[16] Grossman, Gene M. and Elhanan Helpman, "Integration versus outsourcing in industry equilibrium," *Quarterly Journal of Economics*, 117 (2002): 85-120.

[17] Grossman, Gene M. and Elhanan Helpman, "Outsourcing versus FDI in industry equilibrium," *Journal of the European Economic Association* 1 (2003), 317-327.

[18] Grossman, Gene M. and Elhanan Helpman, "Managerial incentives and the international organization of production," *Journal of International Economics*, 63 (2004): 237-262.

[19] Grossman, Gene M. and Elhanan Helpman, "Outsourcing in a global economy," *Review of Economic Studies*, 72 (2005): 135- 159.

[20] Grossman, Sanford J., and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94(4) (1986), 691-719.

[21] Guannel, Boris, and Claire Plateau, "Les échanges internationaux intragroupe dans la mondialisation industrielle en 1999," SESSI (2003).

[22] Hall, Robert E., and Charles I. Jones, "Why do Some Countries Produce So Much More Output per Worker than Others? " *Quarterly Journal of Economics*, 114 (1999), 83-116.

[23] Hart, Oliver, and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98(6) (1990), 1119-58.

[24] Head Keith and Thierry Mayer, "Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade,"(2002), *CEPII Working Paper* 2002-01.

[25] Helpman, Elhanan, Dalia Marin, and Thierry Verdier, *The Organization of Firms in a Global Economy* (Cambridge, MA: Harvard University Press, 2008).

[26] Helpman, Elhanan, Marc Melitz and Yona Rubinstein (2008), " Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, 123(2) (2008), 441-487.

[27] Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "Governance Matters III: Governance Indicators for 1996–2002," *World Bank Working Paper* No. 3106, 2003.

[28] Levchenko, Andrei, "Institutional Quality and International Trade," *Review of Economic Studies*, 74 (2007), 791-819.

[29] Levinsohn James, and Amil Petrin "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70 (2003), 317-342.

[30] Markusen, James, "The Boundaries of Multinational Enterprises and the Theory of International Trade," *Journal of Economic Perspectives*, 9(2) (1995):169-189.

[31] McLaren John, "Globalization and Vertical Structure, " *American Economic Review* 90:5 (2000): 1239-1254.

[32] Marin, Dalia and Thierry Verdier, "Globalization and the New Enterprise," *Journal of the European Economic Association*, 1(2-3) (2003): 337-344.

[33] Marin, Dalia and Thierry Verdier, "Power Inside The Firm and The Market: A General Equilibrium Approach," *Journal of the European Economic Association*, 6(4) (2008), 752-788.

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

[35] Nunn, Nathan, "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade," *Quarterly Journal of Economics*, 122 (2007), 569-600.

[36] Nunn, Nathan, and Daniel Trefler, "The Boundaries of the Multinational Firm: An Empirical Analysis," in Helpman, E., D. Marin, and T. Verdier, *The Organization of Firms in a Global Economy* (Cambridge, MA: Harvard University Press, 2008).

[37] Olley, Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64 (1996), 1263-1297.

[38] Ranjan, Priya and Justin Tobias, "Bayesian Inference for the Gravity Model," *Journal of Applied Econometrics*, 22(4) (2007), 817-838.

[39] Rauch, James E., "Networks versus Markets in International Trade, " *Journal of International Economics*, 48 (1999): 7-35.

[40] Roberts, Mark and James Tybout, "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 87(4) (1997), 545-563.

[41] Spencer, Barbara, "International outsourcing and incomplete contracts," *Canadian Journal of Economics*, 38(4) (2005), 1107-1135.

[42] Tomiura, Eiichi, "Foreign Outsourcing, Exporting, and FDI: A Productivity Comparison at the Firm Level, " *Journal of International Economics*, 72 (2007), 113-127.

[43] Williamson, Oliver, "The Vertical Integration of Production: Market Failure Considerations, " *American Economic Review* 61:2 (1971): 112-123.

[44] Yeaple, Stephen, "Offshoring, Foreign Direct Investment and the Structure of U.S. Trade," *Journal of the European Economic Association*, 4(2-3) (2006): 602-611.

[45] Zeile, William J., "U.S Intrafirm Trade in Goods, " *Survey of Current Business*, 77(2) (1997), 23-38.

Table I: Summary Statistics on μ_p and μ_f

Variable	Observ.	Mean	St. Dev	Min	Max
μ_p	259	0.415	0.448	0	1
μ_f	218	0.373	0.440	0	1

Table II: Summary statistics on firm-level variables

Variable	Obs	Mean	Std. Dev.	Min	Max	Correlation with TFP_i	k_i	η_i^{hq}	η_i^{sk}
TFP_i	247528	3.9955	0.9466	-2.8433	8.7176	1			
k_i	247528	3.6260	0.9761	-8.2213	8.3878	0.17	1		
η_i^{hq}	247528	0.6251	0.1808	0	1	0.08	-0.06	1	
η_i^{sk}	247132	3.0955	0.3108	-6.6951	5.3584	0.31	0.25	0.14	1

Table III: Deviations from industry averages: firm-level variables

Variable	Obs	Mean	Std. Dev.	Min	Max	Correlation with $\Delta_s TFP_i$	$\Delta_s k_i$	$\Delta_s \eta_i^{hq}$	$\Delta_s \eta_i^{sk}$
$\Delta_s TFP_i$	247528	0	0.4487	-2.3820	4.3999	1			
$\Delta_s k_i$	247528	0	0.8246	-13.7051	2.9040	0.16	1		
$\Delta_s \eta_i^{hq}$	247528	0	0.1659	-0.8257	0.5809	0.10	-0.01	1	
$\Delta_s \eta_i^{sk}$	247132	0	0.2770	-9.7548	2.0918	0.46	0.21	0.12	1

Table IV: Firm i-specific determinants of intra-firm trade.

Dependent variable: $y_{i,p,c}$=1 for intra-firm.	1	2	3	4	5	6
TFP_i	0.0411***				0.0291***	0.0324***
	(0.0035)				(0.0041)	(0.0036)
k_i		0.0212***			0.0154***	0.0054***
		(0.0024)			(0.0020)	(0.0019)
η_i^{hq}			0.0345***		0.0117	0.0133*
			(0.0102)		(0.0083)	(0.0078)
η_i^{sk}				0.0595***	0.0172*	0.0156*
				(0.0142)	(0.0097)	(0.0086)
$size_i$						0.0135***
						(0.0012)
NES114 sectoral dummies	yes	yes	yes	yes	yes	yes
Country and product dummies	yes	yes	yes	yes	yes	yes
Number of observations	242,804	242,804	242,804	242,414	242,414	242,414
Pseudo R^2	0.1669	0.1562	0.1314	0.1574	0.1897	0.2264
Log Likelihood	-43,258	-43,809	-45,097	-43,711	-42,037	-40,135

Marginal effects presented. Firm-clustered standard errors in brackets. ***, ** , * denote significantly different from 0 at 1%, 5% and 10% level, respectively.

Table V: Product and Country-specific determinants of intra-firm trade.

Dependent variable: $y_{i,p,c}$	1	2	3	4
k_c	-0.0019***	-0.0109***	-0.0142***	-0.0111***
	(0.0007)	(0.0017)	(0.0051)	(0.0018)
h_c	0.0039	0.0298***	-0.0121	0.0156**
	(0.0027)	(0.0068)	(0.0147)	(0.0065)
μ_p	-0.0077***	-0.0195***	-0.0293**	-0.0086***
	(0.0011)	(0.0025)	(0.0137)	(0.0021)
μ_f		-0.0124**		-0.0080***
		(0.0049)		(0.0017)
Q_c	0.0333***	0.0702***	0.0619*	0.0470***
	(0.0045)	(0.0109)	(0.0353)	(0.0103)
k_p	0.0059***	0.0095***	-0.0127**	0.0012
	(0.0007)	(0.0017)	(0.0060)	(0.0015)
h_p	0.0313***	0.0464***	0.0177	-0.0037
	(0.0019)	(0.0053)	(0.0131)	(0.0040)
Controls				
Tax_c	-0.0241***	-0.0601***	-0.1576**	-0.0439**
	(0.0086)	(0.0211)	(0.0723)	(0.0205)
$OECD_c$	0.0052***	0.0061*	0.0188**	0.0043
	(0.0012)	(0.0032)	(0.0091)	(0.0031)
$Distw_c$	0.0027***	0.0088***	-0.0078***	0.0024
	(0.0007)	(0.0017)	(0.0028)	(0.0017)
$Colony_c$	-0.0035**	-0.0032	0.1573***	0.0070
	(0.0016)	(0.0039)	(0.0522)	(0.0048)
$Language_c$	-0.0025**	-0.0063**	0.0083	-0.0047
	(0.0011)	(0.0029)	(0.0074)	(0.0030)
$Same-leg-orig_c$	0.0007	0.0140***	0.0222**	0.0147***
	(0.0007)	(0.0021)	(0.0102)	(0.0020)
Estimation method	pooled probit	pooled probit	conditional firm fixed effects logit	probit with controls $\mathbf{FC_c, D_{NES}}$
Number of potential observations	1,141,393	366,376	1,141,393	247,528
Number of actual observations used in estimations	989,165	308,666	36,746	209,706
Pseudo R^2	0.0290	0.0197		0.1509
Log Likelihood	-104,176	-56,341	-14,403	-40,129

Marginal effects presented. In the fixed effects logit case, calculations are obtained by setting fixed effects to zero. Firm-clustered standard errors (except for the fixed effects logit) in brackets. ***, ** , * denote significantly different from 0 at 1%, 5% and 10% level, respectively.

Table VI: Share of intrafirm trade at the industry-country level

Dependent variable: $Share_{cs}$	1	2
k_s	0.0002***	
	(0.0000)	
η_s^{hq}	-0.2069***	
	(0.0256)	
η_s^{sk}	0.0039***	
	(0.0005)	
μ_f	-0.0660***	
	(0.0057)	
k_c		0.0078***
		(0.0017)
h_c		0.0324***
		(0.0074)
$Distw_c$		-0.0184***
		(0.0015)
$Colony_c$		0.0034
		(0.0046)
$Language_c$		-0.0157***
		(0.0044)
$Population_c$		0.0064***
		(0.0009)
constant	-0.1314	0.0392
	(0.2171)	(0.0481)
Country dummies	yes	no
Industry dummies	no	yes
Number of observations	10,688	10,680
R-squared	0.1075	0.2668

Robust standard errors in brackets. ***, ** , * denote significantly different from 0 at 1%, 5% and 10% level, respectively.

Table VII: Share of intrafirm trade at the product-country level: Heckman selection model

	First Stage	Second Stage
Dependent variable:	$\widetilde{Share}_{pc}$	$Share_{pc}$
k_p	-0.0045	0.0538***
	(0.0256)	(0.0085)
h_p	0.2945***	0.1731***
	(0.0784)	(0.0254)
μ_p	-0.2018***	-0.1070***
	(0.0408)	(0.0132)
k_c	0.1060***	0.0449***
	(0.0267)	(0.0097)
h_c	0.4221***	0.0329
	(0.1032)	(0.0319)
Q_c	2.5276***	-0.0893*
	(0.1499)	(0.0502)
$Distw_c$	-0.3807***	-0.0270***
	(0.0171)	(0.0064)
$Language_c$	-0.1682***	-0.0682***
	(0.0541)	(0.0145)
$Colony_c$	0.1995***	
	(0.0595)	
$Same-leg-orig_c$	0.3369***	
	(0.0404)	
$Population$	0.3387***	
	(0.0117)	
Mills ratio	0.1915***	
	(0.0173)	
constant	-2.4538***	-0.8520***
	(0.3459)	(0.1182)
Number of observations	9980	3414

Robust standard errors in brackets. ***, ** , * denote significantly different from 0 at 1%, 5% and 10% level, respectively.

Table VIII: The Extensive and Intensive margin of firms' international sourcing: Heckman selection model

	First Stage	Second Stage	
Dependent variable:	y_{ipc}	Intra-firm import value	Outsourcing import value
$1994 Multinational$	0.6114***		
	(0.0101)		
TFP_i	0.4284***	0.3574***	0.5593***
	(0.0110)	(0.0591)	(0.0170)
k_i	0.2161***	0.0929***	0.1926***
	(0.0061)	(0.0325)	(0.0085)
η_i^{hq}	-0.1504***	-0.3354***	-1.2083***
	(0.0262)	(0.1120)	(0.0362)
η_i^{sk}	0.2463***	-0.1833***	-0.1765***
	(0.0180)	(0.0702)	(0.0243)
k_c	-0.1347***	0.1037*	-0.2051***
	(0.0135)	(0.0599)	(0.0192)
h_c	0.1454***	-0.1806	-0.6455***
	(0.0447)	(0.1927)	(0.0663)
μ_p	-0.0670***	-0.1576**	0.8442***
	(0.0141)	(0.0624)	(0.0192)
μ_f	-0.0133	-0.2588***	-0.0896***
	(0.0203)	(0.0904)	(0.0283)
Q_c	0.7075***	-1.9814***	-1.3904***
	(0.0740)	(0.3201)	(0.1126)
k_p	0.0177*	0.2826***	0.1597***
	(0.0091)	(0.0405)	(0.0126)
h_p	-0.0994***	1.0021***	-0.2391***
	(0.0264)	(0.1144)	(0.0381)
Mills ratio		-0.1810**	2.7551***
		(0.0844)	(0.0721)
Tax_c	0.0480	-2.6741***	-1.5106***
	(0.1809)	(0.7772)	(0.2476)
$OECD_c$	0.0099	0.0575	0.6402***
	(0.0237)	(0.0983)	(0.0337)
$Distw_c$	-0.0037	-0.1056***	0.1507***
	(0.0084)	(0.0349)	(0.0126)
$Colony_c$	0.0726***	0.8208***	0.2205***
	(0.0219)	(0.0946)	(0.0334)
$Language_c$	-0.1378***	0.0949	0.5920***
	(0.0147)	(0.1004)	(0.0304)
$Same - leg - orig_c$	0.1599***	-0.2749***	0.0151
	(0.0168)	(0.0751)	(0.0234)
$Constant$	-4.4813***	9.2952***	10.2098***
	(0.1697)	(0.8412)	(0.2412)
Industry dummies	yes	yes	yes
Number of observations	220282	17857	202425

The first stage equation is a probit on the binary variable y_{ipc}. The second stage equation is an OLS regression of the value of imports in a given mode. To save space we present here the results for both modes.
Standard errors in brackets. ***, ** , * denote significantly different from 0 at 1%, 5% and 10% level, respectively.

Table IX: Product and Country-specific determinants of intra-firm trade: Robustness check.

Dependent variable: $y_{i,p,c}$	1	2	3	4
k_c	-0.0025***	-0.0112***	-0.0227***	-0.0108***
	(0.0008)	(0.0018)	(0.0080)	(0.0018)
h_c	0.0060**	0.0307***	-0.0104	0.0140**
	(0.0027)	(0.0066)	(0.0191)	(0.0065)
μ_p	-0.0081***	-0.0199***	-0.0372**	-0.0086***
	(0.0011)	(0.0025)	(0.0179)	(0.0021)
μ_f		-0.0124**		-0.0078***
		(0.0050)		(0.0017)
Q_c	0.0460***	0.1000***	0.1612*	0.0784***
	(0.0051)	(0.0133)	(0.0850)	(0.0122)
k_p	0.0062***	0.0096***	-0.0177**	0.0012
	(0.0007)	(0.0017)	(0.0085)	(0.0015)
h_p	0.0319***	0.0462***	0.0224	-0.0040
	(0.0020)	(0.0052)	(0.0169)	(0.0040)
		Controls		
$Fin - Dev_c$	-0.0081***	-0.0177***	-0.0472**	-0.0169***
	(0.0014)	(0.0034)	(0.0238)	(0.0032)
Tax_c	-0.0218***	-0.0475**	-0.1615*	-0.0328
	(0.0083)	(0.0200)	(0.0841)	(0.0199)
$OECD_c$	0.0054***	0.0057*	0.0215*	0.0038
	(0.0011)	(0.0031)	(0.0111)	(0.0030)
$Distw_c$	0.0044***	0.0122***	-0.0015	0.0056***
	(0.0007)	(0.0018)	(0.0030)	(0.0016)
$Colony_c$	-0.0037**	-0.0028	0.1945***	0.008
	(0.0017)	(0.0041)	(0.0643)	(0.0051)
$Language_c$	-0.0015	-0.0045	0.0133	-0.0028
	(0.0011)	(0.0029)	(0.0105)	(0.0030)
$Same - leg - orig_c$	-0.0009	0.0098***	0.0190*	0.0109***
	(0.0007)	(0.0019)	(0.0099)	(0.0019)
Estimation method	pooled probit	pooled probit	conditional firm fixed effects logit	probit with controls $\mathbf{FC_i}, D_s$
Number of potential observations	1,141,393	366,376	1,141,393	247,528
Number of actual observations used in estimations	949,698	302,239	35,802	205,542
Pseudo R^2	0.0292	0.0211		0.1520
Log Likelihood	-101,661	-55,192	-13,949	-39,310

Marginal effects presented. In the fixed effects logit case, calculations are obtained by setting fixed effects to zero. Firm-clustered standard errors (except for the fixed effects logit) in brackets. ***, ** , * denote significantly different from 0 at 1%, 5% and 10% level, respectively.

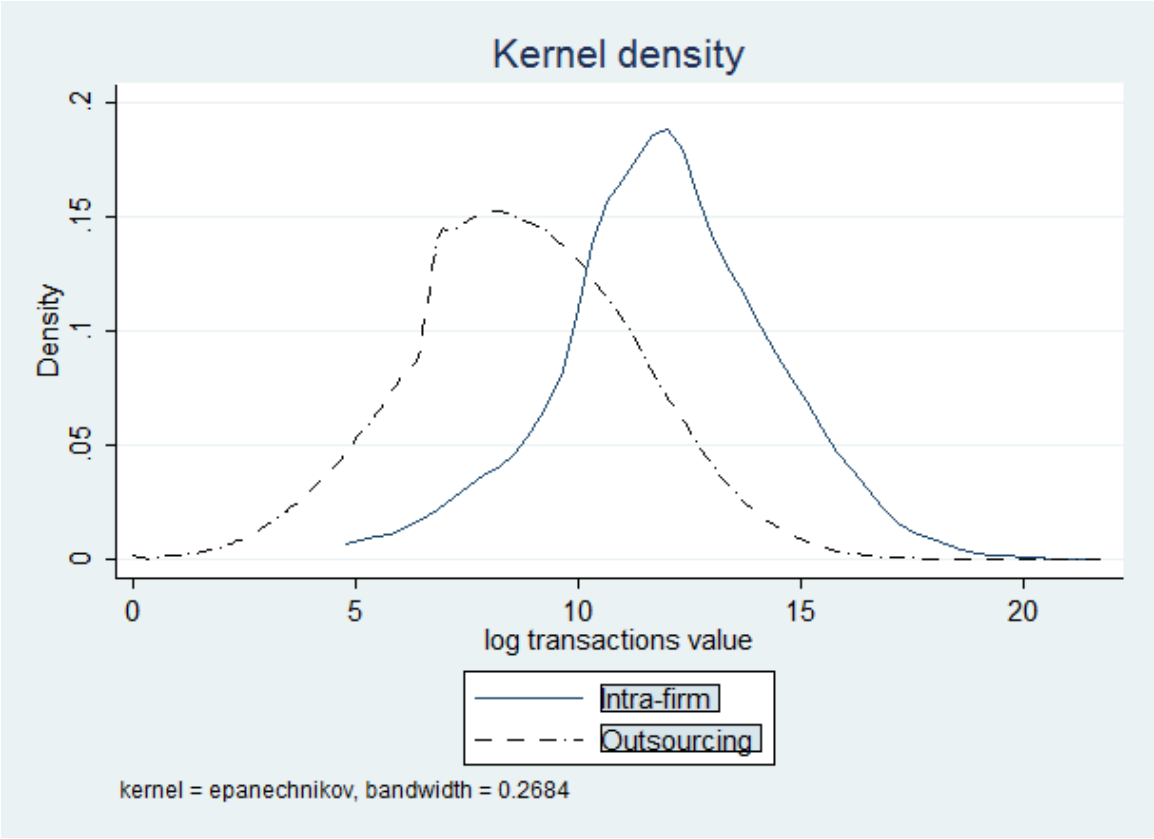

Figure I: Kernel smoothed distribution of log transactions' value