The Boundaries of the Multinational Firm: An Empirical Analysis

Nathan Nunn[*‡]

University of British Columbia and CIAR

Daniel Trefler[*§]

University of Toronto, CIAR and NBER

April 25, 2007

ABSTRACT: Using data on U.S. intra-firm and arm's-length imports for 5,423 products and 210 countries, we examine the determinants of the share of U.S. imports that are intra-firm. Three determinants of this share have been proposed: (1) Antràs (2003) focuses on the share of inputs provided by the headquarter firm. We provide added confirmation and further strengthen the empirical findings in Antràs (2003) and Yeaple (2006). (2) In a model featuring heterogeneous productivities, Antràs and Helpman (2004) focus on the interaction between the firm's productivity level and the headquarter's input share. We find very strong support for this determinant. (3) Antràs and Helpman (2006) add to this the possibility of partially incomplete contracting. We find that consistent with the novel prediction of their model, improved contracting of the supplier's inputs can *increase* the share of U.S. imports that are intra-firm. In short, the data bear out the primary predictions of this class of models about the share of U.S. imports that is intra-firm trade.

Key words: Intra-firm trade; Incomplete contracts; Vertical integration; Outsourcing

JEL classification: F14, F23, L14, L33

[*]Both authors thank the Canadian Institute for Advanced Research (CIAR) for its support. We also thank Pol Antràs, Rob Feenstra, Elhanan Helpman, Gene Grossman, Wilhelm Kohler and Marc Melitz for insightful comments. We are especially grateful to Joel Blit for facilitating construction of the data base.

[‡]Department of Economics, University of British Columbia, 1873 East Mall, Vancouver, BC V6T 1Z1, Canada (e-mail: `nnunn@interchange.ubc.ca`; website: `http://www.econ.ubc.ca/nnunn/`).

[§]Rotman School of Management and Department of Economics, University of Toronto, Toronto, Ontario, M5S 3E6, Canada (e-mail: `dtrefler@rotman.utoronto.ca`; website: `http://www.rotman.utoronto.ca/~dtrefler`).

1. Introduction

This volume is the culmination of a rich and recent research agenda into the determinants of intra-firm trade. Unlike the older literature on international trade in the presence of imperfect competition e.g., Helpman and Krugman (1985), the current literature provides a much more intellectually satisfying notion of what constitutes a firm. It thus provides us with deeper insights into which elements of international trade are done internally to the firm (multinational or intra-firm trade) and which are done outside the boundaries of the firm. Seminal contributions include McLaren (2000), Antràs (2003, 2005), Grossman and Helpman (2002, 2003, 2004, 2005), and Antràs and Helpman (2004).

In this paper we are particularly interested in a strand of the literature that examines the relationship between a multinational firm and its supplier. Each contributes a customized input that is non-contractible. As a result, there is a classic hold-up problem and the multinational must decide whether to vertically integrate its supplier or outsource to its supplier. One narrow strand of the literature – the one we will be dealing with – treats the difference between these two organizational forms as the difference between the outside options of the multinational in the event that the hold-up problem cannot be resolved through bargaining. This treatment of the difference between vertical integration and outsourcing originates with Antràs (2003) and appears again in Antràs and Helpman (2004, 2006).

These papers yield three important insights into the determinants of the share of total U.S. imports that are intra-firm i.e., the share that is imported by U.S. multinationals from their foreign affiliates. First, Antràs (2003) argues that when the U.S. headquarters firm provides the bulk of the non-contractible inputs, underinvestment in inputs is reduced by highly incentivizing the headquarters firm. Vertical integration provides such incentives because it allows the headquarters firm to control at least some of the supplier's inputs even if bilateral bargaining breaks down. In contrast, when the foreign supplier provides the bulk of the non-contractible inputs, the foreign supplier must be highly incentivized. This is done by outsourcing: outsourcing strips the headquarters firm of any control over the supplier's inputs and thus strengthens the bilateral bargaining position of the supplier. In short, the share of U.S. imports that are intra-firm is increasing in the share of (non-contractible) inputs provided by the U.S. headquarters firm. This logic is a specific instance of the larger property rights approach to the firm e.g., Grossman and

Hart (1986).

The second prediction about the share of total U.S. imports that is intra-firm is developed in Antràs and Helpman (2004). They start with the well-known fact that firms display heterogeneous productivities e.g., Bernard and Jensen (1997). Antràs and Helpman (2004) also argue that the fixed costs of producing abroad are lower when outsourcing to a foreign supplier than when using foreign direct investment (vertical integration). Since only the most productive firms capture the market share needed to offset the high costs of vertical integration, not all firms identified by Antràs (2003) as candidates for vertical integration will in fact integrate. Only the most productive will. Thus, the share of U.S. imports that are intra-firm will be large when two conditions are simultaneously satisfied: (*i*) the share of inputs provided by the headquarters firm is large (as in Antràs 2003) and (*ii*) firm productivity is high.

The third prediction about intra-firm trade appears in this volume (Antràs and Helpman, 2006). While Antràs (2003) and Antràs and Helpman (2004) assume that inputs are completely non-contractible, Antràs and Helpman (2006) allow inputs to be partially contractible. This leads to a surprising result. The typical view is that where property rights are strong, outsourcing is more prevalent. Antràs and Helpman (2006) arrive at the opposite conclusion, at least for some parameter values. The logic is simple. As foreign property rights improve so that the supplier's share of non-contractible inputs falls, the party that requires *relatively* more incentives becomes the headquarters firm. These incentives are provided through vertical integration. Improved property rights leads to internalization!

Using data on U.S. intra-firm and arm's-length imports for 5,423 products imported from 210 countries, we examine these determinants of the share of U.S. imports that are intra-firm. Our conclusions mirror the three predictions listed above. (1) In terms of the Antràs (2003) mechanism, we find support for the role of the share of headquarter inputs. This support is stronger than that found in the only two extant empirical studies of the issue, namely, Antràs (2003) and Yeaple (2006).[1] (2) We find strong support for the Antràs and Helpman (2004) prediction that intra-firm trade is largest where headquarter inputs are important *and* productivity is high. (3) We also find support for the Antràs and Helpman (2006) prediction about increased internalization due to an improved contracting environment for the supplier's inputs.

The paper is organized as follows. Sections 2, 3, and 4 examine the predictions of Antràs (2003),

[1] Feenstra and Hanson (2005) also provide support for the property-rights approach in an international trade context.

Antràs and Helpman (2004) and Antràs and Helpman (2006) respectively. Section 5 concludes.

2. The Boundary of the Firm and the Role of η (Antràs, 2003)

We begin by reviewing the salient features of the Antràs (2003) model from the perspective of the empirical work to follow. A U.S. firm produces a brand of a differentiated product. Demand is generated by CES preferences. To produce the good, the firm *must* use two inputs, those produced by the U.S. firm (h for *h*eadquarters) and those produced by a foreign supplier (m for inter*m*ediates). Output of the final good is given by a Cobb-Douglas production function with two key parameters: a Hicks-neutral productivity parameter θ and the cost share of the input provided by the firm η. Specifically,

$$q = \theta \left(\frac{h}{\eta}\right)^{\eta} \left(\frac{m}{1-\eta}\right)^{1-\eta} . \tag{1}$$

The two inputs are entirely customized. Customization raises quality to a threshold which allows the final good to be sold to consumers. Unfortunately for the U.S. firm and its foreign supplier, quality is not observable or contractible. This is modelled by assuming that the investments in customization are non-contractible. Equally unfortunate for the firm and its supplier, customization has no value outside of the relationship. Thus, there is a standard hold-up problem. After the investments in customization have been made there is renegotiation over how the *ex post* quasi-rents from the relationship will be shared.

Let β be the generalized Nash share of the *ex post* quasi-rents from the relationship that go to the U.S. firm. The U.S. firm receives this share plus its outside option. The role of the organizational form (vertical integration versus outsourcing) is that it alters the outside option received by the U.S. firm in the event of a bargaining breakdown in the renegotiation stage. What are the various outside options? If there is no agreement the supplier earns nothing, regardless of the organizational form. The outside option is also 0 for the U.S. firm in an outsourcing relationship. However, for a firm that has vertically integrated with its supplier, no agreement means that the firm can still produce some output by 'forcing' its now-disgruntled supplier to do at least some work. Vertical integration is therefore a way for the firm to improve its outside option in the case of bargaining breakdown.

This difference in the firm's outside options under the two organizational forms leads to a trade-off. Vertical integration allows the firm to grab a larger share of the pie, but it leads to a smaller pie because of underinvestment by the supplier. This is modelled mathematically as follows. Let $k = V,O$ subscripts denote the organizational form with V for vertical integration and O for outsourcing. Recall that β is the share of the *ex post* quasi-rents that goes to the firm. Let R_k be the revenue generated when there is an agreement. If there is no agreement the firm can only sell a portion δ of the final output. With CES preferences and constant markup $1/\alpha$, this generates a revenue of $\delta^\alpha R_V$. Therefore, the firm receives its outside option $\delta^\alpha R_V$ plus a share β of the quasi-rents $(R_V - \delta^\alpha R_V)$. That is, the firm receives $[\delta^\alpha + \beta(1-\delta^\alpha)]R_V$. Let $\beta_V = \delta^\alpha + \beta(1-\delta^\alpha)$ be the firm's share of revenues under vertical integration. Under outsourcing, the outside option is 0 and the quasi rents are R_O so that the firm receives $0 + \beta(R_O - 0) = \beta R_O$. Let $\beta_O = \beta$ be the firm's share of total revenues under outsourcing. The upshot of all this is the central result that the organizational form alters the U.S. firm's share of revenue. In particular, $\beta_V > \beta_O$.

The timing of the game played by the U.S. firm and its foreign supplier is simple. The two match and the U.S. firm chooses the organizational form. Then investments in customized inputs are made. Finally, the initial contract is renegotiated and, if there is agreement, the product is sold.

Both the U.S. firm and the foreign supplier invest and hence each must worry about the other's underinvestment. Where η is large, the surplus generated by the relationship is particularly sensitive to the amount of investment undertaken by the U.S. firm. To reduce the degree of underinvestment by the U.S. firm, the firm must be given a large share of the revenue. This share is largest under vertical integration because $\beta_V > \beta_O$. This is a specific instance of the Grossman and Hart (1986) property-rights theory of the firm where residual control rights are allocated to the U.S. firm. In contrast, when η is small, the surplus generated by the relationship is particularly sensitive to the amount of investment undertaken by the supplier. To reduce supplier underinvestment, the supplier must be given a large share of the revenue. Outsourcing accomplishes this because $1 - \beta_O > 1 - \beta_V$.

Proposition 1 of Antràs (2003) shows that there is a unique value of η — call it η_c — such that the U.S. firm prefers vertical integration for $\eta > \eta_c$ and prefers outsourcing otherwise.

Hypothesis 1 *The exists a unique cut-off η_c with the following property. If $\eta > \eta_c$ then the firm will vertically integrate with the supplier. If $\eta < \eta_c$ then the firm will outsource from the supplier.*

We will refer to this dependence of organizational form on η as the 'Antràs effect.'

Data Sources

To investigate hypothesis 1 we use data on intra-firm and total trade from the U.S. Census Bureau. Importers bringing goods into the United States are required by law to report whether or not the transaction is with a related party. This information allows us to identify whether imports are intra-firm (related party) or at arm's-length (non-related party). See the appendix for details. The trade data are at the 6-digit Harmonized System (HS6) level for the years 2000 and 2005. We are grateful to Andy Bernard for drawing our attention to these data. See Bernard, Jensen and Schott (2005) for an example of how the data have been used.

Our key dependent variable is intra-firm imports as a share of total U.S. imports. Let g index industries and let M_{Vg} be the value of intra-firm U.S. imports in industry g. The V subscript is for vertical integration. Let M_{Og} be the value of arm's-length U.S. imports in industry g. The O subscript is for outsourcing. $M_{Vg} + M_{Og}$ is total U.S. imports and

$$\frac{M_{Vg}}{M_{Vg} + M_{Og}} \tag{2}$$

is intra-firm imports as a share of total U.S. imports of good g.

A drawback of our Census data relative to the Bureau of Economic Analysis (BEA) data on multinationals is that we do not know whether the U.S. importer is the U.S.-owned parent or the foreign-owned affiliate. To address this, we also report results based on a restricted sample of countries. A country is included in the restricted sample if at least two-thirds of intra-firm U.S. imports from the country are imported by U.S. parents. For example, only 3% of U.S. intra-firm imports from Japan are imports by U.S. parents from their foreign affiliates so that Japan is excluded from the restricted sample. Data on intra-firm U.S. imports by country and parent (U.S. versus foreign) are from Zeile (2003) and pertain to 1997. The countries in the full sample are reported in appendix tables 5 and 6. Countries not in the restricted sample are marked with an asterisk.[2]

Data on the inputs provided by the U.S. firm are from the Bartelsman and Gray (1996) data base. For each U.S. 4-digit SIC industry in 1996, the database provides information on capital K_g, employment L_g, capital intensity $\ln K_g/L_g$, nonproduction workers S_g, and skill intensity $\ln S_g/L_g$

[2]It is evident from the table that the countries for which a large share of intra-firm imports are imports from foreign-owned parents are higher income countries. An alternative strategy therefore is to restrict high-income countries from the sample. We have also done this using data are from the *World Development Indicators*. Although we do not report the results here, they are qualitatively identical to the results we obtain from either our full sample or from our restricted sample.

(as in Berman, Bound and Griliches 1994). Note that we use 1996 industry-level data, but the trade data are for 2000 and 2005. In translating the HS6 data into 4-digit SIC data we only keep HS6 codes that go into a unique SIC code. As a result, we are left with 370 of the 400+ possible SIC codes.

Examining Hypothesis 1 (Antràs, 2003): Cross-Industry Analysis

Antràs (2003) examined hypothesis 1 using BEA data on intra-firm U.S. imports as a share of total U.S. imports. He related this share to capital intensity, a proxy for η. Following Antràs and Helpman (2004) we might also want to proxy η by skill intensity. Antràs (2003) worked at the 2-digit SIC level with 28 industries. We start by examining his relationship using the Census data with its 370 industries. In particular, we consider the following cross-industry regression:

$$\frac{M_{Vg}}{M_{Vg}+M_{Og}} = \gamma_0 + \gamma_{K/L} \ln K_g/L_g + \gamma_{S/L} \ln S_g/L_g + \varepsilon_g \tag{3}$$

where $\ln K_g/L_g$ is capital intensity and $\ln S_g/L_g$ is skill intensity. We estimate equation (3) separately for 2000 and 2005, the two years for which the data are available to us.

Estimates of equation (3) appear in table 1. Column 1 reports estimates using trade data from 2000 and column 2 uses data from 2005. Using data from either year, both variables are positive and statistically significant.[3] The capital intensity result confirms the findings of Antràs (2003) and Yeaple (2006) for our sample.[4] Antràs (2003) and Yeaple (2006) do not consider skill intensity: we find that skill-intensive industries tend to import more within firm boundaries. In contrast, the Antràs and Yeaple studies find that the share of intra-firm imports tends to be higher in R&D-intensive industries. At our level of industrial disaggregation there are no measures of R&D intensity. It is possible that our skill-intensity variable is picking up the importance of R&D for the integration versus outsourcing decision.

Columns 3 and 4 report the estimates using our restricted sample of exporting countries to construct industry aggregates. The sample consists of countries for which intra-firm U.S. imports are dominated by U.S. parents. As shown, the results are similar to the results for the larger sample, except that the estimated magnitudes of the coefficients are smaller.

[3] The results are similar if either capital intensity and skill intensity enter separately in the estimating equation.

[4] To remind the reader, the differences between our estimates and those of the Antràs and Yeaple studies are that (1) our data have 370 industries, much more than the 23 used by Antràs and the 51 used by Yeaple and (2) we have import data for 210 exporting countries, much more than the 28 exporters considered by Antràs and the 58 exporters considered by Yeaple. In addition, Yeaple uses only U.S. parents whereas we are forced to include foreign parents. Our restricted sample eliminates most of the foreign parents.

Table 1. The determinants of the share of intra-firm imports in total imports – by industry.

	Full sample		Restricted sample	
	2000	2005	2000	2005
Capital intensity, $\ln K_g/L_g$	.264**	.295**	.169**	.188**
	(.050)	(.049)	(.053)	(.053)
Skill intensity, $\ln S_g/L_g$	.199**	.219**	.105*	.105*
	(.050)	(.049)	(.053)	(.053)
R-squared	.14	.17	.05	.06
Number of observations	370	370	367	367

Notes: This table reports estimates of (3). The dependent variable is $M_{Vg}/(M_{Vg}+M_{Og})$, U.S. intra-firm imports as a share of total U.S. imports. 4-digit SIC industries are the unit of observation. Standardized 'beta' coefficients are reported. Robust standard errors appear in parenthesis. ** and * indicate significance at the 1 and 5 percent levels, respectively. See section 2 'Data Sources' for a description of the restricted sample.

Because we report standardized 'beta' coefficients, one can easily assess and compare the magnitudes of the coefficients for the capital and skill measures. The coefficients for capital suggest that a one standard deviation increase in capital results in a .169 to .295 standard deviation increase in the share of intra-firm imports. This is an economically large effect. The estimated coefficient for skill is somewhat smaller, ranging from .105 to .219.

A problem with the above approach is that it assumes that we can aggregate across exporting countries. Yet as Schott (2004) notes, this may be seriously misleading because an HS6 good produced in a poor country may be very different from an HS6 good produced in a rich country. To address this, we now turn to a different approach than that of Antràs (2003) and Yeaple (2006). Let M_{Vgc} be the value of U.S. intra-firm imports of good g that are imported from country c. Let M_{Ogc} be the corresponding value of arm's-length U.S. imports. Then $M_{Vgc}+M_{Ogc}$ is total U.S. imports of good g from country c and $M_{Vgc}/(M_{Vgc}+M_{Ogc})$ is the intra-firm import share of good g imported from country c. We estimate a regression that pools across industries and countries:

$$\frac{M_{Vgc}}{M_{Vgc}+M_{Ogc}} = \gamma_c + \gamma_{K/L}\ln K_g/L_g + \gamma_{S/L}\ln S_g/L_g + \varepsilon_{gc}. \tag{4}$$

In this regression we control for exporter heterogeneity by allowing for country fixed effects γ_c. Further, g now subscripts HS6 products rather than 4-digit SIC industries.

Table 2. The determinants of the share of intra-firm imports in total imports – by country and industry.

	Full sample		Restricted sample	
	2000	2005	2000	2005
Capital intensity, $\ln K_g/L_g$	.073**	.061**	.055**	.058**
	(.019)	(.016)	(.017)	(.017)
Skill intensity, $\ln S_g/L_g$	.085**	.079**	.077**	.075**
	(.017)	(.015)	(.019)	(.016)
Country fixed effects	Yes	Yes	Yes	Yes
R-squared	.12	.12	.10	.10
Number of observations	110,355	115,781	38,229	41,790

Notes: This table reports estimates of equation (4). The dependent variable is $M_{Vgc}/(M_{Vgc}+M_{Ogc})$, U.S. intra-firm imports as a share of total U.S. imports. An observation is an HS6-country pair. Standardized 'beta' coefficients are reported. Standard errors clustered at the 4-digit SIC industry level appear in parenthesis. ** and * indicate significance at the 1 and 5 percent levels, respectively. All regressions include country fixed effects. See section 2 'Data Sources' for a description of the restricted sample.

The results appear in table 2. Because our variables of interest $\ln K_g/L_g$ and $\ln S_g/L_g$ only vary at the 4-digit SIC industry level while the unit of observation is a country and HS6 good, we report standard errors clustered at the 4-digit SIC level. The estimates, consistent with the cross-industry estimates of table 1, show that the capital and skill intensity of an industry are positively correlated with the share of intra-firm trade. This is true in both 2000 and 2005. It is also true for the full sample and the restricted sample of exporting countries. Overall, these results combined with the cross-industry results of table 1 provide considerable support for hypothesis 1.

3. Productivity Heterogeneity (Antràs and Helpman, 2004)

Examining the data on intra-firm imports one finds that $M_{Vgc}/\left(M_{Vgc}+M_{Ogc}\right)$ is rarely either 0 or 1 as predicted by the Antràs theory. An obvious explanation is aggregation bias. However, simple versions of aggregation bias are not consistent with the data. For one, Bernard et al. (2005) show that even though only one third of U.S. trade is done within the firm, U.S. multinationals account for 90% of all U.S. trade. Thus, it would seem that individual multinationals trade both within the firm and between firms. Further, Bernard, Jensen and Schott (2006) show that at the 10-digit HS level (HS10), multinationals typically conduct both intra-firm and arm's-length trade, sometimes even to the same destination country. Feinberg and Keane (2006) offer additional evidence on this subject. In their study of trade between U.S. firms and their Canadian affiliates they find the following. Only 12% of these relationships are pure horizontal relationships and only 19% are pure vertical relationships. Fully 69% of these relationships involve two-way trade flows, what Helpman (2006) calls 'complex integration strategies'. Thus, simple aggregation bias cannot be the entire story. An explanation in terms of heterogeneity seems more likely.

In particular, Antràs and Helpman (2004) introduce productivity heterogeneity in order to generate entirely novel predictions about the determinants of the share of imports that are intra-firm. To this end they append the Antràs framework onto the Melitz (2003) model. Productivity heterogeneity means that θ in equation (1) varies across firm-supplier relationships. Let $\overline{\pi}_k$ be variable profits for a firm with $\theta = 1$ that uses organizational form $k = V, O$. Then as is well known, profits for a firm with productivity θ that adopts organizational form k are linear in $\theta^{\alpha/(1-\alpha)}$:

$$\pi_k(\theta) = \theta^{\alpha/(1-\alpha)}\overline{\pi}_k - F_k \tag{5}$$

where F_k is the fixed costs of offshoring using organizational form k. Antràs (2003) assumes that all firms have the same productivity ($\theta = 1$) and the same fixed costs ($F_V = F_O$). Under these assumptions the Antràs effect states that $\overline{\pi}_V > \overline{\pi}_O$ if and only if $\eta > \eta_c$ i.e., the U.S. firm prefers vertical integration to outsourcing when the firm's share of inputs is large. Heterogeneity of productivity by itself does not alter this conclusion – it simply magnifies the advantages (or disadvantages) of vertical integration.[5]

However, when the fixed costs of outsourcing vary across organizational forms then productivity heterogeneity matters. How heterogeneity matters depends on whether $F_V - F_O$ is positive or negative. Since the results depend transparently on the sign of $F_V - F_O$, we only present the results under Antràs and Helpman's preferred assumption, namely, $F_V > F_O$. They reason that vertical integration creates a need to supervise the production of intermediate inputs, thus creating managerial overload.

Figure 1 illustrates what happens when heterogeneity is introduced. The figure plots profits under outsourcing $\pi_O(\theta)$ and vertical integration $\pi_V(\theta)$. From Antràs (2003, lemma 3), we know that $\overline{\pi}_V/\overline{\pi}_O$ is increasing in η and equals 1 for $\eta = \eta_c$. This together with equation (5) implies that $\pi_V(\theta)$ is steeper than $\pi_O(\theta)$ for $\eta > \eta_c$ and flatter for $\eta < \eta_c$. From the left-hand panel of figure 1 where $\eta < \eta_c$ and $F_V > F_O$, it must be that outsourcing is always preferred to vertical integration — the Antràs effect and the lower fixed costs of outsourcing both work in favour of outsourcing.

[5] That is, if $F_V = F_O$ then $\pi_V(\theta) > \pi_O(\theta) \Leftrightarrow \overline{\pi}_V > \overline{\pi}_O$ and it remain true that the firm prefers vertical integration if and only if $\eta > \eta_c$.

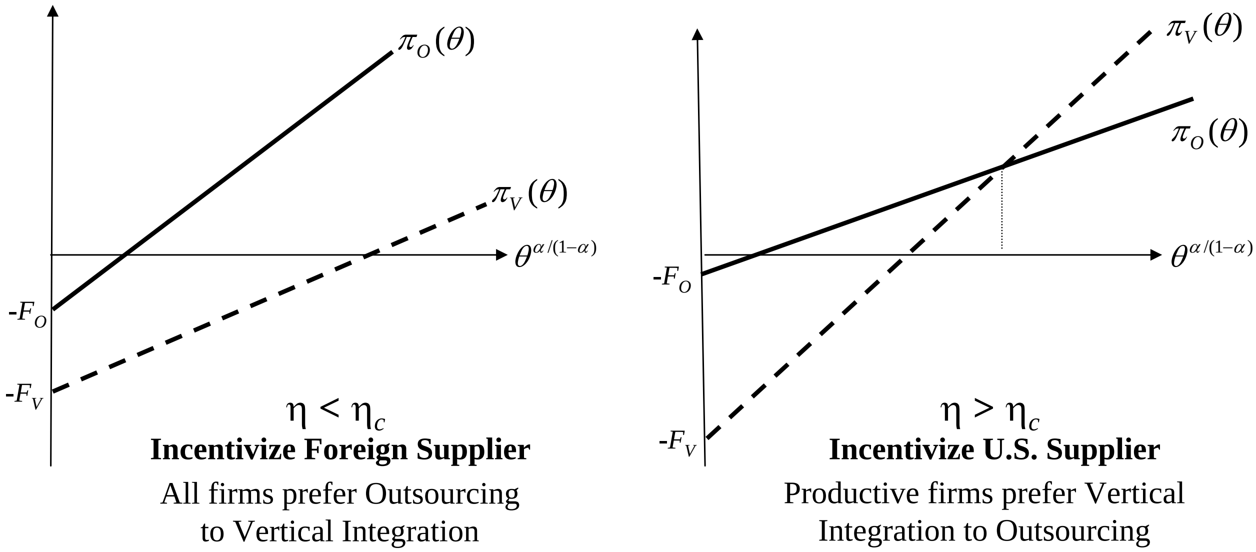

Figure 1. The Outsourcing Decision with Productivity Heterogeneity

When $\eta > \eta_c$, as in the right-hand panel of figure 1, $\pi_V(\theta)$ is steeper than $\pi_O(\theta)$. It follows that the two curves must cross. Firms with productivity to the right of the crossing point will vertically integrate. Firms to the left will outsource. The tension here is that fixed costs push for outsourcing while the Antràs effect pushes for vertical integration. Since the Antràs effect is greatest for the most productive firms, the Antràs effect dominates for productive firms.

All of this leads to an interesting empirical prediction about the share of U.S. imports that are intra-firm i.e., about $M_{Vgc}/(M_{Vgc} + M_{Ogc})$. The share should depend on an interaction of η with θ. When $\eta < \eta_c$, we have that $M_{Vgc} = 0$ so that an increase in θ has no effect on $M_{Vgc}/(M_{Vgc} + M_{Ogc}) = 0$. Where $\eta > \eta_c$, an increase in θ increases $M_{Vgc}/(M_{Vgc} + M_{Ogc})$. This suggests that in industries with firms that tend to have large values of θ we should see a larger share of firms that vertically integrate.

Antràs and Helpman (2004) formalize this by assuming that the distribution of productivities θ within an industry is described by the Pareto distribution i.e., by the cumulative distribution function

$$G(\theta) = 1 - \left(\frac{\theta}{b}\right)^{\lambda}, \quad \theta \geq b > 2. \tag{6}$$

Note that for the Pareto distribution the mean and variance of θ are increasing in λ. Consider the case where $\eta > \eta_c$ as in the right-hand panel of figure 1. Antràs and Helpman (2004, page 573) show that an increase in λ (a rise in the mean and variance of θ) leads to a rise in the share of firms that vertically integrate. Empirically, this means a rise in $M_{Vgc}/(M_{Vgc}+M_{Ogc})$. On the other hand, when $\eta < \eta_c$ as in the left-hand panel of figure 1, an increase in λ has no effect on $M_{Vgc}/(M_{Vgc}+M_{Ogc})$ which is 0 independent of λ.

Moving to an empirical counterpart of λ, let x_g^θ be some measure of productivity dispersion in industry g. We will describe x_g^θ in detail shortly. Helpman, Melitz and Yeaple (2004) and Yeaple (2006) use productivity dispersion x_g^θ as a measure of λ and we follow suit. We can now offer an empirically-oriented hypothesis that comes out of the Antràs-Helpman model.

Hypothesis 2 *Assume $F_V > F_O$. Let x_g^θ be the dispersion of θ across firms within industry g.*

(a) If $\eta < \eta_c$ then dispersion does not affect the intra-firm share of imports: $\frac{\partial M_{Vgc}/(M_{Vgc}+M_{Ogc})}{\partial x_g^\theta} = 0$.

(b) If $\eta > \eta_c$ then dispersion increases the intra-firm share of imports: $\frac{\partial M_{Vgc}/(M_{Vgc}+M_{Ogc})}{\partial x_g^\theta} > 0$.

Examining Hypothesis 2 (Antràs & Helpman, 2004)

To construct our measure of productivity dispersion x_g^θ we follow Helpman et al. (2004) and Yeaple (2006) as closely as possibly given the more limited data available to us. With CES utility and the production function that we have been using, more productive firms have larger sales and exports. Using firm sales as a measure of firm productivity, Helpman et al. (2004) and Yeaple (2006) construct estimates of the dispersion of firm productivity using the standard deviation of firm sales across all firms within an industry. (Their level of industry aggregation allows for only 51 or 52 industries.) We do not have firm-level data. Instead, we construct sales of 'notional' firms from U.S. export data as follows. Let g index HS6 goods. Let $v(g)$ be an HS10 good that feeds into HS6 good g. v stands for the $v$ariety of the HS6 good. Let c index the destination country for U.S. exports. Let l index the location within the U.S. from which the exports were shipped. We define an industry as an HS6 product g and we define the sales of a notional firm as the exports of an HS10 good $v(g)$ exported from U.S. location l to destination country c. Let $X_{v(g),l,c}$ be these exports. Our measure of productivity dispersion within industry g is just the standard deviation of $\ln X_{v(g),l,c}$:

$$x_g^\theta \equiv \sqrt{\mathbf{V}(\ln X_{v(g),l,c})} \tag{7}$$

where $\mathbf{V}$ is the variance operator. The variance is calculated across all triplets $(v(g),l,c)$ that feed into HS6 industry g.[6]

As a robustness check we also use a second method of constructing the sales of a notional firm. Since a single firm may export to more than one country, we aggregate across export destinations and define sales of a notional firm as exports $X_{v(g),l} \equiv \Sigma_c X_{(v(g),l,c)}$. We then construct dispersion as[7]

$$x_g^\theta \equiv \sqrt{\mathbf{V}(\ln X_{v(g),l})} \tag{8}$$

To test hypothesis 2 we estimate how the relationship between productivity dispersion x_g^θ and intra-firm imports varies with headquarter intensity. We do not expect the second derivative to be linear. It will be 0 when η is small and positive when η is large. As well, we do not know where the cut-off level η_c will be. Because of this, we pursue the following estimation strategy. Rank our 370 4-digit SIC industries by headquarters intensity. Headquarters intensity will be measured by either skill intensity or capital intensity. Based on this ranking, divide the 370 industries into five quintiles of 74 industries each. Let $p = 1, \ldots, 5$ index quintiles, with $p = 1$ being the least headquarters-intensive quintile. Finally, let $I_{gp}^\eta = 1$ if industry g is in quintile p and $I_{gp}^\eta = 0$ otherwise.

We consider a regression that allows the relationship between dispersion and intra-firm imports to differ by quintile:

$$\frac{M_{Vgc}}{M_{Vgc} + M_{Ogc}} = \gamma_c + \gamma_{K/L} \ln K_g/L_g + \gamma_{S/L} \ln S_g/L_g + \sum_{p=1}^{5} \gamma_{\eta p}\, I_{gp}^\eta + \sum_{p=1}^{5} \gamma_{\theta\eta p}\, (x_g^\theta \cdot I_{gp}^\eta) + \varepsilon_{gc}. \tag{9}$$

The primary coefficients of interest are the $\gamma_{\theta\eta p}$. Hypothesis 2 states that for low η and hence low p the impact of dispersion should be zero. That is, $\gamma_{\theta\eta p} = 0$ for low p. Hypothesis 2 also states that for high η and hence high p the impact of dispersion should be positive. That is, $\gamma_{\theta\eta p} > 0$ for high p. Since we do not know which quintile p contains the cut-off η_c we cannot be more precise about what 'low' and 'high' p means. We will let the data answer this.

[6]See Helpman et al. (2004, page 307) for an explanation of how the standard deviation of the log of firm sales recovers the parameter λ.

[7]Alternatively, one could approximate a notional firm's exports by exports of an HS10 variety to a particular country from any location in the United States. This would be a more appropriate measure if firms exported goods from multiple locations in the United States. Using this alternative measure produces results that are qualitatively identical to what we report below.

The equation (9) regression also includes capital intensity, skill intensity, headquarter-intensity dummies (I^{η}_{gp}) and country fixed effects.

Table 3 reports results for 2005 intra-firm trade data. The results for 2000 are similar. In columns 1-3 we measure productivity dispersion using equation (7). In columns 4-7 we measure productivity dispersion using equation (8). In the first column we examine the effect of productivity dispersion averaged across all industries, regardless of their headquarter intensities. That is, we impose that all five $\gamma_{\theta\eta p}$ are equal. The average effect is positive and highly significant. The estimated coefficient of .09 suggests that a one standard deviation increase in the dispersion measure increases the proportion of within-firm imports by 2.9 percentage points. Because our variable of interest x^{θ}_{g} is an industry-specific measure, we include country fixed effects and adjust the estimated standard errors for clustering across countries within each HS6 industry.

In the second and third columns, we allow the effect of dispersion to differ depending on the headquarter intensity of the industries. In column 2 we measure headquarter intensity by skill intensity S_g/L_g. In column 3 we measure headquarter intensity by capital intensity K_g/L_g. Consistent with hypothesis 2 we observe a significant jump in the magnitude of the estimated coefficient when moving from the first quintile to the second quintile. After this the estimated coefficient remains essentially constant. That is, we find that $0 \leq \hat{\gamma}_{\theta\eta 1} < \hat{\gamma}_{\theta\eta 2} \approx \hat{\gamma}_{\theta\eta 3} \approx \hat{\gamma}_{\theta\eta 4} \approx \hat{\gamma}_{\theta\eta 5}$. F-tests cannot reject the null hypothesis of the equality of any pair of coefficients among $\hat{\gamma}_{\theta\eta 2}$, $\hat{\gamma}_{\theta\eta 3}$, $\hat{\gamma}_{\theta\eta 4}$, or $\hat{\gamma}_{\theta\eta 5}$. However, F-tests do reject the null hypothesis of equality between $\hat{\gamma}_{\theta\eta 1}$ and either $\hat{\gamma}_{\theta\eta 2}$, $\hat{\gamma}_{\theta\eta 3}$, $\hat{\gamma}_{\theta\eta 4}$, or $\hat{\gamma}_{\theta\eta 5}$. These results provide dramatic confirmation of hypothesis 2. They also suggest that the η_c cut-off is relatively low, somewhere near the first and second quintiles of the distribution of headquarter intensity.

We further examine the robustness of our results in columns 4 to 6 where we report the same results using the alternative measure of productivity dispersion from equation (8). The results using the two different measures of productivity dispersion are very similar.

In the final column of the table, we show that the results are similar when we use our restricted sample of countries. The sample ensures that most of the intra-firm imports are done by firms with U.S. parents. The final column reports estimates using the same specification as in the adjacent column 6. For the other specifications, the restricted sample estimates are also similar to the full sample estimates.

Overall, the evidence from table 3 provides support for hypothesis 2. We find that there is

Table 3. Productivity dispersion as a determinant of the share of intra-firm imports in total imports.

	Productivity dispersion calculated using $X_{v(g),l,c}$			Productivity dispersion calculated using $X_{v(g),l}$			
Headquarter intensity measure:		S_g/L_g	K_g/L_g		S_g/L_g	K_g/L_g	K_g/L_g
Productivity dispersion, x_g^θ	.09**			.08**			
	(.006)			(.005)			
Prod. dispersion interactions:							
Low η, $(x_g^\theta \cdot I_{g1}^\eta)$		.03**	.03**		.04**	.01	.04**
		(.010)	(.009)		(.011)	(.011)	(.011)
Low-mid η, $(x_g^\theta \cdot I_{g2}^\eta)$		.11**	.13**		.10**	.10**	.11**
		(.014)	(.083)		(.009)	(.013)	(.022)
Mid η, $(x_g^\theta \cdot I_{g3}^\eta)$		.11**	.12**		.10**	.11**	.12**
		(.016)	(.014)		(.012)	(.010)	(.017)
Mid-high η, $(x_g^\theta \cdot I_{g4}^\eta)$		.11**	.09**		.09**	.08**	.07**
		(.015)	(.015)		(.011)	(.010)	(.016)
High η, $(x_g^\theta \cdot I_{g5}^\eta)$		.11**	.11**		.09**	.10**	.13**
		(.014)	(.013)		(.011)	(.011)	(.017)
Skill, capital intensity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Headquarter indicators I_{gp}^η	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.12	.13	.13	.12	.13	.13	.11
Number of observations	115,436	115,436	115,436	115,433	115,433	115,433	41,682

Notes: This table reports estimates of equation (9). The dependent variable is $M_{Vgc}/(M_{Vgc}+M_{Ogc})$, the share of U.S. intra-firm imports in total U.S. imports. The unit of observation is an HS6-country pair in 2005. x_g^θ is a measure of productivity dispersion across 'notional' firms in industry g. For the definition of x_g^θ see equation (7) for columns 1-3 and equation (8) for columns 4-7. I_{gp}^η for $p = 1, \dots, 5$ is a dummy for whether the industry's headquarter intensity is in the p^{th} quintile of the headquarter intensity distribution. Headquarter intensity is measured by either skill or capital intensity — see the column headings. All regressions include controls for capital intensity $\ln K_g/L_g$, skill intensity $\ln S_g/L_g$, five headquarter intensity indicator variables I_{gp}^η, and country fixed effects. ** and * indicate significance at the 1 and 5 percent significance levels, respectively. The final column uses the restricted sample that is described in section 2 under 'Data Sources'.

indeed a cut-off level of headquarter intensity. For industries with headquarter intensity greater than this cut-off, productivity dispersion increases the share of intra-firm imports. For industries with headquarter intensity below the cut-off, the estimated relationship is much weaker and close to zero. Because the cut-off is quite low, this suggests that the empirically relevant case from Antràs and Helpman (2004) is the high-η specification.

4. Partially Incomplete Contracting (Antràs & Helpman, 2006)

Antràs and Helpman (2006) introduce the possibility that the inputs supplied by the U.S. firm and its foreign supplier are partially contractible. The input of the foreign supplier (m in equation 1) is assumed to be produced from a continuum of activities. Only a fraction μ_m^S of these activities are contractible. The remaining activities are not contractible. Likewise, the input of the U.S. firm (h in equation 1) is assumed to be produced from a continuum of activities and only a fraction μ_h^S of these activities are contractible. With this set-up all of the previous results carry through, though the cut-off η_c shifts somewhat. Let η_c' be the new cut-off.

The introduction of partial contractibility offers many new insights, one of which deals with the intra-firm share of U.S. imports $M_{Vgc}/(M_{Vgc} + M_{Ogc})$. Antràs and Helpman (2006) show that as the contractibility of the foreign supplier's inputs improves (μ_m^S rises), there are two offsetting effects on the intra-firm share of U.S. imports. To understand the first effect we must introduce the possibility that the U.S. firm can produce in the United States and not just abroad. This is shown in the top panel of figure 2 where we are assuming $\eta > \eta_c'$. (If $\eta < \eta_c'$ then $M_{Vgc} = 0$ as in hypotheses 1 and 2.) There are now two productivity cut-offs. One determines whether the U.S. firm will produce abroad or in the United States. This cut-off occurs where the profits from foreign outsourcing are zero: $\pi_O(\theta) = 0$. The second cut-off is the one we saw in the last section and demarcates the decision whether to outsource abroad or vertically integrate abroad. It occurs where the profits from vertical integration and outsourcing are equal: $\pi_V(\theta) = \pi_O(\theta) = 0$.

Now consider the effects of an improvement in the contractibility of foreign intermediate inputs. This raises the profitability of foreign operations and, in particular, increases the term $\overline{\pi}_k$ in equation (5). Since $\overline{\pi}_k$ is the slope of the figure 2 profit functions $\pi_k(\theta)$, an improvement in foreign contractibility makes the lines in figure 2 rotate counter-clockwise around their fixed vertical intercepts. This is shown in the bottom panel of figure 2. As a result of the improvement in

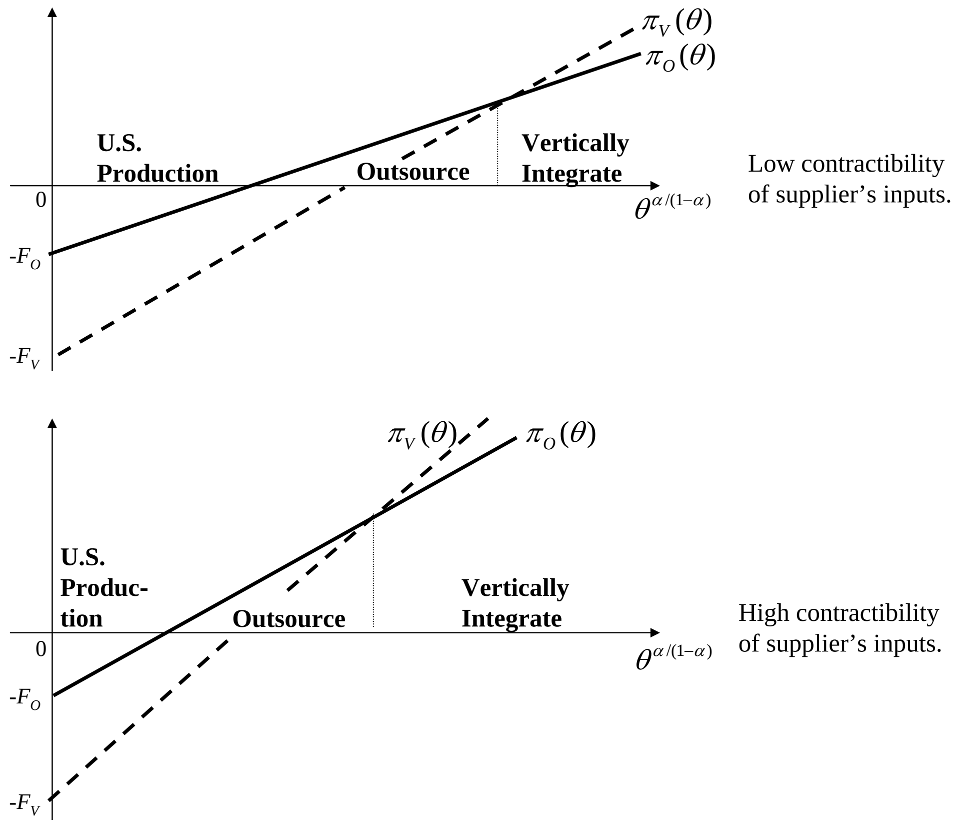

Figure 2. Improved Contractibility of the Foreign Supplier's Inputs

foreign contractibility the two cut-offs shift left. This impacts the intra-firm share of U.S. imports in two ways.

Hypothesis 3 *Assume $F_V > F_O$ and consider an increase in the contractibility of the foreign supplier's inputs (a rise in μ_m^S).*

1. *If $\eta > \eta_c'$ then there are two offsetting effects of improved contractibility on the intra-firm share of U.S. imports.*

   (a) *The 'Standard Effect': U.S. production migrates abroad where it is outsourced. This increases arm's-length U.S. imports and thus lowers the intra-firm share of U.S. imports.*

   (b) *The 'Surprise Effect': The most productive outsourcing relationships become vertically integrated into the U.S. firm's organization. For these relationships, intra-firm imports increase at the expense of arm's-length imports, thus raising the intra-firm share of U.S. imports.*

2. *If $\eta < \eta_c'$ then the intra-firm share of U.S. imports $M_{Vgc}/(M_{Vgc} + M_{Ogc}) = 0$ independent of the contractibility of the supplier's inputs.*

The surprising effect is, well, surprising. However, once stated the insight is simple. With a better contracting environment for inputs produced by the foreign supplier, it is no longer as important to incentivize the foreign supplier. This makes it *relatively* more important to incentivize the U.S. firm. Thus, improved contracting abroad leads the U.S. firm to internalize previously outsourced activities.

Antràs and Helpman (2006) show that the sum of the standard and surprise effects can push the intra-firm share of U.S. imports either up or down. This means that in a regression of the intra-firm share of U.S. imports on foreign contracting institutions, a positive coefficient indicates that the surprise effect is more important than the standard effect while a negative coefficient indicates the opposite.[8]

Examining Hypothesis 3 (Antràs and Helpman, 2006)

In the Antràs and Helpman (2006) model, μ_m^S measures the proportion of contractible "activities" involved in producing the input. We interpret "activities" as the components or intermediate inputs used in the production of the input. Following this logic, we use Nunn's (2007) measure of the proportion of each industry's intermediate inputs that are relationship-specific and therefore susceptible to potential contracting problems. Denote Nunn's measure by z_g. Because we want a measure that is increasing in the completeness of contracts, we use one minus the fraction of inputs that are relationship-specific, which we denote by $(1 - z_g)$.

The relationship-specificity measure from Nunn (2007) provides variation across industries in contractibility, but in Antràs and Helpman's model contractibility varies across industries and countries. We capture variation in contracting across countries using the 'rule of law' variable from the *Governance Matters V* database (Kaufmann, Kraay and Mastruzzi, 2006). The variable measures the enforcement of contracts and the overall quality of a country's legal system. The original measure ranges from -2.5 to $+2.5$ and is increasing in the quality of the contracting environment. We normalize the measure to be between 0 and 1 (like the industry measure z_g) by adding 2.5 and dividing by 5. We denote the normalized variable by r_c.

[8] An alternative interpretation that one can take from the Antràs and Helpman (2006) model is that improved contracting institutions increase μ_m^S, but increase μ_h^S even more. In this case, the standard effect leads to a rise in the intra-firm import share and the surprise effect leads to a fall in the intra-firm import share. In our view, improved contracting for developing countries is best thought of in terms of a disproportional increase in μ_m^S. We thus ignore this alternative interpretation.

To capture the interaction between the relationship-specificity of industries and the overall contracting environment in the country we use the interaction of the industry and country measures:

$$x^{\mu}_{gc} \equiv r_c \times (1 - z_g). \tag{10}$$

This provides a country- and industry-specific measure of contractual completeness that captures in a simple way the notion that both the inherent contractibility of an industry's production process and the overall quality of a country's contracting environment matter. It also captures the intuition that a country's judicial quality matters more in industries with relationship-specific, customized intermediate inputs.

Hypothesis 3 states that in a regression of the intra-firm share of U.S. imports on x^{μ}_{gc} the coefficient should vary across industries. In industries with $\eta < \eta'_c$ the coefficient should be zero. In industries with $\eta > \eta'_c$ the coefficient should be positive or negative depending on the size of the standard and surprise effects. To model this non-linearity we proceed as in the previous section. We group industries into five quintiles based on the value η i.e., on the size of industries' skill intensities $\ln S_g/L_g$ or capital intensities $\ln K_g/L_g$. Let $p = 1, \ldots, 5$ subscripts denote the quintiles and let I^{η}_{gp} be a binary indicator of whether industry g is in quintile p.

We estimate the following equation:

$$\frac{M_{Vgc}}{M_{Vgc} + M_{Ogc}} = \gamma_g + \gamma_c + \sum_{p=1}^{5} \gamma_{\mu\eta p} \, (x^{\mu}_{gc} \cdot I^{\eta}_{gp}) + \varepsilon_{gc} \tag{11}$$

where the $(x^{\mu}_{gc} \cdot I^{\eta}_{gp})$ are the five primary regressors and the $\gamma_{\mu\eta 1}, \ldots, \gamma_{\mu\eta 5}$ are the corresponding coefficients.

The fact that x^{μ}_{gc} varies across goods and countries allows us to include good fixed effects γ_g and country fixed effects γ_c. Of course, with these fixed effects included we can no longer test hypotheses 1 and 2, which are about the industry-specific characteristics $\ln S_g/L_g$, $\ln K_g/L_g$, and x^{θ}_g. As well, because of the good fixed effects, the headquarter indicator variables I^{η}_{gp} are subsumed in the industry fixed effects.

Estimates of (11) appear in table 4.[9] Column 1 reports the average relationship between contractual completeness and intra-firm imports across all industries. We find a positive, marginally significant, relationship which implies that the 'surprise effect' slightly dominates the 'standard effect'.

[9]Note that the measure of contract-intensity is only available at the 6-digit NAICS level, not the HS6 level. We therefore adjust the standard errors for clustering across HS6 goods within 6-digit NAICS industries in our estimating equations.

Table 4. Contractual Incompleteness as a determinant of the share of intra-firm imports in total imports.

	Full sample			Restricted sample		
Headquarter intensity measured by:		S_g/L_g	K_g/L_g		S_g/L_g	K_g/L_g
Contractual completeness, x^{μ}_{gc}	.10*			.12		
	(.043)			(.071)		
Contract. compl. interacted with:						
Low η, $(x^{\mu}_{gc} \cdot I^{\eta}_{g1})$		$-.03$	$-.12$		.05	$-.13$
		(.047)	(.070)		(.077)	(.116)
Low-mid η, $(x^{\mu}_{gc} \cdot I^{\eta}_{g2})$		.01	$-.05$		.04	$-.00$
		(.048)	(.050)		(.079)	(.085)
Mid η, $(x^{\mu}_{gc} \cdot I^{\eta}_{g3})$		.11*	$-.04$*		.13	$-.02$
		(.046)	(.051)		(.079)	(.083)
Mid-high η, $(x^{\mu}_{gc} \cdot I^{\eta}_{g4})$		.09	.10*		.14	.18*
		(.047)	(.048)		(.084)	(.086)
High η, $(x^{\mu}_{gc} \cdot I^{\eta}_{g5})$		.29**	.11*		.24*	.10
		(.065)	(.045)		(.111)	(.073)
Good fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.22	.22	.22	.26	.26	.26
Number of observations	111,768	111,768	111,768	40,428	40,428	40,428

Notes: This table reports estimates of equation (11). The dependent variable is $M_{Vgc}/(M_{Vgc} + M_{Ogc})$, the share of U.S. intra-firm imports in total U.S. imports. The unit of observation is a country and HS6 good in 2005. x^{μ}_{gc} of equation (10) is a measure of contractual incompleteness for industry g in country c. I^{η}_{gp} for $p = 1, \ldots, 5$ is a dummy for whether the industry's headquarter intensity is in the p^{th} quintile of the headquarter intensity distribution. All regressions include HS6 good fixed effects and country fixed effects. Because the contracting regressor is at the NAICS level, standard errors are clustered. ** and * indicate significance at the 1 and 5 percent significance levels, respectively. The restricted sample is described in section 2, 'Data Sources'.

Columns 2 and 3 allow the estimated effects to differ by headquarter intensity. Although, the coefficients are not precisely estimated, the general pattern that emerges is that for low headquarter intensive industries there is no relationship between contractibility and intra-firm imports. For higher headquarter intensive industries, the relationship between contract completeness and intra-firm imports is positive. The relevant cut-off appears to be between the 2nd and 3rd quintiles when headquarter intensity is measured with skill intensity and between the 3rd and 4th quintiles when capital intensity is used.

Columns 4 to 6 report estimates using the restricted sample of exporting countries i.e., the sample dominated by U.S.-owned parents. The results are similar. There appears to be no relationship between contractibility and intra-firm imports in low-headquarter intensive industries, and a positive relationship in high-headquarter intensive industries. Overall, these results are perfectly consistent with Antràs and Helpman's (2006) model.

5. Conclusions

Antràs (2003) proposed that we think of the boundaries of the firm — of the choice between outsourcing on the one hand and vertical integration, foreign direct investment and multinationals on the other — in the property-rights terms of Grossman and Hart (1986). The central assumption of the Antràs approach is that vertical integration allows the U.S. firm to partially control the customized intermediate inputs produced by its foreign supplier. The central implication is that we should see vertical integration in industries that intensively use the headquarter inputs produced by the U.S. firm. We analyzed this implication using Census data on U.S. intra-firm and arm's-length imports of 5,423 products from 210 countries in 2000 and 2005. As predicted by Antràs we found that skill- and capital-intensive industries have a higher ratio of intra-firm imports to total imports. This is true even after controlling for exporter fixed effects. Our results extend those of Antràs (2003) and Yeaple (2006) who used a much smaller number of industries and countries.

Antràs and Helpman (2004) extended the original Antràs model to allow for (*a*) firm-level heterogeneity in productivities and (*b*) fixed costs that are higher for vertical integration than for outsourcing. The extension implies that the intra-firm share of U.S. imports will be highest for firms with two characteristics: (1) high headquarter intensity η and (2) high productivity θ. We found very strong evidence to support this implication.

Antràs and Helpman (2006) extended their earlier model to allow for partial contractibility of the foreign supplier's inputs. An improvement in contractibility has two effects. The standard effect is that it encourages arm's-length transactions. The surprise effect is that it makes the *relative* non-contractibility of the U.S. firm's inputs all the more pressing, thus encouraging vertical integration and more intra-firm imports. This surprise effect should only appear for the most productive firms. This is exactly what we found.

In short, this paper provides rich support for the central predictions in Antràs (2003) and Antràs and Helpman (2004, 2006) about U.S. intra-firm imports as a share of total U.S. imports.

Appendix A. Data Description

Data on intra-firm and total trade are from the U.S. Census Bureau. The trade data are at the 6-digit Harmonized System (HS6) level for the years 2000 and 2005. Each shipment imported into the United States is accompanied by a form which asks about the value of the shipment, the HS10 code and whether or not the transaction is with a related party i.e., whether or not the transaction is intra-firm or at arm's length.

Two parties are related if one owns at least 6% of the other. There is an oft-used alternative data set on intra-firm trade, namely, the BEA's multinationals database. In the BEA data a 10% ownership stake is used to define intra-firm transactions. However, the BEA data provide two pieces of evidence to suggest that a 6% threshold is large enough to ensure a controlling stake. First, only five percent of intra-firm BEA imports involve ownership positions of less than 10%. See table 11 in Mataloni Jr. and Yorgason (2006). Thus, if an ownership position is at least 6% it is likely at least 10%. Second, for a very large proportion of ownership positions in the BEA data, once the position is more than 10%, it is also more than 50%. (Authors' calculations from the BEA data available on the web.) Thus, although the threshold ownership stake is only 6% in our data, it is safe to say that in most cases it is a controlling stake.

The capital intensity $\ln K_g/L_g$ and skill intensity $\ln S_g/L_g$ measures are constructed using data from Bartelsman and Gray (1996). The data are from the United States in 1996, with industries classified at the 4-digit SIC87 level.

We use U.S. export data to construct productivity dispersion measures x_g^θ. The data are from the U.S. Department of Commerce CD *U.S. Exports History: Historical Summary 2000-2004*. x_g^θ is based on 2004 export data for the 2005 regressions and 2000 export data for the 2000 regressions.

The data report the value of U.S. bilateral imports and exports. The data are also disaggregated by the geographic location within the U.S. that the products are exported from. When goods are exported overland, the location is the U.S. Customs port where the surface carrier crosses the border. When goods are exported by sea or by air, then the location is the U.S. Customs port where the merchandise is loaded on to the carrier that is taking the merchandise out of the United States. Finally, when goods are exported by post, the location is the U.S. post office where the merchandise is mailed. There are 46 locations coded in the data.

The industry measures of relationship-specificity are from Nunn (2007). We use Nunn's first measure of relationship-specificity, z_i^{rs1}, and we use the most recent year for which the measure is available, 1997. The measure is classified according to the 1997 I-O classification, which is based on the 6-digit NAICS classification. The country measures of rule of law are from the *Governance Matters V* database (Kaufmann et al., 2006). This measure is from 2005.

The regressions have a maximum of 210 countries, depending on the specific equation being estimated. A list of the countries organized by per capita GDP in 2005 (from the *World Development Indicators*) is given in table 5. Table 6 reports a list of the remaining countries in the sample for which no income data are available. These countries are listed alphabetically in the table.

Table 5. Countries in the sample ordered by 2005 real per capita GDP.

Income	Country	Income	Country	Income	Country
516	Sierra Leone	2,804	Vanuatu*	9,010	Mexico
594	Malawi	2,885	India	9,101	Russia*
620	Tanzania	3,316	Indonesia*	9,140	Botswana
622	Burundi	3,318	Syria	9,444	Malaysia
648	Congo, Dem. Rep. (Zaire)	3,340	Nicaragua	9,993	Chile
663	Guinea-Bissau	3,642	Ecuador	10,286	South Africa
695	Ethiopia	3,769	Armenia	10,710	Latvia*
716	Niger	3,817	Azerbaijan	11,054	Mauritius
788	Madagascar	3,826	Jamaica	11,196	Trinidad and Tobago
807	Yemen	3,870	Egypt	11,204	Croatia*
867	Zambia	3,961	Morocco	11,567	Antigua and Barbuda
898	Eritrea	3,964	Guatemala	11,924	Poland*
899	Congo	4,034	Sri Lanka*	12,046	Lithuania*
917	Mali	4,080	Guyana	12,222	Argentina
1,003	Benin	4,241	Philippines	12,706	Saudi Arabia*
1,006	Central African Republic	4,308	Jordan	13,377	Estonia*
1,047	Kenya	4,423	Paraguay	13,439	Slovak Republic*
1,061	Nigeria	4,575	Albania*	14,024	Oman
1,074	Burkina Faso	4,633	El Salvador	15,304	Seychelles
1,104	Tajikistan	5,158	Samoa*	15,453	Hungary*
1,137	Mozambique	5,182	Swaziland	17,351	Malta*
1,160	Rwanda	5,186	Dominica	17,815	Kuwait
1,359	Uganda	5,219	Peru	17,837	Czech Republic*
1,369	Nepal*	5,264	Cape Verde	18,040	Portugal*
1,412	Togo	5,364	Lebanon	18,840	Korea, South*
1,425	Ivory Coast	5,419	China*	19,078	Bahrain
1,574	Senegal	5,554	Venezuela*	19,244	Slovenia*
1,589	Moldova*	5,575	Fiji*	20,407	Greece*
1,667	Solomon Islands*	5,812	St. Lucia	20,959	Cyprus
1,718	Uzbekistan	5,877	Ukraine	21,518	New Zealand*
1,719	Bangladesh	5,880	St. Vincent and the Gren.	22,109	United Arab Emirates
1,779	Kyrgyz Republic	6,069	Algeria	22,408	Israel*
1,783	Mauritania	6,075	Macedonia*	23,019	Spain*
1,786	Comoros	6,087	Gabon	25,804	Singapore
1,791	Sudan	6,201	Belize	25,899	Italy*
1,796	Laos*	6,406	Belarus*	26,013	Germany*
1,830	Gambia	6,463	Bosnia-Herzegovina*	26,884	Japan*
1,832	Djibouti	6,669	Colombia	26,929	France*
1,889	Mongolia*	6,689	Panama	27,150	Sweden*
1,898	Zimbabwe	6,818	Namibia	27,527	Finland*
1,921	Chad	6,838	Kazakhstan	27,876	Australia*
1,998	Cameroon	6,846	Dominican Republic	28,326	United Kingdom*
2,004	Angola	6,916	Iran	28,327	Hong Kong*
2,004	Guinea	7,125	Turkey	28,579	Belgium*
2,045	Pakistan	7,139	Tunisia	28,732	Canada
2,058	Ghana	7,233	Tonga*	29,216	Netherlands*
2,227	Cambodia*	7,372	Grenada	29,331	Denmark*
2,338	Papua New Guinea*	7,424	Bulgaria*	29,664	Austria*
2,407	Lesotho	7,435	Thailand*	30,365	Switzerland*
2,499	Bolivia	7,531	Brazil	30,376	Iceland*
2,523	Vietnam*	7,793	Romania*	35,341	Norway*
2,614	Georgia	8,658	Uruguay	35,684	Ireland
2,644	Honduras	8,714	Costa Rica	64,299	Luxembourg*

Notes: Countries are ordered by 2005 real per capita GDP from the *World Development Indicators*. An asterisk indicates that the country is not in the restricted sample of countries.

Table 6. Countries in the sample without income data, ordered alphabetically.

Afghanistan	Guadeloupe	New Caledonia*
Andorra*	Haiti	Niue*
Anguilla	Iraq	Palau*
Aruba*	Kiribati*	Qatar
Bahamas	Korea, North*	Reunion
Barbados	Liberia	San Marino*
Bermuda	Libya	Sao Tome and Principe
Bhutan*	Liechtenstein*	Serbia and Montenegro*
British Virgin Islands	Macao*	Somalia
Brunei*	Maldives*	St. Kitts and Nevis
Cayman Islands	Marshall Islands*	Suriname*
Cook Islands*	Martinique*	Taiwan*
Cuba	Micronesia*	Timor, East*
Equatorial Guinea	Monaco*	Turkmenistan
French Guiana	Myanmar*	Turks and Caicos Islands
French Polynesia*	Nauru*	Tuvalu*
Greenland	Netherlands Antilles	West Bank

Notes: Countries are ordered alphabetically. An asterisk indicates that the country is not in the restricted sample of countries.

References

Antràs, Pol, "Firms, Contracts, and Trade Structure," *Quarterly Journal of Economics*, November 2003, *118* (4), 1375–1418.

— , "Incomplete Contracts and the Product Cycle," *American Economic Review*, September 2005, *95* (4), 1054–1073.

— **and Elhanan Helpman**, "Global Sourcing," *Journal of Political Economy*, 2004, *112*, 552–580.

— **and** — , "Contractual Frictions and Global Sourcing," 2006. Mimeo, Harvard University.

Bartelsman, Eric J. and Wayne Gray, "The NBER Manufacturing Productivity Database," Technical Working Paper No. 205, National Bureau of Economic Research, October 1996.

Berman, Eli, John Bound, and Zvi Griliches, "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures," *Quarterly Journal of Economics*, 1994, *109*, 367–397.

Bernard, Andrew B. and J. Bradford Jensen, "Exporters, Skill Upgrading, and the Wage Gap," *Journal of International Economics*, 1997, *42* (1–2), 3–31.

— **,** — **, and Peter Schott**, "Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods," Working Paper No. 11404, National Bureau of Economic Research, 2005.

— **,** — **, and** — , "Transfer Pricing by U.S.-Based Multinational Firms," Working Paper No. 12493, National Bureau of Economic Research, 2006.

Feenstra, Robert C. and Gordon H. Hanson, "Ownership and Control in Outsourcing to China: Estimating the Property-Rights Theory of the Firm," *Quarterly Journal of Economics*, May 2005, *120* (2), 729–761.

Feinberg, Susan E. and Michael P. Keane, "Accounting for the Growth of MNC-Based Trade Using a Structural Model of U.S. MNCs," *American Economic Review*, December 2006, *96* (5), 1515–1558.

Grossman, Gene M. and Elhanan Helpman, "Integration versus Outsourcing in Industry Equilibrium," *Quarterly Journal Economics*, 2002, *117*, 85–120.

— **and** — , "Outsourcing versus FDI in Industry Equilibrium," *Journal of the European Economic Association*, 2003, *1*, 317–327.

— **and** — , "Managerial Incentives and the International Organization of Production," *Journal of International Economics*, 2004, *63*, 237–262.

— **and** — , "Outsourcing in a Global Economy," *Review of Economic Studies*, 2005, *72*, 135–159.

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

Helpman, Elhanan, "Trade, FDI and the Organization of Firms," *Journal of Economic Literature*, 2006, *44*, 589–630.

— **and Paul R. Krugman**, *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*, Cambridge, MA: MIT Press, 1985.

— , **Marc J. Melitz, and Stephen R. Yeaple**, "Export versus FDI with Heterogeneous Firms," *American Economic Review*, 2004, *94* (1), 300–316.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "Governance Matters V: Governance Indicators for 1996Ű2005," Working Paper, World Bank 2006.

Mataloni Jr., Raymond J. and Daniel R. Yorgason, "Operations of U.S. Multinational Companies: Preliminary Results from the 2004 Benchmark Survey," *Survey of Current Business*, November 2006, *86* (11), 37–68.

McLaren, John, "Globalization and Vertical Structure," *American Economic Review*, 2000, *90*, 1239–1254.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71* (6), 1695–1725.

Nunn, Nathan, "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade," *Quarterly Journal of Economics*, May 2007, *122* (2), forthcoming.

Schott, Peter, "Across-Product versus Within-Product Specialization in International Trade," *Quarterly Journal of Economics*, 2004, *119*, 647–678.

Yeaple, Stephen R., "Offshoring, Foreign Direct Investment, and the Structure of U.S. Trade," *Journal of the European Economic Association*, 2006, *4*, 602–611.

Zeile, William J., "Trade in Goods Within Multinational Companies: Survey-Based Data and Findings for the United States of America," 2003. Mimeo, U.S. Bureau of Economic Analysis.