ISSN 2042-2695

CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 1583

Revised April 2019
(Replaced November 2018 version)

Organizing Global Supply Chains: Input-Output Linkages and Vertical Integration

Giuseppe Berlingieri
Frank Pisch
Claudia Steinwender

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
We study whether and how the technological importance of an input – measured by its cost share – is related to the decision of whether to "make" or "buy" that input. Using detailed French international trade data and an instrumental variable approach based on self-constructed input-output tables, we show that French multinationals vertically integrate those inputs that have high cost shares. A stylized incomplete contracting model with both ex ante and ex post inefficiencies explains why: technologically more important inputs are "made" when transaction cost economics type forces (TCE; favoring integration) overpower property rights type forces (PRT; favoring outsourcing). Additional results related to the contracting environment and headquarters intensity consistent with our theoretical framework show that both TCE and PRT type forces are needed to fully explain the empirical patterns in the data.

Key words: vertical integration, supply chains, direct requirements, input output relationship, intrafirm trade
JEL Codes: F10; F14; L16; L23; O14

This paper was produced as part of the Centre's Trade Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

We thank Laura Alfaro, Pol Antràs, Davin Chor, Bob Gibbons, Gianmarco Ottaviano, Veronica Rappoport, Armando Rungi, Catherine Thomas, and John Van Reenen for their helpful comments. We are especially grateful to Paola Conconi and Peter Egger (Pisch's PhD committee), and for discussions at the CEPR Conference on GVCs; NUS Conference on Global Production; CEP Conference; SETC; and presentations at Harvard, MIT, ESSEC, Groningen and St. Gallen. We thank Corcos et al. (2013) and Keiko Ito for sharing data, Jonathan Colmer for code, and Hanwei Huang for help with Chinese data. Access to French data benefited from the use of the Centre d'accès sécurisé aux donnes (CASD), part of the "Investissements d'Avenir" programme (reference: ANR-10-EQPX-17) and supported by a public grant overseen by the French National Research Agency (ANR). Berlingieri acknowledges support by the ESSEC Research Centre and Pisch by the Dr. Heinrich Wachter Foundation.

Giuseppe Berlingieri, ESSEC Business School and Centre for Economic Performance, London School of Economics. Frank Pisch, University of St. Gallen, SIAW and Centre for Economic Performance, London School of Economics. Claudia Steinwender, MIT Sloan, NBER, CEPR and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

Global supply networks, often termed Global Value Chains (GVCs), are the backbone of 21st century manufacturing. Offshoring has brought substantial productivity and welfare gains to firms and consumers in recent decades.[1] Since international transactions are typically conducted in an environment characterized by severe contractual frictions, however, they are often distorted and the full gains from international sourcing remain unrealized.[2] Multinational firms have the option to alleviate these inefficiencies by assigning ownership optimally: for the hundreds, sometimes thousands of different inputs they procure internationally, they can decide whether to trade intrafirm or at arm's length.[3]

Naturally, the input-output linkage involved in a trade relationship is a first order consideration in this context. First, the multinational's trade-off between different organizational arrangements depends on how important each input is for its output. Secondly, making efficient integration decisions can have significant consequences, especially for technologically important inputs: Aggregate productivity gains from offshoring the right inputs can be large – but so are losses due to sub-optimal organizational decisions.[4] Motivated by these considerations, we study whether and how the technological importance of an input affects the multinational's decision of sourcing that input from an international affiliate or from a stand-alone supplier.

We use detailed French firm-product level import data to document that inputs that are technologically important for a multinational's output are substantially more likely to be sourced in-house. To understand this pattern, we construct a stylized incomplete contracting model with both ex ante and ex post inefficiencies that can explain this result when transaction cost economics type forces (TCE; favoring integration) overpower property rights type forces (PRT; favoring outsourcing). Additional empirical results consistent with our theoretical framework show that both TCE and PRT type forces are actually needed to fully explain the empirical patterns in the data. Our work suggests, therefore, that it is necessary and important to incorporate TCE mechanisms into models

[1] Access to foreign intermediates is an important source for productivity gains, for a literature survey see Shu and Steinwender (2019). For welfare effects see, for example, Caliendo and Parro (2015) and Blaum et al. (2018).

[2] See Antras and Staiger (2012).

[3] This organizational margin is important in the aggregate: about 40% of all US trade is intrafirm (Ruhl, 2015).

[4] In addition, the interplay between technological importance of individual inputs and the related organizational decisions is a crucial aspect for how shocks, including those of international trade policy, transmit to the supply and demand side of any economy (e.g., Acemoglu et al., 2012; Bernard et al., 2019).

of multinational activity in GVCs.[5]

In our main empirical exercise, we compare the sourcing strategies for detailed products sourced by the *same* firm. We use a cross section of firm level trade data from France in 1999, which has two particularly useful features: first, product-firm level imports are split into those that are sourced from affiliates and those that are not; and second, we can identify the industry of the buyer. The technological importance of a product will be reflected, among other things, in its cost share. To isolate the variation in cost shares that is driven by technological linkages, we use variation related to input-output coefficients across detailed upstream products within a highly disaggregated downstream *industry* obtained from detailed self-constructed IO tables from France and other countries. Our identification assumption is that these industry level linkages capture fundamental – and thus exogenous to the organizational choice – technological relationships between inputs and outputs.

Applying this empirical strategy, we show that technologically more important inputs are more likely to be sourced from affiliated parties. The role of technological linkages is economically significant: An input at the 75th percentile of the within industry cost share distribution is about 7 percentage points more likely to be sourced in-house than one at the 25th percentile, which amounts to about a quarter of the baseline share of integration. We document that input-output linkages are at least as important in driving make-or-buy decisions of multinationals as previously studied determinants such as capital intensity, skill intensity, or productivity (Corcos et al., 2013).

This is a surprising result, as the property rights frameworks that are typically used in the current literature on multinational firms would predict exactly the opposite: A downstream firm faces stronger incentives to outsource the production of more important inputs to independent suppliers, because it wants to limit underinvestment for these crucial parts and components (PRT type force).[6] In the theoretical part of the paper we show that in order to rationalize our main results, one needs to incorporate the transaction cost economics theory of the firm, as pioneered by Williamson (1985) and Klein et al. (1978), which emphasizes ex post contracting problems between suppliers and their customers.[7] In these models, suppliers of more important inputs cause greater inefficiencies through haggling or mis-coordination, which the downstream firm tries to curb by

[5]For rare applications of TCE ideas in the trade context, see Costinot et al. (2011) or Alfaro et al. (2016).

[6]The property rights framework was developed by Grossman and Hart (1986) and Hart and Moore (1990) to conceptualize firm boundaries. Early contributions that incorporate these models in the international trade literature include Antras (2003), Antras and Helpman (2004), and Grossman et al. (2006).

[7]While there are several other models of firm boundaries, the most influential and empirically successful line of thought is tied to TCE considerations. Therefore, we view it as the natural starting point in our context.

bringing the supplier under its control (TCE type force) — consistent with our empirical findings.

To understand whether PRT forces are just weak on average or completely absent and thus, whether both ex ante and ex post inefficiencies are present in the data, we use our framework to derive additional predictions that hold only when PRT forces are also at work. Specifically, we extend our model to show that when a supply relationship is subject to a better contracting environment ex ante, or when the multinational importer's investment into a relationship is more significant, technological importance makes intrafirm trade even more likely; in other words, our main result is reinforced. Intuitively, when ex ante investments are more contractible, or when the downstream firm has an important investment to make, the incentives for the multinational importer to outsource are generally weaker as underinvestment upstream is less severe. Since underinvestment matters more for technologically more important inputs, they are (even) more likely to be 'made' rather than 'bought'. In the final part of this paper, we provide robust empirical evidence related to the contracting environment and headquarters intensity that are in line with these predictions.

Therefore, both TCE and PRT type forces are needed to fully explain the empirical relationship between technological importance and vertical integration observed in the data. So far, trade economists have predominantly viewed multinational activity through the lens of the property rights theory of the firm. Our results suggest that it is necessary and important to incorporate transaction cost economics mechanisms into trade models and explore their implications. We view this as a main contribution of our paper, which underlines the usefulness of an "integrative framework" as predicted in Gibbons (2005).[8]

Our second contribution to the literature consists of highlighting the technological importance of an input as a causal factor in the sourcing decision.[9] The existing literature has examined various aspects of GVCs that shape intrafirm sourcing decisions. Among these are location or country level characteristics[10] and the technological features of

[8]Other examples of work that draws on multiple branches of the literature on firm boundaries simultaneously include Baker and Hubbard (2003) and, more recently, Kalnins et al. (2018).

[9]Input-output coefficients have featured in empirical work on vertical integration elsewhere: Acemoglu et al. (2010) use direct requirements as measures for relative importance of overall upstream to downstream investment and show that they amplify the effect of RnD intensity on vertical integration. Our focus is different since we examine a given firm and explore how inputs with different cost shares relative to each other are typically sourced. Alfaro et al. (forthcoming) use total requirements as a control in their empirical approach, while we focus on direct requirements in this paper. Finally, Alfaro et al. (2018) find that direct requirements are positively correlated with upstream ownership (which does not necessarily imply intrafirm trade) and note that this is consistent with their delegation model, while we focus on intrafirm trade and how it can be explained with property rights and/or transaction cost economics models.

[10]Examples are the level of (intellectual) property rights protection, judicial quality, or the state of the financial system (e.g., Levchenko, 2007; Nunn, 2007; Acemoglu et al., 2009; Macchiavello, 2012; Carluccio

how individual products are developed, produced, and distributed.[11] Only recently, however, have technological characteristics of different supply chains or networks been researched. Antras and Chor (2013), Alfaro et al. (forthcoming), and Del Prete and Rungi (2017) explore whether multinationals produce in early or late stages of their (purely sequential) supply chains and conclude that they are typically active in activities close to their core business.[12] In this paper we focus on *directly* sourced inputs (as opposed to inputs of suppliers, and inputs of suppliers of suppliers, etc.), and explore which of them are integrated.

Furthermore, we shed new light on the anatomy of intrafirm trade. In particular, we draw attention to and explain the fact that there is not only a skewed distribution of intrafirm sourcing *across* firms (cf. Atalay et al., 2014; Ramondo et al., 2016), but also *within*, and this can be explained by sufficiently detailed cost shares: Multinationals produce only the technologically most significant inputs in-house. Moreover, we complement the findings of Alfaro and Charlton (2009), who show that a large share of FDI is undertaken in vertical supply relationships. According to our findings, such vertical FDI is much more likely to occur along technologically important supply relationships.

We first outline our data, empirical strategy, and main results in Section 2, where we also present a number of robustness checks and our horse race exercise. In Section 3 we introduce our stylized conceptual framework, use it to discuss our main finding, and derive additional implications. We test these predictions in Section 4 and conclude with Section 5.

2 Baseline Empirical Results

In this section we introduce the data, outline our empirical strategy, explain the instrumental variables used, and discuss our main empirical finding.

and Fally, 2012; Eppinger and Kukharskyy, 2017), as well as the tax system (e.g., Flaaen, 2017), trade policy (e.g., Ornelas and Turner, 2011; Diez, 2014; Alfaro et al., 2016), and geography (e.g., Antras and Helpman, 2004; Irarrazabal et al., 2013).

[11] For example, the degree to which both parties of a transaction contribute to it marginally (see, for example, Grossman and Helpman, 2002; Hart and Moore, 1990; Whinston, 2001), whether alternative trading partners are available in case of a break up (e.g., Joskow, 1985; Monteverde and Teece, 1982), and the ease with which comprehensive contracts can be written and enforced (e.g., Acemoglu et al., 2007). For review articles, see Lafontaine and Slade (2007), Bresnahan and Levin (2012), and Legros and Newman (2014).

[12] Recently, Fattorini et al. (2017) explore network centrality in the IO network, while Bolatto et al. (2017) investigate the role of intangible assets in sequential supply chains.

2.1 Data Sets and Summary Statistics

Two features make the French data particularly useful for studying the make-or-buy decision of a firm for specific inputs: first, firm by product level imports are split into those that are sourced from affiliates and those that are not; and second, we can identify the industry and other characteristics of the buyer.

Our main dataset is the Enquête Échanges Internationaux Intragroupe (EIIG), a single cross section in 1999, which covers information about intrafirm trade of French firms.[13] The target population in the survey includes every French firm whose annual trade volume was at least one million Euros and that was owned by a manufacturing group that controls at least 50% of a foreign firm. In other words, the sample is representative of the population of international sourcing flows of multinational firms. Out of this target population (8,236 businesses) roughly half of all firms responded. These 4,305 firms account for about 80% of French trade conducted by French multinational entities.

For each responding firm, the EIIG has information about the value share of imports from related parties for each HS 4 digit product that the firm imports, by country of origin. In our final sample we focus on imports by the EIIG manufacturing firms (ISIC Rev. 3 codes 15 to 37).[14]

We merge these trade data with the Enquête Annuelle d'Entreprise (EAE), which provides us with balance sheet data on all French firms with more than 20 employees and a random sample of smaller firms. We use these data to obtain total expenditure on intermediate inputs, and to identify the industry of the sourcing firm.

Table 1 reports summary statistics for the firms in our data. There are about 3,000 firms in the final sample – we focus on manufacturers and exclude wholesalers and retailers – and the total number of observations at the firm-product-country level is 85,550. Consistent with the target population of the EIIG, our firms are relatively large with respect to employment or sales.

2.2 Empirical Strategy

We want to test whether the technological importance of an input affects whether a firm sources the input from an international affiliate or from a stand-alone supplier.

[13]Other work that uses this data set includes Carluccio and Fally (2012), Corcos et al. (2013), Defever and Toubal (2013), and Carluccio and Bas (2015).

[14]The EIIG survey data was amended with official international trade data by the statistical authorities. This is described in the official documentation and our results are robust to excluding the affected flows.

The technological importance of a product will be reflected, among other things, in its cost share. Our baseline regressor $costshare_{ip}$ is therefore the share of the value of a specific imported product relative to the cost of all other intermediate inputs that the firm buys, domestically or abroad:

$$costshare_{ip} = \frac{\sum_c imports_{ipc}}{totcost_i}, \tag{1}$$

where $imports_{ipc}$ is the total value of all imports by firm i of input p from country c, and $totcost_i$ is total expenditure on intermediate goods by firm i. The regressor varies at the firm i by input p level, where each input is classified according to HS 4 digit (roughly 1,100 product categories). Figure B.1 in Appendix B.3 shows the empirical density of the input cost shares, and Table 1 reports the average and median across all firms.

Our dependent variable is the share of product p imports that firm i acquires from foreign affiliates versus from unaffiliated parties in a specific country c, and is denoted by $intrashare_{ipc}$. Table 1 reports the average import share from affiliated parties across all firms in the sample: The average firm carries out 27% of its transactions inside the boundary of the firm (across products and origin countries). However, the distribution of intrafirm trade is rather skewed towards a few large companies reporting a larger share of intrafirm transactions: The median firm imports only 9% of its transactions from affiliated parties.[15]

Combining this, the baseline structural equation we estimate is

$$intrashare_{ipc} = \beta_1\, costshare_{ip} + \alpha_i + \gamma_{cj} + \phi_{cp} + \varepsilon_{ipc}, \tag{2}$$

where j denotes the industry of the downstream, sourcing firm; and α_i, γ_{cj}, and ϕ_{cp} represent, respectively, firm, country $\times$ downstream-industry, and country $\times$ upstream-input fixed effects.

A few comments regarding our choice of variables are in order. Since we are not using the country variation in the dependent variable in the regression, we could in principle aggregate it across all countries. However, keeping the country variation has two advantages: First, we are able to control for country level sourcing determinants directly in the regression, which reduces noise. Second, we will exploit the country dimension in Section 4, when we test extensions of the stylized model. In any case, we report estimates on the aggregated data in Online Appendix Table B.4; the results are very similar to our baseline specification.

Our results are not sensitive to using other definitions of the dependent variable.

[15] A similar skewness has also been reported for the U.S. (Ramondo et al., 2016).

In particular, we have checked three binary versions: First, we define an integrated (as opposed to outsourced) flow as $intrashare_{ipc} \geq 0.5$. Second, we follow Corcos et al. (2013) in that a flow is within firm iff $intrashare_{ipc} \geq 0.8$ and outside iff $intrashare_{ipc} \leq 0.2$. Finally, we count as fully integrated only those observations that have $intrashare_{ipc} = 1$, while observations with $intrashare_{ipc} = 0$ count as outsourced. We find very similar results with all these dummy variables, which is due to the fact that few products at our highly disaggregated level are sourced with a mix of outsourcing and integration. In our context the *share* of intrafirm trade is conceptually the correct dependent variable, because our theoretical mechanism operates at the finest input level and hence predicts organizational mix at any more aggregate level, such as the HS 4 digit level.

While we believe that our cost share regressor is a good starting point to measure technological importance, input-output linkages at the firm level can be specified in different ways. We have therefore conducted robustness checks in which we use the cost share of an input in *total imports* (instead of total expenditure), or the (log) value of imports as regressors in Online Appendix Table B.5, columns (2) and (3). It turns out that the firm specific intercepts provide an effective normalization, and the parameter of interest is mainly identified by variation in the numerator of (1).

The great advantage of our international trade data is that they allow us to study the properties of very detailed transactions, which is usually impossible to do with domestic data sets. On the downside, this means that we base our measure of technological importance only on the value of internationally sourced inputs. At the highly aggregated 2-digit industry level we can observe both the cost share of imported inputs and the cost share of domestic inputs (using the official French IO table). The two cost shares are highly correlated: the correlation is 0.88 (0.78 if the diagonal in the IO table is excluded). It is therefore plausible to assume that the technological importance of an input is reflected in both cost shares. Our fixed effects and identification strategy, which we will discuss below, has the purpose to isolate the variation in import cost shares that is due to technological importance, rather than other determinants.[16]

The firm fixed effects α_i ensure that we compare the sourcing mode of different products within the same firm i. In this way our estimates are not affected by any firm-

[16] Towards the end of Section 2.4 we present explicit tests in order to understand how well our results could generalize to domestic transactions and whether we should expect a biased estimate from offshoring decisions. First, we show that our estimates are similar when focusing on sourcing from the EU alone, where offshoring costs are low. Second, even within imports from the EU, the effect does not vary significantly with distance from France. These two tests imply that variation in offshoring costs is unlikely to affect our estimates. Finally, we find that the effect of cost shares is the same regardless of whether a firm sources a large or small share of its intermediates from abroad. This shows that significant domestic sourcing that is not captured by our cost shares does not systematically bias our results.

specific characteristic that shapes the sourcing mode across all products, for example the headquarters intensity of the downstream multinational firm — an example that the literature has studied previously.[17] It can be seen in Table 1 that the average firm sources about ten different inputs from abroad, and therefore contributes to our within-firm estimates.

The technological importance of a product will be reflected in its cost share, but there are other determinants of the cost share which may also be related to the sourcing decision. Because our data is so detailed, we are able to add a rich set of additional fixed effects to control for many of these determinants. The fixed effects ϕ_{cp} and γ_{cj} ensure that we do not mistake any other country $\times$ downstream-industry or country $\times$ upstream-input specific characteristics that increase the likelihood of intrafirm sourcing for the effect of the importance of inputs in a firm's production function.

For example, origin country by input fixed effects ϕ_{cp} control for country specific input price related factors that may drive integration decisions (Alfaro et al., 2016), or for trade costs that vary by product and origin country. Country by product fixed effects also absorb separate product p fixed effects and country c fixed effects. Our estimates are therefore not affected by product specific characteristics like the relationship-specificity of a particular input, or the codifiability of tasks required for the production of a certain input; or country specific characteristics such as gravity factors that influence the patterns of FDI, like distance, market size, multilateral resistance etc.

On the other hand, origin country by downstream industry fixed effects γ_{cj} exclude variation that stems from the interaction between financial development of the origin country and financial constraints.[18] They also account for comparative advantage to the extent that different downstream industries need different intermediates that have different factor intensities and some countries have these factors abundantly.

Despite the rich set of fixed effects, one might still be concerned that input cost shares are endogenous to the integration decision. First, firms may substitute towards inputs produced by their foreign affiliates (i.e., reverse causality), for example to trigger increasing returns for them and maximize global profits, or because information frictions are less severe. Second, multinational firms frequently engage in transfer pricing, which distorts input cost shares selectively in integrated relationships.[19] In our setting, firms

[17] Corcos et al. (2013) point to the fact that the EIIG survey suffered from non-response. They also show that this poses a significant problem in their context since their results change meaningfully when they apply a selection correction. Our firm fixed effects absorb their selection correction variable. In the horse race exercise we conduct further below, we drop the firm specific intercepts and apply their selection correction.

[18] See Acemoglu et al. (2009) and Eppinger and Kukharskyy (2017).

[19] There is a substantial body of research that explores the nature and consequences of transfer pricing.

might integrate to avoid paying higher prices due to "double marginalization" in the presence of market power in the upstream markets;[20] conversely, firms in relatively high tax France may charge inflated prices for inputs produced by foreign affiliates in order to artificially reduce their taxable income.[21] Alternatively, transfer pricing is a way of alleviating the burden of tariffs. Third, it is possible that the values recorded in our data do not reflect the economic cost structures of our firms, because inventories may fluctuate significantly as a consequence of demand or supply shocks, and — as we estimate our regressions in a single cross section — inventory states distort input cost shares. Moreover, international trade has been shown to be lumpy due to fixed costs of ordering.[22] Consequently, the cost shares we calculate from international trade data are subject to variation due to shipments arriving early or late with respect to a given accounting year. Finally, cost shares reflect technological input-output relationships only to some extent and depend on many other characteristics that may contribute to measurement error when using input cost shares as proxies for technological importance.[23]

To isolate the variation in cost shares that is driven by technological importance, and therefore exogenous, we use variation from input-output coefficients across detailed upstream products within a highly disaggregated downstream industry obtained from self-constructed IO tables from France and other countries to instrument for the firm level cost shares. The basic identification assumption is that industry-level input-output relationships affect organizational choice only through their effects on input cost shares. This assumption is likely to hold since IO tables capture broad features of the underlying production technology and are, in particular, not likely to be affected by individual firms in a large economy like France. In the following section we provide details on how we construct this instrument.

2.3 Instrument: Input-Output Tables

Our main instrument for the input cost shares is based on a self-constructed French import IO table. In order to achieve a strong first stage we need relatively disaggregated information (1) and in order to ascertain exogeneity we have to make sure that we

For recent examples see Davies et al. (2018), Flaaen (2017), and the citations therein.

[20]See Garetto (2013).

[21]Bernard et al. (2006) find that U.S. multinationals charge on average significantly lower export prices for intrafirm transactions. The average arm's length price is 43 percent higher than the price for intrafirm transactions.

[22]See Alessandria et al. (2010).

[23]For example, a mechanical reason to find a significant relationship is the fact that the denominator in the intrashare variable is a component of the denominator of the cost share variable.

remove our EIIG firms' trade flows from the data we use for an IV (2).

In principle, IO tables are readily available for most countries and France is no exception. However, the most commonly used, official 2 digit ISIC Rev. 3 French table for the year 1999 satisfies neither of the two requirements mentioned. This IO table captures mostly domestic transactions and, together with a high level of aggregation, yields a first stage that is too weak for the demanding identification purposes in this paper. Unfortunately, there are no officially published disaggregated tables available for France (unlike for the U.S.).

We therefore construct our own IO tables for the year 1996 from transaction level import data for the whole of France.[24] To obtain 4 digit NAF 1993 industry codes[25] for all trading firms we rely on the FICUS database, which contains balance sheet and administrative information for the near universe of French enterprises. The customs data are matched to this firm information with a success rate of 91%. We use balance sheet information to compute gross output by NAF industry and calculate the import direct requirements at the NAF industry $\times$ HS 4 digit input level.[26]

To illustrate our approach, imagine the following example. Two car manufacturers in France, PSA and Renault, import chassis of value EUR 1 (PSA) and EUR 2 (Renault) and engines of value EUR 2 (PSA) and EUR 3 (Renault). To construct our IO tables we link all import transactions of chassis and engines (classified according to HS 4 digit) to their respective importers. Summing transactions across all firms in the downstream industry (cars) gives us the total value imported of each HS 4 digit input by the downstream industry. In our example, value(chassis $\rightarrow$ cars) = EUR 3 and value(engines $\rightarrow$ cars) = EUR 5. Now we can use French firm level data to find the gross output for both firms, say EUR 40 (PSA) and EUR 50 (Renault). We can add these up to get gross output at the industry level, namely EUR 90. To compute our import IO direct requirements, we divide the transacted volume by gross output, i.e. dr(chassis $\rightarrow$ cars) = 3/90 and dr(engines $\rightarrow$ cars) = 5/90.

These tables are constructed directly from micro data and therefore we name them "micro" tables. Moreover, since the upstream (product) dimension is classified at the HS 4 digit level and the downstream (industry) dimension follows the much coarser NAF classification, we call these tables "asymmetric" in the sense that one dimension of the

[24] We plan to make our French import IO tables available on our websites for the future use of researchers. Feenstra and Jensen (2012) follow a similar strategy for the U.S.

[25] NAF is the French industry classification and more disaggregated than ISIC or NACE. In our sample, we have 285 NAF manufacturing industries, while there are only 122 in ISIC.

[26] The components we use for gross output are detailed in Appendix B.1.1.

matrix is much longer than the other.[27]

We perform two additional modifications to improve our instrument further. First, when computing the industry level intermediate costs, we leave out a firm's own trade flows, effectively creating firm specific IO tables.[28] Second, we compute the IO table for 1996, three years prior to the date of the regressor: To the extent that import IO tables capture mostly the underlying input-output linkages across inputs (and hence their technological importance), the 1996 direct requirements are good predictors of 1999 input cost shares, while arguably being less suspicious of reverse causality or other problems.[29]

Figure 1, where we present a contour plot of several French IO tables, illustrates why our self-constructed direct requirements are powerful allies. The upper left graph is well known: in the 2 digit level official table by far most of the transaction volume takes place on the main diagonal, while only few, usually proximate sectors are connected off the main diagonal. When we construct our micro IO table at the 2 digit level, this pattern is replicated quite well – an observation we interpret as validation for our approach. As expected, we do find differences between the upper two tables, which relate to the fact that, in contrast to the official IO table, we do not need to make any strong assumption regarding tradability and can simply let the actual trade transactions speak. Furthermore, we focus on external trade only, which improves our first stage.

Constructing the tables at a more disaggregated level has two effects. First, the diagonal becomes relatively "thinner". Second, the elements off the diagonal exhibit more "contrast". In other words, the cells in the lower two plots in Figure 1 have clear borders now and stand out properly from the background. Econometrically, we reduce measurement error and bring the relevant variation to the fore.

The asymmetric IO table at the finest level of disaggregation – our preferred level – exhibits a soft "diagonal", which stems from the fact that product and industry classifications follow a similar ordering. Industry codes are usually assigned on the basis of the product they produce (and vice versa).

[27]Note that in the language of IO tables, we have constructed a USE table, which is not a 'proper' IO table in the inverse Leontief sense, but the meaningful table for the purpose of our estimation. We simplify the exposition here by referring to our tables as "IO tables" and their elements as "direct requirements", acknowledging the fact that this is technically speaking imprecise.

[28]For further robustness, we also removed trade flows of all our EIIG firms from our international trade data when we constructed the import IO table. Our main results are unchanged when we use this IO table, including a sufficiently strong first stage. Moreover, we obtain the same result when we retain all EIIG flows in the data when we construct our IO tables.

[29]Note that our rich set of fixed effects that we include in all specifications already addresses a number of potential endogeneity concerns. For example, upstream product by country effects clean the instrument of many price related considerations.

The actual instrument we use below is not the direct requirement itself, but a categorical variable that indicates quantiles of the direct requirement distribution. Figure 2 shows the empirical density of our self-constructed import requirements. It is very skewed to the left and even the median is relatively small, because we normalize by gross output. The vertical lines indicate quintiles and our preferred instrument is a variable that takes the value 5 whenever the direct requirement of downstream industry j with respect to upstream input p falls into segment V, value 4 if it falls into IV, and so on. In this way we semi-parametrically capture the skewed distribution of the requirements and make the instrument more robust to measurement error.[30]

We argued above that our self-constructed IO table solves the problem of cost shares being endogenous to a firm's integration decision provided the identification assumption is satisfied. One may, however, be concerned that forces that operate in partial or general equilibrium render the exclusion restriction violated despite our efforts. For example, it may be that for some other historical reasons multinationals in industry j are more likely to integrate some products than others. To address this concern, we make use of IO tables from two other countries where such effects are very unlikely to play a first order role. First, we apply the same methodology as above to construct direct import requirements from Chinese micro data. For the year 2006 we have access to the universe of import transactions in goods and can link these to the involved Chinese importer.[31] To identify the industry of the buyer, we link the trade data to the Chinese Annual Industrial Survey (CAIS), which covers all State Owned Firms (SOE) and non-SOEs with sales above 5 million Chinese Yuan. From this data source we obtain the CIC code for every importer and compute gross output at CIC level.[32] Using a crosswalk, we finally concord the CIC downstream industry to ISIC Rev. 3.[33] While this Chinese table, like our French micro table, also has the advantage of being very detailed, it uses variation that is less relevant for France, since the two countries occupy different parts of global

[30]Our main results are fully robust to using other functional forms, for example the direct requirements themselves.

[31]Chinese trade has undergone an historically unprecedented growth spurt in the years following its accession to the WTO in late 2001, which was pronounced at the extensive product margin. In order to limit the number of missings and measurement error in the Chinese IO table, we have chosen a late year in the data available to us.

[32]CIC is the Chinese industry classification.

[33]We note that the need for concording puts a limit on the level of downstream disaggregation, since ISIC Rev. 3 contains significantly fewer codes at the lowest level than the French NAF classification.

value chains and rely on different comparative advantages. The correlation between direct requirements is consequently only .21 (.52 for the instrument), implying a weaker, but still strong enough, first stage.[34]

Second, as an alternative we use the 2002 U.S. benchmark USE table, which captures a sourcing behavior much more similar to the one exhibited by French firms. However, we do not have access to the U.S. micro data, and the officially available table is more aggregate than our micro table, since the inputs are aggregated to the industry level (U.S. IO classification). Furthermore, in order to utilize this table, we had to establish a crosswalk to ISIC Rev. 3. Unfortunately, a concordance directly to NAF was not feasible, and this introduces measurement error. The correlation between US and French direct requirements is .35 (.28 for the instrument). Estimates using the U.S. instrument should be interpreted with these caveats in mind.

One might argue that the organizational choice of a firm itself affects the market structure in the upstream industry, which may result in reverse causality at the industry level, after all. Or one could argue that other factors that influence the integration decision in a whole industry are correlated across countries, therefore even other countries' IO tables cannot be taken as exogenous. While we cannot fully rule this possibility out, below we obtain very similar results using IO tables from three substantially different countries. We therefore think that this alternative explanation for our findings is relatively unlikely: The effect of a firm's sourcing decision on the upstream industry would have to be reflected in the IO tables of all three countries.

2.4 Main Results

Baseline Results

We estimate equation (2) with two stage least squares (2SLS) and allow the error term to be correlated across all observations that belong to the same broad downstream industry and across all observations that belong to the same broad upstream industry – either dimension creates about 50 clusters and is significantly more aggregate than our IV.

The main result of this paper is documented in Table 2. In column (1) we present the unconditional correlation between cost shares and the intrafirm share. In column (2) we add our baseline fixed effects as detailed in equation (2). Finally, columns (3) to

[34] The IO diagonal does not drive these correlations, since omitting them leads to an *increase* in the correlation coefficient. Our IV estimates are also robust to dropping the diagonal transactions for the IVs from all countries. In Appendix Table B.3 we report this robustness check.

(5) report the 2SLS estimates using our preferred asymmetric French table, the asymmetric Chinese table, and the U.S. benchmark table, respectively.[35, 36] The first stage Kleibergen-Paap (KP) statistics are large, so that we are confident about the relevance of all three instruments. As expected given our discussion above, the Chinese and U.S. tables are weaker predictors of French firms' cost shares than the French table. Most importantly, however, our coefficient of interest, β_1, is always estimated to be positive and highly significant. The IV results are significantly larger than the OLS estimates, which is in line with an attenuation bias due to measurement error or a negative bias due to the incentives of firms to integrate more in the presence of tariffs or to avoid double marginalization, which would imply lower cost shares in case of integration (see, for instance, Bernard et al., 2006).

According to our preferred estimate in column (3), an input at the median of the input cost share distribution is about $0.0012 * 11.586 \approx 1.4$ percentage points more likely to be integrated than a wholly insignificant input, over a baseline integration probability of 27 percent. The interquartile difference within a NAF industry is about 7 percentage points, which amounts to about one quarter of the baseline integration probability. These magnitudes are also robust to using variation from Chinese or U.S. sourcing behavior to identify β_1.[37]

Robustness Checks

We next explore how robust these results are to various concerns and report our findings in Table 3, using our preferred instrument, the asymmetric, lagged French micro IO table.[38] The first column repeats the preferred baseline specification on an unrestricted sample.

Since we are concerned with vertical integration in this paper, i.e., the mode of organization for the procurement of inputs, we want to make sure that our results are not driven by horizontal transactions, i.e., sourcing of essentially finished items.[39] As

[35]The first stage and reduced form regressions are reported in Online Appendix B.3, Tables B.1 and B.2.

[36]The number of observations falls from column (1) to (2)-(5) since we exclude singletons. The sample used for Table 2 is conditional on non-missing values for all three instruments.

[37]We also estimate the baseline specification with Logit IV, rather than the linear probability model, to take into account that our dependent variable is largely a binary variable. The results are robust.

[38]The number of observations varies across columns due to different sample restrictions and to different sets of singletons. The number of observations can be higher than in Table 2, since we only condition on the French direct requirement being non-missing, rather than all IO tables.

[39]These could be intermediates that are sold on to other firms or final goods that are sold on to consumers.

an example, a French car manufacturer may assemble a specific model in France, but have foreign production sites in which it assembles other models. If it imports some other models and resells them in France, they will look like inputs in our data that are sourced from affiliates. However, since the car manufacturer resells them, this is horizontal, rather than vertical integration, which we are not interested in. In order to make sure not to capture horizontal integration, we drop all observations in which the downstream importer i is mainly active in the industry j that also produces the good sourced, p. In IO terminology, this means we drop all observations on the IO diagonal. Since our asymmetric table does not have a well defined diagonal, we exclude all observations on the 4 digit ISIC Rev. 3 diagonal. This strategy is conservative in that we do not use the full detail available to us from our micro data and drop many transactions that are in fact purely vertical.

The result for this robustness check is reported in column (2) of Table 3, where the number of observations shrinks by about one third compared to the full sample that is available for our empirical strategy. Even though the KP statistic falls, it is still sufficiently high to take the estimates seriously. Importantly, the point estimate is not significantly different from our baseline result, so that we can conclude that our results are not driven by the high cost shares for horizontally traded products.[40]

In columns (3) and (4) of Table 3 we show that our results are not driven by technical similarity, which may be positively correlated with cost shares. Imagine a producer of foodstuffs (ISIC Rev. 3 code 1549) and a pharmaceutical company (ISIC Rev. 3 code 2423) that source coffee beans (HS code 0901) and formaldehyde (HS code 2912). It should not come as a big surprise that the pharmaceutical company does not buy coffee for its employees (a low cost share input) from its own roastery in Nicaragua and that a foodstuffs company typically refrains from integrating its supplier of machine disinfectant (again a low cost share input). The reason is that firms may be more likely to 'make' inputs that are technically similar to its current production in order to exploit, for example, economies of scale in production or RnD.[41]

To check that our results are not driven by technical similarity, we show in two different ways that they also hold *within* technically similar inputs. We first include 4 digit ISIC Rev. 3 downstream industry $\times$ HS 4 digit upstream product fixed effects in column (3), meaning that we only use variation coming from comparing the same input across very

[40]In appendix Table B.3 we report the specifications from the full baseline Table 2 when we drop the diagonal transactions.

[41]Empirical evidence on a potential role of technical similarity for horizontal integration is provided by Boehm et al. (2017).

similar downstream firms.[42] To use the example above, we compare different producers of ingredients for beverages and examine whether a company for which coffee beans are a more important input is also more likely to source them from its own roastery. Next, we replace the specific intercepts with downstream firm $\times$ upstream 4 digit ISIC Rev. 3 industry fixed effects in column (4), effectively comparing two very similar inputs sourced by the same firm – a pharmaceutical company sourcing formaldehyde and formic acid (HS code 2915), for example. The point estimates fall, but remain significant and large, even when we identify β_1 off very little variation.

We would like to point out that the robustness checks with interacted upstream $\times$ downstream fixed effects serve two additional purposes. When we focus on a particular supply relationship between two industries, we hold their *relative upstreamness*, i.e., their relative distance to each other in the value chain, constant. For example, when comparing different inputs sourced by a firm and produced by the same, highly disaggregated upstream industry (column 4 in Table 3), then these inputs all have the same "distance" in the value chain from the downstream firm. Equally, comparing the same input sourced by different firms in the same downstream industry (column 2 in Table 3), the latter have a very similar "distance" from the upstream product.

The first implication of the fact that our results are still statistically and economically significant is that our finding is not driven by relative upstreamness. Antras and Chor (2013) and Alfaro et al. (forthcoming) show that firms typically integrate the most proximate stages of production relative to their own position in the supply chain. Since our cost shares may be positively correlated with relative downstreamness, this mechanism could potentially account for our finding. The exercises in columns (3) and (4) in Table 3 show, however, that even when relative upstreamness is kept constant, our main finding persists. This is reassuring for technological importance as a separate determinant of integration decisions in supply networks.

More conceptually, with our firm fixed effects, we zoom in on integration decisions *across the legs of a spider*, i.e., into parallel segments of supply chains (Baldwin and Venables, 2013). Previous research has shown that more relatively downstream production stages are integrated on average, and we show that within those (i.e., directly sourced) inputs of a firm, the more important ones are integrated.

A further potential concern relates to our instrument. If the multinationals in our EIIG sample play a dominant role in their respective industries, excluding their own trade flows may not be enough for our purpose. In column (5) of Table 3 we estimate our

[42] Our instrument varies at the 4 digit NAF level, which is significantly more disaggregated than 4 digit ISIC Rev. 3.

baseline specification on the sample of firms in highly competitive industries, i.e. those where the Hirschman-Herfindahl-Index in 1999 was below 0.1.[43] We retain a sizable part of our sample and the cost share coefficient remains virtually unchanged compared to the baseline. We interpret this finding as further support for our identification strategy.

Our results may still be biased by transfer pricing considerations, since cost shares rely on input prices and the potential for within firm adjustments for tax purposes clearly increases with the degree of vertical integration. Even when we rely on our IO tables as a source of exogenous variation, these may still reflect transfer pricing motives at the industry level. To address this concern, we include firm $\times$ origin country fixed effects in our baseline specification and therefore compare only inputs that a firm sources from the same country. Since incentives to transfer price mainly depend on the double taxation treaties signed with the origin country, we abstract from the main portion of worrisome variation in our data. The estimate reported in column (6) of Table 3 shows that our results are robust.[44]

Finally, we explore whether we should expect our results to be influenced by the offshoring decision inherent to international sourcing flows. To this end, we restrict the sample to imports from the EU in column (7) of Table 3. Fixed market access costs are significantly lower within the EU than for origin countries outside it. The comparison between flows inside and outside the EU then reflects a difference in the ease of offshoring, potentially similar to the difference one would expect between domestic sourcing and importing from, say, Germany. We do not find any evidence that the effect of cost shares depends on the difficulty of accessing foreign suppliers.[45]

Quantitative Significance

As a final exercise we want to develop an understanding as to how economically significant technological importance is for the make-or-buy decision *relative to other de-*

[43] We compute these indices from the near universe of French firms in FICUS at the NAF 4 digit level. The result is robust to using other thresholds around 0.1, but the latter is often referred to as a reference point for antitrust considerations.

[44] In Appendix Table B.5, column (1), we also report that the effect of cost shares does not vary with effectively applied ad-valorem tariffs, which would provide a second rationale for distorting internal prices.

[45] In Online Appendix Table B.5, we report two further checks along the same lines. First, we find that the effect of cost shares is the same regardless of whether a firm sources a large or small share of its intermediates from abroad. This shows that significant domestic sourcing that is not captured by our cost shares does not systematically bias our results. Secondly, the effect does not vary significantly with distance within the EU. The additional variable costs that may come with offshoring and distort our cost shares are therefore also unlikely to be a first order concern.

terminants. To do so we compare the magnitude of the cost share effect to those of firm level characteristics that have been shown to correlate strongly with vertical integration. In particular, we include (physical and intangible) capital intensity, skill intensity, and productivity estimates following Corcos et al. (2013), who use very similar data.[46, 47]

To be able to include determinants that vary at the firm level in the regression, we need to drop the firm fixed effects. Furthermore, as mentioned above, the EIIG survey suffered from selected non-response, which can bias estimates in across firm regressions. We therefore follow Corcos et al. (2013) and apply a Heckman procedure to address this potential bias.[48] We report standardized coefficients throughout to enable direct comparisons. Column (1) of Table 4 shows our main result using standardized regressors, dropping firm fixed effects and applying the Heckman selection. In columns (2) to (6) we add the integration determinants that the previous literature has studied one by one.

An input with a cost share that is one standard deviation (SD) higher than the cost share of another input is roughly one fifth of a SD more likely to be sourced from an affiliate. This effect is of a similar order of magnitude across all columns, but larger than those of all other determinants in our horse race, such as skill intensity or productivity — even when we add them all together in column (6).[49] The interquartile effect of cost shares (column (1)) in this sample is 6.5 %, which compares, for example, to 2.7% for intangible capital intensity, 5.5% for skill intensity, and 4.9% for value added per worker. With the caveat that we do not attempt to causally identify the other firm level determinants – which would be well beyond the scope of this article, and which none of the previous literature has attempted to do – we conclude that cost shares are quantitatively at least as powerful, if not more, in shifting vertical integration patterns as other characteristics examined in the literature.

[46] The construction of all variables is detailed in Appendix B.1.1.

[47] In unreported regressions, we investigate the effects of other determinants at the product (contractibility) and country level (institutional quality) as in Corcos et al. (2013). We find input shares being significant throughout and with an effect of similar, if not larger, magnitude compared to the other determinants.

[48] Corcos et al. (2013) estimate inverse Mill's (IM) ratios for survey non-response and add a random sample of large non-multinational importers. We obtain a larger sample than provided by the EIIG and than the one used in Corcos et al. (2013), who aggregate inputs at the coarser CPA level. All further details can be found in appendix B.1.2.

[49] Value added is collinear with capital and skill intensity and hence omitted in the final column.

3 A Stylized Framework

Having presented the main empirical contribution of this paper, we want to explore the mechanism behind our findings. For this purpose we consider the two most successful conceptual frameworks in the literature on multinational firms: the property rights theory (focusing on ex ante inefficiencies) and transaction cost economics (focusing on ex post inefficiencies), and ask how consistent either is with our empirical results. In order to focus on the main trade-offs of the models, in this section we present a stylized framework that incorporates both of these two influential theories — i.e., it incorporates incomplete contracts both ex ante and ex post — but abstracts from certain complications discussed in the literature and makes two reduced form formulations that simplify the exposition.

The framework makes clear that PRT and TCE type forces typically have opposite predictions for the integration of technologically important inputs, and our empirical findings are consistent with a world in which TCE forces are on average stronger than PRT forces. To understand whether PRT forces are just weak on average or completely absent and thus, whether both ex ante and ex post inefficiencies are present in the data, we use our framework to derive additional predictions that hold only when PRT forces are also at work. As we show in the following section, our empirical results imply that PRT forces have to be present in order to fully rationalize the empirical results.

The finding that both PRT and TCE type forces are present in the data has an important implication for researchers: So far, trade economists have predominantly viewed multinational activity through the lens of the property rights theory of the firm. Our results suggest that it is necessary and important to incorporate transaction cost economics mechanisms into trade models and explore their implications.

We want to stress two points here. Of course, we are not the first to employ a TCE style approach in thinking about intrafirm trade (e.g., Costinot et al., 2011). Our results therefore underline the importance of such a line of research. Second, we do not attempt to *test* different models of the boundary of the firm against each other, which is extremely difficult (see Whinston, 2001). Instead, we aim at a simple extension of the PRT workhorse model of multinational firm behavior that incorporates TCE type features, and use it to illustrate the two forces and their interaction with other variables.[50]

[50] Note that while we model TCE type forces in a way consistent with a wide range of TCE approaches, inevitably we are unable to capture the vast breadth of ideas in this literature (e.g., the ideas in Klein et al., 1978).

3.1 Baseline Model

3.1.1 Technology

A downstream multinational firm[51] produces a quantity y of a final good for which it requires a discrete number $N \geq 2$ of inputs.[52] With a slight abuse of notation, we use N to address both the number and the finite set of inputs. The production technology is

$$y = \sum_{n}^{N} m(n)^{\delta(n)}, \quad \delta(n) \in (0,1) \;\; \forall n \in N, \tag{3}$$

where $m(n) > 0$ denotes either the quality or quantity of input n and the inputs may differ by their elasticities with respect to final output (δ).[53] It is this parameter δ, which governs the marginal contribution of an input, that we refer to as technological significance or importance. Moreover, δ is the only source of heterogeneity across inputs and, as we will show below, there will be a positive association between $\delta(n)$ and a firm's cost shares in equilibrium. Note that all inputs are used in production as $\delta(n) < 1$.

Clearly, equation (3) implies that the contributions of the different inputs to overall output are independent from each other and we make this assumption to abstract from interactions between the different supply relationships, as these complicate the model greatly. Indeed, as Nowak et al. (2016) show, any integration pattern across inputs can be optimal in a PRT model with constant technical (and demand) elasticities of substitution across inputs. We abstract from these considerations here to zero in on the main ideas of PRT and TCE mechanisms.

Suppliers of the various inputs can invest into quantity/quality and their costs of producing a unit of input n with quantity/quality $m(n)$ is $c_M < 1$; production upstream is subject to constant returns to quality investments.[54]

3.1.2 Consumer Preferences and the Downstream Market

To simplify matters, the downstream multinational takes the price of its output as given. Two implications follow: first, revenues are proportional to output and we can normalize the price of the final output good $p_y \equiv 1$. Secondly, we have shut down interactions

[51] We label the upstream and downstream firms according to the data available to us for ease of exposition.

[52] Clearly, with fewer than two inputs, the concept of cost shares is not meaningful.

[53] We could generalize this production function to include heterogeneous weights in the basket of inputs. What matters, however, for the make-or-buy decision is elasticity of output with respect to inputs (see Grossman and Hart, 1986). Consequently, we omit these weights from the outset to avoid confusion.

[54] We assume that c_M is sufficiently small so that the surplus from a relationship grows in δ.

between inputs arising from market power and downward sloping demand, which once again helps us focus on the key trade-offs in this stylized framework.

3.1.3 Further Assumptions, Contracting and Timing

There is a continuum of homogeneous suppliers abroad that can potentially produce any given input n, but none of them can produce more than one variety. Contracts are – for now – fully incomplete in the sense that only property rights can be contractually specified and enforced at any point in the game. Crucially, quantity/quality investments and other decisions are non-contractible. The timing of the game is as follows:

1. Contract written that includes arrangements regarding ownership
2. Supplier invests in quantity/quality
3. Output sold and revenues split in bilateral bargaining subject to haggling

In the context of our game we follow the convention that actions, contractual incompleteness, and inefficiencies in stage 1 and 2 are referred to as "ex ante", and in stage 3 as "ex post". While there is no random shock, this nomenclature supports our exposition.

Similar to Antras and Chor (2013), we do not specify the details of the bargaining game in stage 3 to simplify the model and allow for closed form solutions. We assume that property rights over the inputs convey an advantage in outside options – whoever has ownership can sell the input outside of the relationship at a discount, which is tantamount to a better disagreement point and more bargaining power. The share of the surplus appropriated by the downstream firm thus increases in the (endogenous) ownership share it holds in the inputs, $\beta \in [0, 1]$.

Furthermore, we assume that the foreign supplier engages in actions that are designed to increase its bargaining share, which we refer to as "haggling" and which cannot be avoided due to fully incomplete contracts ex post. The original exposition of these activities by Williamson (1985) viewed them as pure opportunism, where firms try to tilt the balance in their favor by wasteful effort; the most obvious examples are perhaps suppliers spending excessive time on price negotiations or over-staffing their factories under cost plus contracting. Alternatively, the downstream firm may have to take costly precautions in the form of inventories or the development of alternative sources for an input (Klein et al., 1978). Since then, other researchers have examined further mechanisms that rely on contracting problems ex post, like adaptation (Forbes and Lederman, 2009) or coordination (Hart and Holmstrom, 2010). Here, inefficiencies arise because upstream and downstream parties have to take ex post actions in response to unforeseen shocks that have to be aligned in some way. To the extent that they are not a productive inefficiency emerges, which is typically disruptive for the buyer's supply chain. What is common to all these frameworks is that ownership by one party fully erases such inefficiencies at the cost of some "governance cost", which remains largely unspecified. Arguably, these TCE considerations are particularly salient in international transactions, where both contractual incompleteness and monitoring problems tend to be more severe.

We capture these TCE mechanisms in a reduced form way by assuming that the part of the total surplus from the relationship that accrues to the multinational in stage 3 is diminished by a factor β^{γ} with $\gamma > 0$.[55, 56] A higher ownership share reduces the ex post inefficiency – until haggling is eliminated with $\beta = 1$ – in line with the idea in Hart and Holmstrom (2010) that ownership can confer residual decision rights: if productive assets are owned by the downstream firm, it can threaten to fire the upstream firm (now an employee) if it behaves in an undesired, opportunistic way. Due to the functional form, the TCE inefficiency increases in strength with γ, and is shut down with $\gamma = 0$.

Furthermore, γ is a function of δ, where $\partial\gamma/\partial\delta > 0$. Intuitively, we thus acknowledge that failure to make aligned decisions with more important suppliers will cause a greater inefficiency or that haggling by crucial upstream suppliers will cause greater disruption. If output responds more elastically to the quantity/quality of an input, the same (proportional) deficiency will lead to a greater loss in output. Consequently, the same threat to hold the downstream firm up will carry more weight in haggling and a supplier of a technologically important input can extract relatively larger rents. Alternatively, mis-coordination in an uncertain environment will cause a relatively higher efficiency loss, for example because of higher precautionary inventories.

Finally, we refrain from modelling an atheoretical "bureaucracy" cost. As will be explicit below, even without it we have a well defined trade-off between vertical integration

[55]While the upstream supplier may face a private cost for haggling in reality, it is immaterial in our setting (the downstream firm allocates property rights unilaterally).

[56]For the most part, TCE models have been notoriously informal. As a consequence, there is, in contrast to PRT, no standard approach even to the basic ingredients of a TCE framework. In our intentionally stylized model a micro-foundation would introduce complications without adding to the insights we are after. For examples of somewhat more explicit structural approaches, see Bajari and Tadelis (2001), Bresnahan and Levin (2013), Legros and Newman (2013), or Alfaro et al. (2016).

and outsourcing.[57], [58], [59]

3.1.4 Solution

We solve the game via backward induction and focus on a single input n (we drop the input indexation from now on to improve the exposition). The surplus generated by adding an input n of quantity/quality m to the final product is equal to m^δ, of which the offshore supplier gets a share $1-\beta$. Consequently, it n has to solve

$$\max_m \; (1-\beta)m^\delta - c_M m,$$

which leads it to optimally invest ($*$ denotes optimal choices)

$$m^* = \left(\frac{(1-\beta)\delta}{c_M}\right)^{\frac{1}{1-\delta}}. \tag{4}$$

This result illustrates the typical insight gained when contracts are incomplete ex ante and can therefore not support a first best solution. Since the supplier only obtains a fraction $(1-\beta)$ of the surplus, and the investment is sunk at stage 3, it will choose to underinvest ex ante. This investment distortion caused by downstream ownership is stronger when δ is large: an important supplier has a larger absolute loss in its marginal investment benefit from downstream ownership and hence limits investment more severely. More important inputs in terms of δ also receive more investment in quality ceteris paribus as the marginal return on these is higher.

The multinational chooses β to maximize its total profits in the first stage, which include the ex post inefficiency as a proportional cost. A simple way to derive predictions regarding β is to consider what Antras and Chor (2013) call the "unconstrained problem", i.e. choosing the value of β freely from $\mathbb{R}$. The total surplus that accrues to the downstream firm is

[57] We can assume a separate cost of vertical integration in the spirit of TCE, but our results are the same, and unnecessarily more complicated, as long as this cost's elasticity to ownership is less sensitive to δ than $\gamma(\delta)$.

[58] A comment regarding ex ante transfers – for example due to ex ante market power of the downstream firm and ensuing take-it-or-leave-it offers – is in order. Allowing for these implies that the multinational firm maximizes the *joint* ex ante surplus of the relationship by picking β. Since it can appropriate all profits through the transfer, there is no incentive to increase its ownership and hence all inputs are outsourced. Clearly, there is no heterogeneity across inputs. This result relies, however, on our assumption that there are no relationship-specific investments to be made downstream. We do not incorporate ex ante transfers here to keep the baseline model as simple as possible.

[59] In our model, we interpret all products as intermediate inputs. Accordingly, we have checked that our baseline result is fully robust to restricting the sample to intermediates.

$$\left(\frac{\delta}{c_M}\right)^{\frac{\delta}{1-\delta}} \underbrace{\beta}_{\text{PRT benefit}} \underbrace{\beta^{\gamma}}_{\text{TCE benefit}} \underbrace{(1-\beta)^{\frac{\delta}{1-\delta}}}_{\text{PRT cost}}. \tag{5}$$

The surplus illustrates the main trade-off that shapes the downstream firm's ownership decision. First, a higher β will directly increase the slice of the total pie it obtains, which we call the "PRT benefit" of ownership. Secondly, higher ownership directly increases the pie through curbing the ex post inefficiency (less haggling). We refer to this aspect as the "TCE benefit" of ownership.[60] The final term in (5) captures the costs of downstream ownership, namely that ex ante investments are distorted under fully incomplete contracts. We call this part "PRT cost" in the spirit of the property rights literature (Grossman and Hart, 1986).

The optimal choice of β is

$$\beta^* = \frac{(1+\gamma)(1-\delta)}{\delta + (1+\gamma)(1-\delta)}. \tag{6}$$

The solution rests on a balance of PRT (ex ante) and TCE (ex post) forces. Setting $\gamma = 0$ shuts down all ex post inefficiencies and allows us to focus on the predictions from a pure PRT model. The PRT force pushes for more important inputs to be outsourced as $\partial\beta^*/\partial\delta|_{\gamma=0} = -1 < 0$. The multinational chooses $\beta = 0$ in order to give maximal investment incentives to the supplier. For $\gamma > 0$, however, the overall effect of technological importance on ownership is the result of the PRT force and the TCE force combined. The latter pushes for vertical integration, since control over the foeign supplier limits haggling and coordination losses. Taking all these insights together, we can derive the following lemma.

Lemma 1 *More important inputs are more likely to be integrated iff*

$$\varepsilon_{1+\gamma,\delta} > \frac{1}{1-\delta},$$

where $\varepsilon_{k,l}$ is the elasticity of k with respect to l.

Proof. The derivative

$$\frac{\partial\beta^*}{\partial\delta} = \frac{[\delta + (1+\gamma)(1-\delta)]\Delta - (1+\gamma)(1-\delta)(1+\Delta)}{[\delta + (1+\gamma)(1-\delta))]^2}$$

[60] In line with the key prediction of TCE – that the costs of outsourcing increase in the total appropriable quasi-rents – the TCE benefit is larger whenever the surplus of the relationship is bigger (see, for example, Monteverde and Teece, 1982; Masten, 1984; Joskow, 1985).

with

$$\Delta \equiv \gamma'(1-\delta) - (1+\gamma).$$

This expression is stricly positive iff

$$\frac{\gamma'\delta}{1+\gamma} > \frac{1}{1-\delta}.$$

■

Intuitively, if the TCE force (embodied in $\varepsilon_{1+\gamma,\delta}$) is relatively strong in an elasticity sense, it can overpower the PRT force (embodied in $1/(1-\delta)$ above) and lead to a result consistent with our main finding.

One final comment is in order. As discussed, if we shut down the TCE force, the highly stylized PRT framework we are left with unambiguously predicts that more important inputs will be *less* likely to be vertically integrated — inconsistently with our empirical findings. While arguably most PRT frameworks in the literature would also have this feature, it is possible to make assumptions that create the opposite prediction. In particular, Nowak et al. (2016) show that, depending on relative parameter values, technological and demand side interactions between inputs that are combined in CES fashion can make integration of more important inputs optimal. Second, if the downstream firm's outside option (not explicitly modelled in our framework) under outsourcing becomes increasingly bad as we move to more important suppliers, it is possible that, at some point, the balance tilts towards vertical integration.[61] Finally, the marginal investment costs could be assumed to depend positively on δ (see Acemoglu et al., 2010). All these assumptions are either very difficult to test empirically or seem unlikely to hold except in special cases, or both. Since TCE type theories of the firm enjoy substantial empirical support, in our view it is more promising to view the world as being shaped by both ex ante *and* ex post forces.[62]

[61] Kohler and Smolka (2018) also construct a PRT framework with multiple suppliers and show that more productive firms integrate a larger share of their suppliers – but they do not show which ones.

[62] The equilibrium payment to the supplier of input n is proportional to

$$[1-\beta(n)^*]\,[m(n)^*]^{\delta(n)} = \left[\frac{\delta(n)^{1+\delta(n)}}{1+(1-\delta(n))\gamma(\delta(n))}\right]^{\frac{1}{1-\delta(n)}} c_M^{\frac{\delta(n)}{\delta(n)-1}},$$

so that for small enough marginal investment costs and therefore high enough contributions to output $m(n)^*$, the cost share increases in technological importance in our stylized conceptual framework.

3.2 Extensions

A simpler alternative to incomplete contracting both ex ante and ex post – and with the same implication regarding vertical integration and cost shares – is of course a pure TCE framework. In the following two subsections we extend our baseline model in two different ways that give us insights about the conditions under which the PRT force may be stronger or weaker relative to the TCE force. We test the predictions derived from these extensions empirically in the subsequent section and thus examine whether PRT forces are in fact needed to understand our main finding – as a large body of literature on multinationals would suggest.

3.2.1 Contracting Environment

We first explore how the effect of technological significance depends on ex ante contractibility. We analyze our baseline model, but introduce the following generalizations. The foreign suppliers no longer choose a single investment under fully incomplete contracts ex ante, but make a continuum of investment choices $x_n(j)$, $j \in [0,1]$, which translate into quantity/quality through $m(n) = exp[\int_0^1 ln\ x_n(j) \mathrm{d}j]$. We assume that all investments $j < \mu$ with $\mu \in [0,1]$ are fully contractible and are chosen by the multinational after ownership has been allocated, but before the offshore supplier has made its investment choices (Acemoglu et al., 2007). All investments with $j \geq \mu$ are fully non-contractible ex ante and the costs of investing are $c_M x_n(j)$.

In sum, μ serves as a parameter that indicates the quality of contracting institutions or the inverse of contract intensity ex ante: If $\mu = 1$ all contracts are fully complete and enforceable ex ante, while with $\mu = 0$ we are back in the case of the baseline model, i.e. with fully incomplete contracts ex ante. Our comparative static of interest will be about this parameter.

The solution of the model proceeds as before and is relegated to the theory Appendix B.2.1. The optimal solution for the ownership share is now

$$\beta^* = \frac{(1+\gamma)[1-\delta(1-\mu)]}{\delta(1-\mu)+(1+\gamma)[1-\delta(1-\mu)]}, \tag{7}$$

and we can, once again, show under what condition we obtain a prediction consistent with our main empirical result.

Lemma 2 *More important inputs are more likely to be integrated iff*

$$\varepsilon_{1+\gamma,\delta} > \frac{1}{1-\delta(1-\mu)}.$$[63]

Proof. The derivative

$$\frac{\partial \beta^*}{\partial \delta} = (1-\mu)\frac{\delta\gamma'[1-\delta(1-\mu)] - (1+\gamma)}{\{\delta(1-\mu) + (1+\gamma)[1-\delta(1-\mu)]\}^2} > 0 \tag{8}$$

iff

$$\frac{\delta\gamma'}{1+\gamma} > \frac{1}{1-\delta(1-\mu)}.$$

■

We are now in a position to state the main prediction about how a better contracting environment ex ante affects our baseline result.

Prediction 1 *A better contracting environment ex ante leads to a stronger relationship between the importance of an input and the probability that it is integrated.*

The proof for this prediction is relegated to the Theory appendix B.2.2, but the intuition is the following. Making ex ante investments more contractible reduces the incentive for the downstream firm to outsource more important inputs (relative to less important ones): Assuming control over an offshore supplier is now less costly in general, as many sub-investments are contractually fixed and there is less underinvestment overall. Since underinvestment is always particularly problematic for more important inputs, these become even more likely to be integrated. In other words, the PRT force that pushes for outsourcing becomes weaker, while the TCE force remains strong.

3.2.2 Headquarters intensity

For our second prediction, we return to our baseline model, but assume that the multinational firm has to make a complementary investment $h(n)$ into every input's quality. These investments can be interpreted as effort to adapt an input to the overall product or how diligently and carefully an input is processed during production. Consequently, the new production function in terms of quantity/quality is

[63] It is easy to verify that there always exists an $a > 0$ such that $\forall \delta \in (0,1)$ this expression is satisfied with $\gamma(\delta) = \exp(a/(1-\delta)) - 1$.

$$y = \sum_{n}^{N} \Big(h(n)^{\eta} m(n)^{1-\eta} \Big)^{\delta}, \tag{9}$$

where η is our parameter of interest. It captures the relative importance of the two investments and is commonly referred to as headquarters intensity.[64] For $\eta = 0$ we are back in our baseline setting. All other baseline assumptions remain intact and we omit the input indexation.

The solution of the model can be found in the Theory appendix B.2.3 again. The optimal ownership share of the multinational is now

$$\beta^* = \frac{(1+\gamma)[1-\delta(1-\eta)]}{\delta(1-\eta)+(1+\gamma)[1-\delta(1-\eta)]}. \tag{10}$$

Lemma 3 once again shows that if the TCE force is stronger than the PRT force, the model with headquarters intensity predicts a relationship between technological importance and integration that is consistent with our main empirical result.

Lemma 3 *More important inputs are more likely to be integrated iff*

$$\varepsilon_{1+\gamma,\delta} > \frac{1}{1-\delta(1-\eta)}.$$

Proof. The derivative

$$\frac{\partial \beta^*}{\partial \delta} = (1-\eta) \frac{\delta\gamma'[1-\delta(1-\eta)] - (1+\gamma)}{\{\delta(1-\eta)+(1+\gamma)[1-\delta(1-\eta)]\}^2} > 0$$

iff

$$\frac{\delta\gamma'}{1+\gamma} > \frac{1}{1-\delta(1-\eta)}.$$

■

Prediction 2 *Higher headquarters intensity leads to a stronger relationship between the importance of an input and the probability that it is integrated.*

The proof for this prediction is fully analogous to the proof for prediction 1 in the Theory appendix B.2.2. If the downstream firm has an important investment to make, the underinvestment costs of vertical integration that pushed for outsourcing of technologically more important inputs (relative to less important ones) in the baseline model

[64] Since it is the multinational that undertakes a backwards vertical integration decision in our setting, it is called "headquarters".

are lower. After all, the multinational can both substitute for the foeign supplier's investment with its own contribution and encourage it through the complementarity implied by Cobb-Douglas technology in (9). As a consequence, the incentives to give ownership to the supplier are reduced and this effect is stronger for more important inputs, which become relatively more likely to be 'made' as opposed to 'bought'.

4 Testing further Empirical Predictions

In this section we implement empirical tests of the two theoretical predictions we derived in the previous part and examine the roles of contracting environment and headquarters intensity in turn. We first briefly outline and discuss the empirical strategy. Then we introduce the additional data sources required for the exercise and finally present our results.

4.1 Contracting Environment

Empirical Specification

To test prediction 1, we interact our main variable of interest, *costshare*, with proxies for contractibility, i.e., the ease with which contracts can be written and enforced. The structural equation for this exercise is

$$intrashare_{ijpc} = \beta_1 \; costshare_{ip} + \beta_2 \; costshare_{ip} \times \mathbb{1}(contractibility) + \alpha_i + \gamma_{cj} + \phi_{cp} + \varepsilon_{ijpc}. \quad (11)$$

We have deliberately omitted the index for our *contractibility* variable in order to highlight that we employ different empirical measures – to be introduced in the next subsection – that vary along various dimensions. Our fixed effects will, however, always absorb its level effect. The interacted indicator variables take the value one whenever a characteristic is above the median of its relevant distribution – for example, if a firm characteristic is above the within 4 digit NAF industry median.

We employ the same identification strategy as for the baseline results in column (3) of Table 2. In addition, we instrument the interaction term with our instrument interacted with the measure for contractibility. In line with Prediction 1, we expect that the sign of our estimates of β_2 will be positive.

Data

For the contracting environment, we first make use of three country level variables, namely the index for property rights protection from V-Dem, a rule of law index from the World Governance Indicators, and an intellectual property rights protection index from Park (2008).[65]

We furthermore take inspiration from Nunn (2007) and compute several different contractibility measures that rely on product differentiation. First, we concord the liberal Rauch index to our HS 4 digit products (using French import values in 1996 as weights) and generate an indicator that equals one whenever a product is not traded on an exchange or reference priced in trade journals. This is our first variable. Secondly, we create a Nunn type contractibility measure at the buyer firm level by calculating the 1996 import value weighted share of homogeneous inputs in total products sourced, which is the second variable we employ. Finally, we calculate the same measures for the downstream industries of our buyer firms to obtain the third variable. We also concord the routineness measure developed by Costinot et al. (2011) to our upstream industries k.[66]

Results

Our results are reported in Table 5. Each column from (1) to (7) introduces a separate interaction variable. Except for the Rauch indicator and firm level contractibility, all variables are significantly positive and very precisely estimated. Overall, we interpret these results as rich evidence in support of Prediction 1, because our measures of contractibility capture several different aspects of the contracting environment in which international trade takes place.

Robustness

Larger, more productive firms tend to be different from their less well performing peers along many dimensions, which is documented by a large literature. It is therefore possible that we mistake mere scale effects – larger firms can afford to establish subsidiaries abroad – for our mechanisms of interest. To address this issue, we control

[65] The country indicator variables are equal to one if a country is among the top 25 in the world.

[66] We cannot rule out the possibility that the proxies we use may also partially capture ex post contractibility. In Appendix B.2.4 we show that higher ex post contractibility would typically work in the opposite direction of prediction 1, i.e. weaken the effect of technological importance. If we do find a positive interaction effect in our exercise here, it is therefore (even stronger) evidence in favor of PRT forces being operative.

for (and instrument) interactions of cost shares with firm size (log employment) and productivity (value added per worker) measures. The results are reported in appendix Table B.6 and show that the patterns we find are fully robust to this concern.

4.2 Headquarters Intensity

Empirical Specification

For testing our prediction 2 regarding headquarters intensity, we estimate equation

$$intrashare_{ijpc} = \beta_1\ costshare_{ip} + \beta_2\ costshare_{ip} \times \mathbb{1}(hqintensity)_j^{down} + \alpha_i + \gamma_{cj} + \phi_{cp} + \varepsilon_{ijpc}. \quad (12)$$

We measure importance of investment at the downstream industry j level and instrument all level and interaction terms. As in the previous exercise, we use dummy variables that indicate when headquarters intensity is larger than the median. According to our theoretical predictions our estimates of β_2 are expected to have a positive sign.

Data

As is common in the literature, we proxy headquarters intensity alternatively by physical or intangible capital intensity, skill intensity, or service intensity at the downstream firm using the EAE as our data source. Service intensity hereby refers to the share of service sector employees in total employment of a firm.[67] Finally, we make use of the RnD intensity variable constructed by Nunn and Trefler (2013) for the whole world.

Results

Results are reported in Table 6. All estimates for the headquarters intensity interactions are positive. We obtain highly significant results for capital intensity, which is very much in line with the findings in Antras (2003). Skill intensity as a proxy for the provision of key conceptual input into the relationship exhibits a similarly significant pattern. In contrast with Acemoglu et al. (2009) and Nunn and Trefler (2013) we do not find a significant impact of RnD intensity, which may stem from the fact that there is considerable measurement error in this variable due to classification crosswalks. Overall, we interpret these findings as evidence in favour of our second prediction.

[67] All variables are described in Appendix B.1.1.

Robustness

Addressing the same concerns about scale effects as above – larger, more productive firms are more capital intensive etc. than their smaller counterparts – we control for labor productivity and employment interactions. The results are reported in appendix Table B.7.

Moreover, the exercise for headquarters intensity is complicated by the fact that what matters for ownership decisions in the theory is the *relative* importance of the upstream to the downstream investment (cf. Acemoglu et al., 2010). In our extended model above, this was captured by the parameters η and $1-\eta$. Empirically, we would not, however, expect that upstream and downstream importance of the (marginal) investment follow a one-to-one relationship. In a final robustness exercise we therefore add upstream investment intensity as a control, i.e., the same interaction variables as before, but for the 4 digit ISIC Rev. 3 industry that produces the product imported into France. The results (appendix Table B.8) are fully robust to this check.[68]

5 Conclusion

Technological importance of an input – in the sense of a cost share – is a major determinant of the make-or-buy decision: Important inputs are significantly more likely to be produced in-house. We use detailed trade and firm level data from France to document this fact and show that it is robust and economically significant. Through the lens of a stylized model with incomplete contracting between a buyer and its suppliers, we interpret this finding as the combination of two effects. In general, firms want to outsource production of their most important inputs to encourage their suppliers' investments into the quality of an input. However, trading high cost share intermediates at arm's length potentially leads to expensive adaptation/coordination failures and opportunistic behavior on the part of the upstream partner, so that vertical integration may be favored. Our baseline estimates are consistent with a world in which the latter, TCE type incentives dominate. We provide additional empirical evidence for this interpretation of our estimates. Consistent with predictions from an extended version of our model, the positive relationship between technological importance and the likelihood of in-house sourcing is stronger if a) contracts are more complete and b) the downstream firm also

[68] Interestingly, upstream investment intensity carries a negative sign almost throughout, which is fully consistent with our theoretical model.

has an important relationship-specific investment to make. Therefore, both PRT *and* TCE type forces are needed to fully explain the empirical relationship between technological importance and vertical integration observed in the data.

Our work highlights two promising avenues for further research on multinational firms. First, we believe that the characteristics of supply networks and of the respective markets have a substantial bearing on how firms organize production. There already exist a few important contributions in this area (e.g., Antras and Chor, 2013; Alfaro et al., forthcoming), but more work is needed to provide the theoretical and empirical evidence sought-after by policy makers and academics alike (e.g., Bresnahan and Levin, 2012). The results in this article make progress on this frontier. Secondly, with a few exceptions, trade economists view multinational activity as shaped by the risk of underinvestment in the spirit of the property rights theory of the firm. While this paradigm delivers explanations for a variety of empirical patterns, its pervasiveness may also lead to a more narrow research agenda. Our results, which complement other work such as Costinot et al. (2011), provide further encouragement to view the international trade landscape through a wider range of conceptual lenses, including transaction cost economics.

References

Acemoglu, Daron, Philippe Aghion, Rachel Griffith, and Fabrizio Zilibotti, "Vertical Integration and Technology: Theory and Evidence," *Journal of the European Economic Association*, 2010, *8* (September), 989–1033.

_ **, Pol Antras, and Elhanan Helpman**, "Contracts and Technology Adoption," *American Economic Review*, 2007, *97* (3), 916–943.

_ **, Simon Johnson, and Todd Mitton**, "Determinants of Vertical Integration: Financial Development and Contracting costs," *Journal of Finance*, 2009, *64* (3), 1251–1290.

_ **, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza TahbazSalehi**, "The Network Origins of Aggregate Fluctuations," *Econometrica*, September 2012, *80* (5), 1977–2016.

Alessandria, George, Joseph P. Kaboski, and Virgiliu Midrigan, "Inventories, Lumpy Trade, and Large Devaluations," *American Economic Review*, 2010, *100* (5), 2304–2339.

Alfaro, Laura and Andrew Charlton, "Intra-Industry Foreign Direct Investment," *American Economic Review*, 2009, *99* (5), 2096–2119.

_ **, Nicholas Bloom, Paola Conconi, Harald Fadinger, Patrick Legros, Andrew F. Newman, Raffaella Sadun, and John Van Reenen**, "Come Together: Firm Boundaries and Delegation," *NBER Working Paper No. 24603*, 2018.

_ **, Paola Conconi, Harald Fadinger, and Andrew F. Newman**, "Do Prices Determine Vertical Integration?," *Review of Economic Studies*, 2016, *83* (3), 855–888.

_ **, Pol Antras, Davin Chor, and Paola Conconi**, "Internalizing Global Value Chains: A Firm-Level Analysis," *Journal of Political Economy*, forthcoming.

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

_ **and Davin Chor**, "Organizing the Global Value Chain," *Econometrica*, 2013, *81* (6), 2127–2204.

_ **and Elhanan Helpman**, "Global Sourcing," *Journal of Political Economy*, 2004, *112* (3), 552–580.

_ **and Robert W. Staiger**, "Offshoring and the Role of Trade Agreements," *American Economic Review*, December 2012, *102* (7), 3140–3183.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson, "Vertical Integration and Input Flows," *American Economic Review*, 2014, *104* (4), 1120–1148.

Bajari, Patrick and Steven Tadelis, "Incentives versus Transaction Costs: A Theory of Procurement Contracts," *The RAND Journal of Economics*, 2001, *32* (3), 387–407.

Baker, George P. and Thomas N. Hubbard, "Make Versus Buy in Trucking:Asset Ownership, Job Design, and Information," *American Economic Review*, 2003, *93* (3), 551–572.

Baldwin, Richard E. and Anthony J. Venables, "Spiders and Snakes: Offshoring and Agglomeration in the Global Economy," *Journal of International Economics*, 2013, *90* (2), 245–254.

Bernard, Andrew B., Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes, "The Origins of Firm Heterogeneity: A Production Network Approach," NBER Working Papers 25441, National Bureau of Economic Research, Inc January 2019.

_ , J. Bradford Jensen, and Peter K. Schott, "Transfer Pricing by U.S.-Based Multinational Firms," *mimeo*, 2006.

Blaum, Joaquin, Claire Lelarge, and Michael Peters, "The Gains from Input Trade with Heterogeneous Importers," *American Economic Journal: Macroeconomics*, October 2018, *10* (4), 77–127.

Boehm, Johannes, Swati Dhingra, and John Morrow, "Product Diversification in Indian Manufacturing," CEP Discussion Papers dp1509, Centre for Economic Performance, LSE November 2017.

Bolatto, Stefano, Alireza Naghavi, Gianmarco I. P. Ottaviano, and Katja Zajc Kejzar, "Intangible Assets and the Organization of Global Supply Chains," Working Papers wp1105, Dipartimento Scienze Economiche, Universita' di Bologna August 2017.

Bresnahan, Timothy and Jonathan Levin, "Vertical Integration and Market Structure," in Robert Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*, Introductory Chapters, Princeton University Press, 2012.

_ and _ , "Transaction Cost Economics," in Robert Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*, Princeton University Press, 2013.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 2015, *82* (1), 1–44.

Carluccio, Juan and Maria Bas, "The Impact of Worker Bargaining Power on the Organization of Global Firms," *Journal of International Economics*, 2015, *96* (1), 162–181.

_ **and Thibault Fally**, "Global Sourcing under Imperfect Capital Markets," *Review of Economics and Statistics*, 2012, *94* (3), 740–763.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Agnes Cornell, Sirianne Dahlum, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Brigitte von Röomer Johannes Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi ting Wang, Tore Wig, Steven Wilson, and Daniel Ziblatt, "Varieties of Democracy (V-Dem) Project," 2018.

Corcos, Gregory, Delphine M. Irac, Giordano Mion, and Thierry Verdier, "The Determinants of Intrafirm Trade : Evidence From French Firms," *The Review of Economics and Statistics*, 2013, *95* (July), 825–838.

Costinot, Arnaud, Lindsay Oldenski, and James Rauch, "Adaptation and the Boundary of Multinational Firms," *The Review of Economics and Statistics*, 2011, *93* (1), 298–308.

Davies, Ronald B., Julien Martin, Mathieu Parenti, and Farid Toubal, "Knocking on Tax Haven's Door: Multinational Firms and Transfer Pricing," *Review of Economics and Statistics*, 2018, *100* (1), 120–134.

Defever, Fabrice and Farid Toubal, "Productivity, Relationship-specific Inputs and the Sourcing Modes of Multinationals," *Journal of Economic Behavior and Organization*, 2013, *94*, 345–357.

Del Prete, Davide and Armando Rungi, "Organizing the Global Value Chain: A Firm-level Test," *Journal of International Economics*, 2017, *109*, 16–30.

Diez, Federico J., "The Asymmetric Effects of Tariffs on Intra-firm Trade and Offshoring Decisions," *Journal of International Economics*, 2014, *93* (1), 76–91.

Eppinger, Peter and Bohdan Kukharskyy, "Contracting Institutions and Firm Boundaries," *mimeo*, 2017.

Fattorini, Loredana, Armando Rungi, and Zhen Zhu, "The Organization of Global Supply Networks," *IMT Lucca EIC Working Paper Series*, 2017, *8.*

Feenstra, Robert C. and J. Bradford Jensen, "Evaluating Estimates of Materials Offshoring from US Manufacturing," *Economics Letters*, 2012, *117* (1), 170 – 173.

Flaaen, Aaron, "The Role of Transfer Prices in Profit-Shifting by U.S. Multinational Firms: Evidence from the 2004 Homeland Investment Act," *mimeo*, 2017.

Forbes, Silke J. and Mara Lederman, "Adaptation and Vertical Integration in the Airline Industry," *American Economic Review*, 2009, *99* (5), 1831–1849.

Garetto, Stefania, "Input Sourcing and Multinational Production," *American Economic Journal: Macroeconomics*, 2013, *5* (2), 118–151.

Gibbons, Robert, "Four formal(izable) theories of the firm?," *Journal of Economic Behavior and Organization*, 2005, *58* (2), 200–245.

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

_ , _ , and Adam Szeidl, "Optimal Integration Strategies for the Multinational Firm," *Journal of International Economics*, 2006, *70*, 216–238.

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

Hart, Oliver D. and Bengt Holmstrom, "A Theory of Firm Scope," *The Quarterly Journal of Economics*, 2010, *125* (May), 483–513.

_ and John Moore, "Property Rights and the Nature of the Firm," *The Journal of Political Economy*, 1990, *98* (6), 1119.

Irarrazabal, Alfonso, Andreas Moxnes, and Luca D. Opromolla, "The Margins of Multinational Production and the Role of Intrafirm Trade," *Journal of Political Economy*, 2013, *121* (1), 74–126.

Joskow, Paul L., "Vertical Integration and Long-term Contracts: The Case of Coal-burning Electric Generating Plants," *Journal of Law, Economics, and Organization*, 1985, *1* (1), 33–80.

Kalnins, Arturs, Stephen F. Lin, and Catherine Thomas, "In-House and Arm's Length: Productivity Heterogeneity and Variation in Organizational Form," *CEP Discussion Paper Series*, 2018, *1579.*

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "The Worldwide Governance Indicators: Methodology and Analytical Issues," *Hague Journal on the Rule of Law*, 2011, *3* (02), 220–246.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian, "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *The Journal of Law & Economics*, 1978, *21* (2), 297–326.

Kohler, Wilhelm and Marcel Smolka, "Productivity and Firm Boundaries," *CESifo Working Paper Series*, 2018, *7214.*

Lafontaine, Francine and Margaret Slade, "Vertical Integration and Firm Boundaries: The Evidence," *Journal of Economic Literature*, 2007, *45* (3), 629–685.

Legros, Patrick and Andrew F. Newman, "A Price Theory of Vertical and Lateral Integration," *Quarterly Journal of Economics*, 2013, *128* (2), 725–770.

_ **and** _ , "Contracts, Ownership, and Industrial Organization: Past and Future," *Journal of Law, Economics, and Organization*, 2014, *30*, 82–117.

Levchenko, Andrei A., "Institutional Quality and International Trade," *Review of Economic Studies*, 2007, *74* (3), 791–819.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–341.

Macchiavello, Rocco, "Financial Development and Vertical Integration: Theory and Evidence," *Journal of the European Economic Association*, 2012, *10* (2), 255–289.

Masten, Scott E., "The Organization of Production: Evidence from the Aerospace Industry," *The Journal of Law & Economics*, 1984, *27* (2), 403–417.

Monteverde, Kirk and David J. Teece, "Supplier Switching Costs and Vertical Integration in the Automobile Industry," *The Bell Journal of Economics*, 1982, *13* (1), 206–213.

Nowak, Verena, Christian Schwarz, and Jens Suedekum, "Asymmetric Spiders: Supplier Heterogeneity and the Organization of Firms," *Canadian Journal of Economics*, 2016, *49* (2), 663–684.

Nunn, Nathan, "Relationship-specificity, Incomplete Contracts, and the Pattern of Trade," *Quarterly Journal of Economics*, 2007, *122* (2), 569–600.

— **and Daniel Trefler**, "Incomplete Contracts and the Boundaries of the Multinational Firm," *Journal of Economic Behavior and Organization*, 2013, *94*, 330–344.

Ornelas, Emanuel and John L. Turner, "Protection and International Sourcing," *Economic Journal*, 2011, *122* (559), 26–63.

Park, Walter G., "International Patent Protection: 1960-2005," *Research Policy*, 2008, *37* (4), 761–766.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi ting Wang, Joshua Krusell, and Farhad Miri, "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data," *University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21*, 2018, *3.*

Ramondo, Natalia, Veronica Rappoport, and Kim J. Ruhl, "Intrafirm Trade and Vertical Fragmentation in U.S. Multinational Corporations," *Journal of International Economics*, 2016, *98* (March), 51–59.

Rauch, James E., "Networks versus Markets in International Trade," *Journal of International Economics*, 1999, *48* (1), 7–35.

Ruhl, Kim J, "How Well is US Intrafirm Trade Measured?," *American Economic Review: Papers & Proceedings*, 2015, *105* (5), 524–529.

Shu, Pian and Claudia Steinwender, "The Impact of Trade Liberalization on Firm Productivity and Innovation," *Innovation Policy and the Economy*, 2019, *19* (1), 39–68.

Whinston, Michael D., "Assessing the Property Rights and Transacion-Cost Theories of Firm Scope," *American Economic Review*, 2001, *91* (2), 184–188.

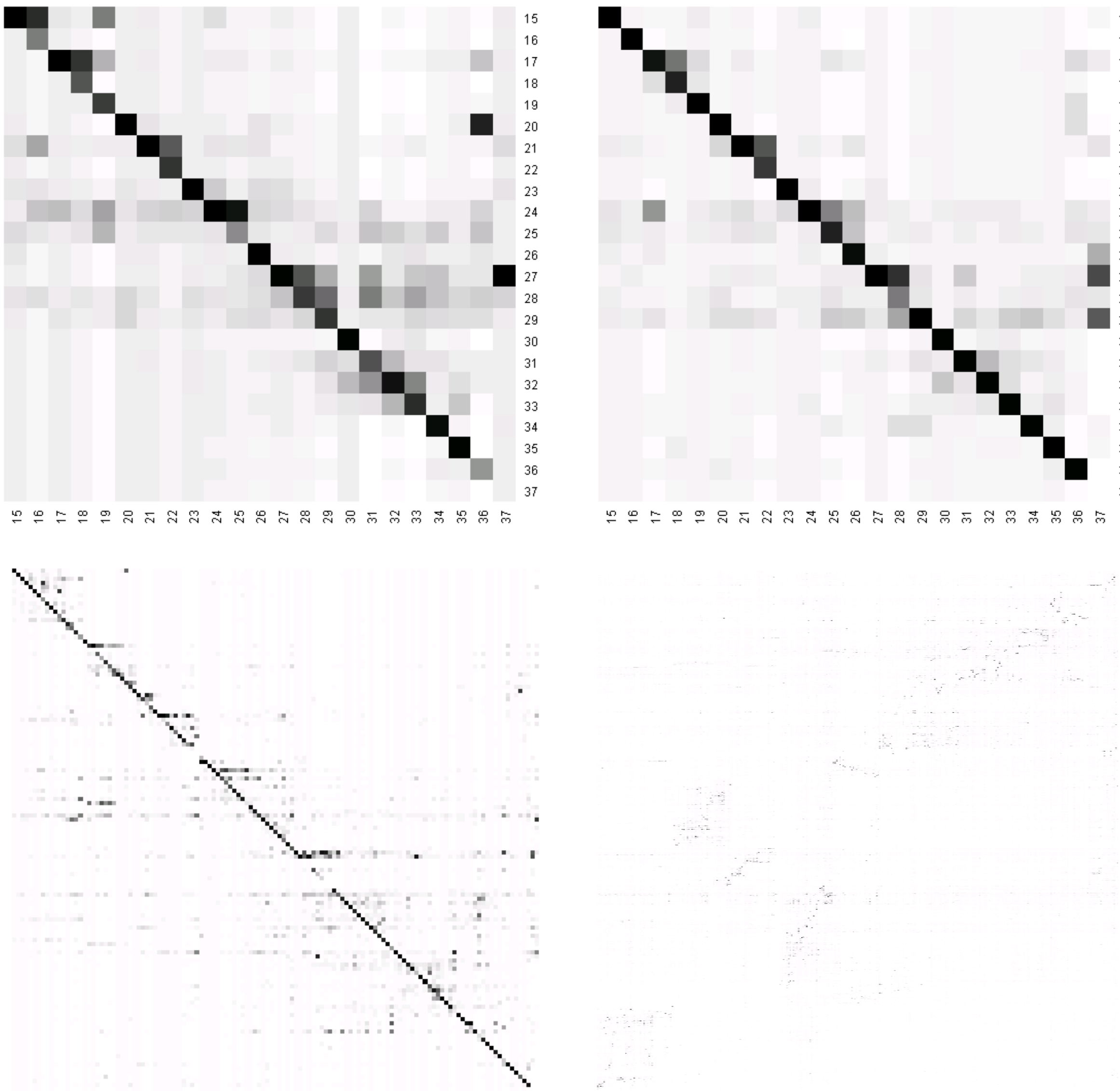
Williamson, Oliver E., *The Economic Institutions of Capitalism*, New York: The Free Press, 1985.

Wooldridge, Jeffrey M., "On Estimating Firm-level Production Functions using Proxy Variables to Control for Unobservables," *Economics Letters*, 2009, *104* (3), 112–114.

Appendices

A Figures and tables for the main text

Figure 1: Contour Plots of Various French IO Tables



In reading pattern starting with upper left: Official 2 digit, 2 digit self-constructed, 4 digit symmetric self-constructed, 4 digit asymmetric self-constructed.

Figure 2: Empirical Density of Direct Requirements (Asymmetric French IO Table in 1996)

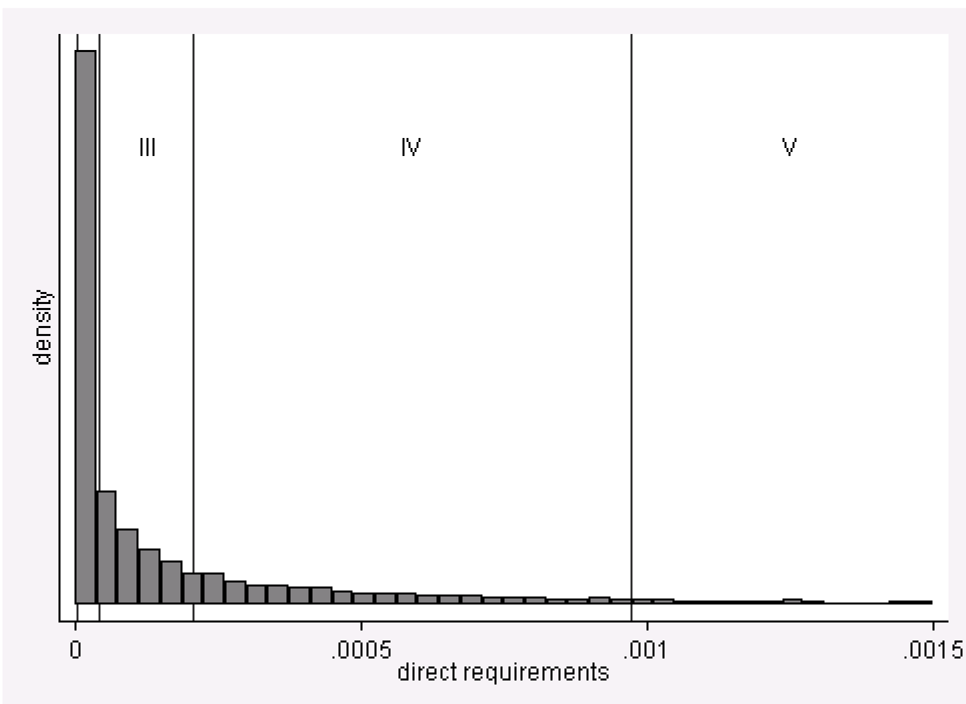


Table 1: Summary Statistics

	mean	median	sd	count
Employment	468	198	1,187	3,101
Sales	160.4k	38.8k	1,137.8k	3,149
Average Input Cost Share	0.0068	0.0012	0.0166	3,101
Average Intrafirm Trade Share	0.27	0.09	0.34	3,151
Average Number of Products	10	7	12	3,151
Capital Intensity	900	445	7103	3,097
Intangible Cap. Int.	106	17	1022	2,965
Skill Intensity	185	172	71	3,097
TFP Wooldridge (ln)	1.53	1.24	1.16	2,997
VA per worker	1,262	650	7,788	3,090

Summary statistics are computed at the firm level and refer to imports only. The average intrafirm trade share is the within firm average computed along the input $\times$ country dimension. The average input cost share is the within firm average computed along the input dimension. The average number of products is the within firm average computed along the input dimension. All variables are explained in Appendix B.1.1.

Table 2: Baseline Estimates

VARIABLES	(1) intra-firm share	(2) intra-firm share	(3) intra-firm share	(4) intra-firm share	(5) intra-firm share
cost share	3.222***	2.479***	11.586***	11.281***	10.246***
	(0.550)	(0.320)	(1.400)	(1.615)	(2.108)
Country*HS4 product FE		YES	YES	YES	YES
Country*NAF 4dig FE		YES	YES	YES	YES
Firm		YES	YES	YES	YES
Observations	76,897	70,016	70,016	70,016	70,016
R-squared	0.015	0.686	0.638	0.641	0.651
Instrument			Micro 1996 excl own firm	Micro China 2006 excl France	Official US 2002 4 digit
KP-stat 1st stage			241.8	103.4	98.66

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The regressor is the input $\times$ firm level cost share in the firm's total expenditure on intermediates. Common sample across columns (2)-(5) and singletons are always removed. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Baseline Robustness

VARIABLES	(1) intra-firm share	(2) intra-firm share	(3) intra-firm share	(4) intra-firm share	(5) intra-firm share	(6) intra-firm share	(7) intra-firm share
cost share	11.577***	10.706***	8.533**	6.229**	10.354***	12.301***	12.497***
	(1.392)	(2.164)	(4.111)	(2.979)	(1.517)	(1.284)	(1.384)
Country*HS4 product FE	YES	YES	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES	YES	YES
Firm	YES	YES	YES		YES		YES
ISIC 4dig FE*HS4 product			YES				
Up ISIC 4dig*Firm				YES			
Country*Firm						YES	
Observations	75,549	55,201	72,024	68,856	42,848	66,385	60,712
Sample	all	(1)	all	all	(2)	all	(3)
R-squared	0.638	0.697	0.736	0.825	0.653	0.754	0.646
Instrument				Micro 1996 excl own firm			
KP-stat 1st stage	231.3	98.2	19.9	40.5	118.5	274.2	217.9

The dependent variable is the input × country × firm level share of intrafirm import value in overall imports of the firm for that input × country pair. The regressor is the input × firm level cost share in the firm's total expenditure on intermediates. Sample (1): We drop all observations in which the downstream importer i is mainly active in the industry j that also produces the good sourced (p). In IO terminology, this means we drop all observations on the IO diagonal. Since our asymmetric table does not have a well defined diagonal, we exclude all observations on the 4 digit ISIC Rev. 3 diagonal. Sample (2): Firms in highly competitive industries, i.e. those where the Hirschman-Herfindahl-Index in 1999 was below 0.1. The Hirschman-Herfindahl-Index $HHI^{downstr}$ is computed within NAF 4 digit industries, using the population of French firms' sales. Sample (3): All import flows from EU countries. The French industry classification NAF contains 285 manufacturing industries, while ISIC is more aggregated with its 122 categories. Number of observations varies due to different sets of singletons (dropped) and availability of tariffs. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Horse Race with Integration Determinants at Firm Level

VARIABLES	(1) intra-firm share	(2) intra-firm share	(3) intra-firm share	(4) intra-firm share	(5) intra-firm share	(6) intra-firm share	(7) intra-firm share
cost share	0.214***	0.216***	0.215***	0.207***	0.209***	0.220***	0.220***
	(0.413)	(0.413)	(0.408)	(0.402)	(0.406)	(0.414)	(0.401)
(log) capital intensity		0.068***					0.042**
		(0.005)					(0.007)
(log) intangible cap. Int.			0.056***				0.025**
			(0.003)				(0.003)
(log) skill intensity				0.101**			0.100***
				(0.061)			(0.042)
(log) VA per worker					0.101***		
					(0.008)		
(log) TFP						0.144***	0.134***
						(0.009)	(0.010)
IM ratio	0.181***	0.169***	0.176***	0.178***	0.173***	0.177***	0.134***
	(0.027)	(0.027)	(0.027)	(0.032)	(0.027)	(0.027)	(0.010)
Country*HS4 product FE	YES	YES	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES	YES	YES
Observations	258,020	258,020	258,020	258,020	258,020	258,020	258,020
R-squared	0.370	0.372	0.372	0.377	0.377	0.372	0.381
Instrument				Micro 1996 excl own firm			
KP-stat 1st stage	137.1	137	134.4	136.6	137.4	136.3	133

The dependent variable is the input × country × firm level share of intrafirm import value in overall imports of the firm for that input × country pair. The main regressors are the firm × input level cost share in total expenditure on intermediates (instrumented) and the IM ratios from a Heckman first stage non-response adjustment (for details, see Appendix B.1.2). Details on how the other variables are constructed can be found in Appendix B.1.1. The French industry classification NAF contains 285 manufacturing industries. Singletons are dropped. Standard errors in parentheses are bootstrapped and two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Contracting Environment

CONTRACTIBILITY PROXY	(1) PR Protect.	(2) Rule of Law	(3) IPR Protect.	(4) Contractibility Product	(5) Contractibility Firms	(6) Contractibility Industry	(7) Upstr. Routineness
cost share	10.048***	9.507***	2.507	12.106***	11.430***	9.735***	8.588***
	(1.623)	(1.796)	(2.269)	(1.363)	(1.315)	(1.784)	(2.236)
$\times$ $\mathbb{1}$(proxy)	3.564***	3.100**	9.827***	0.297	1.559	6.076**	5.369**
	(1.039)	(1.483)	(2.323)	(2.448)	(1.337)	(2.308)	(2.501)
Country*HS4 product FE	YES	YES	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES	YES	YES
Firm	YES	YES	YES	YES	YES	YES	YES
Observations	72,307	72,307	72,307	72,307	72,307	72,307	72,307
R-squared	0.637	0.638	0.640	0.638	0.636	0.627	0.633
Instrument				Micro 1996 excl own firm			
KP-stat 1st stage	108	86.56	23	42.09	111.2	81.93	64.72

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The regressor is the input $\times$ firm level cost share in the firm's total expenditure on intermediates. The interaction variables are described in Appendix B.1.1 and we interact with dummies that equal one if a variable is above the median across the relevant distribution or if a country belongs to the top 25 origins according to an index. Common sample imposed across all columns and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Headquarters Intensity

HQ Intensity Proxy	(1) RnD Intensity	(2) Capital Intensity	(3) Intangible Cap. Intensity	(4) Skill Intensity	(5) Service Intensity
cost share	10.917***	8.607***	11.180***	9.188***	11.635***
	(1.985)	(1.335)	(1.553)	(1.589)	(2.041)
$\times\ \mathbb{1}($ proxy$)$	1.884	7.151***	1.991	4.996*	0.877
	(2.390)	(1.990)	(2.470)	(2.557)	(2.901)
Country*HS4 product FE	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES
Firm	YES	YES	YES	YES	YES
Observations	72,329	72,329	72,329	72,329	72,329
R-squared	0.635	0.630	0.636	0.628	0.636
Instrument			Micro 1996 excl own firm		
KP-stat 1st stage	64.28	80.20	76.62	58.88	102.9

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The regressor is the input $\times$ firm level cost share in the firm's total expenditure on intermediates. The interaction variables are described in appendix B.1.1 and we interact with dummies that equal one if a variable is above the median across the relevant distribution. Common sample imposed across all columns and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Online Appendix

B.1 Data

In this appendix we first describe how we construct our variables. In the second part, we give a detailed account of our replication of the Heckman correction procedure following Corcos et al. (2013).

B.1.1 Variables

In this section we provide the details on how we construct our variables for the empirical analysis.

- *totcost*: Total intermediate costs are computed from EAE and defined as the sum of purchases of goods (R210), purchases of raw materials (R212), and other purchases and charges (R214). In a robustness check we add total labor costs (R216), social contributions (R217), and other charges (R222).
- Gross Output: It is computed from FICUS and defined as the sum of turnover (CATOTAL), change in inventories (PRODIMM, PRODSTO), and other revenues (AUTPREX).
- Capital Intensity: The ratio of the physical capital stock to total employment, where the capital stock is measured as the total of tangible capital assets at end of year (EAE item I150) and total employment is the total number of full time equivalent employees (EAE item E101).
- Intangible Capital Intensity: Same as capital intensity, but uses only the total stock of intangible capital assets at end of year (EAE item I456).
- Skill Intensity: Is defined as average wage, i.e., the ratio of total wage expenses to the employment of the firms, as in Corcos et al. (2013).
- TFP: It is computed using the revised Levinsohn and Petrin (2003) methodology proposed by Wooldridge (2009). The coefficient of a Cobb-Douglas value-added production function are estimated at the 3 digit NACE industry level using intermediate inputs as the proxy for the productivity shock. Real value added is obtained by double-deflation using deflators for output and intermediates from the OECD STAN database. Output is defined as the sum of turnover (R310) and

other sales (R315), while intermediate costs are defined as above. Capital is defined as above and deflated using a deflator for gross fixed capital formation. All variables are logged. TFP at the firm level is then calculated as a residual between the actual and predicted value added using the estimated coefficients.

- Scale: average number of employees over the year (EAE item E101).
- Value added per employee: It is computed from EAE, with value added defined as the difference between turnover (R310) and the sum of purchases of goods (R210) and purchases of raw materials (R212).
- Firm level contract intensity: the variable is constructed using information about firms imports. The firm-level contract intensity is an import value weighted average of the contract intensity of its inputs, where the measure of contract intensity is a dummy equal to one if an input is (liberally) classified as differentiated in Rauch (1999). It is therefore similar to the measure used in in Nunn (2007) and Corcos et al. (2013), except that we weight by import value.
- Industry level contract intensity: same as firm level contract intensity, but weighting by total downstream industry (NAF) imports.
- Service intensity: ratio of workers employed in branches that produce services (NACE codes from 50 to 93) to total employment.
- Property Rights Protection: We use the V-Dem indicator "v2xcl_prpty", which codes the property rights index. The description reads: "Private property includes the right to acquire, possess, inherit, and sell private property, including land. Limits on property rights may come from the state which may legally limit rights or fail to enforce them; customary laws and practices; or religious or social norms. This question concerns the right to private property, not actual ownership of property." The scale is ordinal, but converted to the unit interval. See Coppedge et al. (2018) and Pemstein et al. (2018).
- Rule of Law: We use the Rule of Law index for 1998 provided in the World Governance Indicators, see Kaufmann et al. (2011).
- IPR Protection: We use the IPR protection index provided by Park (2008).
- Routine Task Intensity: We concord the indices (average routineness content of tasks) provided in Costinot et al. (2011) to our NAF industry classification.

- Relationship-specificity: We use the classification in Rauch (1999) and recode every HS 4 input to "relationship-specific" if it is differentiated, while it is "not relationship-specific" if it is traded on exchanges or reference priced (aggregation to HS 4 digit is done by weighting with 1996 import values). Therefore, our variable is a simple indicator at the product level.

B.1.2 Selection Adjustment for Horse Race

As discussed in Corcos et al. (2013), the EIIG survey suffered from significant non-response, which the authors addressed with a Heckman control function approach. All our results with firm fixed effects are fully robust to these concerns due to the fact that we use only within firm variation. Our quantitative explorations in the cross section are liable to sample selection bias, however. Consequently, we implement the two step procedure suggested by Corcos et al. (2013). We are grateful to these authors for giving us access to the target population of EIIG.

In a first step we add the 3,841 firms in the EIIG target population that did not respond to the survey, giving us a total of $4,305 + 3,841 = 8,146$ firm observations in the target population.[1] In line with the focus of our paper as well as of Corcos et al. (2013), we focus on the 5,333 firms that operate mainly in manufacturing. Using the universe of customs data for these firms, we are able to compute three excluded selection variables as in Corcos et al. (2013). In particular, the selection stage of our two step Heckman procedure is

$$\begin{aligned}\mathbb{1}(responded)_i = \alpha + \beta_1\, log(importvalue_i) \\ + \beta_2\, log(\#\ products) + \beta_3\, log(\#\ countries) + \gamma_k + \eta_i,\end{aligned} \tag{B.1}$$

where we regress a dummy for whether or not a firm responded to the survey on its total import value, the number of imported HS 4 digit products, and the number of origin countries for these flows in 1999. γ_k denotes a set of indicators for the firms' three digit ISIC Rev. 3 industry codes in the output market. Having estimated (B.1), we predict inverse Mill's (IM) ratios for our firms.

Again following Corcos et al. (2013) as closely as possible, we extend the EIIG sample with a random sample of non-multinational manufacturing firms who are present in the EAE and who traded more than one million EUR in the previous year (large traders).

[1] Our numbers deviate slightly from those reported in Corcos et al. (2013), since various versions of the EIIG data were published over time.

To do so, we first obtain the full EAE sample of large trader manufacturers by means of the universe of customs data and subtract the EIIG firms. Then we draw a 52.85 percent random sample and add it to the EIIG firms. To run our second stage regressions, we add all necessary international trade and firm level information for this composit sample.[2] In the random sample of non-multinational manufacturers we set the IM ratios equal to zero and assume that all their imported inputs come from third parties – just as in Corcos et al. (2013). We bootstrap all standard errors, clustering two-way at the 3 digit up- and downstream industry level as throughout the paper.

[2]This is equivalent to what Corcos et al. (2013) call their 'large sample', but at the level of HS 4 digit, rather than CPA.

B.2 Theory

B.2.1 Solution for the model with partially incomplete contracting ex ante

The surplus created by the relationship can now be written as

$$\left\{exp\left[\int_0^1 \ln x_n(j)\mathrm{d}j\right]\right\}^{\delta}$$

of which the supplier obtains a share $(1-\beta)$ at the third stage. It hence chooses its non-contractible investments $x_n^{nc}(k)$ in stage 2 to maximize its profits minus the costs $\int_\mu^1 c_M x_n^{nc}(j)\mathrm{d}j$. The optimal choice is

$$x_n^{nc,*}(j) = x_n^{nc,*} = \frac{(1-\beta)\delta}{c_M}(m_n^c)^{\delta}(m_n^{nc,*})^{\delta}, \tag{B.2}$$

where m^c denotes

$$exp\left[\int_0^{\mu} \ln x_n^c(j)\mathrm{d}j\right],$$

i.e. the composite of all contractible investments. Note that non-contractible investments are fully symmetric, so that we can compute the index of non-contractible investments as

$$m_n^{nc,*} = \left(\frac{(1-\beta)\delta}{c_M}\right)^{\frac{1-\mu}{1-\delta(1-\mu)}} (m^c)^{\frac{\delta(1-\mu)}{1-\delta(1-\mu)}}. \tag{B.3}$$

The downstream firm now chooses the levels of the contractible investments in order to maximize its profit from the relationship. The reader should keep in mind that the TCE costs fall exclusively on the downstream firm. Therefore, the profits are

$$\beta\beta^{\gamma}(m^c)^{\delta}(m^{*,nc})^{\delta} - \int_0^{\mu} c_M x_n^c(j)\mathrm{d}j,$$

where we assume that the supplier can recuperate the investment costs fully in its contract with the downstream firm – it therefore does not matter, *who* pays the investment costs.

Optimal contractible investments are

$$x_n^{c,*} = \beta\beta^{\gamma}(1-\beta)^{\frac{\delta(1-\mu)}{1-\delta(1-\mu)}}\left(\frac{\delta}{c_M}\right)^{\frac{1}{1-\delta(1-\mu)}}\frac{1}{1-\delta(1-\mu)}(m_n^{c,*})^{\frac{\delta}{1-\delta(1-\mu)}}. \tag{B.4}$$

Again, we note that all contractible investments are symmetric. Therefore the expression for the index of contractible investments is

$$m_n^{c,*} = \beta^{\frac{\mu[1-\delta(1-\mu)]}{1-\delta}} \beta^{\gamma \frac{\mu[1-\delta(1-\mu)]}{1-\delta}} (1-\beta)^{\frac{\mu\delta(1-\mu)}{1-\delta}} \left(\frac{\delta}{c_M}\right)^{\frac{\mu}{1-\delta}} [1-\delta(1-\mu)]^{-\frac{\mu[1-\delta(1-\mu)]}{1-\delta}}. \tag{B.5}$$

We can now plug the expressions (B.3) and (B.5) into m^c and m^{nc} to obtain the downstream firm's net profits from the relationship

$$\beta\beta^{\gamma}(m^{c,*})^{\delta}(m^{nc,*})^{\delta} = \beta^{\frac{1-\delta(1-\mu)}{1-\delta}} \beta^{\gamma \frac{1-\delta(1-\mu)}{1-\delta}} (1-\beta)^{\frac{\delta(1-\mu)}{1-\delta}} \left(\frac{\delta}{c_M}\right)^{\frac{\delta}{1-\delta}} [1-\delta(1-\mu)]^{-\frac{\mu\delta}{1-\delta}}, \tag{B.6}$$

which it maximizes by choosing β at the first stage of the game. The first order condition is

$$(1+\gamma)[1-\delta(1-\mu)](1-\beta^*) = \delta(1-\mu)\beta^*.$$

The solution to this first order condition is given in expression (7) in the main text.

B.2.2 Proof of Prediction 1

We want to show that the second order derivative of the optimal ownership share (7) is strictly positive. First, note that we can rewrite the derivative $\partial\beta^*/\partial\delta$ in (8) as

$$\underbrace{\frac{1-\mu}{\{1+\gamma[1-\delta(1-\mu)]^2\}}}_{\equiv g(\mu)} \times \underbrace{\{\delta\gamma'[1-\delta(1-\mu)] - (1+\gamma)\}}_{\equiv f(\mu)}.$$

We furthermore defined the two parts of the first order derivative as functions of μ. Restating the goal of the proof, we want to show that

$$\frac{\partial f(\mu)/\partial\mu}{f(\mu)} > -\frac{\partial g(\mu)/\partial\mu}{g(\mu)}.$$

We find that

$$-\frac{\partial g(\mu)/\partial\mu}{g(\mu)} = \frac{1+\gamma[1+\delta(1-\mu)]}{1+\gamma[1-\delta(1-\mu)]}$$

and

$$\frac{\partial f(\mu)/\partial \mu}{f(\mu)} = \frac{\delta\gamma'[1-\delta(1-\mu)]-(1+\gamma)}{\delta\gamma'}.$$

After a few algebraic steps one arrives at the inequality

$$\delta\gamma'[2\delta(1-\mu)-1]-\delta\gamma'\gamma[1-\delta(1-\mu)] > -(1+\gamma)\{1+\gamma[1+\delta(1-\mu)]\} \tag{B.7}$$

Next, we assume that the inequality in Lemma 2, i.e. $\delta\gamma'[1-\delta(1-\mu)] > 1+\gamma$ is satisfied. Note that dividing (B.7) by this inequality gives us a sufficient condition for $(\partial\beta^*)^2/\partial\delta\partial\mu > 0$, namely

$$\frac{2\delta(1-\mu)}{1-\delta(1-\mu)} - \gamma > -\{1+\gamma[1+\delta(1-\mu)]\}.$$

Rearranging this inequality shows that our proof is successful:

$$1 > \gamma[1-\delta(1-\mu)].$$

B.2.3 Solution for the model with headquarters intensity

The surplus of the relationship can be written as

$$h^{\delta\eta}m^{\delta(1-\eta)}.$$

Therefore, in the second stage, the two firms choose their respective investments to maximize

$$(1-\beta)h^{\delta\eta}m^{\delta(1-\eta)} - c_M m$$

in the case of the supplier and

$$\beta\beta^{\gamma}h^{\delta\eta}m^{\delta(1-\eta)} - c_H h$$

in the case of the downstream firm. Solving this system of equations we find the optimal investments, namely

$$m^* = (1-\beta)^{\frac{1-\delta\eta}{1-\delta}}\beta^{(1+\gamma)\frac{\delta\eta}{1-\delta}}\left(\frac{\delta\eta}{c_H}\right)^{\frac{\delta\eta}{1-\delta}}\left(\frac{\delta(1-\eta)}{c_M}\right)^{\frac{1-\delta\eta}{1-\delta}}$$

and

$$h^{*}=(1-\beta)^{\frac{\delta(1-\eta)}{1-\delta}}\beta^{(1+\gamma)\frac{1-\delta(1-\eta)}{1-\delta}}\left(\frac{\delta\eta}{c_H}\right)^{\frac{1-\delta(1-\eta)}{1-\delta}}\left(\frac{\delta(1-\eta)}{c_M}\right)^{\frac{\delta(1-\eta)}{1-\delta}}.$$

Using these two optimal investment levels, we can compute the downstream firm's payoff at stage 1,

$$(1-\beta)^{\frac{\delta(1-\eta)}{1-\delta}}\beta^{(1+\gamma)\frac{1-\delta(1-\eta)}{1-\delta}}\left(\frac{\delta\eta}{c_H}\right)^{\frac{\delta\eta}{1-\delta}}\left(\frac{\delta(1-\eta)}{c_M}\right)^{\frac{\delta(1-\eta)}{1-\delta}}.$$

The solution for the optimal ownership share is given by the first order condition from maximizing the last expression with respect to β.

B.2.4 The model with incomplete contracts ex post

We assume that the ex post inefficiency now depends on μ^p:

$$\beta^{(1-\mu^p)\gamma},$$

while ex ante contracts are fully incomplete without loss of generality. With this specification we assume that the downstream firm can avoid some of the haggling or coordination/adaptation frictions through enforceable contracts. Clearly, more contractibility (higher μ^p) reduces the ex post inefficiency and makes it less sensitive to changes in γ.

The solution of the model proceeds as before in the baseline and imperfect contracting cases. The optimal solution for the ownership share becomes

$$\beta^{*}=\frac{[1+\gamma(1-\mu^p)](1-\delta)}{\delta+[1+\gamma(1-\mu^p)](1-\delta)}. \tag{B.8}$$

We can, once again, show under what condition we obtain a prediction consistent with our main empirical result.

Lemma 4 *More important inputs are more likely to be integrated iff*

$$\varepsilon_{1+\gamma(1-\mu^p),\delta}>\frac{1}{1-\delta}.$$

Proof. The derivative

$$\frac{\partial\beta^{*}}{\partial\delta}=\frac{\delta\gamma'(1-\mu^p)(1-\delta)-[1+\gamma(1-\mu^p)]}{\{\delta+[1+\gamma(1-\mu^p)](1-\delta)\}^{2}}>0 \tag{B.9}$$

iff

$$\frac{\delta\gamma'(1-\mu^p)}{1+\gamma(1-\mu^p)} > \frac{1}{1-\delta}.$$

■

We now show that, in our typical setting, ex post contractibility weakens the effect of technological importance on vertical integration. First, rewrite

$$\frac{\partial\beta^*}{\partial\delta} = \frac{(1-\mu^p)[\delta\gamma'(1-\delta)-\gamma]-1]}{[1+\gamma(1-\mu^p)(1-\delta)]^2}.$$

We want to show that

$$\frac{(\partial\beta^*)^2}{\partial\delta\partial\mu^p} < 0$$

or

$$\delta\gamma'[\gamma(1-\delta)(1-\mu^p)-1] < \gamma\frac{1-2\delta}{1-\delta}. \tag{B.10}$$

In our data, the cost shares of the vast majority of products is very low as can be discerned from the summary statistics presented above. Consequently, δ will in general be quite low and we expect that

$$\gamma(1-\delta)(1-\mu^p)-1 < 0,$$

and

$$\gamma\frac{1-2\delta}{1-\delta} > 0$$

in which case the inequality (B.10) is satisfied.

B.3 Figures and tables

Figure B.1: Empirical Density of Input Cost Shares

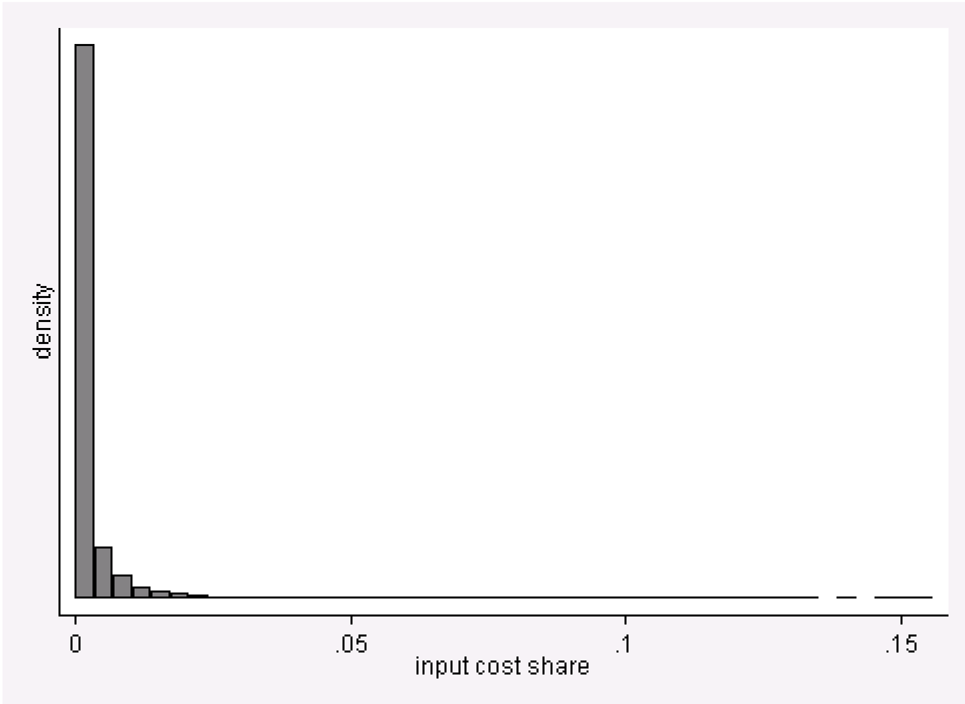


Table B.1: Baseline First Stages

VARIABLES	(1) cost share	(2) cost share	(3) cost share
Quintile Micro Table FR '96	0.002***		
	(0.000)		
Quintile Micro Table CN '06		0.002***	
		(0.000)	
Quintile 4 digit IO table US '02			0.003***
			(0.000)
Country*HS4 product FE	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES
Firm	YES	YES	YES
Observations	70,016	70,016	70,016
R-squared	0.597	0.593	0.594
KP-stat 1st stage	241.8	103.4	98.66

The dependent variable is the firm $\times$ input level cost share in total expenditure on intermediates. The 1996 French import IO table was constructed dropping a firm's own trade flows. The 2006 Chinese IO table was constructed dropping all imports from France. Common sample imposed across all columns and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.2: Baseline Reduced Forms

VARIABLES	(1) intra-firm share	(2) intra-firm share	(3) intra-firm share
Quintile Micro Table FR '96	0.023***		
	(0.002)		
Quintile Micro Table CN '06		0.019***	
		(0.002)	
Quintile 4 digit IO table US '02			0.025***
			(0.005)
Country*HS4 product FE	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES
Firm	YES	YES	YES
Observations	70,016	70,016	70,016
R-squared	0.638	0.641	0.651

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The 1996 French import IO table was constructed dropping a firm's own trade flows. The 2006 Chinese IO table was constructed dropping all imports from France. Common sample imposed across all columns and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.3: Baseline Estimates (diagonal dropped)

VARIABLES	(1) intra-firm share	(2) intra-firm share	(3) intra-firm share	(4) intra-firm share	(5) intra-firm share
cost share	2.373***	1.878**	10.196***	11.741***	18.446**
	(0.706)	(0.756)	(2.191)	(3.092)	(7.495)
Country*HS4 product FE		YES	YES	YES	YES
Country*NAF 4dig FE		YES	YES	YES	YES
Firm		YES	YES	YES	YES
Observations	56,253	50,654	50,654	50,654	50,654
R-squared	0.004	0.718	0.699	0.692	0.643
Instrument			Micro 1996 excl own firm	Micro China 2006 excl France	Official US 2002 4 digit
KP-stat 1st stage			102.7	73.75	16.74

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The regressor is the firm by input level cost share in total expenditure on intermediates. Common sample imposed across columns (2)-(5) and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.4: Baseline Estimates (aggregated across countries)

VARIABLES	(1) intra-firm share	(2) intra-firm share	(3) intra-firm share	(4) intra-firm share	(5) intra-firm share	(6) intra-firm share
cost share	4.347***	3.098***	13.473***	14.372***	14.329***	8.651*
	(0.765)	(0.558)	(1.777)	(2.020)	(3.096)	(4.393)
Firm		YES	YES	YES	YES	YES
HS4 product		YES	YES	YES	YES	YES
NAF 4dig FE*HS4 product						YES
Observations	28,469	28,014	28,014	28,014	28,014	23.945
R-squared	0.018	0.640	0.573	0.561	0.562	0.714
Instrument			Micro 1996 excl own firm	Micro China 2006 excl France	Official US 2002 4 digit	Micro 1996 excl own firm
KP-stat 1st stage			225	119.2	234.4	19.55

The dependent variable is the input × firm level share of intrafirm import value in overall imports of the firm for that input. The regressor is the firm by input level cost share in total expenditure on intermediates. Common sample imposed across columns (2)-(5) and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.5: Additional Robustness

VARIABLES	(1) intra-firm share	(2) intra-firm share	(3) intra-firm share	(4) intra-firm share	(5) intra-firm share
cost share	10.706***			10.206***	12.562***
	(2.164)			(1.261)	(1.421)
$\times$ $\mathbb{1}$(Eff. Appl. Tariff)	1.499				
	(12.830)				
$\times$ $\mathbb{1}$(distance)					-0.155
					(1.232)
import value share		0.611***			
		(0.073)			
(log) import value			0.032***		
			(0.004)		
Country*HS4 product FE	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES
Firm	YES	YES	YES	YES	
Observations	75,533	76,154	70,596	51,348	66,385
Sample	all	all	all	(1)	(2)
R-squared	0.638	0.670	0.684	0.646	0.754
Instrument			Micro 1996 excl own firm		
KP-stat 1st stage	118.5	273.1	236.2	197.9	274.2

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input. The regressor is the firm by input level cost share in total expenditure on intermediates. Data on effectively applied tariffs come from TRAINS; the interaction of cost shares with a dummy equal one if a country $\times$ product tariff lies above the sample median is instrumented with the same interaction with the instrument. The analogous strategy is used for origin country distance to France, which comes from the CEPII gravity data set. Sample (1): Firms whose overall share of imported inputs in overall expenditure on intermediates is above the median across all firms. Sample (2): All import flows from EU countries. Singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.6: Contract Environment: Scale and Productivity Controls

CONTRACTIBILITY PROXY	(1) PR Protect.	(2) Rule of Law	(3) IPR Protect.	(4) Contractibility Product	(5) Contractibility Firms	(6) Contractibility Industry	(7) Upstr. Routineness
cost share	5.616***	5.324***	-1.672	7.813***	7.325***	4.951**	4.895**
	(1.488)	(1.453)	(2.920)	(1.407)	(1.252)	(1.952)	(2.022)
$\times$ $\mathbb{1}$(proxy)	4.514**	3.073**	10.103***	0.293	1.207	6.310***	5.058**
	(1.773)	(1.518)	(2.416)	(2.473)	(1.411)	(2.322)	(2.544)
Country*HS4 product FE	YES	YES	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES	YES	YES
Firm	YES	YES	YES	YES	YES	YES	YES
Observations	72,307	72,307	72,307	72,307	72,307	72,307	72,307
R-squared	0.641	0.643	0.645	0.642	0.641	0.631	0.638
Instrument	Micro 1996 excl own firm						
KP-stat 1st stage	53.80	42.11	12.32	22.64	43.72	39.17	48.59

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The regressor is the input $\times$ firm level cost share in the firm's total expenditure on intermediates. The interaction variables are described in appendix B.1.1, coefficients for employment and value added per worker are omitted, and we interact with dummies that equal one if a variable is above the median across the relevant distribution or if a country belongs to the top 25 origins according to an index. Common sample imposed across all columns and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.7: Headquarters Intensity: Scale and Productivity Controls

HQ Intensity Proxy	(1) RnD Intensity	(2) Capital Intensity	(3) Intangible Cap. Intensity	(4) Skill Intensity	(5) Service Intensity
cost share	9.590***	7.445***	10.478***	8.496***	10.662***
	(2.467)	(1.274)	(1.469)	(1.625)	(2.016)
$\times$ $\mathbb{1}$(proxy)	2.369	7.495***	1.711	4.722*	1.042
	(2.564)	(1.892)	(2.472)	(2.507)	(2.829)
Country*HS4 product FE	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES
Firm	YES	YES	YES	YES	YES
Observations	72,173	72,173	72,173	72,173	72,173
R-squared	0.632	0.627	0.634	0.626	0.634
Instrument			Micro 1996 excl own firm		
KP-stat 1st stage	27.42	40.37	39.29	28.38	50.08

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The regressor is the input $\times$ firm level cost share in the firm's total expenditure on intermediates. The interaction variables are described in appendix B.1.1, coefficients for employment and value added per worker are omitted, and we interact with dummies that equal one if a variable is above the median across the relevant distribution. Common sample imposed across all columns and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.8: Headquarters Intensity: Upstream Controls

HQ Intensity Proxy	(1) RnD Intensity	(2) Capital Intensity	(3) Intangible Cap. Intensity	(4) Skill Intensity	(5) Service Intensity
cost share	12.931***	8.759***	10.641***	11.155***	12.962***
	(2.291)	(1.371)	(1.776)	(1.890)	(2.333)
$\times$ $\mathbb{1}$(proxy downstr)	3.469	7.293***	1.649	6.552**	2.796
	(2.813)	(2.369)	(2.413)	(2.901)	(3.304)
$\times$ $\mathbb{1}$(proxy upstr)	-4.272	-0.409	1.442	-4.704	-4.771
	(3.361)	(2.484)	(2.850)	(2.891)	(2.946)
Country*HS4 product FE	YES	YES	YES	YES	YES
Country*NAF 4dig FE	YES	YES	YES	YES	YES
Firm	YES	YES	YES	YES	YES
Observations	72,329	72,329	72,329	72,329	72,329
R-squared	0.637	0.630	0.635	0.633	0.635
Instrument			Micro 1996 excl own firm		
KP-stat 1st stage	34.91	27.35	36.21	36.38	34.87

The dependent variable is the input $\times$ country $\times$ firm level share of intrafirm import value in overall imports of the firm for that input $\times$ country pair. The regressor is the input $\times$ firm level cost share in the firm's total expenditure on intermediates. The interaction variables are described in appendix B.1.1 and we interact with dummies that equal one if a variable is above the median across the relevant distribution. Common sample imposed across all columns and singletons are dropped. The French industry classification NAF contains 285 manufacturing industries. Standard errors in parentheses are two-way clustered at the 3 digit downstream ISIC Rev. 3 industry and at the 3 digit upstream ISIC Rev. 3 level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1582	Michael Amior	The Contribution of Foreign Migration to Local Labor Market Adjustment
1581	Grant Bickwit Emanuel Ornelas John L. Turner	Preferential Trade Agreements and Global Sourcing
1580	Maria Molina-Domene	Labor Specialization as a Source of Market Frictions
1579	Arturs Kalnins Stephen F. Lin Catherine Thomas	In-House and Arm's Length: Productivity Heterogeneity and Variation in Organizational Form
1578	Emanuel Ornelas Marcos Ritel	The Not-So-Generalized Effects of the Generalized System of Preferences
1577	Matthew Ridlay Camille Terrier	Fiscal and Education Spillovers from Charter School Expansion
1576	John Van Reenen	Increasing Differences Between Firms: Market Power and the Macro-Economy
1575	Margarida Madaleno Max Nathan Henry Overman Sevrin Waights	Incubators, Accelerators and Regional Economic Development
1574	Esteban M. Aucejo Patrick Coate Jane Cooley Fruehwirth, Sean Kelly Zachary Mozenter	Teacher Effectiveness and Classroom Composition
1573	Stephan Heblich Stephen J. Redding Daniel M. Sturm	The Making of the Modern Metropolis: Evidence from London

1572	Johannes Beohm Ezra Oberfield	Misallocation in the Market for Inputs: Enforcement and the Organization of Production
1571	Hans R. A. Koster Edward W. Pinchbeck	How do Households Value the Future? Evidence from Property Taxes
1570	Holger Breinlich Elsa Leromain Dennis Novy Thomas Sampson Ahmed Usman	The Economic Effects of Brexit – Evidence From the Stock Market
1569	J. Vernon Henderson Sebastian Kriticos Dzhamilya Nigmatulina	Measuring Urban Economic Density
1568	Philippe Bracke Silvana Tenreyro	History Dependence in the Housing Market
1567	Ester Faia Sebastien Laffitte Gianmarco Ottaviano	Foreign Expansion, Competition and Bank Risk
1566	Brian Bell Rui Costa Stephen Machin	Why Does Education Reduce Crime?
1565	Richard Murphy Felix Weinhardt Gill Wyness	Who Teaches the Teachers? A RCT of Peer-to-Peer Observation and Feedback in 181 Schools
1564	Paola Conconi Giovanni Facchini Max F. Steinhardt Maurizio Zanardi	The Political Economy of Trade and Migration: Evidence from the U.S. Congress

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE