WWW.DAGLIANO.UNIMI.IT

CENTRO STUDI LUCA D'AGLIANO
DEVELOPMENT STUDIES WORKING PAPERS

N. 267

November 2008

The Determinants of Intra-Firm Trade

*Gregory Corcos**
*Delphine M. Irac***
*Giordano Mion****
*Thierry Verdier*****

* Norwegian School of Economics and Business Administration
** Bank of France
*** Université Catholique de Louvain
**** Paris School of Economics and CEPR

The Determinants of Intra-Firm Trade*

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

First draft: January 26, 2008. This version: May 30, 2008.

Abstract

This paper analyzes the determinants of the intra-firm vs arms'length sourcing mode of imported inputs. We build a unique French dataset of 1,141,393 import transactions at the firm, country and product levels in the year 1999 that allow us to distinguish the different sourcing modes. We study the firms-, country- and product- determinants of intra-firm trade. We confirm a number of theory-based predictions building on the residuals property rights approach and provide some empirical facts that can be used to further refine this as well as alternative theories. In particular, we highlight the fact that firms' heterogeneity needs to be evaluated across different dimensions. Furthermore, we point out that complex goods are more likely to be produced within the firm boundaries suggesting that those material and immaterial attributes that characterized a product play a key role in globalized sourcing strategies.

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

JEL Classification: F23, F12, F19

*Paper prepared within the framework of the FIRB project for the workshop "The International Firm: Access to Finance and Organisational Modes"(Milano), a joint initiative between FIRB, Università degli Studi di Milano, and Centro Studi Luca d'Agliano. We wish to thank our discussant and participants at the workshop for helpful comments and suggestions. The usual disclaimer applies.

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

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

§Corresponding author. CORE, Université catholique de Louvain, 34 Voie du Roman Pays, B1348 Louvain-la-Neuve, Belgium. E-mail: giordano.mion@uclouvain.be.

¶Paris School of Economics, France, and CEPR. E-mail: verdier@pse.ens.fr.

1 Introduction

About every third transaction in international trade occurs within a multinational firm while 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 literature has provided explanations of internalization decisions by looking at the costs of dissipating intangible assets (see a good survey of that approach 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 context[2]. 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, Nunn and Trefler 2008). In this respect, a seminal piece of work is Nunn and Trefler (2008). Using country/product category data, they investigate the determinants of the share of total US imports that is intra-firm. More precisely, they consider the Antràs (2003) and Antràs and Helpman (2004, 2008) empirical predictions explaining a firm's internalization decision by the intensity of downstream production in intermediate inputs. In particular they test the central prediction that ownership rights should be given to the party contributing to the intensive production stage. The empirical analysis provides some empirical support for that prediction.

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,

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

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

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 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 Nunn and Trefler (2008). 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 introduce two additional dimensions to the analysis of Nunn and Trefler (2008). First, we relate internalization decisions to firm-level characteristics such as TFP, as Antràs and Helpman (2004, 2008), among others, have suggested. Second, we introduce determinants of these decisions at the imported product and the final product level. 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).

We take seriously into account the issues of endogeneity and unobserved heterogeneity by means of fixed and/or random effects.

Going down to 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).

Third, we do find that firm productivity increases the likelihood of intra-firm sourcing like in Antràs and Helpman (2004). However, when we interact several measures of headquarter intensity with firm level productivity, we find that the interaction term is not positively correlated with intrafirm status. This remains true even when allowing for discontinuity along the headquarter intensity dimension. This casts doubt on the corresponding prediction by Antràs and Helpman (2004) and the robustness of the Nunn and Trefler (2008) result confirming that prediction on US data. Our results point towards a more general model where headquarter intensity is firm specific

and possibly correlated with productivity.

Fourth, using the intermediate product dimension, we find that firms producing more complex goods are more likely to source their complex intermediate inputs within firm boundaries. This result has no theoretical counterpart because models typically consider only one foreign input.

Fifth, we find that the likelihood of internalizing a transaction increases with distance. There is no obvious reason why variable trade costs should differ by transaction mode. Therefore our preferred interpretation is that the fixed costs of entering arms' length relationships increase more rapidly with distance than the fixed costs of setting up an affiliate. However, further theoretical research is needed to provide explanations for this finding.

Finally, we show that an improvement in the contracting environment does not affect all firms' sourcing decisions uniformly. In particular, it leads to a stronger increase in intrafirm transactions for the most productive, capital-intensive, skill-intensive, and headquarter intensive firms. This result is reminiscent of the ambiguous effect pointed out by Antràs and Helpman (2008). Furthermore, we also find that an improvement in the contracting environment affects disproportionately more contractible goods.

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. While a number of their findings are confirmed, we do not confirm the fact that the most productive firms engage in outsourcing. Indeed we do find that productivity increases the likelihood of intra-firm transaction coherently with Antràs and Helpman (2004). Furthermore, their finding that the interaction between headquarter intensity and productivity matters is also rejected by our data, casting doubt on an important prediction of Antràs and Helpman (2004).

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. Finally section 5 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). They predict the determinants of the share of intrafirm imports in total imports by manufacturing firms, our variable of interest.

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. In particular, 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. This assumption is based on 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 intrafirm imports are more likely in capital-intensive industries, all else equal.

Assume CES preferences with elasticity $\frac{1}{1-\alpha}$ for the final product. The global production function is CRS Cobb-Douglas with constant input shares η and $1-\eta$. 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. Variable profits under integration over variable profits under variable profits under outsourcing equal[3]:

$$\frac{\pi^{VI}}{\pi^{OS}} = \left(1 + \frac{\alpha(1-\phi)\delta^{\alpha}(1-2\eta)}{1-\alpha(1-\eta)+\alpha\phi(1-2\eta)}\right)\left(1+\frac{\delta^{\alpha}}{\phi(1-\delta^{\alpha})}\right)^{\frac{\alpha\eta}{1-\alpha}}(1-\delta^{\alpha})^{\frac{\alpha}{1-\alpha}}$$

This ratio is shown to be monotonically increasing in η, and there exists a unique η^* for which it is equal to unity.

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').

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.

Start by assuming identical θ's and f's. *It is predicted that high-η industries or firms*[4] *are more likely to have a high share of intra-firm imports from a given country* (Antràs 2003).

Assume now that firms 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 that outsourcing, two predictions emerge. First, *within a given industry, firms engaging in intra-firm trade should be more productive than those that outsource.* Second, *intra-firm trade is more likely to occur the higher is the industry specific share η_s of the final good producer in the value-added chain* (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 country's capital/labor*

[3] See Antràs (2003), equation 8 p. 1390. As he points out, the absence of factor prices in this variable profit ratio is an artefact of the Cobb-Douglas specification of technology.

[4] Defever and Toubal (2007) offer the full derivation of this partial equilibrium building block of Antràs (2003) with firm-specific technologies.

ratio. The effect goes through the number of firms in both industries, which increases with capital abundance.

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 their previous model to partial contractibility of production tasks. There they relate organizational choice in offshore operations to a country's contracting environment[5]. 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 low-η sectors or firms[6], that were fully outsourcing their input production, this does not change anything. However, for high-η firms or sectors, we have a different prediction.

Holding headquarter services contractibility constant, an improvement in input contractibility abroad has two effects[7] *on headquarter-intensive firms:*

- *the most productive domestic producers switch to offshore outsourcing (the 'Standard Effect');*
- *the most productive firms resorting to offshore outsourcing insource from foreign affiliates (the '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. In sum, improved contract enforcement in the origin country does not determine organizational choice in itself. As

[5]This 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.

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

[7]Nunn and Trefler term these two effects the 'Standard Effect' and the 'Surprise Effect', respectively.

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.' To take this input-country pair specificity into account, empirical studies should rely on measures of contractibility of the imported input *and* of other inputs used by the headquarter firm, as well as the general quality of the country's judicial system.

Note that in their model with fully non-contractible investments, Antràs and Helpman (2004) found an unambiguously positive effect of 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.

3 Data Sources and Variables Used

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

We build a unique cross-section 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 an 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' import at product level (either CPA96 or HS4 4digit) and their sourcing modes - through independent suppliers and/or affiliates. An intra-firm transaction is defined as a transaction with an affiliate controlled by a single French entity with at least fifty percent of its equity capital.

Aggregating firms' transactions 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

[8]See Appendix A for details.

(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 transaction. 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 number 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 import and export transactions operated by French firms as coming out from custom declarations. For transactions outside the EU15, there is no minimal amount for a transaction to be recorded. Within the EU, only transactions whose total annual amount for a given country-product couple exceeds 250,000 euros per year should be registered. In practice however, many transactions 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' transactions 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 transactions 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 transactions with overseas French territories). The final import flows dataset covers 1,141,393 firm-country-product-sourcing mode combinations, corresponding to 126,926 importers, 201 countries and 272 products (CPA96 4digit).

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=outsourcing).

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 transactions 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 production inputs. We use the revenue-based Levinsohn and Petrin (2003) estimator described in Appendix B. Details on the estimation procedure are provided in Appendix A.

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

[9]The 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 1.5% more observations. See Appendix A for details

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 (intra-firm) or from a third party (outsourcing), 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. 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 final good 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. We have thus built 2 alternative variables measuring the degree of contract completeness of an intermediate product p. Denoting by $R_j^{neither}$ ($R_j^{ref\ priced}$) a dummy variable that takes value 1 if the HS6 product j is neither sold on an organized exchange nor reference priced (not reference priced),[10] 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:[11]

1. $\mu_p^1 = 1 - (\sum_j \theta_{p,j}\ R_j^{neither})$
2. $\mu_p^2 = 1 - (\sum_j \theta_{p,j}\ (R_j^{neither} + R_j^{ref\ priced}))$

Summary statistics on μ_p^1 and μ_p^2 are reported in Table **??**.

[10] We actually use the 'Liberal' Rauch classification. Results are virtually unchanged if we use the 'conservative' one.

[11] See Appendix A for details.

Table 1: Summary Statistics on μ_p^1 and μ_p^2

Variable	Observ.	Mean	St. Dev	Min	Max
μ_p^1	247	0.403	0.449	0	1
μ_p^2	247	0.181	0.331	0	1

Finally, in order to measure the capital and skill intensity of the imported product p, we have constructed the variables k_p and h_p using French technology. k_p equals the average log capital-labor ratio in the industry producing good p, while h_p equals the average of the log of the ratio between total wage expenses and employment in the industry producing p.[12] We acknowledge the fact that these are imperfect measure of the capital and skill intensity of the imported product. However, we will see that some useful insights may be obtained by using these proxies.

3.3.4 Final Product Variables

As explained before 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 will measure the contractibility of the final good f with two alternative measures that are still based on the Rauch classification. Denoting by $R_j^{neither}$ ($R_j^{ref\ priced}$) a dummy variable that takes value 1 if the PRODCOM2002 8 digit product j is neither sold on an organized exchange nor reference priced (not 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:[13]

1. $\mu_f^1 = 1 - (\sum_j \theta_{f,j}\ R_j^{neither})$
2. $\mu_f^2 = 1 - (\sum_j \theta_{f,j}\ (R_j^{neither} + R_j^{ref\ priced}))$

Summary statistics on μ_f^1 and μ_f^2 are reported in Table **??**. These alternative measures proxy for the degree of contract completeness of the final product f. As a matter of fact, μ_f^1 and μ_f^2 are pretty correlated with the equivalent Nunn (2007) measures of overall production complexity

[12] See Appendix A for details.

[13] See Appendix A for details.

relying on the US input-output matrix. When comparing them across the 29 ISIC rev2 3 digit sectors (the only classification for which our data are comparable to those of Nunn) the correlation is, respectively, 0.78 for μ_f^1 and 0.42 for μ_f^2.

Table 2: Summary Statistics on μ_f^1 and μ_f^2

Variable	Observ.	Mean	St. Dev	Min	Max
μ_f^1	218	0.373	0.440	0	1
μ_f^2	218	0.158	0.311	0	1

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. k_c and h_c are (respectively) the log of the capital/labor and human capital/labor ratios provided by Hall and Jones (1999). These variables are available for 115 countries.

Q_c is a measure of quality of institutions and 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. This variable is available for 147 countries.

$Same-leg-orig_c$ is a dummy indicating whether country c adopts a French civil law system. This variable, taken from Djankov et al. (2003), proxy for the degree of legal similarity between the exporting country c and the importing country (France).

The last set of variables comes from CEPII (Centre d'Etude Prospectives et d'Informations Internationales). $Distw_c$ 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). $Colony_c$ is dummy indicating whether country c was a former French colony while $Language_c$ is a dummy indicating whether French is spoken in country c.

4 Empirical evidence

We use a logit 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, 201 countries and 272 products.*[15]

Throughout the analysis we will provide both test of empirical predictions of certain models (Results) as well as some important stylized fact concerning intra-firm and outsourcing (Stylized Facts). Before going into estimations we report a first stylized fact:

- *Stylized 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 some intra-firm transactions (due to non-respondents) that we consider as outsourcing.

There are certainly many possible interpretations of this fact. 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 that 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, 201 countries, and 272 products. These observations still represent more than 60% of French imports from the manufacturing sector.[16] Focusing

[14] Results are robust to the alternative probit specification.

[15] The 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.

[16] In this smaller dataset 5.5% of transactions are intra-firm, representing 34% of the value of imports.

on firm-level determinants, we will use sector, country and product fixed effects 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 **??**, 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. In unreported estimations, we confirm his industry-level result in the French case. 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 **??**). 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 enough, an appealing alternative explanation is that there is within-industry heterogeneity in skill and headquarter intensity. 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 original 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 **??** provide the standard deviations (column 2) and correlations (columns 3 to 6) between 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

[17] 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.

Table 3: Standard deviations and correlations of TFP_i, k_i, η_i^{hq}, and η_i^{sk}

		Correlation with			
Variable	St. Dev	TFP_i	k_i	η_i^{hq}	η_i^{sk}
TFP_i	1.89	1.00			
k_i	0.98	-0.00	1.00		
η_i^{hq}	0.18	-0.01	-0.07	1.00	
η_i^{sk}	0.31	0.11	0.25	0.13	1.00

Table 4: Standard deviations and correlations of $TFP_{i,NES}$, $k_{i,NES}$, $\eta_{i,NES}^{hq}$, and $\eta_{i,NES}^{sk}$

		Correlation with			
Variable	St. Dev	$TFP_{i,NES}$	$k_{i,NES}$	$\eta_{i,NES}^{hq}$	$\eta_{i,NES}^{sk}$
$TFP_{i,NES}$	0.53	1.00			
$k_{i,NES}$	0.82	0.06	1.00		
$\eta_{i,NES}^{hq}$	0.16	-0.15	-0.02	1.00	
$\eta_{i,NES}^{sk}$	0.28	0.25	0.21	0.12	1.00

between TFP and the other three firm-level variables are either negligible or relatively small. 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 $k_{i,NES}$ the value of k_i minus its NES114 industry average (56 industries), and apply the same notation to other variables. We report the standard deviations and correlations of these *within-sector* variables in Table **??**. As one can notice, the standard deviation of productivity within sectors is around 30% of the corresponding value on the whole manufacturing. This means that 70% 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 ($k_{i,NES}$) displays even more heterogeneity than TFP and skill intensity ($\eta_{i,NES}^{sk}$) has a standard deviation within a sector which is not that far from the one of TFP.* The same qualitative pattern emerges if we narrow our definition of industry to go down at the NACE rev1 4 digit level.

Table **??** shows a positive but not dramatically high correlation between $TFP_{i,NES}$ and the other variables *Finally, cross-correlations between $k_{i,NES}$, $\eta_{i,NES}^{hq}$ and $\eta_{i,NES}^{sk}$ 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:

- *Stylized 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 logit model:

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

$$y_{i,p,c} = \begin{pmatrix} 1 & if & y^*_{i,p,c} \geq 0 & (intra-firm) \\ 0 & if & y^*_{i,p,c} < 0 & (outsourcing) \end{pmatrix}.$$

where $\mathbf{D}_{NES}$, $\mathbf{D}_p$ and $\mathbf{D}_c$ stands, respectively, for sector, product and country dummies. Table **??** reports estimations of different specifications of **??**. 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:

- *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 brand 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 brand 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

[18] Even 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 5: Firm i-specific determinants of intra-firm trade. Dependent variable: $y_{i,p,c}$=1 for intra-firm.

Estimated specification	1	2	3	4	5
TFP_i	0.0408***				0.0259***
	(0.0008)				(0.0009)
k_i		0.0225***			0.0151***
		(0.0005)			(0.0005)
η_i^{hq}			0.0360***		0.0101***
			(0.0021)		(0.0018)
η_i^{sk}				0.0654***	0.0235***
				(0.0016)	(0.0019)
NES114 sectoral dummies	YES	YES	YES	YES	YES
Country and product dummies	YES	YES	YES	YES	YES
Number of observations	238,841	238,841	238,841	238,057	238,057
Pseudo R^2	0.1413	0.1299	0.1015	0.1322	0.1668
Log Likelihood	-43,836	-44,271	-45,722	-44,114	-42,345

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

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 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 residual property rights approach literature with incomplete contracts. We will exploit the full import flow dataset (1,141,393 firm-country-product-sourcing mode combinations) using firm fixed effects to control for potentially endogenous unobservable firm characteristics. Quite a lot of firms in the data indeed import different products from many countries under different sourcing modes. These observations provide the source of identification of our conditional fixed effect logit model.

Summary of theoretical predictions. Antràs (2003) predicts that intra-firm trade is more likely with capital abundant countries. Arguably, specific skills are needed in the production of the foreign input and the headquarter firm can even partially provide those skills with a technology transfer. Following the same logic we use both skill abundance and capital abundance (h_c and k_c) as regressors. Previous empirical studies based on aggregate data (like Nunn and Trefler, 2008) have not tested this hypothesis yet.

Another determinant suggested by Antràs and Helpman (2004) is the quality of institutions in the origin country (South). They find that an increase in the quality of institutions Q_c of the exporting country favors unambiguously intra-firm trade relative to offshore outsourcing. But the prediction becomes ambiguous under partially incomplete contracting (Antrà and Helpman 2008). If the quality of contract enforcement rises in the South and improves tasks' contractibility under both sourcing modes then there are two opposite effects. On the one hand, more low-productive firms start outsourcing in the South (Standard Effect); on the hand some high-productivity importers start integrating as there is less need to provide incentives to the foreign supplier (Surprise Effect).

Nunn and Trefler (2008), using industry-level data on intra-firm trade find that the Surprise Effect dominates. We will check here if this finding is robust to firm specific characteristics. In the next Subsection we will try to isolate the two different effects by interacting institutional quality with firms' characteristics, since the theoretical effect comes from two sub-populations of firms.

Testing the theory. In order to test the relationship between contract enforcement quality and incomplete contracting highlighted in Antràs and Helpman (2008) one tempting idea, pioneered by Nunn and Trefler (2008), is to interact the country-specific quality of institutions Q_c with the intermediate product level of contractibility ($\mu_p * Q_c$).[19] However, we do not believe this is a good strategy.

When thinking deeply about what we can really test about the theory, one clear shortcoming comes to mind. A typical firm i in the data imports many intermediate products and not just one as in Antràs and Helpman (2008) model. Therefore, $\mu_p * Q_c$ confounds two things: 1) the choice of an optimal sourcing mode for products with different levels of contractibility; 2) the choice of an optimal sourcing mode for countries with different levels of contract enforcement.

Concerning the issue of optimal sourcing in the case of many inputs, the theory is essentially

[19] Nunn and Trefler (2008) actually interact Q_c with the "average" contractibility of the inputs needed for the production of final product f only (via the US input-output matrix).

silent. We believe that the sign and significance of covariates measuring contractibility of both the final good and intermediate inputs (μ_f and μ_p) will provide useful insights for the design of a more general model dealing with many inputs and outputs.

As for the second point, 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. Empirically, contractibility in the South is a function $f(\cdot)$ of μ_p and Q_c). Therefore to reproduce Antràs and Helpman's comparative statics 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.[20]

Concerning other product and country characteristics, there are a number of variables of interest for which, however, background theoretical models have no clear prediction. Nevertheless, we include these variables in our estimations because we see them as important controls. Furthermore, the sign and significance of these additional regressors will provide stylized facts that can inspire future theoretical work.

The first set of variables are the capital and skill intensity of the intermediate good (respectively k_p and h_p) reconstructed using French technology. Other interesting variables are controls for fixed as well as variable costs of engaging into intra-firm and/or outsourcing activities like log of distance of country c to France ($Distw_c$), having been a French colony ($Colony_c$), sharing French as a common language ($Language_c$), and adopting a French civil law system ($Same-leg-orig_c$).[21]

In order to test the relevance of these country and product level variables we start by estimating the following conditional fixed effects logit model on the whole import flow dataset:

$$\begin{aligned} y^*_{i,p,c} = \alpha + \beta_1\ k_c + \beta_2\ h_c + \beta_3\ \mu_p + \beta_4\ Q_c + \beta_5\ k_p + \beta_6\ h_p + +\beta_7\ Distw_c \\ +\beta_8\ Colony_c + \beta_9\ Language_c + \beta_{10}\ Same - leg - orig_c + i + \varepsilon_{i,p,c} \end{aligned} \tag{2}$$

[20] In unreported regressions we introduce an additional interaction term, $\mu_p * Q_c$, and find that it is not significant. Therefore we rule out more complex specifications of $f(\cdot)$.

[21] In North-South models integrating the intra-firm vs. outsourcing sourcing choice (e.g. Antràs and helpman, 2004), one key variable is the labor cost in the South. As GDP per capita is usually highly correlated with wages, one tempting idea is to use it as a proxy for labor cost. However, we do not believe this is a good choice for the following reasons. First, GDP per capita is at best a poor proxy for labor cost. Wages and productivity vary across countries and what we would really need is a productivity deflated measure of wages for country c (we leave this exercise for future work). Moreover, per-capita GDP (which is an output variable) is very correlated with the determinants of a country productivity like capital/labor ratio, human capital/labor ratio as well as with the quality of institutions. Therefore, for the above reasons, we decide not to use GDP per capita as an additional control.

$$y_{i,p,c} = \begin{pmatrix} 1 & if & y^*_{i,p,c} \geq 0 & (intra-firm) \\ 0 & if & y^*_{i,p,c} < 0 & (outsourcing) \end{pmatrix}.$$

where i is a firm-specific fixed effect potentially correlated with explanatory variables.

Column (1) and (2) of Table **??** report the estimated marginal effects and standard errors of model (**??**). The two set of estimations differs in the measure used for the contractibility of the imported input: μ_p^1 in column (1) and μ_p^2 in column two.

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

To check the robustness of our results we have: 1) Estimated the same model but with firm random effects. This this allow us to introduce μ_f as an additional variable and to exploit the entire dataset with the exception of some observations for which we have no value for μ_f and/or μ_p.[22] Estimations results are provided in column (3) to (4) of Table **??**; 2) Estimated the following logit model on the smaller dataset of 247,528 firm-country-product-sourcing mode combinations for which firm level information is available:

$$\begin{aligned} y^*_{i,p,c} = \alpha + \beta_1\ k_c + \beta_2\ h_c + \beta_3\ \mu_p + \beta_4\ \mu_f + \beta_5\ Q_c + \beta_6\ k_p + \beta_7\ h_p + + \beta_8\ Distw_c \\ + \beta_9\ Colony_c + \beta_{10}\ Language_c + \beta_{11}\ Same - leg - orig_c + \mathbf{X_i}'\mathbf{c} + \varepsilon_{i,p,c} \end{aligned} \tag{3}$$

$$y_{i,p,c} = \begin{pmatrix} 1 & if & y^*_{i,p,c} \geq 0 & (intra-firm) \\ 0 & if & y^*_{i,p,c} < 0 & (outsourcing) \end{pmatrix}.$$

where we add the final good contractibility and we control for observable firms characteristics (firm productivity as well as capital, headquarter and skill intensity) indicated by the vector $\mathbf{X_i}$. Estimations results are provided in column (5) to (6) of Table **??**. To save space we do not report

[22] For some firms, especially those whose primary activity is in the service and/or distribution sector, we cannot associate a level of contractibility for their final product. This is due to the fact that the Rauch (1999) classification, that is building block of our contractibility measure, is concerned essentially with manufacturing, agriculture and mining goods. The same issue apply to imported products.

Table 6: Product and Country-specific determinants of intra-firm trade. Dependent variable: $y_{i,p,c}$=1 for intra-firm (marginal effects presented).

Estimated specification	1	2	3	4	5	6
k_c	-0.0091**	-0.0067**	-0.0170***	-0.0167***	-0.0078***	-0.0077***
	(0.0038)	(0.0028)	(0.0044)	(0.0044)	(0.0009)	(0.0009)
h_c	-0.0277**	-0.0207**	-0.0238	-0.0257	0.0078**	0.0074**
	(0.0139)	(0.0103)	(0.0157)	(0.0159)	(0.0035)	(0.0035)
μ_p^1	-0.0336***		-0.0411***		-0.0094***	
	(0.0046)		(0.0053)		(0.0011)	
μ_p^2		-0.0164***		-0.0182**		-0.0049***
		(0.0052)		(0.0078)		(0.0017)
μ_f^1			-0.1130***		-0.0088***	
			(0.0120)		(0.0016)	
μ_f^2				-0.1214***		-0.0173***
				(0.0168)		(0.0026)
Q_c	0.0897***	0.0660***	0.1016***	0.1032***	0.0447***	0.0447***
	(0.0225)	(0.0166)	(0.0259)	(0.0262)	(0.0057)	(0.0057)
k_p	-0.0123***	-0.0166***	-0.0119***	-0.0270***	0.0012**	-0.0018***
	(0.0029)	(0.0019)	(0.0032)	(0.0029)	(0.0006)	(0.0005)
h_p	-0.0105	-0.0083	-0.0209*	-0.0243**	-0.0095***	-0.0102***
	(0.0097)	(0.0072)	(0.0113)	(0.0116)	(0.0022)	(0.0022)
$Distw_c$	-0.0097***	-0.0073***	-0.0052*	-0.0050*	0.0018***	0.0018***
	(0.0026)	(0.0019)	(0.0030)	(0.0031)	(0.0006)	(0.0006)
$Colony_c$	0.1636***	0.1272***	0.1577***	0.1597***	0.0048***	0.0050***
	(0.0101)	(0.0078)	(0.0156)	(0.0156)	(0.0017)	(0.0017)
$Language_c$	-0.0107**	-0.0081**	-0.0113**	-0.0117**	-0.0072***	-0.0071***
	(0.0046)	(0.0034)	(0.0050)	(0.0051)	(0.0009)	(0.0010)
$Same-leg-orig_c$	0.0139***	0.0102***	0.0169***	0.0168***	0.0099***	0.0099***
	(0.0046)	(0.0034)	(0.0053)	(0.0053)	(0.0011)	(0.0011)
Estimation method	conditional firm fixed effects logit	conditional firm fixed effects logit	random firm effects logit	random firm effects logit	logit with controls $\mathbf{X_i}$	logit with controls $\mathbf{X_i}$
Number of potential observations	1,141,393	1,141,393	1,141,393	1,141,393	247,528	247,528
Number of actual observations used in estimations	36,217	36,217	920,413	920,413	199,870	199,870
Pseudo R^2					0.1492	0.1486
Log Likelihood	-14,227	-14,253	-54,731	-54,790	-39,285	39,313

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

coefficients of controls $\mathbf{X_i}$. Again, few observations are lost because we have no value for μ_f and/or μ_p.

Looking across the different sets of estimates in Table **??** 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 the Antràs (2003) model.[23]
- *Result 4: Intra-firm trade is more likely with countries having good judicial institutions.*

We do not have any particular interpretation of Result 3. Antràs (2003) predicts that capital abundant countries (high k_c) should host more intra-firm trade under the very strong assumptions of capital immobility and factor price equalization. These assumptions are probably unrealistic in our dataset. Interestingly enough, when looking at k_p, one can see that intra-firm trade is more likely to involve imported inputs with a low capital-labor ratio. Overall, our data suggest that firms engage in intra-firm activities in capital-scarce countries in order to produce low capital-intensity inputs.

Result 4 states that the better is a country judicial system (high Q_c), the less likely is that firms engage in arms'length international relationships. This is in line with the predictions of the Antràs and Helpman (2004) model. In the light of the more general Antràs and Helpman (2008) model, we confirm in French data the finding by Nunn and Trefler (2008) that the Surprise Effect dominates. We strengthen this result by controlling for both imported and final good contractibility as well as for firm heterogeneity.

To the extent that a better 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.

Nevertheless our findings are, especially looking at the contract complexity of the goods involved in intra-firm trade (μ_p and μ_f), also coherent with another story. Let us first state the following stylized fact:

[23]This result is robust to considering either k_c as the only explanatory variable or k_c and a subset of the other covariates.

- *Stylized 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 an input complexity and its optimal sourcing mode.

Contract intensive goods are thus more likely to be processed within the firm boundaries and we know, from Result 4, that intra-firm international activities are more likely in countries with good judicial institutions. An alternative to the property-rights approach is to consider that internalization reduces the dissipation of intangible assets. Complex inputs are valuable for firms in part because they embody costly R&D efforts and provide producers with market power. As they require physical and legal protection to prevent imitation, firm boundaries represent a safe place to put these valuable intangible assets. Countries providing better protection of property rights offer more guarantees against imitation.

As far as the 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 may cases.

Finally, one may 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 **??** we have explored the role of firm heterogeneity in explaining the offshore sourcing mode, while in Subsection **??** 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 intensity of the host country, contractibility, quality of institutions, etc. This amounts to

look at interactions between firm and product/country variables.

Table 7: Interaction of Q_c with TFP_i, k_i η_i^{hq}, and η_i^{hq}. Dependent variable: $y_{i,p,c}$=1 for intra-firm (marginal effects presented).

Cross effect of Q_c with:	TFP_i	k_i	η_i^{hq}	η_i^{sk}
1st quintile	-0.0133***	-0.0019	0.0472***	-0.0080***
	(0.0017)	(0.0014)	(0.0124)	(0.0022)
2nd quintile	-0.0079***	-0.0024***	0.0443***	-0.0031
	(0.0016)	(0.0009)	(0.0111)	(0.0020)
3rd quintile	-0.0063***	0.0009	0.0587***	-0.0010
	(0.0016)	(0.0006)	(0.0100)	(0.0018)
4th quintile	-0.0056***	0.0022***	0.0462***	0.0025
	(0.0015)	(0.0004)	(0.0092)	(0.0017)
5th quintile	-0.0017	0.0028***	0.0597***	0.0054***
	(0.0013)	(0.0003)	(0.0076)	(0.0013)
Number of observations	199,870	199,870	199,870	199,870
Pseudo R^2	0.1525	0.1516	0.1527	0.1540
Log Likelihood	-39,132	-39,174	-39,125	-39,065

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 much 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 subpopulation of relatively low productive firms. By contrast, the opposite Surprise Effect comes from high productive firms subpopulation. Thanks to our firm level data we can try to identify the tension between the Standard and the Surprise Effect by looking at the interaction between of productivity (as well as capital, skill and headquarter intensity) with Q_c.

In particular, for each NES114 industry, we have computed the 5 quintiles of the distributions of TFP_i, k_i, η_i^{hq}, and eta_i^{sk} and further constructed five variables taking the the value of the corresponding quintile for each firm-level variable. Finally, we have created a cross product variable between each quintile and Q_c and estimated model **??** adding these additional cross effects. We have performed 4 different estimations for each of the 4 interacting variable.[24]

Results of such estimations are reported in Table **??** 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

[24]In particular we consider the specification with μ_p^1 and μ_f^1.

Table 8: Interaction of μ_p^1 with TFP_i, k_i η_i^{hq}, and η_i^{hq}. Dependent variable: $y_{i,p,c}$=1 for intra-firm (marginal effects presented).

Cross effect of μ_p^1 with:	TFP_i	k_i	η_i^{hq}	η_i^{sk}
1st quintile	-0.0028**	0.0080***	-0.0598***	0.0132***
	(0.0013)	(0.0010)	(0.0150)	(0.0029)
2nd quintile	-0.0005	0.0038***	-0.0578***	0.0137***
	(0.0010)	(0.0008)	(0.0126)	(0.0026)
3rd quintile	-0.0013	0.0040***	-0.0385***	0.0141***
	(0.0010)	(0.0006)	(0.0110)	(0.0025)
4th quintile	-0.0019**	0.0037***	-0.0499***	0.0138***
	(0.0009)	(0.0006)	(0.0096)	(0.0024)
5th quintile	-0.0018***	0.0032***	-0.0220***	0.0113***
	(0.0007)	(0.0005)	(0.0076)	(0.0018)
Number of observations	199,870	199,870	199,870	199,870
Pseudo R^2	0.1495	0.1502	0.1502	0.1497
Log Likelihood	-39,275	-39,240	-39,239	-39,263

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

1st to the 5th quintile. In the case of the interaction between Q_c and TFP_i one can see that the difference between the fifth and the first quintile is positive and (considering twice the sum of standard deviations) is also significant. Actually, values are significantly higher moving up in the quintile scale in all interaction cases except for headquarter intensity. This means that the quality of country c contract enforcement has a stronger positive effect on sourcing via intra-firm, the more productive, capital- and skill-intensive is a firm:

- *Result 5: The 'surprise' effect is significantly stronger for more productive, capital intensive, and skill intensive firms.* 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 **??** and **??** we perform a similar exercise to the one reported in Table **??**. For each NES114 industry, we have again computed the 5 quintiles dummies of the distributions of TFP_i, k_i, η_i^{hq}, and eta_i^{sk}. We have subsequently created a cross product variable between each quintile and μ_p^1 (Table **??**) and μ_f^1 (Table **??**) while estimating model **??** with such additional cross effects.

Table 9: Interaction of μ_f^1 with TFP_i, k_i η_i^{hq}, and η_i^{hq}. Dependent variable: $y_{i,p,c}$=1 for intra-firm (marginal effects presented).

Cross effect of μ_f^1 with:	TFP_i	k_i	η_i^{hq}	η_i^{sk}
1st quintile	-0.0010	0.0106***	0.0670***	0.0343***
	(0.0019)	(0.0010)	(0.0180)	(0.0050)
2nd quintile	-0.0005	0.0067***	0.0784***	0.0361***
	(0.0017)	(0.0008)	(0.0146)	(0.0047)
3rd quintile	-0.0011	0.0020***	0.0878***	0.0354***
	(0.0016)	(0.0007)	(0.0126)	(0.0045)
4th quintile	-0.0016	0.0056***	0.0571***	0.0313***
	(0.0014)	(0.0006)	(0.0111)	(0.0044)
5th quintile	-0.0020	0.0028***	0.0550***	0.0261***
	(0.0012)	(0.0006)	(0.0087)	(0.0036)
Number of observations	199,870	199,870	199,870	199,870
Pseudo R^2	0.1494	0.1516	0.1507	0.1502
Log Likelihood	-39,278	-39,175	-39,217	-39,242

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

Tables inspection reveals that only capital intensity matters. In particular, the higher the capital intensity of the firm the more negative is the impact of contractibility, of both the imported product and the final good, on intra-firm trade. To put differently this original finding:

- *Stylized fact 4: High capital intensive firms are more likely to produce complex goods within firm boundaries.*

5 Conclusion

We have investigated the determinants of the internalization of imports of intermediate products. We have constructed a unique cross-section database of 1,141,393 French import transactions in 1999 (firm-country-product triples) corresponding to 126,926 importers, 201 countries and 272 products . In this dataset we have identified intra-firm transactions and have built a binary variable taking value one when transactions are intra-firm. Although little in number, intra-firm transactions represents a large portion of total imports. We have then conducted a detailed examination of firm-, exporter country- and product- level determinants of intra-firm trade and their interaction in the light of the property-rights models of the international firm boundaries, due to Antràs (2003), Antràs and Helpman (2004), and Antràs and Helpman (2008).

We start by assessing that the the theory overlooks the different dimensions of firms' heterogeneity. Firms are in fact not only different in their productivity. Even in narrowly defined industries, they in fact display a lot of heterogeneity in variables like capital, skill, and headquarter intensity that crucially affects the optimal sourcing mode. However, once recognized their firm-level dimension, the sign and significance of these variables is in line with the basic mechanisms of the residual property rights approach.

We also find that firms engaged in intra-firm trade are more productive and that they value the level of the exporting country quality of enforcement differently. Moreover, we show that intra-firm trade involves low capital intensive inputs imported from low capital intensive countries.

We take a broad view of intra-firm trade and provide some robust empirical evidence that can inspire future theoretical work. In particular we show that complex goods and inputs are more likely to be produced within the firm boundaries and this is particularly true for capital intensive firms. This finding is consistent with a framework in which internalization reduces the dissipation of intangible assets. Complex inputs are valuable for firms in part because they embody costly R&D efforts and provide producers with market power. As they require physical and legal protection to prevent imitation, firm boundaries represent a safe place to put these valuable intangible assets while countries providing better protection of property rights offer more guarantees against imitation.

References

[1] Antràs, Pol, "Firms, Contracts and Trade Structure," Quarterly Journal of Economics, Vol 118 (2003), 1375–1418.

[2] Antràs, Pol, and Elhanan Helpman, "Global Sourcing," Journal of Political Economy, Vol. 112 (2004), 552–580.

[3] 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.

[4] 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).

[5] Djankov, S., R. La Porta, F. Lpez-de-Silanes, and A. Shleifer, "Courts"Quarterly Journal of Economics, Vol. 118 (2003), 453-517.

[6] 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.

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

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

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

[10] Grossman, Gene M. and Elhanan Helpman, Outsourcing versus FDI in industry equilibrium, Journal of the European Economic Association 1: 317-327, 2003.

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

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

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

[14] Gwartney, James, and Robert Lawson, "Economic Freedom of the World: 2003 Annual Report," Mimeograph, Fraser Institute, 2003.

[15] 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.

[16] Head Keith and Thierry Mayer, Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade, (2002) , CEPII Working Paper No.2002-01.

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

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

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

[20] 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

[21] McLaren John, "Globalization and Vertical Structure ", American Economic Review 90:5 (December 2000),pp. 1239-54.

[22] 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

[23] Marin, Dalia and Thierry Verdier, "Market Power Inside the Firm and the Market ", Journal of the European Economic Association, *forthcoming*

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

[25] Nathan, Nunn, and Daniel Trefler Helpman, "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.

[26] Rauch, James E., "Networks versus Markets in International Trade" Journal of International Economics, Vol. 48 (1999), 7–35.

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

[28] 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

An intra-firm transaction is defined in the EIIG database as a transaction with an affiliate controlled by a single French entity with at least fifty percent of its equity capital. The SESSI defines two types of transaction 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 transactions with independent suppliers as outsourcing. In the data 20,952 out of the 81,217 transactions (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,[25] and 10,244 (12.61%) are 'mixed transactions' in the sense that only a share of imports of product p from country c in the year 1999 comes from a foreign affiliate, with the remaining share being imported from a third party. In order to exploit some of the information contained in mixed transactions, we consider them as intra-firm (outsourcing) if the share of intra-firm (outsourcing) in the total transaction value exceeds 80% [26] ending up with 76,364 firm-product-country triples corresponding to 4,193 firms. 31.28% of these transactions are intra-firm corresponding to 55.48% of total imports value in the dataset. 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[27] industries in manufacturing.

Total Factor Productivity of firm i and is as the residual (plus the constant) of a log-linearized

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

[26] Estimation results are not affected by the exclusion of the those mixed transactions that we impute as either intra-firm or outsourcing.

[27] The French NES114 (Nomenclature Economique de Synthèse) sector classification has the advantage of allowing us to merge product and sector information. Its level of detail is roughly between NACE rev1 2 and 3 digit.

three-factor Cobb-Douglas production function, with labor, capital and material inputs as production inputs. We use the revenue added-based Levinsohn and Petrin (2003) estimator. Labor is the full time equivalent average number of employees in a given year. Production is calculated as production sold plus stocks variations. Material inputs are calculated as bought inputs minus stocks variation. Deflators for production 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.[28] 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. Using this methodology, we have built 2 variables measuring the degree of contract completeness of an intermediate product p. Denoting by $R_j^{neither}$ ($R_j^{ref\ priced}$) a dummy variable that takes value 1 if the HS6 product j is neither sold on an organized exchange nor reference priced (not reference priced),[29] 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:

1. $\mu_p^1 = 1 - (\sum_j \theta_{p,j}\ R_j^{neither})$
2. $\mu_p^2 = 1 - (\sum_j \theta_{p,j}\ (R_j^{neither} + R_j^{ref\ priced}))$

As for the 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

[28] Correspondence tables have been obtained using RAMON data available at the web-site: http://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL

[29] We actually use the 'Liberal' Rauch classification. Results are virtually unchanged if we use the 'conservative' one.

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. Denoting by $R_j^{neither}$ ($R_j^{ref\ priced}$) a dummy variable that takes value 1 if the PRODCOM2002 8 digit product f is neither sold on an organized exchange nor reference priced (not reference priced),[30] 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:

1. $\mu_f^1 = 1 - (\sum_j \theta_{f,j}\ R_j^{neither})$
2. $\mu_f^2 = 1 - (\sum_j \theta_{f,j}\ (R_j^{neither} + R_j^{ref\ priced}))$

A.4 Other imported product variables

The other two variables we have constructed are the capital and skill intensity in the production of imported product p denoted, respectively, by k_p and h_p. In order to build such measures, 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.

B Levinsohn and Petrin TFP estimator

The Levinsohn and Petrin (2003) approach consists of a two stages procedure based on the identification of a proxy variable for the unobserved (by the econometrician) productivity component ω_{it}. The identification stems from the assumption that intermediate inputs consumption in production reacts to the observed (by the firm) productivity.

Let us start with a standard log-linearized Cobb-Douglas production function where y is log output, k is log capital stock, l is log employment, m is log intermediate inputs, ω is a productivity shock observed by the firm but not by the econometrician and possibly correlated with inputs, and u is a random term not correlated with any other component of the regression function:

$$y_{it} = \alpha\, k_{it} + \beta\, l_{it} + \gamma m_{it} + \omega_{it} + u_{it}. \tag{A.1}$$

[30] We actually use the 'Liberal' Rauch classification. Results are virtually unchanged if we use the 'conservative' one.

The intermediate input's demand function of firm i at time t is assumed to be $m_{it} = m_{it}(\omega_{it}, k_{it})$ and can be used to generate, once inverted (the invertibility condition is that, conditional on capital, the intermediate inputs demand must be increasing in ω), the proxy-variable:

$$\omega_{it} = g_{it}(m_{it}, k_{it}) \tag{A.2}$$

which, substituted in equation (**??**), yields:

$$y_{it} = \beta\, l_{it} + \Phi_{it}(m_{it}, k_{it}) + u_{it} \tag{A.3}$$

where

$$\Phi_{it}(m_{it}, k_{it}) = \alpha\, k_{it} + \gamma m_{it} + g_{it}(m_{it}, k_{it}) \tag{A.4}$$

The two last equations form a "partially linear" model identifying β. Since the regressors are no longer correlated with the error, β can be now estimated approximating Φ by a third or fourth order polynomial ($\tilde{\Phi}$). However, α and γ are not identified at this stage; in order to yield a consistent estimation, we need to introduce some more structure into the model and to use, in a second stage, the estimated coefficient of labor ($\hat{\beta}$). To see this, net from the output in equation (**??**) the estimated contribution of labor

$$y_{it} - \hat{\beta}\, l_{it} = \alpha k_{it} + \gamma m_{it} + \omega_{it} + u_{it} \tag{A.5}$$

and assume, for simplicity, that ω_{it} evolves according to a first-order Markov process, which implies that $\omega_{it} = E[\omega_{it} \mid \omega_{it-1}] + e_{it}$, where e_{it}, the s.c. "surprise", denotes *innovation* in ω_{it}. Accordingly, capital and material inputs in t adjusts, through investment, to ω_{t-1}, but they do not adjust to e_{it}. Owing to this, $(\hat{\alpha}, \hat{\gamma})$ are obtained by minimizing the following GMM criterion function

$$Q(\alpha^*, \gamma^*) = \min_{(\alpha^*, \gamma^*)} \sum_h \left(\sum_i \sum_{t=T_{i0}}^{Ti1} \widehat{(e_{it} + u_{it})} Z_{iht} \right)^2, \tag{A.6}$$

where: h indexes the elements of $Z_t = (k_t, m_{t-1})$; i indexes firms; T_{i0}, and T_{i1} are, respectively, the first and last period in which firm i is observed; and

$$\widehat{(e_{it} + u_{it})} = y_{it} - \hat{\beta} l_{it} - \alpha^* k_{it} - \gamma^* m_{it} - \widehat{E[\omega_{it} \mid \omega_{it-1}]}. \tag{A.7}$$

According to (**??**), in order to proceed with the minimization of (**??**), we need to know α^*, γ^*, and $\widehat{E[\omega_{it} \mid \omega_{it-1}]}$ ($\hat{\beta}$ is known from the first stage).[31]

[31] The vector

$$E[(e_{it} + u_{it}) \mid Z_t],$$

We can start from calculating the following residuals:

$$y_{it} - \hat{\beta} l_{it} = \hat{\Phi}_{it} \tag{A.10}$$

then, ω_{it} can be obtained using any candidate values α^* and γ^* in the following equation:

$$\hat{a}_{it} = \hat{\Phi}_{it} - \alpha^* k_{it} - \gamma^* m_{it}. \tag{A.11}$$

Using these values, we are able to obtain a consistent approximation to $E[\widehat{\omega_{it} \mid \omega_{it-1}}]$ from

$$E[\widehat{\omega_{it} \mid \omega_{it-1}}] = \delta_0 + \delta_1 \omega_{it-1} + \delta_2 \omega_{it-1}^2 + \delta_3 \omega_{it-1}^3 + \epsilon_{it} \tag{A.12}$$

Finally, given $\hat{\beta}_{it}$, α^*, γ^*, and $E[\widehat{\omega_{it} \mid \omega_{it-1}}]$, the solution of problem (**??**) provides the estimation of capital ($\hat{\alpha}$) and intermediate input ($\hat{\gamma}$) coefficients.

---

at the base of the moment conditions, results from the two following assumptions (which, in turn, represent the conditions under which intermediate input can be thought of as a "perfect proxy" for ω_{it}). The first one is that period t's capital is determined by the investment decisions in the previous period, so that it does not respond to the productivity innovation (e_{it}) in the current period:

$$E\left[(e_{it} + u_{it}) \mid k_{it}\right] = 0 \tag{A.8}$$

The second assumption is that last period's intermediate input choice is uncorrelated with the innovation in the current period:

$$E\left[(e_{it} + u_{it}) \mid m_{it-1}\right] = 0 \tag{A.9}$$