

The Determinants of Intra-Firm Trade*

Gregory Corcos[†] Delphine M. Irac[‡] Giordano Mion[§] Thierry Verdier[¶]

Draft version: April 8, 2009. First version: 26 January 2008

Abstract

This paper analyzes the determinants of the sourcing mode of imported inputs at the firm level. We exploit a unique French dataset of 1,141,393 import transactions spanning across firm, countries and products in 1999, where we observe whether a transaction is intra-firm or at arms' length. We first study which firm-, country- and product-specific factors affect the 'make or buy' choice at the firm level (extensive margin). We confirm a number of theoretical predictions of property-rights models, and provide a number of empirical facts that can be used to refine theory. We finally compare our results with previous findings on more aggregated data, highlighting the importance of separating the extensive and intensive margins of imports.

Keywords: intra-firm trade; outsourcing; firm heterogeneity; incomplete contracts; internationalization strategies; quality of institutions, extensive margin, intensive margin.

JEL Classification: F23, F12, F19

*We thank Andrew Bernard, Thierry Mayer, Dalia Marin, Jim Markusen, Steve Redding, Sébastien Roux, and Alessandro Sembenelli as well as seminar participants at ERWIT 2008 (Appenzell), the Ludwig-Maximilian University of Munich and the workshop "The International Firm: Access to Finance and Organisational 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.

[†]Department of Economics, NHH. E-mail: gregory.corcos@nhh.no.

[‡]Banque de France. E-mail: delphine.irac@banque-france.fr.

[§]Corresponding author. Department of Geography and Environment, London School of Economics; NBB; CEP; CEPR. Address: Houghton St, WC2A 2AE, London, UK. Fax: +44 20 7955 7412. E-mail: G.Mion@lse.ac.uk

[¶]Paris School of Economics; CEPR. E-mail: verdier@pse.ens.fr.

1 Introduction

About every third transaction in international trade occurs within a multinational firm. An additional third has at least one multinational as a party to the transaction. This remarkable current pattern of multinationalization in the global economy has been the source of an important amount of work in international trade theory.

One strand of the theoretical literature has provided explanations of internalization decisions by looking at the costs of dissipating intangible assets (see the surveys in Markusen, 1995 and Barba Navaretti and Venables, 2004). Another more recent line of research emphasizes the costs and benefits of allocating residual rights of control when contracts are incomplete. This rapidly growing literature discusses the determinants of intra-firm trade.¹ Many of these papers have underlined various elements that constitute the internal organization of the firm in an international context². Important theoretical insights have been derived on crucial determinants of international trade transactions that are done internally to the firm (intra-firm trade) and those which are done outside the boundaries of the firm (outsourcing). In particular, a central insight of this literature has been to explain a firm's internalization decision by the intensity of downstream production in intermediate inputs; with as a main prediction the fact that ownership rights should be given to the party contributing to the intensive production stage.

While this theoretical literature has been expanding rapidly, the empirical evidence on these dimensions has been more limited (Antràs 2003, Yeaple 2006, Nunn and Trefler 2008, Bernard et al. 2008) and exclusively focused on the US. All of these studies investigate the determinants of the share of total US imports that is intra-firm. Broadly, they find support for the predictions by Antràs (2003) and Antràs and Helpman (2004, 2008) explaining a firm's internalization decision by the intensity of downstream production in intermediate inputs.

With respect to the theoretical framework, a limitation of this approach is obviously the fact that this type of empirical evidence remains confined to country/sectoral level analyses, while the theory explicitly emphasizes the importance of firm-level determinants. Interestingly, although the availability of firm-level data has led to a great deal of studies of export behavior (as surveyed by Bernard et al. 2007, Mayer and Ottaviano 2007), much less effort has been spent on understanding import behavior, and the boundaries of multinational firms in particular. Hence little micro-evidence has been provided so far on the internalization of international transactions

¹See for instance McLaren (2000), Antràs (2003), Grossman and Helpman (2002, 2003, 2004, 2005), Antràs and Helpman (2004, 2008), Marin and Verdier (2003, 2007)

²Some of the most illustrative recent work along this line of research is published in Helpman, Marin and Verdier (2008).

at a detailed firm/product level. The main purpose of this paper is to fill that gap.

Taking advantage of a dataset documenting imports of manufactured goods by French firms in 1999, this paper aims at identifying the determinants of the internalization of trade flows in intermediate products at the firm, industry, product, and country levels. In particular, we reconsider the validity of the predictions made in recent theoretical studies of multinationals' organizational choice and reassess the empirical results of the above cited papers. Moreover we report a number of additional findings and suggest interpretations that would be useful for further theoretical developments.

An important contribution of our study is to introduce two additional dimensions to the analysis of the above cited papers. First, we relate internalization decisions to firm-level characteristics such as TFP as Antràs and Helpman (2004, 2008), among others, have suggested. We take seriously into account the issues of endogeneity and unobserved heterogeneity by means of fixed and/or random effects. Second, we introduce determinants of these decisions at the country level, as well as the imported and the final product levels. We can therefore exploit more information on the determinants of internalization than studies of intra- and inter-firm trade that rely on US affiliate-level data, such as Feinberg and Keane (2006). Bernard et al. (2008) offer a new index of revealed intermediation to measure the contractibility of imported products. They do not however attempt to measure contractibility at the final product level.

At the level of the firm, the analysis reveals a first very surprising fact that should be taken into account in future theoretical developments. Even at a finely defined sectoral level, one observes substantial factor intensity heterogeneity across firms (about the same degree of magnitude as firm level productivity heterogeneity).

A second striking result is that capital-intensive firms tend to insource labor-intensive goods from labor-abundant countries, contrary to the prediction in Antràs (2003). However, the industry- or product-specific intra-firm import share does vary positively with the capital abundance of the origin country, as in the above cited papers. Our analysis reveals that it is crucial to investigate the intensive margin (volume of transactions) to reconcile our firm-level results with industry- or product-level evidence.

Third, we find 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 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.

Fourth, we find some robust empirical evidence that complex goods and inputs are more likely to be produced within firm boundaries, particularly for capital-intensive firms. Since models typically consider only one foreign input, and property-rights models in particular have little to say about input characteristics, we hope that this result will stimulate further theoretical research.

In independent research, Defever and Toubal (2007) test the predictions of a slightly amended Antràs and Helpman (2004) model with firm-specific (rather than industry-specific) factor intensity differences. Using data on French imports in 1999, they show that the the likelihood of sourcing from an independent supplier is increasing in the intensity in the supplier's input, in the final producer's TFP, in the interaction between the two. Furthermore, they show that the interaction between TFP and institutional quality in the supplier's country favors intra-firm trade.

Our work differs from theirs in several respects. First, our considerably larger dataset (that includes theirs) allows us to include more transactions. We find support for some of their findings but do not confirm the fact that the most productive firms engage in outsourcing, or that the interaction between headquarter intensity and productivity matters for the internalization decision. Second, we introduce several measures of contractibility of imported inputs and inputs used in the production process of the final good, and show that they are significant determinants of the share of intrafirm trade. Third, we introduce additional control variables, such as distance or capital-intensity in final good production, which we find to be significant across all specifications. However, we do not introduce measures of fixed costs at the firm-level.

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 intensive and extensive margins of imports. Finally section 6 concludes and provides avenues for future research.

2 Theoretical Background

In this section, we review three important theoretical contributions that help us understand 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). One of their building blocks 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.³ Antràs (2003) shows that efficiency in labor investments is higher under outsourcing. Therefore his first (partial-equilibrium) prediction is that intrafirm imports are more likely in capital-intensive industries, all else equal.

Assume CES preferences and a global Cobb-Douglas production function with constant shares η and $1 - \eta$ for capital and labor, respectively. Denote by ϕ the final producer's share of surplus in bargaining, and by δ the value of residual rights of control as a share of joint surplus. Antràs (2003) derives the ratio of integration to outsourcing profits, which does not depend on factor prices under the convenient Cobb-Douglas assumption. This ratio is shown to be monotonically increasing in η , the capital intensity parameter. Under plausible parameter restrictions there exists a unique η^* for which both organizational modes are equally profitable. Therefore intra-firm transactions are more likely in more capital-intensive industries.⁴ In Antràs and Helpman (2004, 2008) the same logic applies to the intensity in 'headquarter services', the factor provided by the headquarter firm.

In addition to inter-industry differences in factor intensity these models allow for intra-industry TFP differences. Denote by θ_i the Hicks-neutral TFP parameter of firm i , and denote by f^{VI} and f^{OS} firm i 's fixed costs of resorting to integration and outsourcing, respectively. Assume that firms

³It is assumed that only labor-intensive production can be outsourced, based on stylised facts on US multinationals' internal cost-sharing practices.

⁴Defever and Toubal (2007) extend that prediction to the firm-level when firms differ in their factor intensity.

vary by their θ 's. If fixed costs do not vary across organizational forms, then the previous result is unchanged. If they do, then productivity differences matter. Under the additional assumption that fixed costs are greater under integration than outsourcing, the following prediction emerges. *Within a given industry, firms engaging in intra-firm trade should be more productive than those that outsource (Antràs and Helpman, 2004).*

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 FPE set, he shows that *the share of intrafirm imports increases in the origin 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).

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 the origin country. However, it is unclear whether his theoretical prediction generalises to a model with more industries or more factors.

2.3 Internalization and the Extent of Contract Incompleteness

Antràs and Helpman (2008) extend the model in their 2004 article to partial contractibility of production tasks. There they relate organizational choice in offshore operations to a country's contracting environment⁵. 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. Antràs and Helpman (2008) show that under a certain threshold value of η sectors or firms⁶ keep outsourcing their input production. However, for high- η firms or sectors, we have a different prediction.

⁵This research agenda is motivated by the finding of Nunn (2007) that cross-country differences in contracting institutions explain the variance of trade flows as much as cross-country differences in human capital.

⁶Assuming that the distributions of θ and η are independent.

Holding headquarter services contractibility constant, an improvement in input contractibility abroad has two effects⁷ 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.*

The net effect of an improvement in the contract environment is ambiguous. As explained by Antràs and Helpman (2008), '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 this important insight was absent of the fully incomplete contracts model by Antràs and Helpman (2004), where improved contract enforcement in the origin country unambiguously increased the share of intra-firm imports. We will take into account this new insight in our empirical analysis by using both contractibility of the imported input as well as of the quality of the origin country's judicial system.

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 Euro, owned by manufacturing groups that control at least 50% of the equity capital of a foreign affiliate. The answer 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

⁷Nunn and Trefer term these two effects the 'Standard Effect' and the 'Surprise Effect', respectively.

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.⁸ These data has been used by Defever and Toubal (2007) to test a variant of 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 such value is much higher (55.48%). This raises a serious selection bias issue and in order to solve this problem we make use of another data source.

The second database, coming from French Customs, is the universe of yearly import and export flows operated by French firms as coming out from custom declarations. 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 in the database. This database has been used by Eaton et al. (2004) among others and it is highly representative of aggregate import and export French flows. Aggregating firms' imports in 1999 by country of origin and product (CPA96 4digit) we obtain 1,252,462 observations referring to 126,953 firms. The total value of imports in the database represents 99% of French aggregate imports in 1999 as reported by EUROSTAT.

Since the EIIG database represents reasonably well intra-firm imports, we consider that all trade flows that are reported in the French Customs dataset but not in the EIIG database occur with a third party. 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.

In what follows, we refer to an individual record of our final imports database as a 'transaction'.

⁸See Appendix A for details.

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 activity is in the manufacturing sector (NACE rev1 D category). Firms in the EAE database represents 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 , products by p and countries by c . Our dependent variable, $y_{i,p,c}$, represent the sourcing mode of imported product p from country c chosen by of firm i (1=intra-firm; 0=outourcing).

Our independent variables capture determinants of the internalization decision at the level of the firm, country, imported input and final product (the latter denoted by f).

3.3.1 Dependent Variable

$y_{i,p,c}$ is a binary variable 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 $y_{i,p,c} = 0$.⁹

3.3.2 Firm Level Variables

TFP_i stands for Total Factor Productivity of firm i and is estimated as the residual (plus the constant) of a log-linearized three-factor Cobb-Douglas production function with labor, capital and material inputs as inputs. We use the value-added based Levinsohn and Petrin (2003) estimator (LP).¹⁰ We have also experimented with the Olley and Pakes (1996) estimator (OP) obtaining

⁹The use of a binary response model is justified by the fact that in the data only a very small fraction of transaction is ‘mixed’, i.e firms importing a given product from a given country partly intra-firm and partly from an independent supplier. 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

¹⁰Details on the estimation procedure are provided in Appendix A.

qualitatively identical results. We prefer LP because the OP procedure obliges us to throw quite a few firms due to non-positive values of investments.

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 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.

3.3.3 Imported Products Variables

Unlike previous empirical papers dealing with contract incompleteness and trade like Levchenko (2007), Nunn (2007) and Nunn and Trefler (2008) we are able to identify all international transactions at the product and firm level. Therefore, in our data, a firm typically imports several products (inputs), either from a foreign affiliate or from a third party, that vary a lot in their degree of sophistication.

Because of data constraints, previous empirical studies had to use an ‘average’ measure of the contract complexity of the whole production process, based on the US input-output matrix, of the reporting industry. Levchenko (2007) built a Herfindahl index measuring the degree of variety of inputs needed in the production process. Nunn (2007) and Nunn and Trefler (2008) used the Rauch (1999) classification to proxy for the contract complexity of inputs and reconstructed, using input-output coefficients as weights, an average contract complexity of the importing industry production.

Our data allow us to go one step further and attribute directly a level of contract complexity to each imported product. We analyze how firm i producing a specific final product f chooses to source inputs p of different contract complexity either via intra-firm or via outsourcing. In particular, we follow the idea of Nunn (2007) to attribute to an input a degree of contract completeness that depends on whether the product is sold on an organized exchange, reference priced or neither of the two. Denoting by $R_j^{neither}$ a dummy variable that takes value 1 if the HS6 product j is

neither sold on an organized exchange nor reference priced,¹¹ 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:¹²

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

Finally, we also compute the capital (k_p) and skill intensity (h_p) associated to each imported product p . We construct these measures by matching product categories to French industries, as explained in Appendix A.

3.3.4 Final Product Variables

We can observe detailed records of distinct intermediate inputs p imports used for the production of a specific final good f . Therefore, contrary to previous studies, we can also take into account the contractibility of a final good f , in addition to the contractibility of an imported input.

We measure the contractibility of the final good f using the Rauch classification. Denote by $R_j^{neither}$ a 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 1999:¹³

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

In Table 1 we give summary statistics on both input and final product contractibility measures.

Table 1: 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

Two things are worth noting. The first one is that, looking at transaction patterns, the correlation between μ_p and μ_f is positive (0.28) and significant. This suggest that firm producing final complex goods based on our measure import complex inputs. The second one is that μ_f is

¹¹We actually use the ‘Liberal’ Rauch classification. Results are virtually unchanged if we use the ‘conservative’ one.

¹²Summary statistics for this variable are given in Table 1 below. See Appendix A for more details.

¹³See Appendix A for details.

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 0.78.

We acknowledge the fact that what we should in principle measure is not the degree of contractibility or mode of exchange of the final product f but the contractibility of the tasks performed by the final producer. However, this shortcoming would remain even if we construct a measure based on an input-output matrix like in Nunn (2007).

3.3.5 Origin Country Variables

As to the exporting country c characteristics, we use rather standard variables. Further details can be found in Appendix A.

k_c and h_c stand (respectively) for the capital and skill intensity of country c , while Q_c is a measure of the quality of institutions (rule of law).

Same – leg – orig_c is a dummy indicating whether country c adopts 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 was a former French colony and *Language_c* is a dummy indicating whether French is spoken in country c .

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

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.¹⁴ 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.*¹⁵

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

¹⁴Results are robust to the alternative logit specification.

¹⁵The fact that we restrict our analysis to 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. In other words, the population of interest for us is the population of importing firms, so that we have no sample selection problems in this respect.

going into estimations we report a first fact:

- *Fact 1: There are few intra-firm import transactions, but they are of a high amount.*

Indeed in our data only 2% of transactions are intra-firm but they correspond to 25% of total imports' value. In comparison Zeile (1997) reports a 42.7% share of intra-firm trade in US imports. However, these figures are not comparable because the definition of a foreign affiliate in the French EIIG is much more demanding (50% or more of the equity capital) than the US definition (10%). In addition the EIIG lacks coverage of around 20% of multinationals imports due to non-respondents.

There are certainly many possible interpretations of Fact 1. One possibility, in line with Antràs and Helpman (2004) and (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.¹⁶ 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.

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). Taken literally, the latter two models predict that no intra-firm should take place in component-intensive (low η) industries. This prediction is not valid as such in our data. Intra-firm trade and outsourcing

¹⁶In this smaller dataset 5.4% of transactions are intra-firm, representing 34% of the value of imports.

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 threshold η , 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, 2007)¹⁷, 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 2: 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	0.0000	1.0000	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 3: 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 2 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. Correlation between TFP and the other three firm-level variables

¹⁷In 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.

are either negligible or relatively small like 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 (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 3. 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 much 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 3 shows a weak correlation between $\Delta_s TFP_i$ and the other variables. *Finally, cross-correlations between $\Delta_s k_i$, $\Delta_s \eta_i^{hq}$ and $\Delta_s \eta_i^{sk}$ are never very high and suggest that the 3 variables provide different pieces of information on the technology of the final good producer.*¹⁸ 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:

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

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 4 reports estimations of different specifications of 1. 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. Explanatory variables have always a positive and highly significant coefficient suggesting that:

¹⁸Even the most sophisticated TFP estimation techniques rely on the assumption that input shares are constant across firms. One may thus wonder how reliable TFP estimations are, even within narrowly defined sectors, if firms actually use different technologies. 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 more 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.

Table 4: Firm i -specific determinants of intra-firm trade.

Dependent variable: $y_{i,p,c}=1$ for intra-firm.	1	2	3	4	5
TFP_i	.04116*** (.0008)				.02911*** (.00076)
k_i		.02117*** (.0005)			.01541*** (.0004)
η_i^{hq}			.03423*** (.0020)		.01164*** (.0017)
η_i^{sk}				.05948*** (.0012)	.01724*** (.0011)
NES114 sectoral dummies	yes	yes	yes	yes	yes
Country and product dummies	yes	yes	yes	yes	yes
Number of observations	242,804	242,804	242,804	242,414	242,414
Pseudo R^2	0.1669	0.1562	0.1314	0.1574	0.1897
Log Likelihood	-43,258	-43,809	-45,097	-43,711	-42,037

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

- *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) and is a new empirical result. In fact, in his detailed analysis of Japanese data, Tomiura (2007) is not able to distinguish between intra-firm vs. outsourcing in 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 substantially in line with the residual property rights literature predictions but further suggest that heterogeneity in capital, headquarter and skill intensity needs to be accounted for.

A last comment is in order with respect to Result 1. Using similar French import data, Defever and Toubal (2007) provide the opposite finding. In particular, they show that the likelihood of sourcing from a foreign affiliate is actually *decreasing* in the final producer's TFP. The key element behind such different findings is the fact that we do not rely on the same databases. Defever and Toubal (2007) build on the EIIG survey only while we further complement these data with the universe of French firm-level imports transactions coming from Customs declarations. By construction, firms in the EIIG survey are all multinational having foreign affiliates and they almost all do at least one intra-firm import transaction. Although such firms engage in both intra-firm and outsourcing transactions depending on the specific product and country, the sample is clearly biased towards intra-firm trade firms and it is thus questionable what we can learn from such a specific sub-population.

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-sourcing mode combinations) using various methods to control for firm characteristics.

Antràs (2003) predicts that intra-firm trade is more likely with capital-abundant countries. His prediction can be extended to skill-abundance. Another determinant suggested by Antràs and Helpman (2004) is the quality of institutions in the origin country (South). To test these theories we use capital abundance (k_c), skill abundance (h_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 variables that are not directly related to theoretical models, but that often affect trade flows. In particular we use 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.¹⁹

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

$$y_{i,p,c}^* = \alpha + \beta_1 k_c + \beta_2 h_c + \beta_3 \mu_p + \beta_4 \mu_f + \beta_5 Q_c + \beta_6 k_p + \beta_7 h_p + \beta_8 Distw_c + \beta_9 Colony_c + \beta_{10} Language_c + \beta_{11} Same - leg - orig_c + \varepsilon_{i,p,c} \quad (2)$$

Results are presented in Columns (1) and (2) of Table 5. 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.²⁰ Column (2) presents results for the subset of firms for

¹⁹We 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).

²⁰This is due to the fact that the Rauch (1999) classification, the building block of our contractibility measure, is

which information on μ_f is available.

In order to control for firm characteristics we also estimate a conditional fixed effects logit model.²¹ 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.

$$y_{i,p,c}^* = \alpha + \beta_1 k_c + \beta_2 h_c + \beta_3 \mu_p + \beta_4 Q_c + \beta_5 k_p + \beta_6 h_p + \beta_7 Distw_c + \beta_8 Colony_c + \beta_9 Language_c + \beta_{10} Same - leg - orig_c + i + \varepsilon_{i,p,c} \quad (3)$$

where i is a firm-specific fixed effect potentially correlated with explanatory variables. Results are presented in Column (3) of Table 5.

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 information used in estimations (see the row ‘number of actual observations used in estimations’ in Table 5) 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.

Using fixed effects comes at the cost of losing many observations. An alternative is to estimate the probit model of Equation (2) on the smaller dataset of 247,528 firm-country-product-sourcing mode combinations for which firm-level information is available²², using additional controls:

$$y_{i,p,c}^* = \alpha + \beta_1 k_c + \beta_2 h_c + \beta_3 \mu_p + \beta_4 \mu_f + \beta_5 Q_c + \beta_6 k_p + \beta_7 h_p + \beta_8 Distw_c + \beta_9 Colony_c + \beta_{10} Language_c + \beta_{11} Same - leg - orig_c + \mathbf{X}_i' \mathbf{c} + \mathbf{D}_{NES} + \varepsilon_{i,p,c} \quad (4)$$

where \mathbf{X}_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 provided in column (4) of Table 5. To save space we do not report coefficients of

concerned essentially with manufacturing, agriculture and mining goods. The same issue applies to a lesser extent to the contractibility of imported products.

²¹As is well-known the fixed effects logit estimator is consistent without any assumption on how the fixed effects are related to the independent variables. This property comes from the functional form of the logistic distribution. It is not satisfied in the unobserved effects probit model.

²²The number of observations is reduced because we only have balance sheet data on the universe of French manufacturing firms with at least 20 employees, as mentioned in the previous Section.

Table 5: Product and Country-specific determinants of intra-firm trade.

Dependent variable: $y_{i,p,c}$	1	2	3	4
k_c	-0.0007* (0.0004)	-0.0995*** (0.0105)	-0.0191*** (0.0054)	-0.1000*** (0.0010)
h_c	0.0001 (0.0014)	0.2622*** (0.0407)	-0.0341 (0.0254)	0.0125*** (0.0038)
μ_p	-0.0077*** (0.0005)	-0.2113*** (0.0132)	-0.0514*** (0.0183)	-0.0085*** (0.0012)
μ_f		-0.1350*** (0.0114)		-0.0080*** (0.0017)
Q_c	0.0387*** (0.0023)	0.8301*** (0.0664)	0.1456** (0.0582)	0.0516*** (0.0062)
k_p	0.0059*** (0.0003)	0.1026*** (0.0077)	-0.0232*** (0.0084)	0.0011 (0.0008)
h_p	0.0312*** (0.0007)	0.5046*** (0.0218)	0.0302 (0.0200)	-0.0039* (0.0022)
Controls				
$Distw_c$	0.0024*** (0.0002)	0.0940*** (0.0072)	-0.0148*** (0.0037)	0.0023*** (0.0007)
$Colony_c$	-0.0030*** (0.0006)	-0.0446** (0.0187)	0.2326*** (0.0445)	0.0060*** (0.0020)
$Language_c$	-0.0061*** (0.0004)	-0.1311*** (0.0133)	-0.0176** (0.0080)	-0.0083*** (0.0010)
$Same - leg - orig_c$	-0.0010** (0.0004)	0.1070*** (0.0131)	0.0240** (0.0106)	0.0119*** (0.0012)
Estimation method	pooled probit	pooled probit	conditional firm fixed effects logit	probit with controls \mathbf{X}_i, D_s
Number of potential observations	1,141,393	366,376	1,141,393	247,528
Number of actual observations used in estimations	991979	309307	36,811	210,039
Pseudo R^2	0.0289	0.0193		0.1506
Log Likelihood	-104,332	-56,440	-14,440	-40,191

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

covariates \mathbf{X}_i and industry dummies.

Looking across the different sets of estimates in Table 5 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.
- *Result 4: Intra-firm trade is more likely with countries having good judicial institutions.*

There are several possible interpretations of Result 3. Antràs (2003) predicts that the share of intra-firm trade should increase with k_c . First, his result is derived under the very strong assumptions of capital immobility and factor price equalization. Second, we are looking here at the firm-level decision to undertake intra-firm trade vs outsourcing (i.e. the extensive margin), while Antràs' prediction is concerned with the share of intra-firm trade in *total* imports. To bridge the gap between his and our results one needs to consider two additional margins, namely the impact of k_c on the value of a firm's imports (intensive margin) and the sectoral composition effect across French industries. We come back to this issue in the next Section. Interestingly enough, one can see that there is no robust evidence that intra-firm trade is more likely to involve inputs with a high capital-labor ratio (k_p). It is worth reminding that this variable is computed using French technology and may be prone to measurement error.

Result 4 states that the better 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.²³ We strengthen this result by controlling for both imported and final good contractibility as well as for 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 developed for example in McLaren (2000) and Grossman and Helpman (2002). Incentives based on the optimal allocation of residual property rights over the imported product are the key mechanism that allow Antràs and Helpman (2004) and Antràs and Helpman (2008) to rationalize the Surprise Effect.

Concerning contractibility of both intermediate and final products, we find the following con-

²³This 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 resembles to $\mu_p * Q_c$, 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.

sistent pattern:

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 because these models usually consider only two inputs (one domestic and one foreign) and no general comparative static result can be provided on the relationship between the complexity of an input and its optimal sourcing mode.

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

As far as skill abundance is concerned, neither the product (h_p) nor the country dimension (h_c) provide a clear pattern in distinguishing between intra-firm vs outsourcing. Coefficients are sometimes either positive or negative while being not significant in some cases.

Finally, in our preferred specifications (Columns 3 and 4 in Table 5) one can note the positive impact of colonial ties ($Colony_c$) and sharing a common legal origin ($Same - leg - orig_c$) on the decision to engage in intra-firm. The impact of a common language ($Language_c$) is instead negative while results on distance are ambiguous ($Distw_c$). We do not believe that these coefficients convey much information because they at best provide the relative magnitude at which unobservable fixed and variables costs embedded in our covariates affect the sourcing decision. However, we do believe that these variables represent important controls for our analysis.

4.3 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 can push the analysis of heterogeneity further by looking at whether firms with different productivity and/or capital, skills and headquarter intensity value differently the capital abundance of the host country, contractibility, quality of institutions, etc. This amounts to looking at interactions between firm and product/country variables.

Table 6: Interactions of Q_c with quintiles of TFP_i , k_i , η_i^{hq} , and η_i^{sk} .

Cross effects	TFP_i	k_i	η_i^{hq}	η_i^{sk}
1st quintile	-0.0126*** (0.0020)	-0.0033** (0.0013)	-0.0124 (0.0104)	-0.0060** (0.0025)
2nd quintile	-0.0077*** (0.0019)	-0.0034*** (0.0009)	-0.0137 (0.0089)	-0.0018 (0.0023)
3rd quintile	-0.0074*** (0.0019)	-0.0001 (0.0006)	0.0107 (0.0079)	0.0002 (0.0021)
4th quintile	-0.0055*** (0.0018)	0.0020*** (0.0004)	0.0009 (0.0072)	0.0041** (0.0020)
5th quintile	-0.0014 (0.0016)	0.0032*** (0.0004)	0.0210*** (0.0056)	0.0076*** (0.0016)
Number of observations	210,039	210,039	210,039	210,039
Pseudo R^2	0.1532	0.1532	0.1532	0.1567
Log Likelihood	-40,070	-40,067	-40,067	-39,902

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

This risk of such an kind of 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. Antràs and Helpman (2008) show that the Standard Effect (the quality of institutions favors outsourcing over intra-firm) 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 5 quintiles of the distributions of TFP_i , k_i , η_i^{hq} , and η_i^{sk} . Finally we have created a cross product between each quintile and Q_c and estimated model 4 adding these additional interaction effects. We have performed 4 different estimations, one for each of the 4 interacting variables.

Results of such estimations are reported in Table 6 and, in order to save space, we report only interaction variables' marginal effects and significance. The sign of quintiles is meaningless. What we should check is if these values decrease or increase significantly when moving from the 1st to the 5th quintile. For each type of interaction we find that the difference between the fifth and the first quintile is positive and significant. Actually, in most cases the coefficient of the i^{th} quintile is significantly higher than that of the $(i - 1)^{th}$ quintile. This leads to the following result:

Result 5: The 'surprise' effect is significantly stronger for more productive, capital intensive, headquarter intensive, and skill intensive firms.

Table 7: Interaction of μ_p with quintiles of TFP_i , k_i , η_i^{hq} , and η_i^{hk} .

Cross effects	TFP_i	k_i	η_i^{hq}	η_i^{hk}
1st quintile	-0.0011 (0.0014)	0.0089*** (0.0011)	-0.0413*** (0.0150)	0.0130*** (0.0036)
2nd quintile	0.0001 (0.0011)	0.0039*** (0.0009)	-0.0499*** (0.0124)	0.0127*** (0.0033)
3rd quintile	-0.0018* (0.0011)	0.0037*** (0.0007)	-0.0267** (0.0108)	0.0131*** (0.0031)
4th quintile	-0.0016* (0.0010)	0.0037*** (0.0006)	-0.0452*** (0.0095)	0.0134*** (0.0030)
5th quintile	-0.0016*** (0.0008)	0.0035*** (0.0006)	-0.0132* (0.0070)	0.0122*** (0.0023)
Number of observations	210,039	210,039	210,039	210,039
Pseudo R^2	0.1508	0.1515	0.1518	0.1511
Log Likelihood	-40,182	-40,149	-40,138	-40,167

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

This original result is certainly reminiscent of the heterogeneous impact of Q_c in Antràs and Helpman (2008).

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.

In Tables 7 and 8 we perform a similar exercise to the one reported in Table 6. For each NES114 industry, we interact quintiles of the distributions of TFP_i , k_i , η_i^{hq} , and eta_i^{sk} with respectively μ_p and μ_f .

Inspection of Table 7 reveals no consistent picture with respect to μ_p . This is confirmed by significance tests of the difference between the coefficients of the first and fifth quintiles, as well as across quintiles. However Table 8 and significance tests show a consistent pattern. The higher TFP as well as capital-, skill- and headquarter-intensity, the stronger the negative effect of μ_f on intra-firm trade. To put it differently:

Fact 4: Highly productive, capital-, skill-, headquarter-intensive firms are more likely to produce final complex goods using inputs imported within the firm boundaries. However, there is no heterogeneity across firms in the way imported product contractibility affects the sourcing decision.

Table 8: Interaction of μ_f with quintiles of TFP_i , k_i , η_i^{hq} , and η_i^{hk} .

Cross effects	TFP_i	k_i	η_i^{hq}	η_i^{hk}
1st quintile	0.0008 (0.0021)	0.0116*** (0.0010)	0.0763*** (0.0166)	0.0405*** (0.0057)
2nd quintile	0.0006 (0.0018)	0.0068*** (0.0008)	0.0693*** (0.0137)	0.0415*** (0.0054)
3rd quintile	-0.0011 (0.0017)	0.0017** (0.0007)	0.0904*** (0.0116)	0.0403*** (0.0053)
4th quintile	-0.0011 (0.0016)	0.0054*** (0.0006)	0.0459*** (0.0104)	0.0361*** (0.0051)
5th quintile	-0.0018 (0.0013)	0.0026*** (0.0006)	0.0488*** (0.0074)	0.0317*** (0.0043)
Number of observations	210,039	210,039	210,039	210,039
Pseudo R^2	0.1509	0.1529	0.1524	0.1515
Log Likelihood	-40,179	-40,085	-40,106	-40,150

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

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

A number of previous studies has analyzed different 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 challenge previous findings. The aim of this section is to show that the distinction between the extensive and intensive margin of imports by sourcing mode 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 start by ruling out the possibility that there exists systematic differences in the pattern of observed intra-firm trade in France and the US.²⁴ To this aim we replicate their results on our French data.

Table 9 replicates the cross-industry and cross-country regressions of Antràs (2003). 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^{hk} are sector averages of the corresponding firm-level variables. $Population_c$ is the log of country c 's population in 1999,

²⁴These could come from differences in the specialization of the two economies as well as in the definition of an affiliate (over 50% of equity in France, over 10% in the US).

Table 9: 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		0.0139*** (0.0030)
h_c		0.0426*** (0.0130)
$Distw_c$		-0.0264*** (0.0025)
$Colony_c$		0.0046 (0.0083)
$Language_c$		-0.0256*** (0.0078)
$Population_c$		0.0102*** (0.0016)
constant	-0.1314 (0.2171)	0.0392 (0.0481)
Country dummies	yes	no
Industry dummies	no	yes
Number of observations	10688	8706
R-squared	0.1075	0.2369

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

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 5 where the capital abundance of a country has a negative effect on the intra-firm decision.

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 (i.e. the extensive margin) 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 10 we replicate their estimation strategy. Our excluded variables are $Colony_c$, $Same-$

Table 10: Share of intrafirm trade at the product-country level: selection model

	First Stage	Second Stage
Dependent variable:	\widehat{Share}_{pc}	$Share_{pc}$
k_p	-0.0052 (0.0256)	0.0535*** (0.0085)
h_p	0.2952*** (0.0784)	0.1735*** (0.0254)
μ_p	-0.2013*** (0.0408)	-0.1069*** (0.0132)
k_c	0.1060*** (0.0267)	0.0447*** (0.0097)
h_c	0.4221*** (0.1032)	0.0332 (0.0319)
Q_c	2.5276*** (0.1499)	-0.0892* (0.0502)
$Distw_c$	-0.3807*** (0.0171)	-0.0270*** (0.0064)
$Language_c$	-0.1681*** (0.0541)	-0.0682*** (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*** (0.3459)	-0.8504*** (0.1182)
Number of observations	9980	3414

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

$leg - orig_c$ and $Population_c$. We find broadly similar results. In particular we also find that the quality of institutions (Q_c) has a positive effect in the selection equation but a negative effect on the share of intra-firm trade in the second-stage equation. This result suggests that the quality of institutions positively affects the participation rate while decreasing the relative value of intra-firm vs. outsourcing imports.

5.2 Determinants of Sourcing Modes at the Intensive and Extensive Margins

We showed in the previous section that, in line with the theory, firm heterogeneity is crucial in understanding the decision to undertake intra-firm trade or offshore outsourcing, i.e. the extensive margin. By studying the extensive margin we have also outlined the impact of product contractibility and country characteristics while trying to link them to theory. In contrast, theory

offers little guidance on how firm-level characteristics affect the intensive margin of imports. In order to both reconcile our findings with aggregate evidence and offer new insights to theory we use a Heckman procedure to properly identify the determinants of the intensive margin of imports. Bernard et al. (2008) represents the first attempt to deal with selection issues, but we go one step further here by going down to the firm level.

We first run a selection equation, based on our preferred specification described in Table 5 (Column 4). We then run two separate second-stage regressions, one for intra-firm import values and one for outsourcing import values.²⁵

We use multinational status of a firm in 1994 as the excluded variable.²⁶ The persistency of multinational status, that we indeed find in the French data, suggests the presence of substantial sunk costs of creating a foreign affiliate. It is our contention that, conditional on our 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.²⁷

In Columns 3 and 4 of Table 11 k_c has a negative and significant coefficient for the outsourcing intensive margin, while the effect on the intra-firm intensive margin is not significantly different from zero. Therefore while capital abundance of a country decreases the likelihood to engage in intra-firm trade, it has a positive impact on the relative value of intra-firm vs. outsourcing imports. As long as the second effect dominates the share of intra-firm trade value at the sector level can indeed be increasing with capital abundance (as in Table 9). Note that this result holds within a sector (industry dummies are included) so that composition effects across sectors are not needed to reconcile firm-level with industry-level evidence.

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 intra-firm imports and not significant for outsourcing import values. This is reminiscent of the result in Bernard et al. (2008) as well as in our Table 10 based on product-country intra-firm shares. Finally the contractibility of the imported product μ_p has a negative effect on both the extensive and intensive margins of

²⁵Because we focus on the intensive margin we now include firm-country-product observations with mixed sourcing modes. We break down these observations into an intra-firm and an outsourcing transaction, while attributing the corresponding import values.

²⁶We use data from the LIFI (“Liaisons Financières”) database. The LIFI is collected by the French national institute of statistics (INSEE) and provides information on ownership relationships across firms that have a legal entity in France.

²⁷In our dataset the correlation between multinational status in 1994 and 1999 is 0.38. The correlation between multinational status in 1994 and y_{ipc} is 0.25. Both figures suggest that persistency is at work but is not the only determinant of engaging in intra-firm trade.

Table 11: Heckman selection procedure

Dependent variable:	First Stage	Second Stage	
	y_{ipc}	Intra-firm import value	Outsourcing import value
1994 <i>Multinational</i>	0.6115*** (0.0101)		
TFP_i	0.4289*** (0.0110)	0.3515*** (0.0591)	0.5603*** (0.0170)
k_i	0.2158*** (0.0061)	0.0879*** (0.0325)	0.1928*** (0.0085)
η_i^{hq}	-0.1506*** (0.0262)	-0.3244*** (0.1120)	-1.1948*** (0.0362)
η_i^{sk}	0.2452*** (0.0180)	-0.1823*** (0.0702)	-0.1744*** (0.0243)
k_c	-0.1270*** (0.0119)	0.0695 (0.0540)	-0.0474*** (0.0165)
h_c	0.1482*** (0.0446)	-0.2389 (0.1868)	-1.1619*** (0.0613)
μ_p	-0.0665*** (0.0141)	-0.1532** (0.0624)	0.8476*** (0.0192)
μ_f	-0.0128 (0.0203)	-0.2609*** (0.0904)	-0.0858*** (0.0283)
Q_c	0.6994*** (0.0737)	-1.7439*** (0.3106)	-0.6383*** (0.1034)
k_p	0.0174* (0.0091)	0.2805*** (0.0405)	0.1520*** (0.0126)
h_p	-0.1017*** (0.0264)	0.9973*** (0.1144)	-0.2387*** (0.0381)
Mills ratio		-0.1866** (0.0844)	2.7490*** (0.0719)
<i>Distw_c</i>	-0.0030 (0.0084)	-0.1027*** (0.0349)	0.1202*** (0.0123)
<i>Colony_c</i>	0.0693*** (0.0219)	0.7412*** (0.0946)	0.3435*** (0.0304)
<i>Language_c</i>	-0.1451*** (0.0147)	-0.0056 (0.0660)	0.1449*** (0.0201)
<i>Same – leg – orig_c</i>	0.1602*** (0.0142)	-0.3905*** (0.0640)	-0.1264*** (0.0200)
<i>Constant</i>	-4.5315*** (0.1596)	8.8060 (0.8129)	8.7115*** (0.2245)
Industry dummies	yes	yes	yes
Number of observations	220645	17857	202788

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

intra-firm trade while positively affecting outsourcing import values. Not surprisingly we thus find that product contractibility decreases the share of intra-firm trade at the product-country level. This result is consistent with Bernard et al. (2008) despite the fact that 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-, exporter country- and product- level determinants of intra-firm trade, and their interactions.

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 significantly 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). We believe that our result should stimulate further theoretical research involving firm characteristics other than TFP.

To be sure, specialists of organisations have emphasised the importance of tailoring theoretical predictions to a firm's context in empirical work. They have also pointed out that since most theories of the firm are non-mutually exclusive, their prevalence may vary from firm to firm (see for example Baker and Hubbard, 2001 and Gibbons, 2005). Our claim, however, is different. If we find strong empirical regularities at the firm level rather than pure idiosyncrasies, then there is a need for at least one new theory involving firm-level differences in factor intensity, for example.

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, as in US studies the share of French intra-firm imports is positively correlated with capital abundance at the industry-country and product-country levels. In fact, capital abundance has a positive impact on the value of intra-firm transactions relative to outsourcing imports, and that effect dominates in share regressions.

Theory does not guide us in understanding this result on transaction values (as acknowledged in Antràs, 2003). In addition, several caveats should be noted. First, cross-country differences in

transfer pricing policies (possibly related to tax evasion) may matter. Second, Antràs' assumption that endowments are in the factor price equalization set is unlikely to hold for all the countries we consider. Confirmation of our results in other firm-level datasets (in the US and elsewhere) would be desirable. Nevertheless, our result suggests 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.

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, which is the relevant margin for these particular firms.

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 judicial uncertainty has a stronger effect on trade flows inside than outside firm boundaries (as our Table 11 suggests) is an interesting question for future research.

Fourth, we find some robust empirical evidence that complex goods and inputs are more likely to be produced within firm boundaries, particularly for capital-intensive firms. The property-rights approach is silent about product characteristics and we hope that our finding can inspire future theoretical work. Our finding, however, is consistent with transaction-cost or dissipation of intangible assets explanations. Complex inputs embody costly R&D efforts or the use of other intangible assets, and protection against imitation is easier and cheaper within firm boundaries.

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

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*, Vol. 118 (2003), 1375-1418.
- [3] Antràs, Pol, and Elhanan Helpman, "Global Sourcing," *Journal of Political Economy*, Vol. 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] Baker, G. P. and T. N. Hubbard, “Empirical Strategies in Contract Economics: Information and the Boundary of the Firm,” *The American Economic Review*, Vol. 91, No. 2, *Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association* (May, 2001), 189-194
- [6] Bernard, Andrew B. and J. Bradford Jensen, “Why Some Firms Export,” *The Review of Economics and Statistics*, Vol. 86, No. 2 (2004), 447-464.
- [7] Bernard, Andrew B., J. Bradford Jensen, Stephen Redding and Peter K. Schott, “Firms in International Trade”, *Journal of Economic Perspectives*, 21(3), 2007, 105-130.
- [8] Bernard, Andrew B., J. Bradford Jensen, Stephen Redding and Peter K. Schott, “Intra-Firm Trade and Product Contractibility” (mimeo 2008)
- [9] Defever, Fabrice, and Farid Toubal, “Productivity and the Sourcing Modes of International Firms: Evidence from French Firm-Level Data” *CEP Discussion Paper No 842* (2007).
- [10] Djankov, S., R. La Porta, F. López-de-Silanes, and A. Shleifer, “Courts”, *Quarterly Journal of Economics*, Vol. 118 (2003), 453-517.
- [11] Eaton, Jonathan, Samuel Kortum, and Francis Kramarz “Dissecting Trade: Firms, Industries, and Export Destinations,” *American Economic Review Papers and Proceedings*, Vol. 94 (2004), 150-154.
- [12] Feinberg, Susan and Michael Keane, “Accounting for the Growth of MNC-based Trade using a Structural Model of U.S. MNCs,” *American Economic Review*, Vol. 96(5): 2006, 1515-1558
- [13] Gibbons, R., Four formalizable theories of the firm. *Journal of Economic Behavior and Organization*, Vol. 58 (2005), 200-245
- [14] Griliches, Z. and Mairesse, J., *Production functions: the search for identification*. In: Strm, S., Editor, 1998. *Econometrics and Economic Theory in the 20th Century*, Cambridge Univ. Press, Cambridge.
- [15] Grossman, Gene M. and Elhanan Helpman, Integration versus outsourcing in industry equilibrium, *Quarterly Journal of Economics* 117: 85-120, 2002.
- [16] Grossman, Gene M. and Elhanan Helpman, Outsourcing versus FDI in industry equilibrium, *Journal of the European Economic Association* 1: 317-327, 2003.
- [17] Grossman, Gene M. and Elhanan Helpman, Managerial incentives and the international organization of production, *Journal of International Economics* 63: 237-262, 2004
- [18] Grossman, Gene M. and Elhanan Helpman, Outsourcing in a global economy, *Review of Economic Studies* 72: 135- 159, 2005
- [19] Guannel, Boris, and Claire Plateau, “Les échanges internationaux intragroupe dans la mondialisation industrielle en 1999,” *SESSI* (2003).

- [20] Gwartney, James, and Robert Lawson, “Economic Freedom of the World: 2003 Annual Report,” Mimeograph, Fraser Institute, 2003.
- [21] Hall, Robert E., and Charles I. Jones, “Why do Some Countries Produce So Much More Output per Worker than Others?” *Quarterly Journal of Economics*, Vol. 114 (1999), 83-116.
- [22] Head Keith and Thierry Mayer, Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade, (2002) , CEPII Working Paper No.2002-01.
- [23] Helpman, Elhanan, Marc Melitz and Yona Rubinstein (2008), “ Estimating Trade Flows: Trading Partners and Trading Volumes,” *The Quarterly Journal of Economics*, MIT Press, vol. 123(2) (2008), 441-487
- [24] Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, “Governance Matters III: Governance Indicators for 1996–2002,” Working Paper No. 3106, World Bank, 2003.
- [25] Levchenko, Andrei, “Institutional Quality and International Trade,” *Review of Economic Studies*, Vol. 74 (2007), 791-819.
- [26] Levinsohn James, and Amil Petrin “Estimating Production Functions Using Inputs to Control for Unobservables,”*Review of Economic Studies*, Vol. 70 (2003), 317-342.
- [27] Markusen, James, “The Boundaries of Multinational Enterprises and the Theory of International Trade” , *Journal of Economic Perspectives*, Vol. 9, No. 2. Spring, 1995, pp. 169-189
- [28] McLaren John, “Globalization and Vertical Structure ”, *American Economic Review* 90:5 (December 2000),pp. 1239-54.
- [29] Marin, Dalia and Thierry Verdier, ”Globalization and the New Enterprise”, *Journal of the European Economic Association*, vol. 1, 2-3, April-May, 337-344, 2003
- [30] Marin, Dalia and Thierry Verdier, “Power Inside The Firm and The Market: A General Equilibrium Approach,”*Journal of the European Economic Association*, MIT Press, vol. 6(4) (2008), 752-788
- [31] Nathan, Nunn, “Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade,” *Quarterly Journal of Economics*, Vol. 122 (2007), 569-600.
- [32] Olley, Steven and Ariel Pakes, “The Dynamics of Productivity in the Telecommunications Equipment Industry,” , *Econometrica*, Vol. 64 (1996), 1263-1297.
- [33] Nathan, Nunn, and Daniel Trefler, “The Boundaries of the Multinational Firm: An Empirical Analysis,” forthcoming in Helpman, E., D. Marin, and T. Verdier, *The Organization of Firms in a Global Economy*, Harvard University Press, 2008.
- [34] Ranjan, Priya and Justin Tobias, “Bayesian Inference for the Gravity Model” *Journal of Applied Econometrics*, Vol. 22, No. 4 (2007), 817-838
- [35] Rauch, James E., “Networks versus Markets in International Trade” *Journal of International Economics*, Vol. 48 (1999), 7-35.

- [36] 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.
- [37] Tomiura, Eiichi, “Foreign Outsourcing, Exporting, and FDI: A Productivity Comparison at the Firm Level ” *Journal of International Economics*, Vol. 72 (2007), 113-127.
- [38] Yeaple, Stephen, “Offshoring, Foreign Direct Investment and the Structure of U.S. Trade,” *Journal of the European Economic Association*, 2006, 4(2-3), 602-611
- [39] Zeile, William J., “U.S Intrafirm Trade in Goods, ” *Survey of Current Business*, Vol. 77(2) (1997), 23-38.

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 contract relationships. 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,²⁸ 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 starting 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.

²⁸In 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.

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 bought inputs 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 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). In our empirical analysis we work with the CPA96 4digit classification (490 products). However, the Rauch data cover 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.²⁹ 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 for the final product contractibility, we have first used a correspondence table from 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.

²⁹Correspondence tables have been obtained using RAMON data available at the web-site: http://ec.europa.eu/eurostat/ramon/reasons/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 comes from the “rule of law” 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. *Same – leg – orig_c* is taken from Djankov et al. (2003).

The last set of variables (*Distw_c*, *Colony_c*, and *Language_c*) comes from CEPII (Centre d’Etudes Prospectives et d’Informations Internationales, Paris). *Distw_c* is the log of distance of country c to France. The distance is calculated starting from regional distances which are then aggregated at the country level using region populations as weights. Further details may be found in Head and Mayer (2002).