Chapter 8

Internalization: Empirical Evidence

In the last two chapters, I have reviewed two of the leading theories of the boundaries of the firm and I have shown how to embed them into our benchmark model of global sourcing with heterogeneous firms. In this last chapter of the book, I will describe how these internalization theories can be taken to the data. The empirical literature on this topic is still budding and has yet to provide fully convincing empirical tests of these models. Several well-crafted papers have offered different pieces of evidence that are consistent with one or more of those models, but the power of such tests remains fairly low, as I will try to explain below. The goal of this chapter is thus not only to overview and replicate past work, but rather to try to highlight some of its limitations and suggest avenues for future research in this area.

In great part, the current limitations of the empirical literature on multinational firm boundaries are due to the fact that empirically testing internalization theories poses at least two important challenges. First, data on the integration decisions of firms are not readily available, and thus researchers are often left to test these theories with specific industry- or product-level data. Second, the predictions from these models are associated with subtle features of the environment (such as the relative value of the marginal return to non-contractible, relationship-specific investments) that, by their own nature, are generally unobservable in the data (see Whinston, 2003).

As already explained in Chapter 1, the first limitation is not specific to the study of *international* internalization decisions. There are very few comprehensive datasets that allow researchers to measure the extent to which firms control the different agents involved in their production processes. Indeed, the pioneering empirical literature in the 1980s that implemented tests of the transaction-cost theory relied on records of the integration decisions of a handful of firms in quite specific industries. For instance, Monteverde and Teece (1982) employed data for 133 components purchased by two U.S. automakers (General Motors and Ford). Masten (1984) focused on a single firm in the aerospace industry that procured a large number (1,887 to be precise) of inputs. The classical studies of Joskow (1985, 1987) studied the ownership and contractual relationships between U.S. coal suppliers and electric utilities. Even the more recent studies testing alternative theories of the boundaries of the firm have relied on rather peculiar sectors, such as Baker and Hubbard's (2003) or Gil's (2007) studies of the trucking and movie industry, respectively. This approach has also pervaded the international business literature, where the transaction-cost theory has been frequently tested using a small sample of specific internalization decisions of multinational firms in certain industries and countries.[1]

The second limitation of the empirical literature on firm boundaries, namely the limited ability of researchers to proxy for the key objects in those theories, has also been a recurrent concern when assessing the literature on integration decisions. Product complexity, contractual incompleteness and relationship-specificity might have a precise definition in economic models, but they are much harder to gauge in the data. Admittedly, the existing contributions to the empirical literature on multinational firm boundaries have not made an awful lot of progress in addressing this second measurement hurdle.

With regard to the first challenge on data availability, however, an advantage of researchers in the International Trade field is that Customs Offices keep a detailed record of the exchange of goods crossing political borders. To give a precise example, while it would be hard to imagine that a researcher would gain access to information on each domestic input purchase of General Motors, the U.S. Bureau of Customs and Border Protection keeps a record of each international trade transaction involving a U.S. based firm, including all of General Motor's *imported* input purchases. Furthermore, each U.S. import transaction includes various pieces of information, such as the identity of the foreign entity exporting goods into the U.S. and whether that entity is related (in an ownership sense) to the U.S. buyer. In sum, the U.S. customs data contain rich information on the integration decisions of *every* U.S. firm with regards to its foreign suppliers of components.

[1] For instance, Davidson and McFetridge (1984) studied 1,376 internal and arm's-length transactions involving high-technology products carried out by 32 U.S.-based multinational enterprises between 1945 and 1975. See also Mansfield, Romeo and Wagner (1979), Mansfield and Romeo (1980), and Kogut and Zander (1993) for related contributions.

Gaining access to these type of firm-level data is not a simple matter, however.[2] In practice, most researchers (including myself) need to rely on product-level data that aggregate the purchases of that particular product by firms and consumers in the importing country. Crucially, for some countries, these product-level data also contain information on the extent to which these aggregated import transactions are transacted between related parties or non-related parties. In this chapter, I will make extensive use of the U.S. Related-Party Trade database collected by the U.S. Bureau of Customs and Border Protection and managed by the U.S. Census Bureau. This dataset is publicly available from the U.S. census website (http://sasweb.ssd.census.gov/relatedparty/) and provides information on related and non-related-party U.S. imports and exports at the six-digit Harmonized System (HS) classification (which consists of over 5,000 categories) and at the origin/destination country level. This is exactly the same dataset I used in the empirical tests of the global sourcing model performed in Chapter 5, except that I will now be exploiting the related-party information in the data.[3]

The remainder of this chapter is structured as follows. I will first overview some key features and limitations of the U.S. Related-Party Trade database, and will discuss how it can be used to construct tests of the transaction-cost and property-rights theories of multinational firm boundaries. In the process, I will illustrate how one can extend these models to a multi-country environments to better exploit the variation in the intrafirm trade share both across products and countries. Towards the end of the chapter, I will briefly describe alternative data sources that have been and are being used to shed light on the internalization decisions of multinational firms. When doing so, I will put particular emphasis on the availability of firm-level datasets (with different levels of representativeness) that contain detailed information on the sourcing strategies of *firms* in different countries. I will conclude the chapter by offering some thoughts on future avenues for empirical research in the area.

[2]In the U.S. case, one needs to first obtain security clearance and then have a project approved by the U.S. Census data administrators. Most young researchers in the field first access the U.S. Census and Customs data by serving as research assistants to other researchers with approved projects using those data.

[3]The U.S. Related-Party Trade database is in fact available at the more disaggregated six-digit Harmonized System (HS) industrial classification. This dataset is not freely downloadable but can be purchased from the U.S. Census at a fee. Although I have not used these richer data in the tests performed in this book, I have made it available for download at http://scholar.harvard.edu/antras/books.

The U.S. Related-Party Trade Database

Because the U.S. Related-Party Trade database will feature prominently in this chapter, it is important to devote some space to discussing its main advantages and disadvantages. To some extent, this discussion will reiterate some arguments that were already presented in Chapters 1 and 5, but it is worth repeating them here for completion.

Several features of U.S. Related-Party Trade database make it particularly attractive to empirical researchers. First, the database is publicly available and easily downloadable from the U.S. census website. Second, the data are of high quality and are not subject to sampling error, since (i) several quality assurance procedures are performed and (ii) the data offer a *complete* picture of the sourcing transactions of U.S. firms. Third, there is a large amount of variation in the data: the share of U.S. intrafirm imports over total U.S. imports is very large (close to fifty percent), but varies widely across products and origin countries. In Chapter 1, I documented this variation through various figures. Fourth, by including information on *all* industrial sectors, rather than a single sector, these data make it easier to spot certain *fundamental* factors that appear to shape whether or not international transactions are internalized independently of the sector one studies. This is particularly relevant because the models I have developed in this book are highly stylized, and do not aspire to capture the precise workings of any specific sector. A fifth advantage of using these comprehensive datasets is that by covering a wide range of sectors, countries and time periods, they offer the potential to exploit exogenous changes in sector characteristics (due perhaps to technological change) or in institutional characteristics of exporting or importing countries (due, for instance, to institutional reforms) to better identify some of the effects predicted by the theories. I will speculate on this last potential use of these data at the end of this chapter.

Let us next turn to some of the limitations of using the U.S. Related-Party Trade database. These largely overlap with the limitations described in Chapter 5, when I used this same database to study the global location decisions of U.S. firms. First, there is an obvious tension in using aggregated product-level data to test the validity of theories of firm boundaries. Second, the data are reported based on the sector or industry category of the good being transacted and do not contain information on the sector that is purchasing the good. Third, the dataset does not distinguish between imports of intermediate inputs and imports of finished products. Fourth, in related-party transactions, the data do not typically report which firm is owned by whom, that is whether integration is backward or forward, and whether trade occurs within U.S.-based or within foreign-based multinationals. Fifth, the data provide no information on the extent to which parties are related with each others, such, as for instance, an equity share of the parent company in the affiliate. A sixth and final concern is that U.S. data can only capture those sourcing decisions that entail goods being shipped back to U.S. headquarters or affiliates, while in practice some large firms ship parts and components across foreign locations (within and across firm boundaries) and then only ship back to the U.S. fully assembled products (as is the case of the iPad 3 discussed in Chapter 1).

I will defer addressing the second, third and fourth limitations until I present the empirical tests below, but let me briefly confront the other three concerns upfront. With regards to the first limitation, it is important to point out that, as in the case of the tests performed in Chapter 5, the specifications considered below are derived from the models by aggregating the individual producers' ownership decisions into product-level intrafirm trade shares. Thus, *product*-level data are used to test *product*-level predictions. This is not to say, of course, that firm-level data would not be enormously useful in testing these models, as emphasized later in the chapter. The fifth concern regarding the lack of information on equity shares is particularly worrisome given that the threshold equity stake of 6% for recording a transaction as involving related parties is very low. As already mentioned in Chapter 1, however, extracts from the confidential foreign direct investment dataset collected by the Bureau of Economic Analysis indicate that well over 90% of intrafirm trade appears to involve majority-owned affiliates. The sixth and final limitation concerning global value chains implies that U.S. intrafirm imports generally underrepresent the involvement of U.S. multinational firms in global production networks. This is indeed a reason for concern, but it is not obvious how this phenomenon biases the results of empirical studies using these data. An active literature in international trade is attempting to shed light on global value chains through the use of the recently constructed World Input Output tables (see Timmer, Erumban, Los, Stehrer and de Vries, 2014, for a review). Unfortunately, this data source is too aggregated to adequately complement the other sources of data used in this book, and they contain no information on the extent to which global value chains exchange goods within or across firm boundaries.

Cross-Industry Tests: Model Predictions

Having discussed the pros and cons of the U.S. intrafirm trade data, let us now put them to work. I will begin by implementing empirical tests of some cross-industry implications of the two-country transaction-cost and property-rights models developed in Chapters 6 and 7. Because I do not have information on the extent to which U.S. firms source domestic inputs within firm boundaries or at arm's-length, I will focus on the predictions of these models for the share of overall foreign input purchases that are imported within firm boundaries. Furthermore, I will largely concentrate on the case in which contractual frictions in domestic transactions are relatively small, and we can ignore domestic integration as an equilibrium sourcing mode. I do so for three reasons. First, because it seems a sensible assumption to make when the domestic economy is the U.S., which has a legal system that ensures a high degree of contract enforcement. Second, because the findings of Atalay, Hortacsu and Syverson (2013) suggest that intrafirm shipments of physical goods indeed account for a very small share of overall domestic shipments of U.S. establishments. And third, because ruling out domestic integration will significantly simplify our overview of the empirical predictions emanating from the models. In any event, towards the end of the chapter, I will discuss the implications of re-introducing domestic integration into the framework.

From the results in Chapters 6 and 7, a succinct way to express the share of overall foreign input purchases that are imported within firm boundaries is

$$Sh_{i-f} = \frac{\Psi_{OV}/\Gamma_{OO}}{\left[\left(\frac{\tilde{\varphi}_{OV}}{\tilde{\varphi}_{OO}}\right)^{\kappa-\sigma-1} - 1\right] + \Psi_{OV}/\Gamma_{OO}} \tag{8.1}$$

where

$$\frac{\tilde{\varphi}_{OV}}{\tilde{\varphi}_{OO}} = \left[\frac{f_{OV} - f_{OO}}{f_{OO} - f_{DO}} \times \frac{1 - (w_N/\tau w_S)^{-(1-\eta)(\sigma-1)}\,\Gamma_{DO}/\Gamma_{OO}}{\Psi_{OV}/\Gamma_{OO} - 1}\right]^{1/(\sigma-1)}, \tag{8.2}$$

and

$$\Psi_{OV} = \begin{cases} \lambda^{1-\sigma} & \text{in the Transaction-Cost Model;} \\ \Gamma_{OV} & \text{in the Property-Rights Model.} \end{cases} \tag{8.3}$$

More specifically, in the transaction-cost model, this corresponds to equations (6.9) and (6.12), while in the property-rights model, it follows from equation (7.9) and a variation of equation (7.10) incorporating domestic contractual frictions (i.e., with an extra term $\Gamma_{DO} < 1$). Equations (8.1), (8.2) and (8.3) are useful in highlighting both the common and distinct predictions of the transaction-cost and property-rights models. It is worth fleshing out these predictions one last time before discussing the evidence.

Let us begin with the common predictions. Note first that, in both models, we have that the share of intrafirm imports is decreasing in κ and $w_N/\tau w_S$, and increasing in Γ_{DO}/Γ_{OO}. This is because these terms are tightly related to the sorting pattern of firms by productivity into sourcing modes, according to which firms engaged in intrafirm trade are more productive than those conducting offshore outsourcing. On the one hand, this sorting pattern implies that $\tilde{\varphi}_{OV} > \tilde{\varphi}_{OO}$, and this delivers the negative effect of κ (or positive effect of productivity and size dispersion) on the share of intrafirm trade in (8.1). The intuition behind this result is identical to that of any model with heterogeneous firms and a Pareto distribution of productivity, as described in Chapters 2 and 4. On the other hand, the equilibrium sorting pattern also implies that when the effective marginal cost of foreign sourcing decreases (either because trade costs fall, wage differences increase, or relative offshore contractual frictions decrease), some firms are led to select into offshoring, but these firms necessarily do so via offshore outsourcing, thus reducing the intrafirm trade share. To fix ideas, I will refer to this mechanism as the *selection into offshoring* channel, and it is captured by the terms $(w_N/\tau w_S)^{-(1-\eta)(\sigma-1)}$ and Γ_{DO}/Γ_{OO} in equation (8.2).

Leaving aside these common features, the key distinction between the two models resides in the ratio Ψ_{OV}/Γ_{OO}, which captures the relative organizational efficiency of intrafirm and outsource offshoring. This ratio is important because it governs (i) the selection of firms into intrafirm trade, and (ii) the relative demand for inputs by firms integrating and outsourcing abroad.[4] In the transaction-cost model, the numerator Ψ_{OV} is a function of exogenously given governance costs λ, i.e., $\Psi_{OV} = \lambda^{1-\sigma}$. Instead, in the property-rights model, Ψ_{OV} is shaped by the determinants of the contractual efficiency of intrafirm foreign sourcing Γ_{OV}, and the share of intrafirm trade depends on the relative contractual efficiency of foreign integration and outsourcing. As I have discussed in detail in Chapters 4, 6, and 7, Γ_{OV} and Γ_{OO} can be mapped to several primitive features of the models, such as the level of headquarter intensity, the degree of contractual incompleteness and relationship-specificity, demand and input substitution elasticities, and so on.

In order to empirically test and discriminate across these two models a crucial question is then, how are the ratios $\lambda^{1-\sigma}/\Gamma_{OO}$ and Γ_{OV}/Γ_{OO} shaped by these deep parameters of the model? In Table 8.1, I provide a summary of some of the key comparative statics I discussed in Chapters 6 and 7. For completeness, I also include comparative statics related to the *selection-into-*

[4]The first effect is captured by the term Ψ_{OV}/Γ_{OO} in (8.2), while the second one corresponds to the terms Ψ_{OV}/Γ_{OO} in (8.1).

offshoring ratio Γ_{DO}/Γ_{OO}, which was discussed at length in the empirical tests in Chapter 5, and which is common for both models.

Table 8.1. Effect of Parameters on Ψ_{OV}/Γ_{OO} and Γ_{DO}/Γ_{OO}

Transaction-Cost Model	σ	η	ϕ	μ_{hS}	μ_{mS}	ϵ_h	ϵ_m	ρ
Ψ_{OV}/Γ_{OO}	Ambiguous	Ambiguous	$-$	$-$	$-$	$+$	$+$	$-$
Γ_{DO}/Γ_{OO}	$+$	Ambiguous	$-$	$-$	$-$	$+$	$+$	$-$
Property-Rights Model	σ	η	ϕ	μ_{hS}	μ_{mS}	ϵ_h	ϵ_m	ρ
Ψ_{OV}/Γ_{OO}	Ambiguous	$+$	$-$	$-$	$+$	$+$	$-$	$-$
Γ_{DO}/Γ_{OO}	$+$	Ambiguous	$-$	$-$	$-$	$+$	$+$	$-$

Table 8.1 indicates that both models have identical qualitative implications for the following five parameters: the elasticity of demand σ, the level of financial contractibility ϕ, the degree of contractibility μ_{hS} and of relationship-specificity ϵ_h of headquarter services, and the elasticity of substitution across inputs ρ. Thus, empirically testing these predictions is useful for validating or rejecting the two models, but *not* for discriminating among them.

A first difference between the two models is in the role of headquarter intensity in shaping the ratio Ψ_{OV}/Γ_{OO}. While in the transaction-cost model, such dependence was generally ambiguous and depended in subtle ways on the environment, in the property-rights model the predicted sign is unambiguously and robustly positive.[5] When it comes to assessing the overall effect of headquarter intensity on the share of intrafirm trade Sh_{i-f}, one cannot forget, however, its effect via the selection into offshoring channel. This selection effect can in turn be broken down into two components. On the one hand, there is the direct effect of a higher η in equation (8.2), which reduces the relevance of cross-country wage differences for profits, thereby hindering selection into offshoring and increasing Sh_{i-f} on that account. On the other hand, there also exists an effect of η via the term Γ_{DO}/Γ_{OO} governing relative offshore contractual frictions; as indicated in Table 8.1, such an effect is generally ambiguous. Where do all these different effects leave us? A cautious way to summarize the above discussion is that one could interpret a positive dependence of the share of intrafirm trade on headquarter

[5] To be precise, in Chapter 7, we have encountered one violation of this result, under certain parameter values, when dealing with the extension with relationship-specificity and generalized Nash bargaining.

intensity as (weak) supportive evidence for the property-rights model, but without such a dependence necessarily leading one to reject the validity of the transaction-cost model.

A second key difference between the two models relates to the effects of μ_{mS} and ϵ_m on the ratio Ψ_{OV}/Γ_{OO}. As discussed in Chapter 6, in the transaction-cost model, any increase in contractibility or decrease in relationship-specificity of manufacturing inputs tends to reduce the relative profitability of integration. Instead, we have seen in Chapter 7, that in the property-rights model, increases in μ_{mS} or reductions in ϵ_m actually tend to *increase* the relative efficiency of vertical integration. The presence of the selection into offshoring channel again complicates matters, because these same changes in the parameters also increase the relative efficiency of offshore outsourcing relative to domestic outsourcing, thereby leading to a decrease in the share of intrafirm trade on that account. Nevertheless, I would argue that evidence of a positive effect of input contractibility on the share of intrafirm trade or evidence of a negative effect of input specificity on this same share can be interpreted as supporting the property-rights model over the transaction-cost one.

Cross-Industry Tests: Data and Benchmark Results

Let us next turn to the empirical implementation of these tests. The general strategy I will follow here is very simple. I will attempt to find valid empirical proxies for the key parameters in Table 8.1 and see how they shape the share of intrafirm imports, as measured using the U.S. Related-Party Trade database. I will begin by using the raw six-digit NAICS dataset, which is available for 390 manufacturing industries and 12 years. Although this should lead to a total of 4,680 observations, the volume of total U.S. imports is zero for 29 of those observations. Thus the share of intrafirm trade, defined as the ratio of related-party imports to the sum of related and non-related imports, is only available for 4,651 industry-year observations.[6] It would be more satisfactory to have a richer econometric model that jointly attempted to explain the existence of positive import volumes as well as their breakup into intrafirm and arm's-length imports, but I will not attempt to do so in this book. I do not think that this should be a huge matter of concern in regressions using industry-year level data, but I admit that it may be less

[6] As noted in Chapter 1, a very small share of the volume of imports is categorized as "nonreported." Defining the share of intrafirm imports as the ratio of related-party imports to total imports makes virtually no difference for the results presented in this chapter.

immaterial in the specifications developed below which exploit the source-country variation in the U.S. Related-Party Trade database, which contains many observations with zero trade flows.

Before diving into the econometrics, the left panel of Table 8.2 reports the ten industries with the lowest average intrafirm import share Sh_{i-f} over the period 2000-11. The right panel of this same table lists the ten industries with the highest intrafirm import share over the same period. The set of industries in the right panel generally appear to involve more complex production processes than those in the left panel, but there are important exceptions, such as the presence of "Guided missiles and space vehicles manufacturing" in the left panel or of "Asphalt shingle and coating materials manufacturing" in the right panel. Inspection of the table also raises the key concern that many of the sectors in the list appear to produce almost exclusively final goods. As in Chapter 5, we will work below to refine our sample to restrict the analysis to imports of intermediate inputs.

Table 8.2. The Ten Industries with the Lowest and Highest Intrafirm Trade Shares

10 with Lowest Intrafirm Trade Shares		10 with Highest Intrafirm Trade Shares	
.012	Guided missile & space vehicle ma.	.949	Automobile manuf.
.022	Motor home manuf.	.945	Heavy duty truck manuf.
.026	Manufactured mobile home manuf.	.854	Photographic film & chemical manuf.
.037	Rubber & plastics footwear manuf.	.844	Irradiation apparatus manuf.
.038	Other footwear manuf.	.814	Asphalt shingle/coating materials ma.
.039	Cut stone & stone product manuf.	.807	Electronic capacitor manuf.
.043	Canvas and related product mills	.805	Medicinal and botanical manuf.
.053	Infants' cut and sew apparel manuf.	.799	Other aluminum rolling and drawing
.053	Poultry processing	.797	Computer storage device manuf.
.058	Women's footwear manuf.	.794	Pharmaceutical preparation manuf.

Sources: U.S. Census Related-Party Trade Dataset, 2000-11

Leaving these caveats aside for the time being, Table 8.3 presents a set of benchmark regressions in which the share of intrafirm imports is correlated with various industry-level variables. The specifications and variables used in the estimation are almost identical to those in Table 5.2 in Chapter 5, except that the dependent variable is now the intrafirm import share rather than the offshoring share.

The first three columns of Table 8.3 focus on the role of headquarter intensity in shaping intrafirm trade shares. I begin by proxying headquarter intensity with standard measures of R&D, skill and physical capital intensity of U.S. manufacturing firms. These variables were discussed in Chapter 5 and interested readers can consult the Data Appendix for details. The use of physical capital intensity to proxy for headquarter intensity can be motivated by appealing to my own work in Antràs (2003). Remember that in that framework, I assumed that the investments provided by headquarters were more physical-capital intensive than those provided by suppliers. Furthermore, I assumed that all investments were noncontractible and fully relationship-specific and thus the model generated a positive correlation between *unobservable* headquarter intensity and observable physical capital intensity. The assumptions needed to make that connection are strong, so I will work on relaxing them below.

The first column of Table 8.3 documents a positive correlation between the three benchmark measures of headquarter intensity and the share of intrafirm trade. R&D and physical capital intensity appear to be particularly important in shaping this share. Both of these coefficients are highly statistically significant and the magnitude of the coefficients is large. The table shows beta coefficients, and thus an increase of one standard deviation in R&D intensity or physical capital intensity increases the share of intrafirm trade by 0.385 or 0.274, respectively. The effect of skill intensity is significant at the 10% confidence level, but its magnitude is much smaller. In Antràs (2014), I provide scatter plots of the partial correlations between the share of intrafirm imports and each of these measures of headquarter intensity and demonstrate that they are not driven by a few outliers.

Early papers using intrafirm trade data to shed light on the empirical determinants of multinational firm boundaries have typically interpreted correlations of the type shown in column (1) as providing support for the property-rights theory. This interpretation is explicit, for instance, in Antràs (2003), Yeaple (2006), and Nunn and Trefler (2008). There are, however, various reasons why one should be cautious in interpreting the results in that manner. First, the statistical power of these tests is low; as mentioned before, these positive correlations are consistent with the property-rights theory but they are not necessarily inconsistent with alternative theories of firm boundaries, such as the transaction-cost theory. Second, U.S. physical capital, skill, and R&D intensity measures are imperfect proxies for headquarter intensity as they only capture imperfectly the relative importance of the noncontractible, relationship-specific investments carried out by headquarters and their suppliers. Nunn and Trefler (2013*b*) point out, for instance, that standard measures of capital intensity embody several investments that are fairly easy to contract on or that are not particularly relationship-specific. If the property-rights theory is correct, one would then expect investments in specialized equipment to be much more relevant for the integration decision than investments in structures or in non-specialized equipment (such as automobiles or computers), which tend to lose little value when not used in the intended production process.

Table 8.3. Determinants of U.S. Intrafirm Trade Shares

Dep. Var. $\frac{\text{Intrafirm Imp}}{\text{Total Imports}}$	(1)	(2)	(3)	(4)	(5)	(6)
Log(R&D/Sales)	0.385**	0.361**	0.328**	0.301**	0.085**	0.337**
	(0.047)	(0.046)	(0.052)	(0.048)	(0.015)	(0.057)
Log(Skilled/Unskilled)	0.091^{+}	0.097*	0.192**	0.061	0.006	-0.146*
	(0.051)	(0.049)	(0.064)	(0.055)	(0.015)	(0.074)
Log(Capital/Labor)	0.274**					
	(0.042)					
Log(Capital Struct/Labor)		-0.256**	0.007	-0.253**	-0.060**	-0.126^{+}
		(0.076)	(0.069)	(0.078)	(0.023)	(0.074)
Log(Capital Equip/Labor)		0.529**		0.554**	0.106**	0.303**
		(0.073)		(0.076)	(0.022)	(0.082)
Log(Autos/Labor)			-0.250**			
			(0.050)			
Log(Computer/Labor)			-0.012			
			(0.049)			
Log(Other Eq./Labor)			0.290**			
			(0.066)			
Freight Costs				-0.173**	-0.104**	-0.076*
				(0.055)	(0.014)	(0.038)
Tariffs				0.007	-0.010*	-0.049
				(0.028)	(0.004)	(0.041)
Productivity Dispersion				-0.019	-0.013	-0.059
				(0.050)	(0.016)	(0.055)
Elasticity of Demand				0.036	-0.021^{+}	0.136^{+}
				(0.060)	(0.011)	(0.073)
Weighting	None	None	None	None	None	Imports
Fixed Effects	Year	Year	Year	Year	Ctr/Year	Ctr/Year
Observations	4,651	4,651	4,651	4,651	312,884	312,884
R-squared	0.312	0.343	0.344	0.369	0.170	0.585

Standard errors clustered at the industry level. $^{+}, ^{*}, ^{**}$ denote 10, 5, 1% significance.

In columns (2) and (3) of Table 8.3, I explore these ideas and confirm the empirical findings of Nunn and Trefler (2013*b*) when using disaggregated measures of capital intensity. More specifically, in column (2) of Table 8.3, I find that the positive effect of physical capital on the share of intrafirm trade is concentrated in equipment capital, while structures actually have a negative and significant effect on integration. A further decomposition using data from the Annual Survey of Manufactures (see the Data Appendix) unveils that the effect of equipment capital intensity is *not* driven by expenditures on computers and data processing equipment or on automobiles and trucks, which would be problematic for the theory. In fact, the effect of expenditures on automobiles and trucks appears to have a statistically significant *negative* effect on the share of intrafirm trade, a result which is tempting to map to the negative effect of higher headquarter service contractibility (or lower headquarter services relationship-specificity) on the integration decision predicted by both the transaction-cost and property-rights models. The scatter plots provided in Antràs (2014) confirm again that these partial correlations are not driven by a handful of outliers.

The fourth column of Table 8.3 reverts back to the specification with capital equipment being a composite category but incorporates proxies for (i) freight costs and U.S. tariffs to capture trade frictions τ; (ii) a measure of within-industry productivity dispersion $1/\kappa$; and (iii) a proxy for the elasticity of demand σ. The sources of these variables are Peter Schott's website, the World Integrated Trade Solution (WITS) database, Nunn and Trefler (2008), and Broda and Weinstein (2006), respectively, as documented in Chapter 5 and the Data Appendix.[7] The reason for including these variables in our first set of results is that they are predicted to shape the share of intrafirm imports even in the Benchmark versions of the transaction-cost and property-rights models, featuring totally incomplete contracts, full relationship specificity and bilateral contracting with a single supplier.

Of these four additional variables, only freight costs appears to have predictive power for the intrafirm trade share, but the sign of this dependence is the opposite one than the theories would predict. The selection into offshoring mechanism would tend to associate higher freight costs with fewer firms offshoring and higher intrafirm import shares, yet the coefficient on this variable is negative, highly significant and sizable in economic terms. Coupled with the negative (though statistically insignificant) effect of pro-

[7]As in the case of Chapter 5, the entire dataset and Stata program codes used in the empirical analysis in this chapter are available for download at http://scholar.harvard.edu/antras/books.

ductivity dispersion, this result casts doubt on the empirical validity of the sorting pattern underlying the models of multinational firm boundaries developed in Chapters 6 and 7. We will return to this issue towards the end of the chapter, when discussing studies using firm-level data on intrafirm versus arm's-length global sourcing decisions of firms.

Columns (5) and (6) of Table 8.3 exploit the full cross-sectoral *and* cross-country variation of the intrafirm import data. I first compute sectoral intrafirm trade shares at the exporter-country level, by computing the ratio of related-party imports to the sum of related- and non-related-party imports from a particular country j. By including source-country-year fixed effects into the regressions, I continue to exploit purely cross-product variation, but this specification better isolates the effect of sectoral-level characteristics by controlling for unobservable country characteristics that might shape both the types of products the U.S. imports from those countries as well as whether those transactions are internalized or not. Because at the country-industry level there are many more observations with zero import volumes than in the purely cross-sectoral data, the number of observations in columns (5) and (6) falls very short of the potential 1,085,760 observations corresponding to 390 sectors, 232 countries and 12 years of data. In particular, only 312,884 of those observations feature positive imports for the sum of related and non-related import values. This number is extremely close to the number of observations in column (6) of Table 5.2 in which we restricted the sample of offshoring shares to those featuring a positive value, with the small discrepancy being explained by a few observations in which only import flows with non-reported relatedness are positive.

The only difference between columns (5) and (6) is that in the latter column I follow Antràs and Chor (2013) in weighting each data point by the value of total imports for that industry-country-year. This is motivated by possible measurement error introduced into the intrafirm trade share by the presence of trade flows whose related-party status was not reported to the U.S. Census Bureau, an issue of particular concern for observations with small trade volumes. Indeed, the raw correlation between the share of "unreported" trade and the log of total imports is negative and very large (-0.52). Notice also that the weighted regression features a substantially higher R-squared (0.585) than the unweighted one (0.170).

The qualitative nature of the results in columns (5) and (6) is similar to that in the regressions with the aggregated cross-industry data. High levels of R&D and equipment capital intensity continue to be associated with significantly higher intrafirm trade shares, though the effects are quantitatively smaller when not weighting the observations in column (5). The troublesome negative effect of freight costs on the intrafirm trade share also appears to be robust to the use of the country-level related-party information, though its statistical significance is greatly reduced by weighting observations by import volumes. Furthermore, in column (5) I now also find that U.S. tariffs have a statistically negative effect on intrafirm trade shares, though this effect is no longer significant in the weighted regression in column (6).[8] Finally, the effect of skill intensity appears significantly negative in column (6), suggesting perhaps that this variable is not an appropriate proxy for headquarter intensity.

Cross-Industry Tests: Refined Benchmark Tests

As pointed out in Chapter 5 and again earlier in this chapter, there exist at least six serious limitations associated with using U.S. import data to construct a measure of the relative propensity to integrate foreign suppliers. Earlier, I elaborated on the first, fifth and sixth concerns, so I can now focus on the second, third and fourth ones. Fortunately, there is a close parallel in the way that I address this concerns here and how I dealt with them in Chapter 5, so I can swiftly work through them.

Remember that our second concern with product-level U.S. import data is that they do not identify the industry or sector purchasing the imported goods. Consistently with the bulk of the literature, the industry-level controls in Table 8.2 corresponded to data on the industry of the product being imported. This seems justified when studying the effect of freight costs and tariffs, but it is clearly invalid when exploring the role of the final-good producer's elasticity of demand σ. Furthermore, using this approach when constructing measures of headquarter intensity is only consistent with models of multinational firm boundaries under restrictive assumptions.[9] As argued in Chapter 5, a more satisfactory approach is to construct measures of headquarter intensity and demand elasticities of the average industry *buying* those inputs using information from Input-Output tables, as first proposed by Antràs and Chor (2013). Although one could similarly advocate the con-

[8] It should be noted that this result contrasts with that obtained by Díez (2014), who using similar data instead finds a positive association between the prevalence of intrafirm trade and U.S. tariffs. He also finds a negative correlation between U.S. intrafirm imports and foreign tariffs and shows that it can be reconciled with a variant of the Antràs and Helpman (2004) framework.

[9] For the case of capital intensity, it can be justified in Antràs' (2003) framework due to the unrealistic assumption that factors of production are internationally immobile so the headquarter's capital investments are undertaken in the location of the supplier division or firm and embodied in the imported good.

struction of a buyer version of our productivity dispersion measure, I have argued in Chapter 5 that this methodology cannot be suitably applied to measures of dispersion.

As in Chapter 5, building buyer versions of some variables leads me to switch from the NAICS six-digit industry classification (at which the raw data are reported) to 2002 Input-Output industry codes (IO2002). The number of sectors in the sample is thus reduced to 253 for a total of 3,036 observations. Column (1) of Table 8.4 reports results analogous to those in column (4) of Table 8.3 but with the IO2002 classification instead of NAICS classification. The change in industry classification leads to relatively small changes in the qualitative and even quantitative nature of the results. Relative to column (4) in Table 8.3, skill intensity now has a positive and significant effect on the intrafirm import share, while the effect of capital structures now appears to be positive (when before it was negative). Still, as the next few specifications will demonstrate, these two effects are not robust to further refinements of the sample. A more robust result that is already visible in column (1) of Table 8.4 is the fact that not only freight costs but also U.S. tariffs now appear to be significantly negatively correlated with the intrafirm trade share.

Buyer versions of the elasticity of demand and of the proxies for headquarter intensity are introduced in column (2) of Table 8.4. There are three main consequences of this change in these independent variables. First, the positive effects of skill intensity and capital structures disappears, consistently with our NAICS results in Table 8.3. Second, the magnitude of the positive coefficients on R&D and capital equipment intensity increases markedly. Third, the positive effects of productivity dispersion and of the elasticity of demand become significant at standard confidence levels. Both of these effects are consistent with the predictions of both the transaction-cost and property-rights models, though as mentioned before, these theories do not ensure an unambiguously positive effect of the elasticity of demand.

In column (3) of Table 8.4 I tackle what I earlier identified to be the third limitation of U.S. product-level import data, namely the fact that it conflates intermediate input and finished goods imports. Although the models developed above are not inconsistent with headquarters importing fully assembled goods from abroad, it is important to attempt to purge finished products out the data for at least two reasons. First, the models developed in this book emphasize input transactions and thus, at the very least, one should check that the results continue to hold when focusing on those type of transactions. Second, I would guess that a significant share of finished goods entering the U.S. are imported by wholesalers and retailers, and these types of firms are not represented in the industry-level manufacturing database that I am using to construct the buyer versions of headquarter intensity and the elasticity of demand.

Table 8.4. Refined Determinants of U.S. Intrafirm Trade Shares

Dep. Var. $\frac{\text{Intrafirm Imp}}{\text{Total Imports}}$	(1)	(2)	(3)	(4)	(5)	(6)
Log(R&D/Sales)	0.164**	0.222**	0.240**	0.251**	0.052**	0.246**
	(0.058)	(0.064)	(0.072)	(0.072)	(0.017)	(0.068)
Log(Skilled/Unskilled)	0.174*	0.009	0.036	0.025	-0.031	-0.182
	(0.072)	(0.081)	(0.082)	(0.082)	(0.023)	(0.113)
Log(Capital Struct/Labor)	0.199**	-0.105	-0.027	-0.031	-0.013	-0.032
	(0.066)	(0.105)	(0.121)	(0.121)	(0.038)	(0.089)
Log(Capital Equip/Labor)	0.144**	0.392**	0.232*	0.235*	0.071*	0.149^{+}
	(0.046)	(0.099)	(0.117)	(0.118)	(0.032)	(0.077)
Seller Freight Costs	-0.231**	-0.221**	-0.254**	-0.240**	-0.131**	-0.081
	(0.069)	(0.075)	(0.089)	(0.087)	(0.020)	(0.068)
Seller Tariffs	-0.076*	-0.070**	-0.104**	-0.102**	-0.022**	-0.079^{+}
	(0.031)	(0.025)	(0.021)	(0.021)	(0.006)	(0.044)
Seller Dispersion	0.039	0.120^{+}	0.043	0.046	0.035^{+}	0.060
	(0.077)	(0.073)	(0.081)	(0.082)	(0.018)	(0.038)
Elasticity of Demand	0.105	0.163*	0.186*	0.184*	-0.011	0.085**
	(0.078)	(0.065)	(0.080)	(0.081)	(0.011)	(0.025)
Sample Restrictions	None	None	W	W+NT	W+NT	W+NT
Weighting	None	None	None	None	None	Imports
Fixed Effects	Year	Year	Year	Year	Ctr/Year	Ctr/Year
Buyer vs. Seller Controls	Seller	Buyer	Buyer	Buyer	Buyer	Buyer
Observations	3,036	3,036	2,480	2,478	148,947	148,947
R-squared	0.348	0.359	0.322	0.313	0.194	0.526

Standard errors clustered at the industry level. $^{+}$, *, ** denote 10, 5, 1% significance.

As in Chapter 5, I adopt the methodology developed by Wright (2014) in order to attempt to isolate intrafirm and arm's-length imports of intermediate inputs. This methodology was briefly discussed in Chapter 5 and it is reviewed in detail in the Data Appendix, so I will not elaborate on it here. I will simply note that this correction lead us to drop 39 industries that exclusively produce final goods, but it also modifies different sectors differentially because the discount factor applied to the data is constructed starting with highly disaggregated (i.e., HS ten-digit) product and country-level import data. The intrafirm import share will be reduced in sectors in which, relative to arm's-length imports, related-party imports originate from countries that tend to export finished goods to the U.S., as deduced from the disaggregated product-level data. In practice, however, these smooth adjustments are small and the correlation between the Wright-adjusted intrafirm trade shares and the raw ones is very high (0.919). Hence, the largest effect of the Wright adjustment on the intrafirm trade shares used in the regressions stems from dropping the 39 industries from the sample.[10] Comparing columns (2) and (3) of Table 8.4, it is clear, however, that this sample correction only has a minor effect on the estimates, with the main result being that the effect of productivity ceases to be statistically significant at standard confidence levels.

The fourth concern we raised regarding the use of the U.S. Related-Party database was that it did not distinguish between trade within U.S. multinationals (which might map better to backwards integration) and trade within foreign multinationals operating in the U.S. (which perhaps better reflects forward integration). With that in mind, in column (4) of Table 8.4, I follow Nunn and Trefler (2013*b*) in checking the robustness of the results to a restricted sample that better fits the spirit of our global sourcing model. In particular, I drop from the sample those U.S. imports originating from five countries (Iceland, Italy, Finland, Liechtenstein, and Switzerland) for which shipments from foreign headquarters to their U.S. affiliates are likely to be predominant, relative to shipments to U.S. parents from their foreign affiliates in those countries. More details on how these countries are identified and on robustness to alternative sets of dropped countries are available in Chapter 5 and in the Data Appendix. The results of implementing this sample restriction are shown in column (4) of Table 8.4, and it is evident that the impact on the estimates is very modest.

Although the Wright and Nunn-Trefler sample corrections have not produced a sizeable impact on the estimated coefficients, I still view these adjustments worthwhile performing given the nature of the models that are being taken to the data. Furthermore, in order to illustrate their significance, in Table 8.5 I report the ten sectors with the lowest and highest Wright- and Nunn-Trefler-corrected intrafirm import shares averaged over the period 2000-11. Comparing these rankings with those in Table 8.2, it is encouraging to see that some of the problematic sectors in Table 5.1 (such as "Guided missiles and space vehicles manufacturing") are no longer listed. It is also

[10] The number of observations in column (3) drops by more than $39 \times 12 = 468$ because for 88 additional industry-year observations, the Wright adjustments sets *total* intermediate input imports to zero.

interesting to note that these adjustments have a nontrivial effect on certain intrafirm import shares, such as the case of the "Automobile manufacturing" sector, whose intrafirm import share drops from 0.949 to 0.800.

Table 8.5. The Ten Input Industries with the Lowest and Highest *Corrected* Intrafirm Trade Shares

10 with Lowest Intrafirm Trade Shares		10 with Highest Intrafirm Trade Shares	
.019	Footwear manufacturing	.901	Heavy duty truck manuf.
.047	Cut stone and stone product manuf.	.845	Irradiation apparatus manuf.
.058	Other leather manufacturing	.832	Asphalt shingle/coating materials manuf.
.063	Primary smelting/refining copper ma.	.809	Lighting fixture manufacturing
.066	Institutional furniture manuf.	.800	Automobile manufacturing
.078	Prefabricated wood building manuf.	.797	Computer storage device manuf.
.080	Fiber, yarn, and thread mills manuf.	.765	Pharmaceutical preparation manuf.
.084	Household & institutional furniture man.	.750	Electro-medical & -therapeutic appl. ma.
.094	Seafood product preparation	.737	Travel trailer and camper manuf.
.112	Paper bag & treated paper manuf.	.734	Tire manufacturing

Sources: U.S. Census Related-Party Dataset Same plus a sample adjustment based on Wright (2014)

In the last two columns of Table 8.4, I exploit the full variation of the intrafirm trade data across both products and countries, while applying the Wright and Nunn-Trefler corrections to trade flows. The specifications include country-year fixed effects, so the variation being exploited is again cross-sectoral, but I am now controlling for time-varying unobserved country characteristics. Columns (5) and (6) only differ in that, in the latter specification, I weight observation by the total volume of U.S. imports to alleviate measurement error concerns. This last set of estimates in column (6) is broadly consistent, both qualitatively as well as quantitatively, with the aggregated cross-industry specification in column (4). An important difference is that the magnitude of the negative effect of the two trade costs variables is greatly reduced, and the coefficient on freight costs, in particular, is no longer significant. Similarly, the effect of productivity dispersion is positive but is only significant at the 12% level.

Cross-Industry Tests: The Role of Contracting

So far, I have focused on an empirical analysis of the predictions of the 'Benchmark' transaction-cost and property-rights models developed in Chapters 6 and 7. The robust positive effect of R&D and equipment capital intensity we have documented is often interpreted as providing support for the property-rights model, since η is a key determinant of the optimal allocation of ownership rights in that model. Still, when working with product-level data that aggregates the individual decisions of firms, the share of intrafirm imports is also shaped by the selection into offshoring effect, which in both types of models tends to generate a positive correlation between intrafirm import shares and measures of headquarter intensity.

I will next explore more elaborate tests that exploit some of the novel predictions that emerged when studying the various extensions of these benchmark models. More specifically, in the next two tables, I will build on the insights in Chapters 6 and 7 – and summarized in Table 8.1 –, and incorporate into the specifications in Table 8.4 proxies for financial constraints (ϕ), contractibility (μ_{hS}, μ_{mS}), relationship-specificity (ϵ_h, ϵ_m), input substitutability ρ, and downstreamness.

As a first step, in Table 8.6, I present results of specifications analogous to those in Table 8.4, but that include eight new regressors: two proxies for the importance of financial constraints in a given sector, four related to the degree of contractibility, one measure capturing the relationship-specificity of investments, and a final one related to the degree of input substitutability in production. These variables are the same ones included in Table 5.6 of Chapter 5, so I refer the reader to that chapter (and to the Data Appendix) for a discussion of the underlying sources. All of these variables are constructed based solely on information of the product being imported into the U.S., so the proxies for contractibility and relationship-specificity are more closely related to the parameters μ_{mS} and ϵ_m, respectively, than to μ_{hS} and ϵ_h.

In the first three columns of Table 8.6, I report the results of introducing these eight variables *one at a time* into the Wright- and Nunn-Trefler-corrected regressions in columns (4), (5) and (6) of Table 8.4. Although eight coefficients appear on each column, it should be understood that these coefficients are obtained by running eight separate regressions. To save space, I do not report the coefficients on the variables already included in Table 8.4, but these coefficients are only modestly affected by the inclusion of these new eight variables.[11]

These first three columns of the table provide broad support for the notion that larger financial frictions (i.e., higher financial dependence or lower asset tangibility) are associated with higher intrafirm trade shares, with the size and statistical significance of these results being particularly high when ex-

[11] The whole set of regression coefficients can be obtained by accessing the data and programs available online at http://scholar.harvard.edu/antras/books.

ploiting the cross-country dimension of the data while weighting observations by total import volumes. Similarly, all four measures of product contractibility are negatively associated with the extent to which foreign input purchases are internalized, and again the magnitude and statistical significance of these effects is highest when introducing these measures into our preferred weighted specification with country-industry-year data. Finally, the evidence points towards a positive effect of specificity and a negative effect of input substitutability on intrafirm import shares, though these coefficients are generally insignificant except for the case of our preferred specification in column (3).

Table 8.6. Contractual Determinants of U.S. Intrafirm Trade Shares

Dep. Var. $\frac{IntrafirmImp}{TotalImports}$	(1)	(2)	(3)	(4)	(5)	(6)
Financial Dependence	0.186*	0.028	0.206**	0.182*	0.029	0.196**
	(0.087)	(0.019)	(0.045)	(0.088)	(0.019)	(0.041)
Asset Tangibility	-0.124	-0.015	-0.256**			
	(0.078)	(0.019)	(0.062)			
Nunn Contractibility	-0.084	-0.012	-0.166*	-0.073	0.000	-0.121+
	(0.070)	(0.019)	(0.070)	(0.076)	(0.021)	(0.073)
Levchenko Contractibility	-0.124+	-0.054**	-0.176**			
	(0.073)	(0.019)	(0.055)			
Costinot Contractibility	-0.131+	-0.001	-0.131*			
	(0.071)	(0.018)	(0.063)			
BJRS Contractibility	-0.191*	-0.056**	-0.085+			
	(0.078)	(0.021)	(0.046)			
Specificity	0.044	0.020	0.180*	0.006	0.017	0.055
	(0.070)	(0.019)	(0.074)	(0.074)	(0.021)	(0.067)
Input Substitutability	-0.014	-0.016	-0.078+	-0.000	-0.014	-0.014
	(0.042)	(0.017)	(0.047)	(0.043)	(0.017)	(0.028)
Sample Restrictions	W+NT	W+NT	W+NT	W+NT	W+NT	W+NT
Fixed Effects	Year	Ctr/Year	Ctr/Year	Year	Ctr/Year	Ctr/Year
Weighting	None	None	Imports	None	None	Imports
Observations	2,478	148,947	148,947	2,478	148,947	148,947
R-squared	$\simeq$ 0.322	$\simeq$ 0.194	$\simeq$ 0.548	0.336	0.195	0.582

Standard errors clustered at the industry level. $^{+}$, *, ** denote 10, 5, 1% significance.

In the last three columns of Table 8.6, I run regressions analogous to those in columns (1), (2) and (3), but in which a single proxy for financial constraints, a single proxy for contractibility, and the proxies for specificity and input substitutability are all included *in the same* specification. As in Table 5.6 of Chapter 5, in those columns I select Rajan and Zingales' financial dependence measure and Nunn's measure of contractibility because they are particularly popular in the literature. As is clear from the table, the simultaneous inclusion of these variables reduces the impact of each of them individually, but the sign of these effects is the same as in columns (1), (2), and (3), and the effects of financial dependence and Nunn contractibility continue to be statistically significant in the final preferred specification in column (6).

If one refers back to Table 8.1, one will quickly verify that the signs of the coefficients in Table 8.6 exactly correspond to the predictions of the transaction-cost model. We cannot, however, invoke these results to discard the property-rights model because, if the selection into offshoring effect is powerful enough, these same patterns could be generated by that model (see Table 8.1).

Discriminating Between Models

Is there then any hope in discriminating between the property-rights and transaction-cost theories of multinational firm boundaries? In Table 8.7, I experiment with two alternative ways to do so. First, I exploit the rich but diametrically opposite implications of both models for the effect of downstreamness on the integration decision. As readers may recall, we showed in Chapter 6 that, in the transaction-cost model, downstreamness had a negative effect on integration whenever inputs are sequential complements, while it had a positive effect on integration when inputs are sequential substitutes. Instead, in the property-rights model of Chapter 7, the opposite is true: downstreamness has a positive effect on integration in the sequential complements case, while it has a negative effect on integration in the sequential substitutes case.

Which of these two predictions is most consistent with available data? In order to answer this question, one needs to first take a stance on (i) how to measure downstreamness, and (ii) how to proxy for whether the integration decision corresponds to the sequential complements or sequential substitutes case. In the first two columns of Table 8.7, I make progress on these fronts building on the approach in Antràs and Chor (2013). First, they define the downstreamness of the product being imported into the U.S. as a weighted index of the average position in the value chain at which an industry's output is used (i.e., as final consumption, as direct input to other industries, as direct input to industries serving as direct inputs to other industries, and so on), with the weights being given by the ratio of the use of that industry's output in that position relative to the total output of that industry.[12] Second, in order to distinguish between the cases of sequential complements and substitutes, they use the U.S. import demand elasticities estimated by Broda and Weinstein (2006) and data on U.S. Input-Output Tables to compute a weighted average of the demand elasticity faced by the *average* buyer of the product being imported into the U.S. The idea is that for sufficiently high values of this average demand elasticity, one can be relatively confident that input substitutability is lower than the demand elasticity, with the converse being true for sufficiently low values of this average demand elasticity.

In column (1) of Table 8.7, I include two interactions of downstreamness with dummy variables for high and low σ sectors into the cross-industry specification of column (4) of Table 8.6. The 'High-σ' dummy variable takes a value of 1 if the average buying industry of the imported product sector features a Broda-Weinstein demand elasticity above the median one in the sample, while the 'Low-σ' dummy variable takes a value of 1 when a product's average buyer demand elasticity is below the median one. In column (2) of Table 8.7, I add these same interactions to the weighted specification with cross-country and cross-industry specification in column (6) of Table 8.6.[13] As is clear from the results in these two tables, there is robust evidence of a differentially more positive effect of downstreamness on integration in high-σ sectors (i.e., in the complements case) than in low-σ sectors. This is consistent with the property-rights model, but inconsistent with the transaction-cost model. In fact, in column (1), the signs of these coefficients are exactly as predicted by the property-rights model and opposite to those implied by the transaction-cost model, though the coefficient on the second interaction is not statistically different from zero.

Antràs and Chor (2013) show that this differential positive effect of downstreamness on integration in high-σ sectors is robust to alternative measures of downstreamness and various specifications. Furthermore, when looking at the effect for different quintiles of the distribution of σ, the positive effect is consistently concentrated in the highest quintiles of σ, while the effect is often negative in the lowest quintiles of σ. The differential effect is also apparent even without controlling for other factors, as illustrated in Figure 8.1. In the figure, for the subset of industries with above-median average buyer demand

[12]This measure was developed independently by Fally (2012), and its properties were further studied in Antràs, Chor, Fally and Hillberry (2012). More details on this measure are available in the Data Appendix.

[13]In both columns, I also add a dummy variable for whether the average buying industry σ is above the median one, to be able to better interpret the interaction coefficients.

elasticities (labeled as "Complements"), the average 2005 U.S. intrafirm import share increases as we move from the lowest tercile of $DownMeasure$ to the highest. When considering those industries facing below-median average buyer demand elasticities (labeled as "Substitutes"), the pattern is exactly reversed, with the intrafirm import share steadily declining across terciles of downstreamness. These patterns exactly line up with those predicted by the property-rights model of global sourcing, while they contradict those predicted by the transaction-cost model.

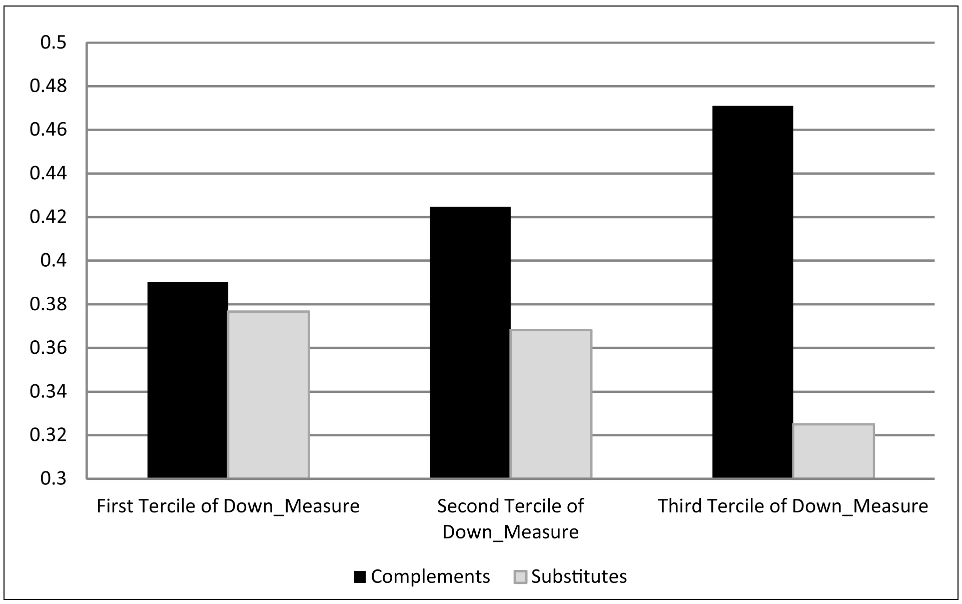

Figure 8.1: Downstreamness and the Share of Intrafirm Trade

A second promising way to discriminate between the property-rights theory and the transaction-cost one consists of exploiting the implications of these theories for the effect of contractibility and relationship-specificity on the share of intrafirm trade. Remember, in particular, that in the property-rights model, the effect of these variables on the prevalence of integration depends crucially on the extent to which contractual incompleteness stems from noncontractibilities or specificities in the inputs controlled by the final-good producer or by his or her suppliers. If production processes in certain sectors are particularly noncontractible or feature high specificities because of the nature of the investments carried out by headquarters, then the theory would predict that the share of intrafirm trade should be negatively affected by the level of these sectors' contractibility and positively affected by specificity. Conversely, if the source of noncontractibilities or specificity stem from the nature of the supplier's activities, the theory may instead predict a positive correlation between the share of intrafirm trade and contractibility and a negative correlation with specificity.[14] The latter results would be hard to reconcile with transaction-cost theories of multinational firm boundaries.

Table 8.7. Further Contractual Determinants of U.S. Intrafirm Trade Shares

Dep. Var. $\frac{IntrafirmImp}{TotalImports}$	(1)	(2)	(3)	(4)	(5)	(6)
Downstreamness x High σ	0.291^{+}	0.330**	0.296^{+}	0.344**	0.291*	0.321**
	(0.150)	(0.060)	(0.150)	(0.058)	(0.148)	(0.052)
Downstreamness x Low σ	-0.159	0.099	-0.155	0.100	-0.165	0.040
	(0.138)	(0.078)	(0.139)	(0.077)	(0.137)	(0.074)
Seller Nunn Contractibility	-0.059	-0.026	-0.027	0.138	-0.046	0.033
	(0.068)	(0.057)	(0.092)	(0.085)	(0.070)	(0.053)
Buyer Nunn Contractibility			-0.051	-0.185*		
			(0.096)	(0.075)		
Seller Nunn Specificity	-0.015	-0.011	-0.028	-0.038	-0.090	-0.176**
	(0.078)	(0.061)	(0.083)	(0.064)	(0.092)	(0.068)
Buyer Nunn Specificity					0.124	0.284**
					(0.116)	(0.060)
Sample Restrictions	W+NT	W+NT	W+NT	W+NT	W+NT	W+NT
Fixed Effects	Year	Ctr/Year	Year	Ctr/Year	Year	Ctr/Year
Weighting	None	Imports	None	Imports	None	Imports
Observations	2,478	148,947	2,478	148,947	2,478	148,947
R-squared	0.357	0.614	0.358	0.620	0.362	0.632

Standard errors clustered at the industry level. $^{+}$, *, ** denote 10, 5, 1% significance.

Although the property-rights theory generates sharp predictions for how the source of noncontractibilities or specificity affects the share of intrafirm trade, a natural challenge for empirical work is finding appropriate proxies for these different types of noncontractibilities and specificity. In the last four columns of Table 8.7, I experiment with a simple approach to attempt to separate those effects. In particular, I argued above that because the Nunn measure of contractibility is based solely on the product being imported, it seems natural to relate it to the parameters μ_{mS} in the model. In columns

[14] The qualifier "may" in the previous sentence is necessary because via the selection into offshoring mechanism, improvements in manufacturing input contractibility may reduce the share of intrafirm trade on that account.

(3) and (4), I add a measure of the contractibility of the average sector buying the good entering the U.S. (as inferred from Input-Output tables) to the specifications in columns (1) and (2), respectively. I interpret that average buyer contractibility as reflecting the contractibility of headquarter services, i.e., the parameter μ_{hS}. Adding this variable to the cross-industry specification in column (1) of Table 8.7, has little effect on the estimates, but when doing the same to our preferred weighted specification in column (2), notice that the sign of the coefficient on seller contractibility (i.e., μ_{mS}) becomes positive and is very close to being significant at the 10% level, while the sign of the coefficient on buyer contractibility (i.e., μ_{hS}) is negative and significant at the 5% level. These patterns are precisely those predicted by the property-rights model.

In columns (5) and (6), I repeat the same exercise, but this time focusing on our measure of specificity. The results in this case are even more supportive of the property-rights model. Buyer and seller specificity shape the intrafirm trade share in opposite directions in both specifications. Buyer specificity is positively associated with integration, but the opposite is true for the case of seller specificity, a result that is not easily reconcilable with transaction-cost models, but that is predicted by the property-rights theory (at least under a wide range of parameter values). Furthermore, in our preferred weighted regression in column (6), both coefficients are highly statistically significant.

Limitations and Alternative Approaches

The empirical results presented so far in this chapter are broadly supportive of some of the key features of the internalization models presented in Chapters 6 and 7. The property-rights model, in particular, seems to fare especially well in the data. We have found robust evidence of a positive effect of headquarter intensity on the share of intrafirm trade, particularly when attempting to isolate the relative intensity of the noncontractible and relationship-specific investments carried out by suppliers. Furthermore, the evidence also points to a positive effect of productivity dispersion and financial constraints on the integration decision, and to a negative effect of input substitutability on this share, although the statistical significance of these effects has generally been found to be weaker. Finally, the differential effect of downstreamness on high- versus low-σ sectors, as well as the contrasting role of buyer versus seller measures of contractibility and relationship-specificity on the integration decision have provided sharper evidence permitting a discrimination between the property-rights and transaction-cost theories.

As encouraging as these results might appear to be for the property-rights theory, there are various reasons for taking them with a grain of salt. At a broad level, it is clear that the independent variables used in the empirical analysis are only imperfect proxies for the key primitive parameters shaping the internalization decisions of firms in the models. For instance, the results in columns (4) and (6) are highly suggestive of the opposite effects of μ_{hS} and μ_{mS} and of ϵ_h and ϵ_m on the integration decision, but they fall short of tightly identifying those effects. A specific reason to treat these results on contractibility with caution is that the buyer and seller versions of the contractibility and specificity variables are highly correlated with each other; the correlation between the buyer and seller (Nunn) contractibility and specificity measures are 0.834 and 0.814, respectively.[15]

A second reason to remain skeptical about the empirical validity of the property-rights model is the fact that we have found fairly robust evidence of some effects that run counter to some of the key predictions of the model, at least in their benchmark versions. Most notably, freight costs and U.S. tariffs appear to be negatively correlated with intrafirm import shares, with these effects being statistically significant at standard confidence levels, except in our weighted specifications exploiting both the cross-country and cross-industry data (consistently with the results reported in Table 8.4). Remember that this negative effect of trade barriers is inconsistent with the predictions of both models, at least under the derived equilibrium sorting pattern of firms by productivity into organizational forms. A key question, however, is whether this equilibrium sorting pattern is consistent with available firm-level evidence. I will return to this issue below.

A third concern with the tests performed so far is that, even if one is persuaded with the way I have proxied for the key parameters in the models, it would be hard to claim that these test convincingly *identify* the role of these parameters in shaping the internalization decisions of firms. One might worry, for instance, that I have omitted certain sectoral characteristics that are crucial for integration and that may be correlated with the industry variables included in the regressions above. Readers might recall that, in Chapter 5, I alluded to a similar reasoning when justifying the poor performance of contractual variables in explaining the cross-sectional variation in offshoring shares. This in turn led me to explore alternative approaches that exploited the idea that industry characteristics should have a differential effect on the propensity of firms to offshore from particular countries,

[15] This is due in turn to the disproportionate weight of the diagonal (within-industry) elements in the Input-Output tables.

depending on characteristics of these countries.

I will next apply a similar strategy to regressions explaining the intrafirm trade share, but before doing so it is important to emphasize two points. First, although I remain worried about omitted variable biases, I believe that this concern is much lessened in regressions explaining the intrafirm import share relative to regressions explaining offshoring shares. The reason for this is that it is much easier to envision omitted factors that shape differentially domestic versus foreign input purchases than omitted factors that are relevant for the relative propensity to import inputs from abroad within or across firm boundaries. To substantiate that claim, simply notice that the R-squared obtained in the cross-industry regressions in this Table 8.6 or 8.7 are more than twice as large than those obtained in the analogous Table 5.6 in Chapter 5. Similarly, the R-squared of the weighted regressions using cross-industry and cross-country data are very large (reaching 0.632 in column (6) of Table 8.7) and are more than three times as large as the highest R-squared in the offshoring share regressions in Table 5.6.

A second point worth highlighting is that, as I will soon overview, the property-rights model of Chapter 7 does not offer particularly sharp predictions for how the interaction of country and industry characteristics should shape intrafirm trade shares. Hence, although exploring alternative strategies is worthwhile, it is less clear in the current context that they will be as useful as they were in Chapter 5 in empirically validating or rejecting the theoretical models developed in this book.

As in the case of Chapter 5, before turning to these richer tests, I briefly describe how to extend the transaction-cost and property-rights models to a multi-country environment in order to provide a semi-structural interpretation of the tests to be performed below.

Internalization Theories in a Multi-Country World

Let us then return to the multi-country framework first introduced toward the end of Chapter 2 and further expanded to include contractual frictions in Chapter 5. As in the analysis in Chapter 5, I simplify matters by focusing on a version of the model in which each final-good producer procures only one input (as in the two-country model) and in which the firm-level extensive margin of offshoring is not operative.

Remember that in our multi-country global sourcing model, final-good producers learned the productivity with which they could source inputs from any given country $j \in J$ only after paying the fixed cost f_{ij} of sourcing from that country j. I then defined the global sourcing strategy $\mathcal{J}_i(\varphi) \subseteq J$

of a firm with productivity φ as the set of countries in which a firm from country i with productivity φ had paid the associated fixed costs of offshoring f_{ij}. A simple way to extend the framework to incorporate a choice between intrafirm and arm's-length sourcing is to redefine the global sourcing strategy as choosing a set $\tilde{\mathcal{J}}_i(\varphi) \subseteq J \times K$, where $K = \{V, O\}$ is an indicator function capturing whether input provision is vertically integrated (V) or outsourced (O). In plain words, the firm not only decides whether or not to invest in being able to source from any country $j \in J$, but also chooses whether its intermediate input purchases should come from an integrated affiliate in j, from an arm's-length supplier in j, or from both of them. In a general version of the model, the fixed costs associated with these different options would be allowed to vary and a natural counterpart of our assumption in the two-country model would be to assume that $f_{ijV} > f_{ijO}$. This would lead relatively larger firms to be more likely to select into intrafirm trade, and might also explain (when allowing for multiple inputs) why firms often buy inputs from foreign countries both within and across firm boundaries.

As mentioned before, however, I will simplify matters by shutting down these selection effects and assume that the fixed costs of sourcing f_{ijV} and f_{ijO} are small enough such that all firms from i find it profitable to incur these costs and draw a productivity parameter $1/a_{mjk}$ from each country $j \in J$ and for each organizational form $k \in \{V, O\}$. I will further assume that although the values of $1/a_{mjV}$ and $1/a_{mjO}$ are firm-specific, they are drawn *independently* from each other (and also independently from the draws in other countries) from a Fréchet distribution:

$$Pr(a_{mj} \geq a) = e^{-T_j a^\theta}, \quad \text{with } T_j > 0 \text{ and } \theta > \sigma - 1.$$

These are obviously strong and unrealistic assumptions, but attempting to relax them here would lead me too far astray.

Given this setup, it is then a simple matter to follow closely the steps in Chapter 5 to verify that the share of all intermediate inputs purchased by firms in i that originate from country $j \in J$ and are transacted under the ownership structure $k \in \{V, O\}$ is given by

$$\chi_{ijk} = \frac{T_j \left(\tau_{ij} w_j \Psi_{ijk}^{1/(1-\eta)(1-\sigma)}\right)^{-\theta}}{\sum_{l \in J} \sum_{k' \in \{V,O\}} T_l \left(\tau_{il} w_l \Psi_{ilk'}^{1/(1-\eta)(1-\sigma)}\right)^{-\theta}} \text{ for } j \in J \text{ and } k \in \{V, O\}$$

where Ψ_{ijk} summarizes the transaction-cost efficiency of sourcing from j under organizational form k. The value of Ψ_{ijk} under outsourcing, i.e., Ψ_{ijO}, captures the contractual efficiency with which firms from i can outsource from country j. In our simpler two-country model, this corresponds to the terms Γ_{DO} and Γ_{OO} for domestic and offshore outsourcing in both the property-rights and transaction-cost models. On the other hand, the determination of Ψ_{ijV} is distinct in the two models. In the transaction-cost model, we simply have $\Psi_{ijV} = \lambda^{1-\sigma}$, where $\lambda > 1$ captures the governance costs of running an integrated structure. This parameter λ could easily be allowed to vary across country-pairs. In the property-rights model, Ψ_{ijV} corresponds to the contractual efficiency of vertical integration, which in the two-country model we denoted by Γ_{DV} and Γ_{OV} in domestic and offshore sourcing relationships.

With our additional recurring assumption that, regardless of ownership structure, input purchases are priced such that they constitute the same multiple of operating profits in all countries and under all organizational forms, we can then conclude that when looking at the intermediate input imports from any country j, the share of those imports that is transacted within firm boundaries is simply given by

$$Sh_{i-f} = \frac{(\Psi_{ijV}/\Psi_{ijO})^{\theta/(1-\eta)(\sigma-1)}}{1 + (\Psi_{ijV}/\Psi_{ijO})^{\theta/(1-\eta)(\sigma-1)}}. \tag{8.4}$$

Hence, the share of intrafirm imports at the origin-country-level is crucially shaped by the ratio Ψ_{ijV}/Ψ_{ijO}, which was thoroughly studied in Chapters 6 and 7. Obviously, equation (8.4) is much simpler than our earlier equations (8.1), (8.2) and (8.3). Notice, in particular, that it is not a function of wage rates or trade frictions, and that it only depends on two indices of efficiency corresponding to integration and outsourcing in a particular sourcing country j, rather than on indices in multiples countries, including the home one. Naturally, the reason for this simplicity is that we have ignored the extensive margin of firms and this, of course, is not a virtue of this variant of the model, but rather a limitation. For this reason, in some of the specifications below, I will include some variables (in particular, freight costs and tariffs) that would almost surely affect the intrafirm import share if we allowed firms to face a nontrivial choice regarding whether or not to invest in being able to source inputs from a given country within or across firm boundaries (or both).

Empirical Implementation of the Multi-Country Model

Leaving aside these important caveats, I will next leverage the extremely simple form of (8.4) to motivate an empirical specification that studies the effect of interactions of industry and country characteristics on the propensity of firms to engage in intrafirm trade. With that in mind, I first note that if instead of focusing on the intrafirm import share Sh_{i-f}, I instead compute the ratio of intrafirm imports to arm's-length imports, equation (8.4) becomes log-linear. Reintroducing input subscripts v and taking logs, we can then express (8.4) as:

$$\ln\left(M_{ijv}^{if}/M_{ijv}^{nif}\right) = \frac{\theta}{(1-\eta_v)(\sigma_v - 1)} \ln\left(\Psi_{ijVv}/\Psi_{ijOv}\right) + \varepsilon_{ijv}, \tag{8.5}$$

where ε_{ijv} is an error term assumed to satisfy all the necessary orthogonality conditions. Obviously the ratio Ψ_{ijVv}/Ψ_{ijOv} is not something we observe in the data, but one can follow the guidance of the models developed in Chapters 6 and 7 – as summarized in Table 8.1 – to write this ratio as a function of empirical proxies of the deep parameters of those models.

A tricky issue in making that mapping is deciding whether certain parameters of the model are (i) industry-specific but common across countries, (ii) country-specific but common across industries, or (iii) industry- and country-specific. In Chapter 5, I argued that on conceptual grounds and also due to data limitations, it was natural to treat the elasticity of demand σ, headquarter intensity η, relationship-specificity ϵ, and input substitutability ρ as deep industry parameters unaffected by the particular country from which U.S. firms source. Conversely, I treated the degree of financial constraints ϕ and contractibility μ as either an industry characteristic (in Table 5.6), a country characteristic (in Table 5.8), or an interaction of a sector-specific component and a country-pair-specific component (in Table 5.7). Maintaining these assumptions, we can then express equation (8.4) in a general form as

$$\ln\left(M_{ijv}^{if}/M_{ijv}^{nif}\right) = \frac{\theta}{(1-\eta_v)(\sigma_v - 1)} \Phi\left(\sigma_v, \eta_v, \epsilon_v, \rho_v, \mu_v, \phi_v, \mu_{ij}, \phi_{ij}\right) + \varepsilon_{ijv}. \tag{8.6}$$

Note that this equation is closely related to our regressions above using both cross-product and cross-country variation. More specifically, suppose that the country-year fixed effects included in those specifications appropriately controlled for the terms μ_{ij} and ϕ_{ij} (and any other unobserved country-specific determinant of integration). In such a case, one could indeed invoke (8.6) when claiming to estimate the partial effect of the industry-specific variables σ_v, η_v, ϵ_v, ρ_v, μ_v, and ϕ_v on the ratio $M_{ijv}^{if}/M_{ijv}^{nif}$, a ratio that is tightly related to the intrafirm import share.[16]

[16] A key difference relative to the previous regressions is that trade barriers and final-

Unfortunately, it is pretty clear from the formulae for the ratio Ψ_{ijVv}/Ψ_{ijOv} derived in Chapters 6 and 7 that the effects of country-level variables and industry-level variables interact with each other, and thus demeaning intrafirm trade shares or $\ln\left(M_{ijv}^{if}/M_{ijv}^{nif}\right)$ within countries (and years) will not absorb the effects of these country-level variables. This is precisely the same reasoning we used in Chapter 5 to motivate the inclusion of both country-year *and* industry fixed effects when explaining the level of offshoring. We shall now follow a similar approach in empirical specifications attempting to explain the relative prevalence of intrafirm transactions in offshoring.

Before diving back into the empirics, there is one more issue worth discussing. One might wonder, in particular, why in light of equation (8.6) I have focused above on linear specifications in which the share of intrafirm imports – rather than $\ln\left(M_{ijv}^{if}/M_{ijv}^{nif}\right)$ – was the dependent variable. It should be stressed, however, that in order to reach the simple specification (8.6), I have had to shut down selection effects, which in our two-country model were essential for generating certain comparative static results, such as those related to trade barriers τ or productivity dispersion κ on the intrafirm trade shares. These selection effects are key features of the data. In the raw NAICS data, of the total 1,085,760 product-country-year observations, only 313,152 (28.8%) feature positive import volumes, and of those only 189,340 (17.4%) feature positive intrafirm imports. Furthermore, for another 13,816 (1.27%) observations, intrafirm imports are positive, but non-related-party imports are not. Thus, in regressions in which the left-hand-side variable is $\ln\left(M_{ijv}^{if}/M_{ijv}^{nif}\right)$ a significant share of observations in which the intrafirm trade share is not strictly between zero and one will be discarded. This is the reason why I have chosen to present results explaining the *level* of the intrafirm import share up to this point in the chapter. And it is precisely for this reason that the results to be presented below explaining $\ln\left(M_{ijv}^{if}/M_{ijv}^{nif}\right)$ should be treated with caution, in the same way that I argued that one should take the results with the logarithm of imports as a dependent variable in Chapter 5 with a grain of salt.[17]

good producers' productivity dispersion play no role in equation (8.6). This is of course due to the fact that selection effects have been neutralized in the multi-country model.

[17] I have also experimented with log-linear regressions in which the share of intrafirm imports is the dependent variable in the regressions in Tables 8.8 and 8.9, and the results I obtained were very similar. Nevertheless, the regressions with $\ln\left(M_{i-f}/M_{ni-f}\right)$ as dependent variable featured a higher R-squared.

Results of the Difference-in-Difference Specifications

As in Chapter 5, I will now explore specifications including industry and country fixed effects as well as interactions of industry and country characteristics. In particular, I express (8.6) as

$$\ln\left(M_{USjv}^{if}/M_{USjv}^{nif}\right) = \alpha_v + \alpha_j + \boldsymbol{\beta}\mathbf{Z}_j\mathbf{z}_v + \varepsilon_{jv}, \tag{8.7}$$

where $\mathbf{Z}_j$ and $\mathbf{z}_v$ are vectors of source-country and industry variables, while α_v and α_j are industry and country-year fixed effects (as in Chapter 5, I omit time subscripts for simplicity).

I begin in Table 8.8 by including the same set of interaction terms as in Table 5.7 of Chapter 5. Again, the inclusion of these variables is not motivated on theoretical terms (at least not based on the models developed in this book), but rather because they have featured prominently in the recent literature on trade and institutions. More specifically, Table 8.8 includes nine interaction terms: two Heckscher-Ohlin interactions related to physical capital and skilled labor intensity and relative abundance, four interactions of industry-level contractual 'intensity' and country-level contractual enforcement, two interactions of financial 'dependence' and financial development, and a final interaction capturing the differential role of rigid labor markets across sectors. I refer the reader to Chapter 5 and to the Data Appendix for more details on the source of these variables.

In column (1) of Table 8.8, I begin by presenting a bare-bones specification that includes only the two Heckscher-Ohlin interactions, which as pointed out in Chapter 5, had a positive and significant effect on the propensity of U.S. firms to offshore. The table indicates a positive but negligible effect of the physical capital interaction, but a negative and significant effect of the skilled labor interaction. The latter result suggests a lower propensity to integrate skill-intensive production processes in relatively skill-abundant countries. In the remaining columns of Table 8.8, I introduce the seven 'institutional' interaction terms, first one at a time in column (2), and then jointly in the remaining columns. All columns use the Wright and Nunn-Trefler corrected data and the last three columns only differ in the set of additional controls included in the regression. As in Table 5.7, column (3) includes no additional controls (other than the industry and country-year fixed effects), column (4) includes interactions of the seven institutional industry variables with GDP per capita, and column (5) includes interactions of sector dummies with GDP per capita. These different specifications were rationalized in Chapter 5 and the same arguments apply to the current context.

Table 8.8. Contractual Determinants of U.S. Intrafirm Trade Shares

Dep. Var.: $\ln(M^{if}_{USjv}/M^{nif}_{USjv})$	(1)	(2)	(3)	(4)	(5)
K Intensity × K Abund.	0.019		-0.215	-0.205	-3.518**
	(0.183)		(0.213)	(0.257)	(0.708)
Skill Inten × Skill Abund	-0.344*		-0.426^{+}	-0.207	0.203
	(0.175)		(0.230)	(0.265)	(0.322)
Nunn × Rule		0.134*	0.068	-0.044	-0.098
		(0.067)	(0.066)	(0.118)	(0.103)
Levchenko × Rule		0.060^{+}	0.057	0.064	0.027
		(0.032)	(0.035)	(0.070)	(0.059)
Costinot × Rule		-0.046	-0.184*	-0.341*	-0.294*
		(0.071)	(0.075)	(0.144)	(0.146)
BJRS × Rule		0.083	0.159*	0.137	0.069
		(0.079)	(0.066)	(0.125)	(0.123)
Rajan-Zingales × Credit/GDP		0.102*	0.302^{+}	0.874**	0.238
		(0.145)	(0.167)	(0.288)	(0.152)
Braun × Credit/GDP		0.220	0.290^{+}	0.318*	0.301*
		(0.154)	(0.156)	(0.152)	(0.131)
Firm Volatility × Labor Flexibility		-0.275^{+}	-0.374*	-0.332^{+}	-0.344*
		(0.154)	(0.171)	(-0.188)	(0.161)
Sample Restrictions	W+NT^{+}	W+NT^{+}	W+NT^{+}	W+NT^{+}	W+NT^{+}
Ctr/Year & Ind Fixed Effects	Yes	Yes	Yes	Yes	Yes
Interactions with GDP pc	No	No	No	Yes	No
Industry Effects × GDP pc	No	No	No	No	Yes
Observations	89,669	≃ 88,000	84,738	77,307	77,307
R-squared	0.732	≃ 0.73	0.738	0.745	0.769

Standard errors clustered at the country/ind. level. $^{+}, ^{*}, ^{**}$ denote 10, 5, 1% significance.

Because the specifications in this table are not motivated by any of the models developed above, one should not infer too much from the results to be discussed. It should be noted, however, that to the extent that the included institutional interactions capture different sources of transaction costs, one would expect based on the transaction-cost theory that the sign of the coefficient of these interactions should now be the opposite to the one obtained in Chapter 5. Intuitively, if intrafirm trade circumvents all contractual inefficiencies, any contracting-related interaction term that positively predicts the overall level of offshoring, should now necessarily reduce the share of this offshoring conducted within firm boundaries. Judged by this criterion, how well does the transaction-cost model fare against the evidence?

Consider first the interactions related to contract enforcement. In Table 5.7 we found robust evidence of a negative effect of these interactions on U.S. intermediate input imports. In words, the positive effect of source country contract enforcement on U.S. firms' input purchases was lower, the more contractible inputs are. Based on the transaction-cost model, one would thus expect these same interactions to have a positive effect on intrafirm import shares. Although the effect of the Nunn and Levchenko contracting interactions is indeed positive and significant when introduced alone in column (2), their significance disappears in the columns that include all interactions simultaneously. Furthermore, the only interaction term whose significance survives is the Costinot interaction, but it does so with the same negative sign obtained in Table 5.7.

On a more positive note, Table 8.8 indicates that the Braun and Cuñat-Melitz interactions do indeed appear to shape intrafirm import shares with the opposite sign with which they affected overall offshoring levels. In words, those results indicate that an improvement in source-country financial or labor-market institutions will tend to disproportionately decrease the propensity of U.S. firms to internalize transactions in sectors with hard access to finance and in sectors with higher needs for labor reallocations across firms. Finally, it is worth pointing out that the effect of the Heckscher-Ohlin interactions appear to be negative in most specifications, but their size and significance differs dramatically across specifications, as reflected for instance in the implausibly large negative coefficient on the physical capital interaction in column (5).

In Table 8.9, I turn to a specification that is more closely related to the global sourcing models in Chapters 6 and 7. In the same manner that Table 8.8 basically replicated the results of Table 5.7 but with a different dependent variable, in Table 8.9 I build closely on the specification of Table 5.8. In particular, the specification is again equation (8.7) but the vector of interaction terms now includes: (i) the two Heckscher-Ohlin interactions mentioned above; (ii) measures of freight costs and U.S. tariffs that vary both across sectors and countries; and (iii) five institutional interactions inspired by the global sourcing models. All but one of the latter institutional regressors constitute interactions of the source-country level of contract enforcement μ_j (as proxied by their rule-of-law index) with empirical proxies for some of the key 'industry' parameters of the model: input substitutability (ρ_v), demand elasticity (σ_v), specificity (ϵ_v), and headquarter intensity (η_v). The last interaction is the product of headquarter intensity η_v and financial development ϕ_{ij}. The data sources are the same ones as detailed in Chapter 5 (see the

Data Appendix), where recall that headquarter intensity η_v is measured as the first principal component from a factor analysis of the buyer versions of R&D, equipment capital and skill intensity variables.

Before discussing the empirical results, let us briefly review what the two models of internalization developed in this book would predict for the sign of these different interactions terms. For reasons discussed at length in previous chapters and also reviewed earlier in this chapter, both the transaction-cost and property-rights model would predict a positive effect of trade barriers (freight costs, U.S. tariffs) on the intrafirm trade share. Remember that the selection into offshoring channel was key for these effects, so it is not too surprising that trade barriers do not feature in our multi-country specification in (8.6). Still, I believe it is sensible to include these in the empirical specifications. As in our previous results in this chapter, Table 8.9 will demonstrate that these proxies for trade barriers appear to have a negative effect on the intrafirm import share, a finding that is hard to reconcile with the benchmark sorting pattern assumed in the models.

The empirical relevance of the selection into offshoring channel also justifies the inclusion of the two Heckscher-Ohlin interactions. More precisely, if we were to close the models in such a way that the relative wage costs were affected by relative factor endowments, we would typically obtain that the relative wage w^N/w^S would be lower, the more abundant is the source country in physical capital and skilled labor. As a result, we would expect the intrafirm trade share to be higher for inputs originating in more capital- and skill-abundant countries, with the effect being disproportionate for less headquarter intensive sectors. In sum, a more realistic general equilibrium version of our models would predict a negative sign on the two Heckscher-Ohlin interactions. This is consistent with the results in Table 8.8, although we have seen that those coefficients are not particularly stable. The coefficients on these interactions in Table 8.9 are very similar in magnitude to those in Table 8.8, so I do not report them to save space (interested readers can inspect the dataset and programs downloadable at http://scholar.harvard.edu/antras/books).

We are then left with the five interaction terms capturing institutional factors. What are the predicted signs for these interactions predicted by the models in Chapters 6 and 7? To answer this question for the case of the transaction-cost model, it suffices to refer to the predictions of the model for the arm's-length component of U.S. intermediate input imports, and then simply invert the sign of those predictions. Building on equation (5.14), we can write this in succinct form as

$$\ln\left(M^{if}_{USjv}/M^{nif}_{USjv}\right) = \Phi_{TC}\left(\underset{+}{\mu_{ij}\times\rho_v}, \underset{-}{\mu_{ij}\times\sigma_v}, \underset{-}{\mu_{ij}\times\epsilon_v}, \underset{\text{ambiguous}}{\mu_{ij}\times\eta_v}, \underset{-}{\phi_{ij}\times\eta_v}\right). \tag{8.8}$$

For the case of the property-rights model, deriving sharp predictions for the effects of these interactions on the intrafirm import share proves to be much more challenging. In the Theoretical Appendix, I discuss in more detail these comparative statics. The bottom line, however, is that although in some cases the numerical examples appear to point towards a particular sign for these effects, it is often possible to construct numerical examples in which the opposite sign applies for a region of the parameter space. The analytical results derived in the Appendix can be summarized as follows:

$$\ln\left(M^{if}_{USjv}/M^{nif}_{USjv}\right) = \Phi_{PR}\left(\underset{\text{ambiguous}}{\mu_{ij}\times\rho_v}, \underset{\text{ambiguous}}{\mu_{ij}\times\sigma_v}, \underset{\text{ambiguous}}{\mu_{ij}\times\epsilon_v}, \underset{-}{\mu_{ij}\times\eta_v}, \underset{\text{ambiguous}}{\phi_{ij}\times\eta_v}\right), \tag{8.9}$$

implying that only the interactions $\mu_{ij} \times \eta_v$ has an unambiguous predicted sign. It should be noted, however, that in most numerical examples, I have found the interaction $\mu_{ij} \times \rho_v$ to affect the intrafirm trade share negatively, while the interaction of $\mu_{ij} \times \sigma_v$ can be shown to affect integration negatively when contract enforcement in country j affects only the contractibility of headquarter services, but positively when it affects only the contractibility of manufacturing production (see the Theoretical Appendix).

Comparing the predicted signs in (8.8) and (8.9), it should be clear that it will be extremely difficult to discriminate between the two theories based on this difference-in-difference approach. In particular, for none of the interactions I have experimented with, can one state that the effect of that interaction on the relative propensity to integrate is unambiguously of opposite sign in the two theories. Hence, even if one finds robust evidence for some of the predictions of one of the models, this same evidence cannot possibly be used to refute the other theory. In sum, for the goal of discriminating across models, I view the type of cross-industry results in Tables 8.6 and especially 8.7 as being much more useful.

With these caveats in mind, let us then turn to Table 8.9 to study how these different interaction terms affect the relative prevalence of intrafirm imports in the data. Column (1) implements the same specification as in column (1) of Table 8.8, but now expanded to include measures of freight costs and U.S. tariffs (the Heckscher-Ohlin interactions are not reported to save space). As anticipated before, and consistently with our findings throughout this chapter, both of the coefficients on trade barriers are negative and significant, which constitutes a challenge for the models.

Table 8.9. Testing the Transaction-Cost and Property-Rights Models

Dep. Var.: $\ln(M^{if}_{USjv}/M^{nif}_{USjv})$	(1)	(2)	(3)	(4)	(5)
Freight Costs	-0.201**		-0.162**	-0.142**	-0.086
	(0.058)		(0.059)	(0.055)	(0.059)
Tariffs	-0.156*		-0.197**	-0.185**	-0.168**
	(0.067)		(0.059)	(0.055)	(0.048)
Input Substit. × Rule		0.036	0.033	0.073	0.057
		(0.028)	(0.024)	(0.048)	(0.048)
Demand Elasticity × Rule		-0.062**	-0.062**	-0.109*	-0.060
		(0.012)	(0.012)	(0.043)	(0.073)
Nunn Specificity × Rule		-0.196**	-0.150*	1.697^+	-0.005
		(0.063)	(0.060)	(0.906)	(0.115)
Headquarter Intensity × Rule		0.015	-0.006	0.069	0.105*
		(0.030)	(0.027)	(0.053)	(0.051)
Headquarter Intensity × Credit/GDP		0.018	-0.003	-0.004	-0.007
		(0.042)	(0.050)	(0.044)	(0.041)
Sample Restrictions	$W+NT^+$	$W+NT^+$	$W+NT^+$	$W+NT^+$	$W+NT^+$
Ctr/Year & Ind Fixed Eff	Yes	Yes	Yes	Yes	Yes
Interactions with GDP	No	No	No	Yes	No
Industry Effects × GDP	No	No	No	No	Yes
Observations	89,393	$\simeq$ 88,000	87,298	79,654	79,654
R-squared	0.737	$\simeq$ 0.74	0.744	0.749	0.770

Standard errors clustered at the country/ind. level. $^+, ^*, ^{**}$ denote 10, 5, 1% significance.

In column (2), I introduce the five interactions motivated by our model, *one at a time*, though for compactness I report all five coefficients in a single column. Note that the sign of the coefficients is consistent with the predictions of the transaction-cost model for four of the five coefficients, with the exception being $\phi_{ij} \times \eta_v$, whose effect is positive though indistinguishable from zero. Nevertheless, only two of these five interaction terms appear to have a statistically significant effect on $\ln\left(M^{if}_{USjv}/M^{nif}_{USjv}\right)$.

In the last three columns of Table 8.9, I present results in which the five institutional interactions are included in the *same* regression, together with the Heckscher-Ohlin interactions and the trade cost measures. The three columns differ in that column (4) includes interactions between the industry-level institutional variables and GDP per capita, while column (5) includes a whole vector of interactions of sectoral dummies with GDP per capita.

The results in column (3) are very similar to those in column (2). Although now *all* the coefficients on the institutional interactions are consistent with the transaction-cost model, it continues to be the case that only two of them are significant at the 5% confidence level. The inclusion of interactions with GDP per capita in column (5) has a bigger impact on the coefficients, particularly on the interaction of specificity and the rule of law, which changes dramatically from being negative and significant at the 5% level to being positive, large in magnitude and significant at the 10% level. Controlling for interactions of sectoral dummies and GDP per capita in column (6) also turns out to affect the estimates in a sizable manner. The coefficient on freight costs is no longer significant, while only U.S. tariffs and the interaction of headquarter intensity with the rule of law remain statistically significant at standard levels. The sign of both of these coefficients is, however, opposite to the one predicted by the property-rights model.

Overall, the difference-in-difference approach in Tables 8.8 and 8.9 has yielded much poorer results than the same approach applied to explaining offshoring shares in Chapter 5. It might be argued that part of the higher success in Chapter 5 is associated with the fact that the inclusion of country-year and industry fixed effects in those specifications still left a large share of the variation in U.S. intermediate input imports unexplained, while their inclusion in Tables 8.8 and 8.9 leaves less variation in the share of intrafirm imports to be explained. Indeed, the R-squared of a simple regression of $\ln\left(M_{USjv}^{if}/M_{USjv}^{nif}\right)$ on country-year and industry fixed effects is 0.731, while the analogous R-squared in regressions with log U.S. offshoring shares as a dependent variable is 0.616. This might be a factor in justifying the higher statistical significance of the coefficients in the regressions in Chapter 5, but it cannot possibly explain the fact that our global sourcing model *correctly* predicted the sign of the coefficient on the institutional interactions.

The results in Table 8.9 are not only weak in statistical terms, but they do not generally support the qualitative predictions of the internalization models either, particularly in the case of the property-rights model. Nevertheless, as argued above, the property-rights model does not offer sharp predictions for how the interaction of country and industry characteristics should shape intrafirm trade shares. And even when some analytical results can be obtained, one needs to work sufficiently hard to derive them to worry that they might rely quite heavily on the chosen functional forms. For these reasons, I view the specification tests in Tables 8.3, 8.4, 8.6 and 8.7, with their own admitted limitations, as more transparent tests of the models of multinational firm boundaries.

Other Sources of Product-Level Data

So far, I have restricted attention to product-level tests employing intrafirm import data from the U.S. Related-Party Trade database. The focus on the U.S. data is partly justified by the fact that many of the key contributions to this literature conduct their analysis with that dataset. Recent examples include the work of Nunn and Trefler (2008, 2013*b*), Bernard, Jensen, Redding and Schott (2010), Costinot, Oldenski and Rauch (2011) and Díez (2014), while an early use of the same dataset can be found in Helleiner and Lavergne (1979).[18]

Even while maintaining the focus on U.S. trade flows, many researchers have constructed alternative measures of intrafirm trade by employing direct investment data from the U.S. Bureau of Economic Analysis, which can be downloaded from: http://www.bea.gov/iTable/index_MNC.cfm. Zeile (1997) and Ruhl (2013) provide insightful descriptions of this database, which essentially aggregates the intrafirm exports and imports of U.S. foreign affiliates and of U.S. affiliates of foreign companies. The BEA database has the advantage of providing more information on the extent to which the ownership link between importers and exporters is large enough to convey effective corporate control, in line with the models of internalization developed in previous chapters. Another useful feature of this dataset is that it allows one to distinguish between trade within U.S. multinationals from trade within foreign multinationals operating in the United States. As readers will recall, I made use of this aspect of the BEA data to refine my estimates above, in line with the approach in Nunn and Trefler (2013*b*). The main drawback of this data source is that it is not as comprehensive as the U.S. Related-Party database: it only covers a limited number of fairly aggregated sectors and is only recorded in a handful of benchmark years, forcing researchers to extrapolate from those years to create yearly series for U.S. intrafirm imports. Another limitation of the data is that the BEA records intrafirm imports by U.S. affiliates of foreign parents as originating from the country of ownership (i.e., of the parent company) rather than from the country in which the shipment actually originated. Among others, this source of data has recently

[18] To be precise, Bernard, Jensen, Redding and Schott (2010) had access to the confidential firm-level U.S. census data on related-party trade, but their regression analysis was performed at the sectoral level.

been used by Antràs (2003) and Yeaple (2006) to test the property-rights theory, though it had also been used by Lall (1978) and Siddharthan and Kumar (1990) to test the transaction-cost theory.

To the best of my knowledge, the United States is the only country that collects detailed customs-level trade statistics distinguishing trade between related parties (intrafirm trade) from trade between non-related parties (arm's-length trade). In that sense, the U.S. Related-Party Trade database is unique.[19] For other countries, one can use a similar methodology as is used with the BEA dataset to create measures of intrafirm trade based on survey information on the trade transactions of multinational firms and their affiliates. For example, the OECD Activities of Foreign Affiliates (AFA) database contains intrafirm data for nine countries (Canada, Israel, Italy, Japan, Netherlands, Poland, Slovenia, Sweden and the United States), but coverage is far from complete, as described by Lanz and Miroudot (2011). Perhaps for this reason, this source of data has not been employed so far to shed light on the determinants of the internalization decisions of firms.

Some contributions to the empirical literature on multinational firm boundaries have used product-level export data from the Customs General Administration of the People's Republic of China. These data do not distinguish between intrafirm and arm's-length trade, but they do contain detailed information on whether the exporter is a foreign-owned plant or not. As such, the data are suitable for an analysis of the determinants of foreign ownership of suppliers in China, regardless of the identity of the buyers of their goods. A very interesting feature of the Chinese data is that not only do the data differentiate between ordinary trade and processing trade, but in addition processing trade is categorized under different types of customs regimes depending on whether the plant in China is in charge of importing inputs or that responsibility falls on a foreign producer. The former type is referred to as 'import-and-assembly', while the latter is labelled 'pure-assembly'. Feenstra and Hanson (2005) and Fernandes and Tang (2010) exploit this feature of the data to test rich variants of the property-rights theory of multinational firm boundaries (see also Feenstra, 2011). More recently, Li (2013) has also employed the Chinese processing trade data to test a transaction-cost model of firm boundaries emphasizing the role of communications costs in hindering outsourcing transactions.

[19] This will hopefully change in the future as the United Nations Statistics Division has explicitly recommended the collection of intra-firm trade data in customs-based merchandise trade statistics (see United Nations, 2010).

Empirical Evidence Using Firm-Level Data

The transaction-cost and property-rights theories are theories of firm boundaries and thus *firm*-level data would appear to be the ideal laboratory to employ when testing these theories. To give a precise example, some of the comparative statics results derived in the transaction-cost and property-rights models in Chapters 6 and 7 were tightly related to the equilibrium sorting pattern of firms by productivity into organizational forms. Is that sorting pattern consistent with available data? To answer this question one obviously needs to make use of firm-level data.

Unfortunately, firm-level data on the global sourcing decisions of firms are not readily available. Although most countries maintain census or survey data on the firms operating in their economy, there are many fewer datasets that provide sufficiently detailed information on the global sourcing practices of these firms, and more specifically on whether they import intermediate inputs within or across firm boundaries. I will next describe five firm-level datasets that have been used by researchers in recent years, while emphasizing both their main advantages and limitations and outlining some of the results that have been obtained when exploiting these data sources.

An early paper using firm-level data to shed light on the firm boundaries decisions of multinational firms is Tomiura (2007), who uses data from the Basic Survey of Commercial and Manufacturing Structure and Activity in Japan. The survey covers 118,300 Japanese manufacturing firms and is regarded as an accurate overall representation of the Japanese manufacturing sector. Unfortunately, the survey was carried out only in one year, 1998. The survey contains a variety of information on the operations of firms (sales, employment, capital expenditures, exports, foreign direct investment) and crucially also asks firms whether they "contract out manufacturing or processing tasks to other firms overseas." Hence, the survey can be used to explicitly distinguish firms that are engaged in foreign outsourcing versus those that are engaged in foreign direct investment. On the downside this Japanese dataset does not appear to contain information on the volume (i.e., the intensive margin) of foreign insourcing and outsourcing. Tomiura (2007) uses the dataset to show that, consistently with the equilibrium sorting of the transaction-cost and property-rights models (see for instance Figures 6.4 and 7.2), firms that are engaged in FDI and intrafirm trade are significantly more productive than firms that are engaged in foreign outsourcing, which in turn are more productive than firms sourcing domestically. An interesting feature of the data, which many other datasets have unveiled as well, is that most firms are neither "pure FDI" firms nor "pure outsourcing" firms, which is suggestive of the relevance of models in which firms follow hybrid global sourcing strategies (such as the model in Antràs and Chor, 2013).

A second line of papers, most notably Carluccio and Fally (2011), Defever and Toubal (2013) and Corcos, Irac, Mion and Verdier (2013) have used French firm-level data from the EIIG (Échanges Internationaux Intra-Groupe).[20] The EIIG is a survey conducted in 1999 by the SESSI (Service des Études et des Statistiques Industrielles), which documents the extent to which firms import foreign inputs from related or non-related parties. Interestingly, the data identify those input purchases that are related to contract manufacturing (classified as "work based on plans"), and Defever and Toubal (2013) focus their analysis on these purchases because they better map to our global sourcing models. Another important feature of the EIIG survey is that it provides a firm-level break-up between intrafirm and arm's-length offshoring by source country and by four-digit HS input product codes. Thus, the data are similar in scope to the U.S. Related-Party Database, but they are at the firm-level and also better capture the intermediate input component of imports. The main drawback of the dataset is its coverage. The survey includes only French firms that traded more than 1 million euros in 1999, and more importantly, it is restricted to firms that are owned by manufacturing groups that control at least fifty percent of the equity capital of an affiliate based outside France. Though not all firms responded to the survey, the 4,305 respondent firms represent more than 80% of total exports and imports of French multinationals in 1999. Still, the fact that only firms with at least one affiliate outside France were sampled raises serious concerns about sample selection biases. Corcos, Irac, Mion and Verdier (2013) acknowledge this problem and complement the dataset with data coming from the French Customs Office, documenting the universe of yearly imports and exports flows in 1999 at the firm, origin country and product level, hence allowing them to offer a more representative picture of the foreign outsourcing operations of French firms.[21]

The goals and scope of the papers using the EEIG dataset are somewhat different, but they all find supportive evidence of a positive correlation between headquarter intensity and the relative importance of intrafirm trade, with the measures of headquarter intensity in Corcos, Irac, Mion and Verdier (2013) being *firm-level* measures (namely, capital intensity, skill in-

[20] This dataset has also been used recently by Carluccio and Bas (2014) to study the link between intrafirm trade and labor market institutions.

[21] The EEIG survey can in turn be matched with another SESSI database, the EAE (Enquête Annuelle Entreprise), which provides balance sheet data on manufacturing firms with at least 20 employees.

tensity and the ratio of value added over sales of the importing firm) based on the importer's operational data. Defever and Toubal (2013) and Corcos, Irac, Mion and Verdier (2013) find instead conflicting evidence regarding the relative productivity of firms engaged in foreign outsourcing and foreign integration. More specifically, Defever and Toubal (2013) find that in the sample of French firms that own at least one affiliate abroad, firms engaging in foreign outsourcing are on average *more* productive than firms engaging in FDI. They rationalize this finding by appealing to the notion that for entities belonging to a multinational network, the fixed costs of procuring inputs from arm's-length suppliers external to the network are likely to be higher than the fixed costs of doing so from suppliers within the network. This reversal of the ranking of fixed costs relative to the one assumed in our models naturally leads to an analogous reversal in the equilibrium sorting pattern, with the average productivity of firms that engage in offshore outsourcing now being higher than that of firms engaging in offshore insourcing. Corcos, Irac, Mion and Verdier (2013) show, however, that when incorporating into the dataset the many firms in France that do not own affiliates abroad, the productivity advantage of FDI firms over foreign outsourcers predicted by the our global sourcing models is restored (see, for instance, their Figure 1).

A third firm-level dataset available to test models of the boundaries of multinational firms is a panel dataset of Spanish manufacturing firms collected by the Fundación SEPI. The ESEE (Encuesta sobre Estrategias Empresariales) surveys approximately 2,000 Spanish firms with at least ten employees on a yearly basis since 1990. The ESEE provides information on the income and balance sheet statistics of firms, and also on a variety of firm-level organizational variables. A notable characteristic of the ESEE is its representativeness, which is ensured by the careful statistical criteria used in the initial year of the sample and the special attention that has been given to account for entry and exit of firms of different sizes in subsequent years.[22] For the purposes of testing global sourcing theories, a particularly relevant feature of the data is that they allow one to compute the overall spending on intermediate inputs by firms and their breakup into (i) domestic purchases from independent suppliers, (ii) domestic purchases from affiliated parties, (iii) imports from foreign independent suppliers, and (iv) imports from foreign affiliates. Hence, one can easily map some of the variables of the survey

[22] Details on the survey characteristics can be downloaded from the following website: http://www.fundacionsepi.es/esee/sp/presentacion.asp. The data are accessible for a relatively modest fee to any researcher regardless of nationality. Among others, it has been used in other contexts by Delgado, Fariñas and Ruano (2002) and Guadalupe, Kuzmina and Thomas (2012).

into the four key organizational forms emphasized by the models in Chapters 6 and 7. An important disadvantage of this Spanish dataset is that it only distinguishes between domestic and foreign input purchases, with the latter not being disaggregated by country of origin.

Using the ESEE dataset, Kohler and Smolka (2013, 2014) find that, conditional on the location of sourcing (domestic or foreign), firms that purchase inputs from integrated suppliers appear to be significantly more productive than those procuring inputs from arm's-length suppliers. This confirms the findings of Tomiura (2007) for Japan and of Corcos et al. (2013) for France, and further demonstrates that the superior performance of integrating firms holds also when focusing on domestic sourcing. This is consistent with the sorting pattern assumed in the models in Chapters 6 and 7. Conversely, Kohler and Smolka (2009) find mixed evidence regarding the relative productivity of firms outsourcing abroad and firms integrating in Spain.

Using the same data that Kohler and Smolka (2009) used, Figure 8.2 plots the distribution of (Olley-Pakes) total factor productivity of firms according to the organizational form they adopt when sourcing. The probability density functions in Figure 8.2 confirm that domestic outsourcers are (on average) the least productive firms, while foreign integrators are (on average) the most productive firms. More interestingly, it appears that the distribution of productivity of firms that engage in domestic integration is a shift to the right of that of firms that outsource abroad. This sorting pattern is inconsistent with the one assumed in the benchmark versions of the global sourcing models, but is in line with the alternative ranking of fixed costs $f_{OV} > f_{DV} > f_{OO} > f_{DO}$ we explored in Chapter 6. As discussed in that chapter, this alternative sorting pattern delivers the *exact* same comparative statics as our benchmark version of the transaction-cost model. In the Theoretical Appendix I show, perhaps surprisingly, that the same is not true in the property-rights model. In fact, through the lens of the property-rights model, the sorting pattern implied by Figure 8.2 may well help rationalize the robust negative effects of trade frictions on intrafirm trade shares unveiled in the empirical work developed earlier in this chapter. Intuitively, reductions in trade barriers not only lead firms that were sourcing domestically to select into outsourcing, but now also push the most productive among the domestic integrating firms to select into foreign integration. As a result, the predicted overall effect of reductions in trade barriers on the share of intrafirm imports can now well be negative, at least in the property-rights model. In the Theoretical Appendix, I show in addition that the overall effect of trade frictions is particularly likely to be negative when the contractual insecurity associated with offshoring stems largely from lower contractibility of input manufacturing.

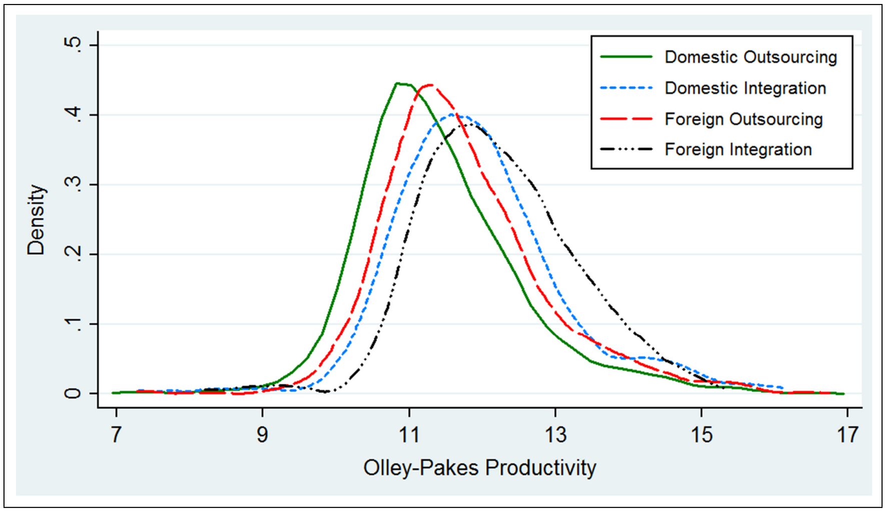

Figure 8.2: Organizational Sorting in Spain

One might worry that the sorting pattern unveiled in Figure 8.2 is specific to the Spanish data. The work of Federico (2010) suggests, however, that a similar sorting is observed when using firm-level data from Italy, this being the fourth of the five datasets I will review here. Federico (2010) uses firm-level data from the Survey on Italian Manufacturing Firms, conducted every 3 years by Mediocredito Capitalia (MCC). The MCC survey combines some of the virtues of the French EEIG and Spanish ESEE datasets. On the one hand, the survey provides information on the extent to which Italian firms make use of contract manufacturing, thus excluding the purchase of standardized inputs.[23] On the other hand, the data allow researchers to distinguish between four types of suppliers, corresponding to the same four organizational forms in the models and in the Spanish data. The main drawback of this dataset is that its coverage is far less complete than that of the other surveys described above, and severely undersamples small firms. Using these data, Federico (2010) finds a sorting pattern qualitatively similar to that in Figure 8.2, with firms with medium-low productivity choosing foreign outsourcing, and firms with medium-high productivity choosing domestic integration.

[23] Contract manufacturing is defined as "a contract by which an entrepreneur engages itself on behalf of the buying company to carry out workings on semifinished products or raw materials, or to supply products or services to be incorporated or used in the buying company's economic activity or in the production of a complex good, in conformity with the buying company's projects, techniques, technologies, models or prototypes" (see Federico, 2010).

A fifth international firm-level dataset that has been used to shed light on the boundaries of multinational firms is the Dun & Bradstreet's WorldBase (DB) database. This massive dataset contains information on hundreds of thousands of establishments in more than 200 countries and territories. In contrast to the four surveys described above, the DB dataset does not contain comprehensive operational data related to these plants, but it does offer a comprehensive picture of firm boundaries across borders. In particular, the DB database contains detailed information on the location, ownership (e.g., its domestic or global parent) and industry classification of plants worldwide. Alfaro, Conconi, Fadinger and Newman (2014) use these data to document a positive association between higher tariffs on final products (as measured by MFN tariffs at the four-digit SIC industry level for all WTO members) and an index of domestic vertical integration constructed using the ownership information in the DB database as well as input-output tables. Their empirical exercise exploits both cross-section and time-series variation in trade policy, but also considers China's entry into the WTO in 2001 as a quasi-natural experiment. The authors take these results as empirically validating the model of organizational design in Legros and Newman (2013) and Conconi, Legros and Newman (2012).

Conclusions and the Road Ahead

The empirical evidence discussed in this chapter has offered broad support for some of the key predictions of the internalization models developed in Chapters 6 and 7. The property-rights model, in particular, has fared particularly well against the evidence. It would be premature, however, to interpret the evidence as being conclusive because, for the most part, the tests performed up to today have relatively low statistical power. In my view, successful testing of the theory will need to be based on one of the following three approaches.

A first possibility is to better exploit the large variation in the relative prevalence of integration retrievable from the U.S. Related-Party database or from some of the firm-level datasets described above. In particular, I believe that the cross-industry and cross-country specifications that I have experimented with earlier in this chapter are interesting and informative, but they cannot convincingly identify a causal effect of headquarter intensity or of product contractibility (even when appropriately measured) on the share of intrafirm imports. A potential avenue for future research is to use narrower slices of the data, perhaps (i) focusing on the patterns in a single industry, but exploiting exogenous changes in sector characteristics driven by technological or demand-driven shocks, in the spirit of Baker and Hubbard (2003), or perhaps (ii) performing analyses exploiting within-country variation stemming from changes in the institutional characteristics of countries, such as observable changes in the quality of institutions or in restrictions on foreign ownership in those countries.

A second approach would consist of using the available firm-level datasets to better identify the various channels via which contractual factors shape the intensive and extensive global sourcing decisions of firms. Many of the comparative static results described at length in Chapters 6 and 7 followed from a combination of selection channels and input demand channels. When working with product-level data, one can indeed only test the outcome of the balance of these effects. Nevertheless, firm-level dataset should in principle allow researchers to effectively separate these channels, which in turn could prove useful in better discriminating among available theories. Some aspects of the empirical analysis of Corcos, Irac, Mion and Verdier (2013) with the French EEIG dataset constitute a valuable first attempt along these lines, but more work is needed on this front. In a similar vein, there are surely many untapped sources of firm-level data on the global sourcing strategies of firms. For instance, the empirical studies of global sourcing in Fort (2014) and Antràs, Fort and Tintelnot (2014) make use of a rich dataset that links the U.S. Census Bureau's 2007 Economic Censuses (EC), Longitudinal Business Database (LBD), and Import transaction database. The latter is the basis of the U.S. Related-Party database and thus contains information on whether firm-level import transactions are integrated or not. Another interesting aspect of the data, which Fort (2014) exploits in her empirical analysis, is that the data can be used to identify U.S. firms' decisions to contract for manufacturing services from domestic or foreign suppliers, which constitute a better match for the models developed in this book than overall input purchases.

A third fruitful area of future research would involve a more structural usage of available firm-level datasets. At present, little work has been devoted to structurally estimating the models I have discussed in this book. This is partly due to the stylized nature of these frameworks, and partly due to the under-utilization of this type of empirical techniques in the international trade field. An exception is the work of Antràs, Fort and Tintelnot (2014), who structurally estimate a version of the multi-country global sourcing model developed in this book, but focusing on its complete-contracting version in Chapter 2. I would anticipate that future theoretical developments in this area will likely produce much richer frameworks of the internalization decisions of multinational firms. These frameworks should prove to be more amenable for structural work. Structural techniques have of course their own limitations, but their main appeal is that they open the door for a quantitative evaluation of the models. How do the organizational decisions of multinational firms shape firm-level performance? How do they affect labor markets, product markets, and social welfare? These are key questions that have so far not been sufficiently explored in the literature.