

Organizing the Global Value Chain: a firm-level test

Davide Del Prete[†]
(Sapienza University of Rome)

Armando Rungi[‡]
(IMT Lucca)

July 2015

Abstract

In the last two decades, technological progress and a decrease in trade barriers fostered the formation of global value chains, for which different sequences of production stages, previously performed in close proximity, can now be unbundled globally. In this contribution we test at the firm level the optimal allocation of ownership rights along a productive sequence, as in the framework set by Antràs and Chor (2013). For this purpose we exploit an own-built dataset made of 4,214 parents which have acquired or established at least one affiliate in the period 2004-2012. Overall, they control 104,720 affiliates and operate in 185 countries. Assuming a technological orientation from the final consumer upwards, we positively test that incentives to integrate suppliers vary systematically with: i) the relative upstream or downstream position of the affiliate with respect to the parent; ii) the elasticity of demand faced by the parent. However, after the VIII decile of the affiliates' downstreamness these results are no more statistically significant, as we argue horizontal integration strategies could prevail. Further, we find scope for the role played by firm heterogeneity, as more productive and bigger parent companies prefer picking affiliates next to final consumers. Once controlling for the complexity of the internal supply chain at the moment the investment decisions occur, we find that already big internal chains show a lower propensity to integrate at the margin, probably discounting increasing coordination costs.

Keywords: global value chains, vertical integration, property rights theory, multinational enterprises, downstreamness, heterogeneous firms, business groups.

JEL codes: F14, F23, D23, G34, L20.

We would like to thank Daniel Treffer, Joanne Oxley, Kunal Dasgupta, Gianmarco Ottaviano, Dalia Marin, Carlo Altomonte, Giorgia Giovannetti, Anna Giunta, Giuseppe De Arcangelis and Pierluigi Montalbano, as well as participants at the 'European Trade Study Group 2014', 'Explaining Economic Change 2014 - Rome', 'Warsaw International Economic Meeting 2015', and at various seminars for useful comments. Financial support was provided by CrisisLab PNR Project, at IMT Institute for Advanced Studies, Lucca.

[†] Mail to davide.delprete@uniroma1.it. Department of Economics and Social Sciences, Piazzale Aldo Moro, 5 - 00185 Rome - Italy.

[‡]Mail to armando.rungi@imtlucca.it. Laboratory for the Analysis of Complex Economic Systems, piazza San Francesco 19 - 55100 Lucca - Italy.

1 Introduction

Since the late 80s, technological progress and a decrease in trade barriers have fostered the formation of cross-border sequences of productive stages undertaken by firms along virtually international assembly lines¹. From the product design to the distribution to consumers, all intermediate stages of production can involve networks of firms that are dispersed in several countries. Each production stage can be eventually organized by a company in two alternative ways, either keeping the input production within its boundaries, in case of standard vertical integration, or outsourcing it and engaging in arm's length contracts. Thereby, the ordered sequence of all production stages establishes a chain with some suppliers that are integrated or not within one or more companies' productive boundaries.

The aim of this contribution is to test for the organization of these Global Value Chains (GVCs) at the firm-level, as sequences of intermediate stages that are subject to contractual frictions. In this context, the final and optimal allocation of ownership rights along the sequence can depend both on the relative position of each intermediate producer, and on the surplus that can be extracted from the sale of the final output, on which all the producers along the chain can rely. Here we adopt and test the theoretical framework set by Antràs and Chor (2013), whose main proposition predicts that if the demand elasticity for the final output is sufficiently elastic, vertical integration occurs for the stages of the supply chain that are more proximate to the final consumer. We find confirmation of this main prediction, but we also provide evidence of some non-linearities, once the full range of demand elasticity is exploited and the parent company approaches the bottom of the supply chain. Results are robust to several specifications, also controlling for the simultaneity bias that could possibly arise after including in the same sample newly integrated affiliates and long-established corporate structures.

Moreover, we are able to detect firm-level heterogeneity in vertical integration strategies along the value chain. We find that the more productive and larger parents are able to pick affiliates located relatively downstream. Also, conditional on past integration choices, we find that more productive parents integrate less likely further stages of production. Similarly, parents already controlling complex supply chains show a lower propensity to integrate further intermediate producers as they have to discount increasing internal coordination costs.

For the purpose of our empirical analysis we exploit an own-built database of domestic and multinational Business Groups (BGs), following the methodology set in Altomonte and Rungi (2013), for which the observation unit is a parent company that organizes several affiliates through a form of hierarchical control. Hence, we complement our firm-level data with metrics of downstreamness and demand elasticity sourced respectively by Antràs and Chor (2013) and Broda and Weinstein (2006).

For example, consider the case of two multinational business groups present in our dataset: Sony and Johnson & Johnson. We report these two case studies in Tables 1 and 2. The first is a group originated in Japan and primarily focused on electronics manufacturing, whereas the second is a US multinational producing medical devices and pharmaceutical products. From our data both exhibit similar degrees of parent downstreamness, but face very different average demand elasticities and hence very different vertical integration propensities. Both BGs have a similar group size, as they control 405 and 353 affiliates respectively and, when looking at their parent outputs, they are among the most downstream in our sample (.87 and .92), but final consumers tend to be much less price sensitive in the case of Sony (elasticity around 4.79) than in the case of Johnson & Johnson (elasticity around 12.72). Accordingly, the average affiliate-level downstreamness is .50 for Sony and .71 for Johnson & Johnson. That is, when the demand is sufficiently elastic, vertical integration preferably occurs more downstream, leaving potential outsourcing of inputs relatively more upstream².

Our empirical framework allows the decision making center, i.e. the parent firm, to locate in

¹See Baldwin and Lopez-Gonzalez (2014) and previously Hummels et al. (2001) for a discussion of the relevance of the phenomenon and of the structural economic changes it entails.

²In the following analyses we will consider the affiliate-level downstreamness, going beyond average positioning. However, these correlations are systematic in our data. For example, when we take two smaller groups. Seachange International Inc., primarily active in computer and peripheral equipment manufacturing, controls 11 affiliates and has a downstream parent (.92) that faces a relatively low demand elasticity (4.77). Ashok Leyland, an Indian automobile manufacturing, controls 10 affiliates and still has a very downstream parent (.94), but in this case it faces a much higher demand elasticity (84.19) for the final output. Consistently with the above results the average affiliates' downstreamness is relatively lower for the former (.53) than for the latter (.68).

Table 1: A case of sequential substitutes (low demand elasticity)

0	0.50	Sony
0	0.50	0.87
<i>Integration (Affiliate)</i>		<i>Outsourcing</i>

Table 2: A case of sequential complements (high demand elasticity)

0	0.71	Johson & Johnson
0	0.71	0.92
<i>Outsourcing</i>		<i>Integration (Affiliate)</i>

any of the production stages, as we collect all firms' activities along the chain, controlling also for potential firm-level heterogeneity from financial accounts. Thus, we can make reference to the relative positioning as opposed to the absolute position of each parent company in relationship to vertical integration choices. Indeed, based on preliminary evidence from our sample, we find that parent companies can be often located up in the supply chain, in violation of the assumption by the Antràs and Chor (2013) model that places the decision making centers always at the bottom, where consumers buy a final product. Nonetheless, we argue that our empirical evidence shows that the main tenet of the theoretical framework is still valid, but until the VII or VIII deciles of the affiliate positioning, depending on the downstreamness metrics we adopt.

The remainder of this paper is organized as follows. In Section 2 we briefly review previous literature. Section 3 introduces the construction of our sample and the variables used in our econometric investigations. In Section 4 we present different empirical specifications and robustness checks. Finally Section 5 concludes.

2 Literature review

Several works investigated the determinants of cross-border vertical integration, i.e. the global decision to 'make or buy', as giving rise to intra-firm transactions. Among others, Antràs (2003) revealed why intra-firm trade is mainly concentrated in capital intensive industries and between capital abundant countries, while Antràs and Helpman (2004) argued that only the most productive among them are able to sustain higher sunk costs of international vertical integration in a context of heterogeneous firms, and that would explain the positive correlation existing between intra-firm trade and productivity dispersions. These theoretical models were generalized by Antràs and Helpman (2008) in order to accommodate for varying degrees of contractual frictions, finding that better contracting institutions in the suppliers' country of origin let offshoring strategies prevail³. With a broader perspective, Acemoglu et al. (2007) were the first to consider the possibility that unique headquarters have to

³For a review of firms' organization strategies and trade, see Antràs and Yeaple (2014).

commit to contracts with several suppliers, in this way extending in scope the one-shot *'make or buy'* decision. Eventually, they show that a greater contractual incompleteness leads to the adoption of less advanced technologies, even more when intermediate inputs are highly complementary.

Based on empirical findings, we can distinguish two main strands of literature.

On one hand, Nunn and Treffer (2008) and Nunn and Treffer (2013) have a primary industry/product focus when examining the determinants of U.S. imports share that occurs intra-firm. They confirm a logic of property-rights theories in the case of multinational firms: vertical integration prevails when non-contractible headquarters inputs are more relevant and productivity is higher. Similarly, Bernard et al. (2010) find that intra-firm trade mainly occurs for skill-intensive products from developing countries but for capital-intensive products from advanced countries.

On the other hand, when exploiting firm level data, Tomiura (2007) and Kohler and Smolka (2012) document the well-known productivity premia associated with multinational enterprises, this time differentiating by sourcing strategies: foreign outsourcers and exporters tend to be less productive than the firms active in FDI or involved in multiple globalization modes, but they are more productive than purely domestic firms. There is also evidence that for the most productive multinationals the chance of trading with an independent supplier is higher when they intensively use relationship-specific inputs, as in Defever and Toubal (2013). In addition, Corcos et al. (2013) support the prediction that intra-firm imports are more likely in capital- and skill-intensive firms, in highly productive firms and from countries with a good rule of law, whereas Acemoglu et al. (2009) register a greater propensity to vertically integrate in hosting countries where one can find both higher contracting costs and a more favorable financial environment. Also, as shown by Alfaro and Charlton (2009), a great share of intra-firm trade occurs already from within the same industry and Alfaro et al. (2013) ascertain that higher market prices lead to a greater propensity to vertically integrate.

However, none of the previous works consider the sequential nature of production as affecting location and organizational decisions by multinational enterprises. First efforts in this direction are made by Harms et al. (2012), who consider a particular sequence of stages and a non-monotonic variations of transportation costs along the chain to show the trade-off faced by firms between offshoring and domestic production. Moreover, Costinot et al. (2013) develop a theoretical model for which a supply chain is ordered following the ideal standardization content of production, with more standardized stages entailing a lower content of knowledge and hence a higher country growth potential. Their aim is to prove how country patterns of specialization can have consequences on income distribution on a world scale.

We instead choose to exploit the framework proposed by Antràs and Chor (2013), where the authors draw a property-rights model of firms' boundaries, dissecting the optimal allocation of ownership rights in a context where production processes are sequential and contracts between a final producer and its suppliers are all potentially incomplete. Differently from previous works, they introduce a technological ordering in production stages, so that one stage can commence only when intermediate inputs from upstream stages are complete. This is a noteworthy advance in the comprehension of the organization of Global Value Chains, since their model design is capable to proxy an actual productive environment in which a firm and its suppliers have to bargain sequentially, but on the basis of an expected surplus that is realized only after the sale of the final good, at the end of all production stages. In this sequential context the authors introduce the classic trade-off faced by each supplier, who has to undertake a relation-specific investment in order to provide a customized input that is partially non-contractible. Hence, contractual incompleteness leads the final good producer, who owns residual control rights in case of integration, to enhance its bargaining power in case of a contractual breach. However, as in similar frameworks, the strategy of vertical integration reduces the suppliers' incentives to invest in the productive relationship.

The key novelty is that there exists a dependence among all production stages because the relation-specific investment made by upstream suppliers affects also the incentives to invest in more downstream stages. If an investment made by a supplier increases the value of the marginal product, production stages can be labelled sequential complements, and it occurs when the price elasticity of final-good demand is higher than the value of the elasticity of substitution across different inputs. Conversely, if a supplier perceives a diminished product value, we have the case of sequential substitutes, in which

the demand elasticity of the final good is sufficiently low. Therefore, inputs are sequential complements (substitutes) only if the elasticity of final good demand is higher (lower) than the elasticity of substitution across the intermediates provided by the different suppliers.

At the time of writing, a preliminary successful attempt to test at the firm-level this framework is done in Alfaro et al. (2015), although with more limited information in terms of sample coverage, not controlling for affiliate-level heterogeneity, as they aggregate affiliates' information at the industry-level. More interesting, from a theoretical point of view, Alfaro et al. (2015) are extending the original model by Antràs and Chor (2013) in order to introduce asymmetry and heterogeneity in inputs, after mapping differences in marginal input productivity across stages to differences in input contractibility.

In the following analysis, we positively test the propensity towards vertical integration along sequential production stages by parents acquiring control over companies, identifying complements and substitutes cases and introducing controls for firm-level heterogeneity. We assess the role of firm heterogeneity. Then we relax the assumption of independence between past and present choices of vertical integration, to check for the endogenous formation of integrated supply chains. Thereafter, we detect some non-linearities, at the bottom of the supply chain, once the full distribution of affiliates' downstreamness and parent elasticity is exploited.

3 Sample construction

As we are interested in studying the determinants of organization of production stages by a parent company, the sample we build comprises networks of affiliates whose economic activity is coordinated by a single ultimate owner, that we consider as headquarters. In this, we rely on a methodological framework set by Altomonte and Rungi (2013), according to which the observation unit is a set of affiliates including its headquarters, whose units can be located either in the country of origin of the parent or abroad, and that are controlled after reaching an absolute majority threshold for direct and indirect equity participation. The benefit of adopting a network approach resides in the possibility to take into account both direct control, when majority of votes is reached directly at the parent-level, and indirect control, when affiliates exert control on other sub-affiliates. The phenomenon of cross-participations is often neglected in other empirical works⁴. Here we follow international standards (OECD, 2005; EUROSTAT, 2007; UNCTAD, 2009) to set a control threshold at absolute majority of stakes ($\geq 50.01\%$). Hence, for our purpose, we elaborate data from two original sources: we retrieve M&A deals occurred in the period 2004 to 2012 from the Zephyr Database and we complement this information with data on networks of affiliates existing in 2012, as sourced from the Ownership Database contained in Orbis⁵, from where we collect also firm-level financial accounts, when available, and incorporation dates.

After combining information from acquisitions and takeovers, on the basis of incorporation and acquisition dates, we are able to separate between a stock and a flow of affiliates that can help us in determining the choice of integrating production stages also as a function of the already established internal supply chains, possibly developed over a very long span of time.

As we report in Table 3, we end up with a main sample of 4,214 manufacturing parents controlling a total of 104,720 affiliates in 185 countries. Of these, 71,011 affiliates are incumbent, whereas 33,709 are affiliates newly integrated after 2004 and until 2012.

Since the main target of our analysis is the organization of production sequences before reaching a final consumer, we exclude from our analysis parent companies that are mainly active in primary or services industries, i.e. their final output is not a manufactured product⁶. Still, vertically integrated

⁴ Affiliates can participate with their own portfolio of stakes in determining the total allocation of property rights on top of the parent company. Eventually, the object of analysis becomes a Business Group, made of at least two formally autonomous legal entities that are coordinated by a unique parent, which can develop very complex organizational structures. For more details on the procedure and its benefits, see Altomonte and Rungi (2013) and Altomonte et al. (2015).

⁵ Both sources we use are compiled by Bureau van Dijk, a consultancy firm collecting companies' information for business intelligence. All in all, Zephyr contains information on over one million deals, while Orbis includes up-to-date information for over 120 million firms worldwide.

⁶ We detect about 223,167 parents that are mainly involved in primary or services industries, which can have invested

supply chains in our final sample can rely on service inputs, provided by affiliates throughout the entire production network.

The rich structure of our data allows controlling for a possible endogenous formation of internal value chains, with parents integrating activities but conditional on already established productive networks. That is, we can test if and how past and present integration choices are conditionally independent. For this purpose we will introduce a control function approach, taking advantage of the information about affiliates that were integrated before 2004, as reported in the third column of Table 3. Indeed, many groups of firms present in our data can report even century-old traditions of productive activity, going farther back any measured variable we can include as a control in our empirical specifications.

Table 3: Sample coverage I: Parents and affiliates.

	Parents	All affiliates (A+B)	Affiliates as before 2004-2012 (A)	Newly integrated affiliates (B)
Manufacturing parents (investors in 2004-2012)	4,214	104,720	71,011	33,709

Table 4: Sample coverage II: Main geographic regions.

	Parents	All affiliates (A+B)	Affiliates as before 2004-2012 (A)	Newly integrated affiliates (B)
OECD	3,371	88,041	60,867	27,174
non-OECD	843	16,679	10,144	6,535
European Union	1,421	58,508	40,620	17,888
United States	1,407	18,005	10,592	7,413
Rest of the world	1,386	28,207	19,799	8,408
<i>of which:</i>				
Japan	285	7,079	5,986	1,093
Latin America	50	2,539	1,727	812
Middle East	47	1,002	642	360
China	154	2,113	1,293	820
Africa	24	1,448	1,072	376
ASEAN	211	2,748	2,035	713
<i>Total</i>	<i>4,214</i>	<i>104,720</i>	<i>71,011</i>	<i>33,709</i>

In Table 4 we provide a geographic coverage of our main sample by some countries/areas where we validate the existence of a high flows' concentration between capital-abundant countries, a crucial feature detected also in several works when looking at related-parties trade data. Parents are classified by their home country in the second column, while in the third column we report the total number of affiliates owned by those parents, either as a stock before 2004 or as investment operations undertaken after 2004, respectively reported in column 4 and in column 5 of Table 4.

or not in the period of our analyses (2004-2012). Preliminary analysis shows that a relative majority of them includes small finance and consulting firms, construction and transportation companies.

As expected, the majority of our internal supply chains originate in OECD countries (80%), whose affiliates represent around 85% (88,041 out of 104,720) of our final sample. Of these, 27,174 are new acquisitions. In addition, EU countries report the largest number of parents (34%) and affiliates (55%), with two thirds of them already integrated in the past, while in US we find about 1,400 parents (32%) which control 18,005 affiliates⁷.

From our sample, a representative parent controls on average 24.87 affiliates, it is active in 6 countries and started its activity on average in 1983. However, distributions are rather skewed since the median parent controls only 5 affiliates present in 2 countries. About 63% of our sample is made of multinational groups, i.e. networks coordinated by a parent that report at least one foreign affiliate. Yet, we keep also entirely domestic groups, for which all affiliates are located in the parents' country of origin, as a further control group, assuming that parents can choose to develop supply chains entirely at home.

For the purpose of our analysis we link firms' activity information with industry-level metrics of demand elasticity and downstreamness. The first we source from Broda and Weinstein (2006), assuming that for sufficiently high (low) values of this average demand elasticity the corresponding elasticity of substitution across inputs is low (high)⁸.

As for downstreamness metrics, we source directly from Antràs and Chor (2013)⁹, according to which the relative location of an industry in production processes is measured as the distance from final consumers, thus giving an orientation to technological processes over different stages of production, eventually leading to production of final goods. Hence, downstreamness metrics are normalized on a range in $[0, 1]$, where 1 represents full proximity to final demand and 0 is the beginning of a production line.

In absence of original information on actual shipments of intermediate inputs, these metrics turn to 2002 US Input-Output tables produced by US Census Bureau in order to obtain average measures of the relative position of each industry in the production processes¹⁰. We exploit both alternative metrics of downstreamness: the first is built as the ratio of the aggregate direct use of an input to the aggregate total use of that industry (*DuseTuse*), whereas the second weighs for the average position of that industry in the supply chain at which an industrial output is used (*DownMeasure*).

After merging with our firm-level sample, in Table 5 we report the ten highest and lowest values of *DuseTuse* and *DownMeasure* across the 473 manufacturing industries we observe at the 6-digit of the NAICS rev.2007. The industries featuring the lowest downstreamness values tend to be in raw materials processing (aluminium, petrochemical, or copper), whereas industries with highest values are footwear and automobile manufacturing. In Table 5, the two alternative measures share five out of ten bottom industries and six out of top ten industries in our firm-level sample.

Thus, for each parent and affiliate along the control chain we have industry affiliations at the 6-digit NAICS rev.2007 classification, including both primary and secondary activities. After matching downstreamness industry metrics with the 6-digit NAICS rev.2007 firms' activities, we average over parents' and affiliates' primary activities to obtain their positioning along the supply chain. Indeed, given the conglomerate nature of some groups of firms included in our sample, we may want exploit information on the full set of activities performed by a company, which can be active in more than one business or can be multiproduct in nature. Nevertheless, 90% of parent firms and 70% of affiliates in our sample present a single primary activity. In Table 6 and in Figure 1 and 2 we report

⁷To validate our dataset we can mainly rely on Altomonte and Rungi (2013), from which we borrow the data methodology applied to the same data sources. There, the authors exploit a full dataset of all 270,474 active headquarters controlling more than 1,500,000 (domestic and foreign) affiliates worldwide in 2010. They check for value-added generation, then they validate the full sample after matching against the corresponding figures provided by (UNCTAD, 2011). Correlations by country for parents and affiliates are respectively .94 and .93.

⁸As in the industry-level investigation by Antràs and Chor (2013), we assume that any existing cross-sectoral variation in input substitutability is largely uncorrelated with the elasticity of demand faced by the parent.

⁹These metrics for distance from the final consumer, based on input-output tables, were first proposed by Antràs et al. (2012), and inspired by the first effort made by Fan and Lang (2000), where however no technological sequence was derived for the relative position of industries over the supply chain.

¹⁰The use of the US Input-Output table for the whole set of countries is justified by the assumption of a correlation in inputs patterns across countries, assuming a common technology frontier, a Leontief production function and cross-country factor price equalization (Acemoglu et al., 2009).

Table 5: Highest and lowest values of downstreamness metrics from sample

NAICS code	Industry label	DuseTuse	NAICS code	Industry label	Down
<u>Lowest 10 values</u>			<u>Lowest 10 values</u>		
331314	Secondary smelting and alloying of aluminum	.0000	325110	Petrochemical manufacturing	.2150
325110	Petrochemical manufacturing	.0599	331411	Primary smelting and refining of copper	.2296
331411	Primary smelting and refining of copper	.0741	331314	Secondary smelting and alloying of aluminum	.2461
336112	Light truck and utility vehicle manufacturing	.0814	325191	Gum and wood chemical manufacturing	.2595
325211	Electric housewares and household fan manufacturing	.1205	325192	Cyclic crude and intermediate manufacturing	.2595
325910	Printing ink manufacturing	.1325	325193	Ethyl alcohol manufacturing	.2595
311119	Other animal food manufacturing	.1385	325199	All other basic organic chemical manufacturing	.2595
333220	Plastics and rubber industry machinery manufacturing	.1420	331312	Primary aluminum production	.2622
331311	Alumina refining	.1447	331311	Alumina refining	.2622
331312	Primary aluminium production	.1447	325311	Nitrogenous fertilizer manufacturing	.2658
<u>Highest 10 values</u>			<u>Highest 10 values</u>		
337122	Nonupholstered wood household furniture	.9922	336213	Motor home manufacturing	.9879
339116	Dental laboratories	.9942	316211	Rubber and plastics footwear manufacturing	.9927
332992	Small arms ammunition manufacturing	.9955	316212	House slipper manufacturing	.9927
332993	Ammunition (except small arms) manufacturing	.9955	316213	Men's footwear (except athletic) manufacturing	.9927
316211	Rubber and plastics footwear manufacturing	.9967	316214	Women's footwear (except athletic) manufacturing	.9927
316212	House slipper manufacturing	.9967	316219	Other footwear manufacturing	.9927
316213	Men's footwear (except athletic) manufacturing	.9967	337121	Upholstered household furniture manufacturing	.9928
316214	Women's footwear (except athletic) manufacturing	.9967	337122	Nonupholstered wood household furniture manufacturing	.9948
316219	Other footwear manufacturing	.9967	336112	Light truck and utility vehicle manufacturing	.9995
336111	Automobile manufacturing	.9997	336111	Automobile manufacturing	.9997

descriptive statistics and visualizations of downstreamness distributions from our firm-level sample. The correction made originally from input-output tables on the *DownMeasure*, for the length of the supply chain, is particularly evident when comparing distributions of either affiliates or parents with the values of *DuseTuse* metrics. The first has a different support, at least from our firm-level sample, as it starts at a higher value of .21. Also, both parents and affiliates show a thicker left tail in the *DownMeasure* distribution. As we will see in the following empirical analyses, this difference will not reflect a substantial change in results for relative positioning of a parent with respect to its integrated affiliates.

In fact, we can already notice an important descriptive finding. In contradiction with the original assumption by Antràs and Chor (2013) model, in general a company does not perform as a final-good producer. As we detect in Table 6, an average parent has a downstreamness around .6. That is, a parent's initial and absolute positioning on a supply chain can be high and far from the final consumer before starting any sourcing decisions. Indeed, if we take as reference the unweighted downstreamness (*DuseTuse*), it does range from a minimum of .07 to a maximum of .99 (from .23 to .99 for the alternative *Down* metrics). Thus, we can argue that what really matters is the relative (not absolute) positioning of the parent with respect to integrated stages of production, once taking as exogenous the orientation of technological stages towards demand. On the other hand, parents tend to own affiliates that can operate both in industries that are downstream or upstream with respect to the parent firm's industry. Among others, this result is consistent with what Ramondo et al. (2015) find, for the huge span of activities possibly integrated by multinational corporations. In any case, a generalization of the Antràs and Chor (2013) setup seems to be valid, since on average, the sample of affiliates show in Table 6 a slightly smaller downstreamness than sample parents in all moments of sample distributions. In the following analyses we will see how our empirical strategies do confirm these preliminary weak findings, paving the way for a systematic analysis in terms of relative distances between parents and

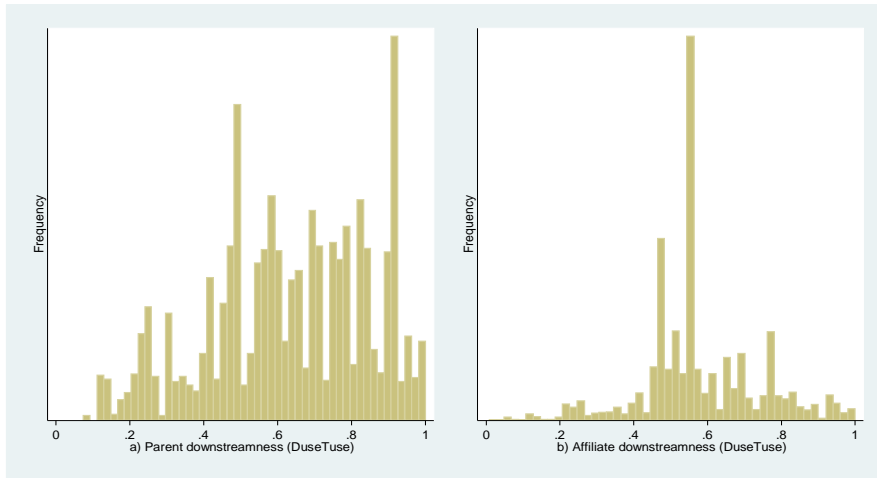
affiliates downstreamness and demand elasticity.

In Table 6 we further include descriptives of the demand elasticity after merging with firm-level information, borrowing from Broda and Weinstein (2006)¹¹. Here we elaborate on demand elasticity to reproduce the notions of sequential substitutes and sequential complements already proposed by Antràs and Chor (2013) in an industry-level analysis. Accordingly, we split the sample in two subsets by the median elasticity value (i.e. 5.97) for all 437 manufacturing sectors present in our data. Hence, we classify the above-median industries as complements and the below-median as substitutes cases.

Table 6: Downstreamness and elasticity across firms

		Obs.	Mean	Median	St. dev.	Min	Max
Parents	<i>DuseTuse</i>	4,201	.63	.65	.22	.07	.99
	<i>Down</i>	4,201	.57	.56	.21	.23	.99
	<i>Rho</i>	4,201	8.85	5.97	9.98	1.30	108.50
Affiliates	<i>DuseTuse</i>	90,928	.58	.54	.16	.01	1
	<i>Down</i>	90,928	.52	.53	.18	.21	.99

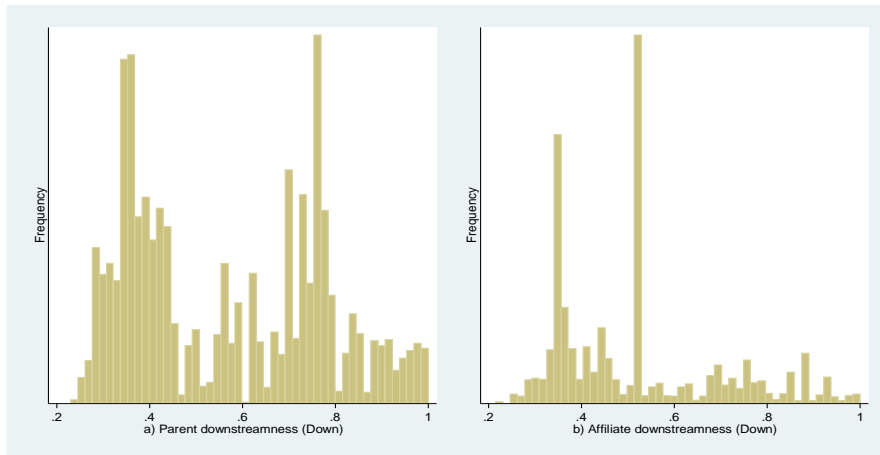
Figure 1: Parent and affiliate downstreamness from sample (DuseTuse metrics)



Finally, in Table 7 we report summary statistics for a set of measures that in literature have been identified as systematic determinants of the propensity to transact within firm boundaries. First, we add parent level variables: labor productivity allows controlling for the regularity according to which the most productive firms have the highest probability to invest abroad, as suggested by Antràs and Helpman (2004) and Helpman et al. (2004), and so tend to integrate a larger interval of production stages. Capital intensity of the parent is usually positively associated with intra-firm trade (Antràs, 2003). Firm size, age and number of already established affiliates allow to control whether the largest parents and the oldest parents are more prone to face the sunk costs of a vertical integration (Blomström and Lipsey, 1991), also in the case they hold already a multinational status (Greenaway and Kneller, 2007). We further include the degree of contractibility for affiliates' activities (i.e. the sellers'

¹¹The use of the US import demand elasticities for the whole set of countries is justified by the assumption, made also in other works, of broadly similar consumers' preferences across countries. Further, we apply the same matching procedure to firms' activities as for the downstreamness metrics.

Figure 2: Parent and affiliate downstreamness from sample (Down metrics)



industries), as based on the underlying share of products from an industry that are transacted on organized exchanges or are reference-priced according to Rauch (1999) classification, hence potentially calling for an easier deal to be reached when performed within a firm’s boundary (Nunn and Trefler, 2008). Two other country level variables sourced from the World Bank are instead associated to affiliates’ location: Rule of Law reflects perceptions on the extent to which agents have confidence in and abide by the rules of the economy, and Entry Cost represents cost to start a business, as a percentage of income per capita. In Data Appendix we provide further details for sources and usage of these variables.

Table 7: Firm controls

	Obs.	Mean	St. dev.	Min	Max
<u>Parents</u>					
(Log of) labor productivity	2,438	4.32	1.10	-3.82	11.72
(Log of) capital intensity	2,623	4.90	1.30	-3.18	13.39
(Log of) size	3,140	6.80	2.35	0	12.81
Age	4,214	30	28	1	824
Number of affiliates	4,214	25	3	1	1071
<u>Affiliates</u>					
Contractibility	46,630	.03	.12	0	1
Rule of Law	86,349	.74	.16	.13	.89
Entry Cost	85,564	1.50	1.56	-6.91	7.63

4 Empirical strategies

4.1 Baseline

Our aim is to explain the position of an integrated affiliate along a value chain as a function of a parent’s downstreamness and its output demand elasticity. We start by testing the following baseline equation augmented with firm, industry and country level information:

$$X_{i(j)c} = \beta_0 + \beta_1 X_j + \beta_2 \rho_j^* + \beta_3 (\rho_j^* \cdot X_j) + \beta_4 Z_{i(j)} + \beta_5 W_c + \varepsilon_{i(j)c} \quad (1)$$

where $X_{i(j)c}$ is the i th affiliate downstreamness, alternatively measured as *DuseTuse* or *DownMeasure* as from Antràs and Chor (2013), integrated by the j th parent and operating in country c . Among independent variables, X_j stands for j th parent downstreamness and ρ_j^* is a latent variable for parent demand elasticity, that we interact $(\rho_j^* \cdot X_j)$ with the j th parent position along the supply chain. Briefly, with this baseline equation we start following the procedure set by Antràs and Chor (2013), when making use of the original demand elasticity sourced by Broda and Weinstein (2006), in order to split the sample into industries that can be considered sequential complements (with elasticity above the median, $\rho_j > \rho_{med}$) and industries that can be considered sequential substitutes (with demand elasticity below the median, $\rho_j < \rho_{med}$) over the value chain. As a result, our indicator ρ_j^* is a dummy taking the value 1 in the complements cases and 0 for substitutes. Accordingly, we expect $\beta_3 > 0$, as parents would show a greater propensity to integrate suppliers that enter further downstream when the final demand is sufficiently elastic. $Z_{i(j)}$ collects firm level control measures in logarithmic scale, namely parent labor productivity, capital intensity, size, total number of controlled affiliates, together with a dummy variable for multinational status and finally an affiliate’s degree of contractibility based on its industrial activity. In W_c we collect two host country c ’s control measures, namely Rule of Law and Entry Cost, sourced by the World Bank¹².

Nested results for the baseline specification are reported in Table 8, where we find a confirmation of the theoretical prediction we were looking for: the higher the demand elasticity faced by a parent the more downstream are its integrated suppliers when its relative positioning is taken into account. In all specifications we indeed show that the effect of parent downstreamness is positive and significant at the 1% level in case of sequential complements. The finding is robust to inclusion of both alternative metrics for downstreamness.

Further, we find that parents tend to control production stages proximate to their main activity. Each parent and its suppliers tend to be relatively close along a value chain, as the higher the parent downstreamness the higher also the affiliates’ downstreamness.

This novel result is confirmed also when we include firm-level controls and host country fixed effects. We argue that what we find here can be discussed also in connection with Atalay et al. (2014), who report that after the acquisition an affiliate and a parent are very similar along multiple dimensions. In the US case, they find that vertical ownership does not seem to increase *ex post* the level of actual shipments among productive establishments. Hence, the authors suggest that an omitted variable bias could be responsible for their counterintuitive finding, as it is not possible to control for unobserved exchanges of intangible inputs, which can instead be the rationale for new vertical acquisitions. We rather argue that the finding of Atalay et al. (2014) could suffer from another different omitted variable bias that the authors do not consider, namely the affiliate’s relative positioning along the chain that can be the same before and after the acquisition. Indeed, if the sequential setup of the Antràs and Chor (2013) model is true, and we believe in our previous finding, both an affiliate and its proximate parent would show similar levels of downstreamness, before and after the acquisition, as the overall propensity to exchange physical intermediate inputs would be already given by the *ex ante* peculiarity of the production stage they perform.

¹²For further details on data sources and elaborations on variables, see Data Appendix.

Table 8: Baseline estimations, all affiliates

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
affiliate downstreamness (<i>DuseTuse or Down</i>)	OLS	OLS and firm controls	OLS fe and firm controls	OLS	OLS and firm controls	OLS fe and firm controls
parent DuseTuse	.150*** (.005)	.224*** (.009)	.231*** (.008)			
parent Down				.212*** (.005)	.330*** (.010)	.327*** (.008)
complements	-.040*** (.004)	-.055*** (.007)	-.062*** (.006)	-.022*** (.003)	-.007 (.006)	-.023*** (.005)
complements*parent DuseTuse	.060*** (.005)	.102*** (.011)	.108*** (.009)			
complements*parent Down				.025*** (.006)	.016 (.012)	.043*** (.010)
(log of) labor productivity		.009*** (.002)	.007*** (.001)		.015*** (.002)	.013*** (.002)
(log of) capital intensity		-.011*** (.002)	-.008*** (.001)		-.021*** (.002)	-.020*** (.001)
(log of) size		.006*** (.001)	.006*** (.001)		.004*** (.001)	.002*** (.001)
(log of) established affiliates		-.007*** (.001)	-.006*** (.001)		-.002** (.001)	-.001 (.001)
(log of) age		-.006*** (.001)	-.004*** (.001)		-.007*** (.001)	-.008*** (.001)
multinational group		-.034*** (.006)	-.042*** (.005)		-.033*** (.006)	-.024*** (.005)
contractibility		-.191*** (.010)	-.205*** (.009)		-.120*** (.007)	-.123*** (.006)
rule law		-.012 (.008)			.006 (.008)	
entry cost		-.004*** (.001)			.001 (.001)	
Constant	.487*** (.003)	.518*** (.012)	.206*** (.013)	.400*** (.003)	.415*** (.012)	.345*** (.011)
Observations	90,279	30,342	39,811	90,279	30,342	39,811
Adjusted R-squared	.082	.158	.174	.073	.157	.174
Country fixed effects	NO	NO	YES	NO	NO	YES

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Interestingly, productivity and size are positively related to our dependent variable, which reveals that the more productive and larger parents pick more likely affiliates located next to the final consumer, while older parents seem more prone to integrate upstream stages of production. We will see however in Section 4.2 that the correlation with parent age is not resilient to controls for simultaneity bias, when mixing up together old established affiliates and newly integrated ones. Similarly, also the negative correlation between affiliate downstreamness and total number of integrated affiliates will vanishes when we will perform an endogenous treatment model.

Since sample construction, we kept on purpose both domestic and multinational groups of firms, as we rationally assume a company can also choose to develop a supply chain entirely in the country of origin. Hence, a dummy variable for the multinational status controls whether a different vertical strategy occurs when crossing national borders. Here we find that groups having at least one activity abroad are more likely to integrate upstream production stages. In addition, we check for different combinations of factors of production that eventually relates to value chain positioning, by adding a measure of headquarters' capital-intensity. Our central result does not change, but we find that a higher capital intensity is associated with stages of production farther up from final consumers.

Nonetheless, the company's choice of organizational mode can be affected by external institutional

Table 9: Variants on sample composition

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS with DuseTuse	OLS with Down	OLS with DuseTuse	OLS with Down	OLS with DuseTuse	OLS with Down	OLS with DuseTuse	OLS with Down
affiliate downstreamness <i>(DuseTuse or Down)</i>								
	excluding horizontal	excluding horizontal	manufacturing affiliates only	manufacturing affiliates only	upstream affiliates only	upstream affiliates only	new affiliates only	new affiliates only
parent DuseTuse	.063*** (.009)		.314*** (.011)		.502*** (.012)		.167*** (.018)	
parent Down		.203*** (.010)		.428*** (.011)		.387*** (.010)		.260*** (.018)
complements	-.080*** (.007)	-.019*** (.007)	-.050*** (.008)	-.011* (.007)	-.051*** (.010)	-.037*** (.006)	-.077*** (.013)	-.028** (.012)
complements*parent DuseTuse	.124*** (.011)		.090*** (.013)		.094*** (.014)		.140*** (.021)	
complements*parent Down		.013 (.012)		.026** (.013)		.043*** (.012)		.051** (.022)
(log of) labor productivity	.002 (.002)	.012*** (.002)	.010*** (.002)	.012*** (.002)	.003 (.002)	.005*** (.002)	.011*** (.002)	.017*** (.003)
(log of) capital intensity	-.009*** (.002)	-.023*** (.002)	-.009*** (.002)	-.016*** (.002)	-.013*** (.002)	-.012*** (.002)	-.012*** (.002)	-.020*** (.003)
(log of) size	.005*** (.001)	.004*** (.001)	.008*** (.001)	.002** (.001)	.002** (.001)	.001 (.001)	.007*** (.001)	.003** (.001)
(log of) established affiliates	-.005*** (.001)	-.002* (.001)	-.007*** (.001)	-.001 (.001)	-.007*** (.001)	.001 (.001)	-.012*** (.002)	-.005*** (.002)
(log of) age	.001 (.001)	-.004*** (.001)	-.008*** (.001)	-.008*** (.001)	-.007*** (.001)	-.001 (.001)	-.002 (.002)	-.005** (.002)
multinational group	-.024*** (.007)	-.027*** (.008)	-.038*** (.006)	-.020*** (.006)	.001 (.008)	-.006 (.007)	-.002 (.007)	-.005 (.007)
contractibility	-.252*** (.011)	-.169*** (.007)	-.117*** (.010)	-.093*** (.008)	-.095*** (.011)	-.030*** (.006)	-.203*** (.020)	-.132*** (.013)
rule law	-.011 (.009)	.019** (.009)	.001 (.009)	-.001 (.008)	.003 (.009)	.028*** (.008)	.009 (.013)	.014 (.014)
entry cost	-.005*** (.001)	.000 (.001)	-.005*** (.001)	-.002*** (.001)	-.005*** (.001)	.001 (.001)	-.004*** (.001)	-.001 (.002)
Constant	.605*** (.013)	.481*** (.014)	.448*** (.015)	.375*** (.014)	.241*** (.015)	.218*** (.012)	.508*** (.021)	.412*** (.021)
Observations	25,352	25,352	23,293	23,293	13,217	13,818	8,404	8,404
Adjusted R-squared	.080	.084	.199	.240	.372	.318	.157	.125
Country fixed effects	NO	NO	NO	NO	NO	NO	NO	NO

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

environments, especially when its boundaries extend across different countries and/or industries. To test if our results are robust over geographic and industrial extensions, we included both country proxies of contractual environment and implicit contractual frictions for affiliates' industries. Not surprisingly, we find that higher levels of inputs' contractibility are negatively correlated with affiliates' downstreamness. As from the original calculation of the contractibility proxy by Rauch (1999), upstream stages are more likely to be reference-priced and/or traded on an organized exchange, thus they are more easily contractible¹³. This is consistent with an increased propensity by the parent to integrate upstream stages if parties could specify better their respective rights and duties (Antràs and Helpman, 2008). Less intuitive seems the negative correlation between affiliates' downstreamness and host countries' entry cost.

Lastly, we find that the overall quality of institutions, here proxied by the country rule of law, is not significantly related to affiliates' positioning along the chain.

All our results are robust in significance and signs when we introduce the alternative metric for downstreamness (*Down*), which takes into account the 'length' of the chain after adopting a weighing

¹³We report results for inputs' contractibility using the conservative classification in Rauch (1999), where inputs are differentiated and/or reference-priced. The magnitude of the affiliates' contractibility coefficient is higher under the more liberal classification the author proposes. Results are available upon request.

system by production stages, in columns 4, 5 and 6 of Table 8.

Further, main results are also robust to sample composition effects. In Table 9 we test our full specification first after excluding purely horizontal strategies, for affiliates positioned at the same level of the parent (i.e. affiliates reporting the same level of downstreamness), then restricting our sample to manufacturing inputs in order to avoid measurement problems in services industries. Additionally, in columns 5 and 6 of Table 9, we proceed testing if our results are also robust when we take into account only the affiliates that are located more upstream than the parent, whereas in the last two columns we exclusively control for newly established affiliates. In all these cases we find confirmation of our main prediction, according to which parents tend to integrate more downstream affiliates once facing sequential complement case. Only in the case of excluding stages that are more downstream than the parent, we register shaky results for correlations with parent size and multinational status. The latter should not come as a surprise, as we are deliberately excluding from our sample an important part of the affiliates' actual distribution over the supply chain.

4.2 Conditional dependence of integration choices

We now turn to investigate the possible endogenous formation of value chains, when including in the same sample investment decisions undertaken over a long span of time.

We argue that this empirical problem can be worked out with a binary control function approach¹⁴, eventually exploring how new integration choices can be affected by parent features and the extent of already established supply chains. We believe that such an approach better catches the sequential and dynamic nature of vertical integration choices, whereas simple least squares estimates can suffer from a too strong assumption of conditional independence among them.

For this purpose, we are able to separate from our sample the investment choices made in the period 2004 to 2012, for which we have contemporary observations of parent companies' characteristics, from the affiliates already controlled by the same parent companies but before the additional investment decisions were undertaken.

We make use of a control treatment function at a first step represented by a probit equation as follows:

$$d_{i(j)} = \beta_0 + \beta_1 Z_j + \varepsilon_{i(j)} \quad (2)$$

where $d_{i(j)}$ is a binary variable equal to 1 if the i th affiliate is integrated after 2004 by the j th parent and 0 otherwise. Z_j represents parent-level control measures, namely number of already established affiliates, labor productivity, capital intensity, size and age, all in logarithmic scale.

Thereby, the outcome equation borrows from our baseline in eq. 1, as follows:

$$X_{i(j)c} = \beta_0 + \beta_1 X_j + \beta_2 \rho_j^* + \beta_3 (\rho_j^* \cdot X_j) + \beta_4 Z_{i(j)} + \beta_5 W_c + \beta_6 d_{i(j)} + \varepsilon_{i(j)c} \quad (3)$$

to which we add the same binary variable $d_{i(j)}$ as in 2, this time as a further regressor. Its coefficients will eventually return us the average treatment effect (ATE) on affiliate positioning over the supply chain. Our specification allows netting out from the outcome equation the simultaneous effect of endogenous parents' characteristics, as they can already be the result of past integration choices. Errors are clustered by parent companies and we perform maximum likelihood estimation. In Table 10 and 11 we report results respectively for treatment and outcome equations, for both alternative measures of downstreamness, with and without hosting country fixed effects.

¹⁴For a comparison among different possible treatment-outcome estimators, see Angrist (2001), Vella and Verbeek (1999) or Heckman et al. (2003). Although consensus is not reached on the most promising among competing approaches, authors agree on the similarity of estimates that are eventually obtained. Here we follow the methodology originally introduced by Heckman (1997).

Table 10: Conditional dependence, past and present choices of integration (to be continued)

	(1)	(2)	(3)	(4)
	ET	ET and country fe	ET	ET and country fe
<i>Treatment equation</i>				
Dependent variable: Newly integrated yes/no				
(log of) established affiliates	-.039** (.016)	-.050*** (.016)	-.034** (.015)	-.047*** (.015)
(log of) labor productivity	-.114*** (.030)	-.088*** (.026)	-.123*** (.029)	-.097*** (.025)
(log of) capital intensity	.027 (.030)	.039 (.025)	.041 (.029)	.051** (.024)
(log of) size	-.118*** (.013)	-.095*** (.013)	-.122*** (.013)	-.098*** (.013)
(log of) age	-.065*** (.018)	-.075*** (.016)	-.054*** (.015)	-.066*** (.015)
Constant	1.386*** (.144)	1.124*** (.126)	1.329*** (.135)	1.560*** (.021)

(to be continued)

Preliminarily, we observe that the choice of controlling for the conditional dependence among vertical integration choices is successful as the estimated correlation (ρ) in Table 11, between the errors of the treatment equation and the errors of the outcome equation is statistically significant and high in magnitude, relatively higher when we use the *DownMeasure* metrics. The Wald test rejects the null hypothesis of the independence of the outcome from the treatment equations.

Nonetheless, we can exploit some further information from Table 10, with respect to parent-level characteristics and their sequential investment decisions. We find that already bigger and more productive parents integrate less likely new stages of production at the margin. Similarly, headquarters that established already complex supply chains, after collecting a higher number of affiliates, show a lower propensity to integrate further stages of production, probably after discounting higher internal coordination costs than smaller supply chains developed by parents collecting less affiliates. This is an important result that was concealed in previous least squares estimates. It further shows that a role in the formation of integrated supply chains is also played by parent-level heterogeneity and supply chain complexity, which the original theoretical model by Antràs and Chor (2013) does not take into account. At the moment we are writing, a first interesting attempt is made in a preliminary version of a work by Alfaro et al. (2015), in order to introduce firm-level heterogeneity on the input side and relate it to the degree of contractibility over the supply chain. In any case, the main theoretical prediction of the framework set by Antràs and Chor (2013) is still confirmed, as in the outcome equation the interaction term between parent downstreamness and its demand elasticity is positive and statistically significant. The higher the demand elasticity faced by a parent, the more downstream are its integrated suppliers, when its relative positioning is taken into account. Only in the third specification, its coefficient and the coefficient on the complements' dummy variable are not significant, albeit they align to previous estimates when controlling for country fixed effects in the last column of Table 11.

Interestingly, even after controlling for simultaneity bias, we confirm that more productive and bigger parents are able to reach more downstream production stages, proximate to the final consumer.

Rather, as expected, some parent characteristics are now not significantly correlated with affiliate positioning on the chain. Especially in the case of parent age and network size measured by number of affiliates, an endogenous mechanism is particularly evident. An older group had time to become bigger, collecting more affiliates as a consequence of past choices of vertical integration. Eventually, in Table 11, we find that the newly integrated affiliates show on average a higher downstreamness (i.e. they are on average located more proximate to the final consumer), as the coefficient on the treatment dummy for new and old investment operations is positive and significant.

Table 11: Conditional dependence, past and present choices of integration (continues)

(continues)	(1)	(2)	(3)	(4)
	ET	ET and country fe	ET	ET and country fe
<i>Outcome equation</i>				
Dependent variable:				
Affiliate downstreamness (<i>DuseTuse</i> or <i>Down</i>)				
parent <i>DuseTuse</i>	.224*** (.036)	.231*** (.032)		
parent <i>Down</i>			.316*** (.026)	.316*** (.026)
complements	-.055** (.028)	-.062*** (.025)	-.010 (.017)	-.026*** (.005)
complements*parent <i>DuseTuse</i>	.101** (.044)	.108*** (.040)		
complements*parent <i>Down</i>			.018 (.035)	.047*** (.009)
(log of) labor productivity	.014*** (.005)	.011** (.041)	.026*** (.005)	.022*** (.004)
(log of) capital intensity	-.013** (.005)	-.010** (.005)	-.025*** (.005)	-.024*** (.004)
(log of) size	.011*** (.003)	.009*** (.003)	.013*** (.003)	.010*** (.003)
(log of) established affiliates	-.006 (.004)	-.004 (.003)	-.001 (.004)	.003 (.003)
(log of) age	-.003 (.004)	-.002 (.004)	-.002 (.004)	-.005 (.004)
multinational group	-.016 (.010)	-.020** (.009)	-.015 (.010)	-.004 (.010)
contractibility	-.192*** (.031)	-.204*** (.027)	-.118*** (.024)	-.120*** (.020)
rule law	-.014 (.013)		.004 (.012)	
entry cost	-.004*** (.001)		-.001 (.001)	
newly integrated (Yes/No)	.120*** (.021)	.103*** (.020)	.244*** (.016)	.242*** (.016)
Constant	.392*** (.052)	.348*** (.074)	.177*** (.035)	.181*** (.055)
rho	-.391***	-.331***	-.703***	-.698***
Observations	30,340	39,808	30,340	39,808
Log pseudolikelihood	-7,859.082	-11,340.352	-7,159.00	-11,275.876
Wald test for independent equations	35.30***	31.14***	211.33***	221.96***
Country fixed effects	NO	YES	NO	YES

Maximum likelihood estimates. Clustered standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.3 Exploring non-linearities along the supply chain.

In this section we want to explore possible non-linearities along the supply chain, after untying parent demand elasticity over affiliates' downstreamness and controlling for parents' positioning on the supply chain, in this way abandoning the elasticity cutoff that Antràs and Chor (2013) identify for identification of sequential inputs' complementarity and substitutability.

Therefore, adopting a quantile regression for the baseline eq. 1, we find that at the conditional median (see Table 12) the main theoretical prediction is confirmed, but only when we make use of one of the alternative downstreamness metrics that we have at disposal, *DuseTuse*, whereas the adoption of the *DownMeasure* leads to opposite results. The coefficients on parent demand elasticity and its interaction with parent downstreamness show a reversal of sign.

As we already noted in Figures 1 and 2, these two measures reported a different statistical support and skewness in both affiliates' and parents' distributions. The *DownMeasure* has a narrower standard deviation and is more skewed to the left. This difference in statistical distributions comes after the introduction of a weighing system, from input-output tables, in order to correct for the differences in

length of productive chains, as originally made by Antràs and Chor (2013).

Table 12: Results at the conditional median

Dependent variable:	(1)	(2)	(3)	(4)
affiliate downstreamness (<i>DuseTuse</i> or <i>Down</i>)	Median reg and firm controls	Median reg fe and firm controls	Median reg and firm controls	Median reg fe and firm controls
Parent <i>DuseTuse</i>	.326*** (.010)	.361*** (.008)		
Parent <i>Down</i>			.592*** (.013)	.615*** (.011)
Parent demand elasticity	-.002*** (.001)	-.002*** (.001)	.003*** (.001)	.002*** (.001)
Parent demand elasticity* <i>DuseTuse</i>	.003*** (.001)	.002*** (.001)		
Parent demand elasticity* <i>Down</i>			-.005*** (.001)	-.004*** (.001)
(log of) labor productivity	.010*** (.003)	.010*** (.002)	.010*** (.002)	.009*** (.002)
(log of) capital intensity	-.013*** (.002)	-.010*** (.002)	-.015*** (.003)	-.014*** (.002)
(log of) size	.006*** (.001)	.005*** (.001)	.001 (.001)	.001 (.001)
(log of) established affiliates	-.005*** (.002)	-.002* (.001)	-.001 (.002)	-.001 (.001)
(log of) age	-.005*** (.002)	-.003** (.001)	-.002 (.002)	-.003 (.002)
multinational group	-.046*** (.009)	-.056*** (.007)	-.034*** (.010)	-.026*** (.009)
contractibility	-.174*** (.014)	-.179*** (.011)	-.118*** (.015)	-.120*** (.014)
rule law	-.024** (.012)		-.003 (.013)	
entry cost	-.006*** (.001)		-.001 (.001)	
Constant	.481*** (.018)	.325** (.143)	.239*** (.019)	.245*** (.011)
Observations	30,342	39,811	30,342	39,811
Country fixed effects	NO	YES	NO	YES

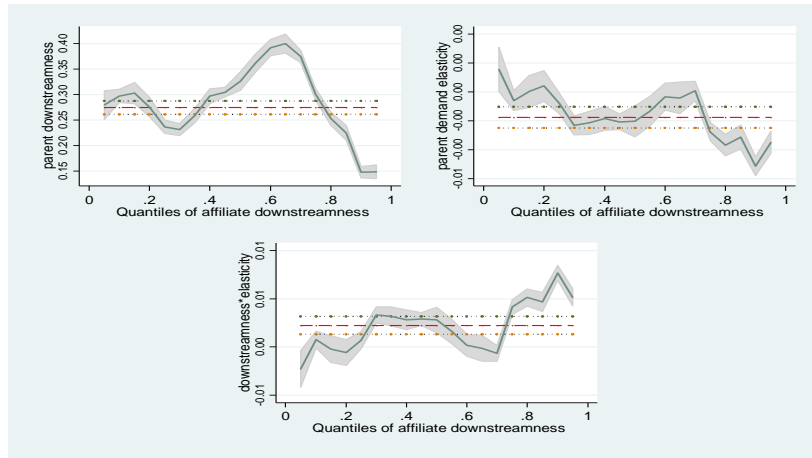
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Hence, to check that our results are not due to a different quality of one metrics over another, we go beyond conditional median regression and look at single quantiles. In this way, we are also able to identify when the sign reversal occurs. We report visual evidence of estimates for marginal effects in Figure 3 (Azevedo, 2011). Note that for parsimony we only draw visual estimates for coefficients on parent downstreamness, the elasticity of demand of its final output and their interaction term, assuming *ceteris paribus* the other control variables, over affiliates' downstreamness along the horizontal axis. A reference line is also reported for comparison for simple least squares estimates. Narrow bands of 95% pointwise confidence intervals are reported for both quantiles and least squares specifications.

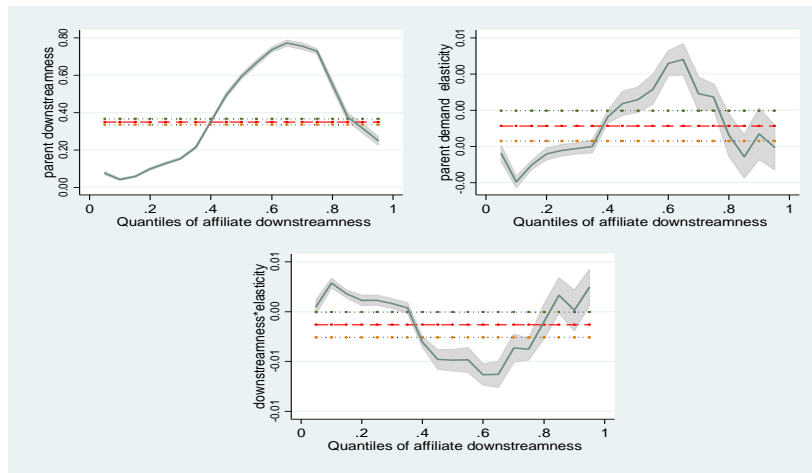
We find that all three terms of the regression equation can register non-linearities over affiliates' downstreamness, whatever metrics we adopt, in some cases even crossing the zeros. These non-linearities were concealed by any linear estimates we performed before.

We can now compute the threshold at which our correlations of interest reverse signs, considering the first derivative of elasticity of substitution on affiliate downstreamness.

Figure 3: Detecting non-linearities: quantile regressions on affiliates' downstreamness



a) affiliate and parent downstreamness as DuseTuse



b) affiliate and parent downstreamness as Down

Summing up the marginal effect of the elasticity of demand alone and its interaction term with downstreamness, while considering the level of parent downstreamness at each quantile, we estimate that a sign reversal occurs at values around .8, whether we adopt *DuseTuse* or *Down* metrics. In the case of *Down* metrics coefficients are also statistically non significant in that region of the distribution.

That is, we have a confirmation of the main theoretical prediction when affiliates have a value of downstreamness below .8, which is valid for about 70 percent of our affiliate-level sample. However, above this threshold, when affiliates are located at the bottom of the supply chain, the elasticity correlations can switch signs and/or lose statistical significance.

Certainly, we cannot exclude that a measurement problem can occur when exploiting imperfect proxies for both the technological ordering of the sequence, sourced from Antràs and Chor (2013), and the elasticity of substitutions of final outputs, sourced from Broda and Weinstein (2006). Still, we can also think of some reasonable economic explanations for the peculiarity of companies that are located at the very bottom of a supply chain. We believe that approaching final consumers it is more likely that we find horizontal integration strategies, mixed together with vertical integration

strategies, whose determinants the original theoretical model by Antràs and Chor (2013) is not able to catch. Alternatively, we argue, there could be cases for which affiliates at the bottom of the chain can be involved in completely different lines of business, reflecting the conglomerate nature of a business group to which they belong, or we can be in presence of multiproduct firms.

5 Conclusion

In this contribution we tested at the firm-level the optimal allocation of ownership rights along Global Value Chains. We find empirical evidence for the theoretical framework set by Antràs and Chor (2013), whose main prediction is that vertical integration choices over a sequence of production stages depend both on the relative position of each intermediate producer, and on the surplus that can be extracted from the sale of the final output, on which all the producers along the chain can rely. Indeed, we find that if the final demand elasticity is sufficiently high, parent companies prefer to engage in vertical integration of downstream production stages.

Further, we find that parent companies and affiliates tend to be located in proximity over a supply chain, whereas we detect some non-linearities, especially close to the final consumer. After the VIII decile of affiliates' downstreamness distribution results on demand elasticity are no more statistically significant. We presume that this is due to the presence of horizontal integration strategies, for which determinants of optimal allocation of property rights can follow different patterns.

More interesting, we are able to show that an important role is also played by firm-level heterogeneity, as we find that more productive and bigger headquarters prefer to pick downstream affiliates. This is an important empirical insight that could pave the way for further theoretical work, as it shows that there is a certain degree of asymmetry of different stages over the supply chain. As far as we know, a preliminary and promising theoretical effort is at the moment under way by Alfaro et al. (2015), extending the original model by Antràs and Chor (2013), to map differences in marginal input productivity across stages to differences in input contractibility.

Moreover, we obtain another interesting result regarding firm heterogeneity over supply chains, after controlling for conditional dependence of present choices of vertical integration on *ex ante* parent characteristics and supply chain size. We find that big and more productive parent companies are less likely to integrate at the margin.

On the other hand, already complex supply chains, with a large number of affiliates, show a lower propensity to integrate further production stages. We believe that this is due to the rise of internal coordination costs, when the extent of the production network becomes too difficult to manage.

A Data Appendix

Downstreamness: computed by Antràs and Chor (2013) from the 2002 U.S. I–O Tables, after using the detailed Supplementary Use Table after conversion tables issued by the BEA 2002, to obtain average measures of the relative position of each industry in U.S. production processes. Antràs and Chor (2013) propose two measures of downstreamness. The first measure is the ratio of the aggregate direct use to the aggregate total use ($DUseTUse$) of a particular industry i 's goods, where the direct use for a pair of industries is the value of goods from industry i directly used by firms in industry j to produce goods for final use, while the total use is the value of goods from industry i used either directly or indirectly in producing industry j 's output for final use. A high value of $DUseTUse$ thus suggests that most of the contribution of input i tends to occur at relatively downstream production stages close to final demand. The second measure of downstreamness ($DownMeasure$) is a weighted index of the average position in the value chain at which an industry's output is used, with weights given by the ratio of the use of that industry's output in that position relative to the total output of that industry. These measures have been finally averaged over the 6-digit NAICS rev.2007 parent and affiliate primary activities.

Demand Elasticity: computed by Antràs and Chor (2013) from U.S. HS10 products import demand elasticities (Broda and Weinstein, 2006). These are merged with a comprehensive list of HS10 codes from Pierce and Schott (2012). For each HS10 code missing an elasticity value, they assigned a value equal to the trade-weighted average elasticity of the available HS10 codes with which it shared the same first nine digits. This is done successively up to codes that shared the same first two digits, to fill in as many HS10 elasticities as possible. Using the IO-HS concordance provided by the BEA with the 2002 U.S. I-O Tables, they then take the trade-weighted average of the HS10 elasticities within each IO2002 category. At each stage, the weights used were the total value of U.S. imports by HS10 code from 1989 to 2006, calculated from Feenstra et al. (2002). There remain 13 IO2002 industries without elasticity values after the above procedure. For these, they assign a value equal to the weighted average elasticity of the IO2002 codes with which the industry shared the same first four digits, or (if the value was still missing) the same first three digits, using industry output values as weights. This yielded import elasticities for the industry that sells the input in question (ρ). These measures have been finally averaged over the 6-digit NAICS 2007 parent primary activities.

Complements: computed from the Demand Elasticity measure. We identify as Sequential Complements industries those with ρ above the cross-industry median (5.37 for the 437 manufacturing sectors present in our data), while Sequential Substitutes are the industries below the cross-industry median.

Capital intensity: computed at parent-level from Orbis database. It is the ratio between fixed assets over number of employees for parent firm j .

Age: computed at parent-level from Orbis database. It is the age of the parent at the year when choices of integration were undertaken.

Size: computed at parent-level from Orbis database. It is the number of consolidated employees by parent firm j .

Productivity: computed at parent-level from Orbis database. It is the value added over number of employees for each parent firm j .

Affiliates' number: computed at parent-level from the Orbis database. It is the number of already established affiliates as controlled by parent firm j before choices of integration occurred.

International Group: computed at parent-level from the Orbis database. It is a dummy variable equal to 1 if the parent firm j owns at least one foreign affiliate and 0 otherwise.

Investment: computed at parent-level from the Orbis database. It is a dummy variable equal to 1 if parent firm j has integrated an affiliate from 2004 to 2012 and 0 otherwise.

Contractibility: we source from Antràs and Chor (2013), who computed from the 2002 U.S. I–O Tables, following the methodology of Nunn (2007) For each IO2002 industry, they first calculate the fraction of HS10 constituent codes classified by Rauch (1999) as neither reference-priced nor traded on an organized exchange, under Rauch’s “liberal” classification. The original Rauch classification is at the level of SITC Rev. 2 products. These were associated with HS10 codes using a mapping derived from U.S. imports in Feenstra et al. (2002). The authors took one minus this value as a measure of the own contractibility of each IO2002 industry. This measure we average over the 6-digit NAICS 2007 affiliate i primary activities.

Entry costs: we source country entry costs from the World Bank - Doing Business. They built it using data on the number of procedures, number of days, and cost (as a percentage of income per capita) required to start a business. We average over 2003-2005 for each affiliate i and host country c .

Rule of law: from the Worldwide Governance Indicators (Kaufmann et al., 2011). The annual index is linearly rescaled from its original range $[-2.5, 2.5]$, in order to lie in a range $[0, 1]$, and averaged over the period 2004-2010 for each affiliate’s country of origin. This reflects perceptions on the extent to which agents have confidence in and abide by the rules of society, especially in the quality of contract enforcement, property rights, the role of police, and the courts, as well as the likelihood of crime and violence.

References

- Acemoglu, D., P. Antràs, and E. Helpman (2007). "Contracts and Technology Adoption". *American Economic Review* 97(3), 916–943.
- Acemoglu, D., S. Johnson, and T. Mitton (2009). "Determinants of vertical integration: financial development and contracting costs". *The Journal of Finance* 64(3), 1251–1290.
- Alfaro, L., P. Antràs, D. Chor, and P. Conconi (2015). "Internalizing Global Value Chains: A Firm-Level Analysis". *mimeo*.
- Alfaro, L. and A. Charlton (2009). "Intra-industry Foreign Direct Investment". *American Economic Review* 99(5), 2096–2119.
- Alfaro, L., P. Conconi, H. Fadinger, and A. Newman (2013). "Do Price Determine Vertical Integration? Evidence from trade policy". *NBER Working Paper 16118*.
- Altomonte, C., I. Colantone, A. Rungi, and T. Sonno (2015). "Global value networks". In J. Amador and F. di Mauro (Eds.), *The Age of Global Value Chains: Maps and Policy Issues*, VoxEu, CEPR Policy Portal.
- Altomonte, C. and A. Rungi (2013). "Business Groups as Hierarchies of Firms: Determinants of Vertical Integration and Performance". *ECB Working Paper No. 1554*.
- Angrist, J. D. (2001). "Estimation of limited dependent variable models with dummy endogenous regressors". *Journal of business & economic statistics* 19(1).
- Antràs, P. (2003). "Firms, Contracts, And Trade Structure". *The Quarterly Journal of Economics* 118(4), 1375–1418.
- Antràs, P. and D. Chor (2013). "Organizing the Global Value Chain". *Econometrica* 81(6), 2127–2204.
- Antràs, P., D. Chor, T. Fally, and R. Hillberry (2012). "Measuring the Upstreamness of Production and Trade Flows". *The American Economic Review* 102(3), 412–416.
- Antràs, P. and E. Helpman (2004). "Global Sourcing". *Journal of Political Economy* 112(3), 552–580.
- Antràs, P. and E. Helpman (2008). "Contractual Frictions and Global Sourcing". *The Organization of Firms in a Global Economy*, 9–54.
- Antràs, P. and S. Yeaple (2014). "Multinational Firms and the Structure of International Trade". *Handbook of International Economics* 4(55).
- Atalay, E., A. Hortaçsu, and C. Syverson (2014). "Vertical integration and input flows". *The American Economic Review* 104(4), 1120–1148.
- Azevedo, J. (2011). "GRQREG: Stata module to graph the coefficients of a quantile regression.". *Statistical Software Components*.
- Baldwin, R. and J. Lopez-Gonzalez (2014). "Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses". *The World Economy*.
- Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2010). "Intra-firm trade and product contractibility". *Peterson Institute for International Economics Working Paper (10-3)*.
- Blomström, M. and R. E. Lipsey (1991). "Firm size and foreign operations of multinationals". *The Scandinavian Journal of Economics*, 101–107.
- Broda, C. and D. Weinstein (2006). "Globalization and the Gains From Variety". *Quarterly Journal of Economics* 121(2), 541–585.

- Corcos, G., D. M. Irac, G. Mion, and T. Verdier (2013). "The determinants of intrafirm trade: Evidence from French firms". *Review of Economics and Statistics* 95(3), 825–838.
- Costinot, A., J. Vogel, and S. Wang (2013). "An elementary theory of global supply chains". *The Review of Economic Studies* 80(1), 109–144.
- Defever, F. and F. Toubal (2013). "Productivity, relationship-specific inputs and the sourcing modes of multinationals". *Journal of Economic Behavior & Organization* 94, 345–357.
- EUROSTAT (2007). "Recommendations Manual on the Production of Foreign Affiliates Statistics (FATS)". *European Commission*.
- Fan, J. P. H. and L. H. P. Lang (2000). "The measurement of relatedness: An application to corporate diversification". *The Journal of Business* 73(4), 629–660.
- Feenstra, R., J. Romalis, and P. K. Schott (2002). "U.S. Imports, Exports and Tariff Data, 1989-2001". *Working Paper 9387, NBER*.
- Greenaway, D. and R. Kneller (2007). "Industry differences in the effect of export market entry: learning by exporting?". *Review of World Economics* 143(3), 416–432.
- Harms, P., O. Lorz, and D. Urban (2012). "Offshoring along the production chain". *Canadian Journal of Economics/Revue canadienne d'économie* 45(1), 93–106.
- Heckman, J. (1997). "Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations". *Journal of Human Resources*, 441–462.
- Heckman, J., J. L. Tobias, and E. Vytlačil (2003). "Simple estimators for treatment parameters in a latent-variable framework". *Review of Economics and Statistics* 85(3), 748–755.
- Helpman, E., M. Melitz, and S. Yeaple (2004). "Export versus FDI with heterogeneous firms". *American Economic Review* 94.
- Hummels, D., J. Ishii, and K.-m. Yi (2001). "The nature and growth of vertical specialization in world trade". *Journal of International Economics* 54, 75–96.
- Kaufmann, D., A. Kraay, and M. Mastruzzi (2011). "The worldwide governance indicators: methodology and analytical issues". *Hague Journal on the Rule of Law* 3(02), 220–246.
- Kohler, W. K. and M. Smolka (2012). "Global Sourcing: Evidence from Spanish Firm-Level Data". In *Quantitative Analysis of Newly Evolving Patterns of International Trade*, Volume Volume 18 of *World Scientific Studies in International Economics*, pp. 139–193. WORLD SCIENTIFIC.
- Nunn, N. (2007). "Relationship-specificity, incomplete contracts, and the pattern of trade". *Quarterly Journal of Economics* 122(2), 569–600.
- Nunn, N. and D. Trefler (2008). "The Boundaries of the Multinational Firm: An Empirical Analysis". In E. Helpman, D. Marin, and T. Verdier (Eds.), *The Organization of Firms in a Global Economy*. Cambridge University Press.
- Nunn, N. and D. Trefler (2013). "Incomplete contracts and the boundaries of the multinational firm". *Journal of Economic Behavior & Organization* 94, 330–344.
- OECD (2005). "Guidelines for Multinational Enterprises". *OECD, Paris*.
- Pierce, J. and P. Schott (2012). "A concordance between ten-digit US Harmonized System Codes and SIC/NAICS product classes and industries". *Journal of Economic and Social Measurement* 37(1), 61–96.
- Ramondo, N., V. Rappoport, and K. J. Ruhl (2015). "Intrafirm Trade and Vertical Fragmentation in US Multinational Corporations". *mimeo*.

- Rauch, J. E. (1999). "Networks versus markets in international trade". *Journal of International Economics* 48(1), 7–35.
- Tomiura, E. (2007). "Foreign outsourcing, exporting, and FDI: A productivity comparison at the firm level". *Journal of International Economics* 72(1), 113–127.
- UNCTAD (2009). "Training Manual on Statistics for FDI and the Operations of TNCs Vol. II.". *United Nations*.
- UNCTAD (2011). "World Investment Report 2011". *United Nations*.
- Vella, F. and M. Verbeek (1999). "Two-step estimation of panel data models with censored endogenous variables and selection bias". *Journal of Econometrics* 90(2), 239–263.