HARVARD | BUSINESS | SCHOOL

Come Together: Firm Boundaries and Delegation

Laura Alfaro
Nick Bloom
Paola Conconi
Harald Fadinger
Patrick Legros
Andrew F. Newman
Raffaella Sadun
John Van Reenen

Working Paper 18-051

Come Together: Firm Boundaries and Delegation

Laura Alfaro
Harvard Business School

Nick Bloom
Stanford University

Paola Conconi
Université Libre de Bruxelles

Harald Fadinger
University of Mannheim

Patrick Legros
Université Libre de Bruxelles

Andrew F. Newman
Boston University

Raffaella Sadun
Harvard Business School

John Van Reenen
Massachusetts Institute of Technology

Working Paper 18-051

Come Together: Firm Boundaries and Delegation*

Laura Alfaro, HBS and NBER
Nick Bloom, Stanford, CEP, NBER and CEPR
Paola Conconi, Université Libre de Bruxelles (ECARES), CEPR and CESifo
Harald Fadinger, Mannheim and CEPR
Patrick Legros Université Libre de Bruxelles (ECARES), Northeastern and CEPR
Andrew F. Newman, Boston University and CEPR
Raffaella Sadun, HBS, CEP, NBER and CEPR
John Van Reenen, MIT, CEP, NBER and CEPR

May 2019

Abstract

Little is known about the relationship between firm boundaries and the allocation of decision rights within firms. We develop a model in which final good producers choose which suppliers to integrate and whether to delegate decisions to integrated suppliers, when they are ex-ante uncertain about their ability. In this setting, integration has an option value: ownership rights give producers authority to delegate or centralize production decisions, depending on the realized ability of suppliers. To assess the evidence, we construct measures of vertical integration and delegation for thousands of firms in many countries and industries. Consistent with the model, we find that (i) integration and delegation co-vary positively; (ii) firms delegate more decisions to integrated suppliers of more valuable inputs; and suppliers are more likely to be integrated if (iii) they produce more valuable inputs and (iv) operate in industries with greater productivity dispersion.

JEL classifications: D2, L2.

Keywords: Vertical integration, decentralization, real options, supply assurance.

*An earlier version of this paper was circulated under the title "All Together Now: Integration, Delegation, and Management." We are extremely grateful for comments from Philippe Aghion, Ricardo Alonso, Jacques Crémer, Wouter Dessein, Luis Garicano, Bob Gibbons, Oliver Hart, Francine Lafontaine, Rocco Macchiavello, Marian Moszoro, Jawwad Noor, Andrea Prat, Mike Powell, Veronica Rappoport, Birger Wernerfelt, and participants in the FOM Conference, the Fifth CEPR IMO Workshop, the MIT Empirical Management Conference, ASSA, the Paris-Berkeley Organizational Economics Workshop, EEA-ESEM, and seminars at UCLA, NUS, Oxford, Collegio Carlo Alberto, MIT, Mannheim, Nova SBE, PUC-Rio, Unversitat Autònoma de Barcelona, LSE, King's College London, Southampton, and HSE Moscow. The authors gratefully acknowledge financial support from the FNRS (Conconi), the DFG - CRC TR224 (Fadinger), and the European Research Council under the European Union's Seventh Framework Programme - ERC Grant Agreement N. 339950 (Legros).

1 Introduction

Why do firms integrate suppliers? A key benefit, of course, is to have control over production decisions: in the presence of contracting frictions, ownership rights allow top management to improve productive efficiency by imposing costly investments on suppliers, such as adjusting their inputs, re-training, or even relocating their plants.[1]

Equally compelling, if less appreciated, integration confers greater control over the firm's internal organization. Among the residual decision rights bundled with owning an asset is the ability to re-assign its use or control to others. Top management can seamlessly re-allocate decision-making to or from the suppliers that the firm owns, particularly in response to new information. If a supplier turns out to be highly informed or capable, delegating production decisions to him is an option. If instead the supplier emerges as less competent, management can choose centralization, making those decisions itself. An independent supplier, by contrast, always maintains control as part of his ownership rights. Firm boundaries and the allocation of decision-making inside the firm are thus intrinsically linked.

The "control over control" that comes with ownership helps guarantee the firm a minimum quality and quantity of inputs, and thereby introduces a novel mechanism of "supply assurance" as a rationale for integration: the advantage of ownership is not so much that it can be used to force a supplier to provide an input that he might otherwise sell to someone else, but that it allows the firm to deploy control to the party best suited to using it. Outsourcing offers no such flexibility.

Failure to appreciate the interdependence between the interim allocation of decision rights and the extent of the firm boundaries can be disastrous, as Boeing's infamous 787 Dreamliner fiasco starkly illustrates. Boeing outsourced the design and manufacture of key components of the 787 (e.g., fuselage, wings, stabilizers) to independent suppliers. Unprecedented problems in design and compatibility ensued, often due to poor quality sub-components, which led to years of delivery delays and billions of dollars in cost overruns.[2]

Despite their evident connection, the interplay between vertical integration and delegation has scarcely been explored. This paper brings these twin organizational design decisions together, both theoretically and empirically. It first develops a model to jointly study vertical

[1] A well-known example is the acquisition of the Fisher Body by General Motors, which relocated the auto-body production facility close to the GM Buick factories in Flint, Michigan.

[2] In sharp contrast to Boeing's prior practice of providing all designs and performing intermediate as well as final assembly, the change of ownership structure left each major supplier with responsibility "for managing its own [sub-component] subcontractors," which "operated largely out of Boeing's view." Resulting delays ran to 40-months, and overruns to more than $10 billion. Part of Boeing's remedial reorganization for the Dreamliner was to acquire some of the major suppliers in order to have more direct control on the production of its inputs (see Tang and Zimmerman, 2009; Zhao and Xu, 2013; and McDonald and Kotha, 2015).

integration and delegation. It then assesses the evidence in light of the model, combining information on vertical integration and delegation for thousands of firms in multiple countries and industries.[3]

In the theoretical model, firm boundaries and the internal allocation of control are endogenous, the result of optimizing behavior by a headquarters (HQ) producing a final good. HQ has an exogenous "productivity," interpretable as a measure of entrepreneurial ability, or product demand. Production of the final good can use "generic" or "adapted" inputs. Inputs (e.g., the seats in an airplane, or a section of its fuselage) are more valuable if they are tailored to the final product (e.g., planes intended for sale to different carriers need different seats, which in turn vary by class of service; fuselage parts must be mutually adjusted with utmost precision in order to assemble a functioning aircraft). The nature and means of such investments are often difficult to specify contractually, because they are complicated to fully describe and often obscure until late in the course of production.

If the transaction is at arms length, HQ has neither contracts nor authority to see the investments through, so only the generic version of the input is feasible. By contrast, if the supplier is integrated, HQ can exercise authority to elicit adaptation investments from the supplier. Generic inputs rely only on the supplier's direction to produce, while adapted ones require coordinated follow-up investments by both HQ and the supplier. HQ and integrated suppliers have different views about how to carry out the adaptation process.[4] The supplier has low variable stakes in the enterprise profit, but bears private costs of investments. HQ is the main residual claimant on profit, but also bears some investment costs.

HQ first chooses which of its suppliers to integrate. She is ex-ante uncertain about the capability of suppliers to adapt inputs and only learns this after she has made the integration decision. At that point, she can decide whether to delegate decision making to integrated suppliers or centralize, retaining control over their production decisions. In the case of a non-integrated supplier, there is no option to delegate or centralize production decisions, but his capability still plays a role, determining the value of the delivered generic input.[5]

[3]Though the logic of the model also applies to lateral integration, involving goods sold in separate markets that are complementary either in production or consumption, data limitations make it difficult to construct firm-level measures of lateral integration: this would require information on firms' sales by product line for narrowly defined industries, which we do not observe in our data.

[4]Tensions about the way operations should be carried out within firms can arise because of differences in background, training, corporate culture, or managerial vision (e.g. Van den Steen, 2005; Hart and Holmström, 2010).

[5]This is a stark, and admittedly highly simplified, depiction: in reality, inputs are often "adapted" across firm boundaries, not just within them. The model could be modified to allow for some adaptation across firm boundaries, at the cost of some complexity. The key point is that integration offers more control than other ownership structures, both over certain key decisions (represented by the initial fixed investment in the model) and the allocation of control over other decisions (represented by the follow-up investments).

Since non-contractibility prevents HQ from internalizing the supplier's costs once the relationship begins, when she retains control, she always chooses the maximum possible adaptation investments for him, regardless of her own productivity or his capability. The cost is that she foregoes the supplier's expertise in case he happens to be highly capable. Delegation takes advantage of the supplier's capability, but there is a cost: an incentive problem generated by the imperfectly aligned interests of HQ and the supplier. Higher productivity attenuates this incentive problem, as the private costs of coordination weigh less heavily relative to the benefits in decision makers' calculations: delegation becomes less costly as productivity rises. Delegation will therefore increase with the productivity of the HQ.

A more productive HQ will also have stronger incentives to integrate suppliers. As HQ's productivity increases, integration becomes relatively more productive than non-integration (both because adapted inputs are more valuable than generic ones, and because of the incentive response of delegation) and the costs of integration decline (because centralization becomes less likely). For a more productive HQ, the efficiency gains of integration are thus more likely to offset the costs, in line with the "value theory" of integration.[6]

Since both the propensity to integrate suppliers within the firm boundaries and the propensity to delegate decisions to integrated suppliers increase in the productivity of the HQ, the first testable prediction is that *integration and delegation should co-vary positively*. This result underscores a fundamental conceptual distinction between delegation and non-integration. Delegation is a non-contractible act of relinquishing control that can in principle be revoked at will by managerial fiat. Non-integration, by contrast, is the result of a formal sale of assets (Baker, Gibbons, and Murphy, 1999). "One-dimensional" organizational models that focus on the allocation of control have a hard time distinguishing between complete non-integration and complete delegation: both would seem to put decisions as far removed from the "center" as possible. From the perspective of such models, it would seem that integration and delegation ought to covary negatively. Contrary to this presumption, the model predicts a positive correlation between them.

The value theory logic also predicts that both *integration and delegation should depend on the value contribution of the inputs*: suppliers that contribute more to enterprise value should be more likely to be integrated; among the integrated suppliers, more decisions should be delegated to those that provide more valuable inputs.

To assess the validity of the model's predictions, we construct firm-level measures of vertical integration, as well as proxies for the extent of delegation within firms. To measure vertical integration, we combine information from the WorldBase dataset on firms' produc-

[6] See Legros and Newman (2013, 2017) for the theory and Alfaro *et al.* (2016) for evidence.

tion activities with detailed input-output data to identify integrated and non-integrated inputs. To measure delegation, we use World Management Survey (WMS) data from interviews with plant managers in 20 countries on the degree of autonomy granted to them by central headquarters. In the empirical analysis, we use a matched Worldbase-WMS sample (with data on both decentralization and vertical integration) consisting of 3,444 plants, and a broader Worldbase sample (with information on vertical integration only) containing 67,111 plants.[7]

In line with the key prediction of the theoretical model, we find that plant-level delegation is robustly positively correlated with firm-level vertical integration. This result holds up in the baseline regressions and in a series of robustness checks (e.g., including different sets of fixed effects and controls, using different samples of firms).

The empirical results also confirm the role of the technological importance of the inputs for integration and delegation choices, in line with the second and third prediction of the theoretical model. We find that final good producers are indeed more likely to integrate suppliers of more valuable inputs. Among integrated suppliers, more autonomy is granted to those producing more valuable inputs. In terms of magnitude, our estimates indicate that increasing the input-output coefficient by one standard deviation increases delegation by around 0.05 standard deviations and increases the probability of integration by 0.6 percentage points (which corresponds to a 60 percent increase given the baseline probability of one percentage point).

In the model, integration creates an option value: if a supplier turns out to be of low capability, HQ is able to ensure at least a minimal level of input contribution/quality of input by directing the production process herself; by contrast, if the supplier turns out to be of high capability, HQ can get a higher quality input by delegating the production decision to him. Such an option is not available under non-integration wherein the producer is entirely reliant on the supplier's capabilities. Following the intuition from option theory, this observation suggests that characteristics of the distribution of supplier capability should influence integration decisions. In line with this intuition, we find that the probability that firms integrate a particular input within their boundaries increases with the dispersion in the capability of suppliers in that input industry. In terms of magnitude, a one-standard-deviation increase in our preferred measure of input risk increases the probability of integrating a supplier by around 0.34 percentage points (which corresponds to a 34% increase relative to the baseline integration probability).

We see the theoretical model as a useful benchmark for understanding how elements of organizational design that were previously considered separately may fit together in theory and practice. We believe that this model is a plausible interpretation of the patterns we ob-

[7] The integration and delegation measures are respectively from Alfaro *et al.* (2016) and Bloom, Sadun, Van Reenen (2012).

serve in the data. We discuss alternative theories that could account for some (but not all) of our empirical findings.

The structure of the paper is as follows. Section 2 briefly reviews the related literature. Section 3 presents the theoretical model. Section 4 describes the datasets and the key variables used in the empirical analysis. Section 5 presents the empirical results. Section 6 discusses alternative theoretical explanations for these results. Section 7 offers some concluding comments about the implications of our findings for the theory of the firm.

2 Related Literature

Organizational economists generally agree that the diverse elements of organizational design interact with each other and must work in concert for optimal performance (Milgrom and Roberts, 1990; Roberts, 2007). Yet, the economics of firm organization itself is starkly split into separate divisions (Gibbons and Roberts, 2013). There are theories of what determines the boundaries of the firm. Then there are theories of how a firm organizes itself internally, for example in the degree to which decisions are delegated from top- to mid-level managers.

Although some studies have emphasized the conceptual difference between integration and delegation (Baker, Gibbons, and Murphy, 1999; Hart and Holmström, 2010), there has been little theoretical work to operationalize these differences. And, to the best of our knowledge, there is no systematic empirical work along those lines. Understanding the functioning of complex organizations has become even more important in light of recent studies emphasizing how organization affects aggregate and firm-level performance (e.g., Hortaçsu and Syverson, 2007; Forbes and Lederman, 2010; Bloom, Sadun, and Van Reenen, 2016; Akcigit, Alp, and Peters, 2018).

Our work is mainly related to two streams of literature, which focus on each of the organizational choices we bring together in this paper. First, we build on the vast literature on firm boundaries. Theoretical studies have looked at the technological and contractual determinants of vertical integration (e.g., Coase, 1937; Grossman and Hart, 1986; Hart and Moore, 1990; Holmström and Milgrom, 1991; Hart and Holmström, 2010). The view of integration in our model is similar to that of Williamson (1975), and puts it in the "ex-post non-contractible" branch of incomplete-contracts economics (e.g., the 2002 version of Hart and Holmström, 2010; Aghion, Dewatripont, and Rey, 2002; Legros and Newman, 2008; Dessein, 2014). Another strand focuses on market determinants of vertical integration; beside the value theory papers already mentioned, this includes McLaren (2000), Grossman and Helpman (2002); and Conconi, Legros, and Newman (2012). Earlier theoretical approaches include the resource-based view of Wernerfelt (1984), the routines-based theory of Nelson and Winter (1982), and the knowledge-based explanation of Kogut and Zander (1992). Lafontaine and Slade (2007, 2013) provide an excellent overview of the empirical literature on firm boundaries. Some studies have tried to shed light on the determinants of vertical integration using firm-level data within specific industries (e.g., Joskow, 1987; Woodruff, 2002; Baker and Hubbard, 2003), countries (e.g., Acemoglu *et al.*, 2010), or across industries within countries (e.g., Acemoglu, Johnson, and Mitton, 2009), In addition to exploring the determinants of firm boundaries, the literature has examined the consequences of vertical integration. For example, Chipty (2001), Hortaçsu and Syverson (2007), and Forbes and Lederman (2010) study the impact on production efficiency and competition (respectively in the cable TV, ready-mix concrete, and airline industries).

Second, we build on the literature on delegation. Theoretical studies include Holmström (1984), Aghion and Tirole (1997), Dessein (2002), Hart and Moore (2005), Alonso, Dessein, and Matouschek (2008), Alonso and Matouschek (2008), Marin and Verdier (2008), Dessein, Garicano, and Gertner(2010). Much (but by no means all) of this literature views delegation as a means of selecting the best decision by assigning the decision right to (ex-ante) best informed parties; often this helps to incentivize the delegate to become better informed in the first place. In our simplified model of delegation, the assignment of control is a response to (symmetric) information: the (ex-post) most capable (or possibly least time constrained) party gets it. The two approaches are complementary – the adaptation decisions could involve the acquisition of further information – and our approach is mainly for tractability. On the empirical side, contributions include Acemoglu *et al.* (2007), Guadalupe and Wulf (2010), McElheran (2014), Caliendo, Monte, and Rossi-Hansberg (2015), Wu (2017), and Katayama, Meagher, and Wait (2018).

A number of papers have studied pairwise interactions of organizational design elements from the theoretical point of view. Examples include Holmström and Tirole (1991), Holmström and Milgrom (1991, 1994), Dessein, Garicano, and Gertner (2010), Rantakari (2013), Friebel and Raith (2010), Van den Steen (2010), Dessein (2014), and Powell (2015). As far as we are aware, only Baker Gibbons, and Murphy (1999) and Hart and Holmström (2010) consider delegation and firm boundaries together, and only from a theoretical perspective.

As already mentioned, the model builds on the "value theory" of integration. Consistent with the model's predictions, we find that suppliers of inputs that contribute more value to the production of a firm's output, as proxied by input-output coefficients, are more likely to be integrated. This finding is in line with the results of previous studies on firm boundaries (e.g., Acemoglu *et al.*, 2010; Alfaro *et al.*, 2019; Berlingieri, Pisch, and Steinwender, 2018). In this paper, we show theoretically and empirically that input value also affects delegation choices within the firm boundaries: top management delegates more decisions to suppliers of more valuable inputs.

Finally, our paper is related to the literature on supply assurance motives for integration (e.g., Carlton, 1979; Bolton and Whinston, 1993; Baker, Gibbons, and Murphy, 2002). Those papers tend to focus on demand uncertainty and/or the ability for non-integrated suppliers to sell their inputs to other buyers. The supply assurance in our model derives from uncertainty about the production process, or more precisely, the capabilities of the firm's members to solve attendant problems. In this sense it is related to work on the design of knowledge hierarchies (Garicano, 2000) and referrals (Garicano and Santos, 2004) though in those papers the allocation of control is decided contingently through contracts rather than managerial authority.

3 The Model

3.1 Production

Consider a production process in which a final good j is produced with n inputs indexed by i. An enterprise is composed of an HQ, who produces the final good, and n suppliers, S_i. HQ has "productivity" $A > 0$, an index of the profitability of her product appeal or her entrepreneurial ability. The expected value of the enterprise is

$$A \sum_{i=1}^{n} \pi_{ij} \, \mathbb{E}v_i, \tag{1}$$

where the contribution of supplier i depends on the technologically determined value-added share π_{ij} in producing good j, augmented by $\mathbb{E}v_i$, the expected amount that is generated by the supplier. As discussed below, this supplementary value will depend partly on whether the input is adapted to HQ's specific needs, as well as on investments and production decisions that are determined by the organizational environment. For now, consider the relationship between HQ and a typical supplier; the index notation is suppressed.

Inputs can either be *generic* or *adapted*. If the input is generic, the value enhancement generated by the supplier is $v = \lambda y$, where $\lambda < 1$ measures the potential gain from adapting the input to the HQ's output (lower λ reflects lower quality of the generic input), and $y \geq 0$ is the supplier's "capability," a random variable with distribution $F(y)$ supported on $[0, \overline{y}]$ ($\overline{y}$ is possibly infinite; in any case the expectation $\mathbb{E}y$ is finite). A generic input therefore contributes value $A\pi\lambda y$ to the enterprise.

To adapt an input, the supplier must first make a fixed investment at private cost ϕ. For example, he may go through lengthy meetings and plant visits to learn about specific features of the final good, take training courses that familiarize him with the final good producer's style, brand or reputation, or move close to the HQ's premises. For simplicity, assume this investment is irreversible: once the process has begun, for better or for worse, the input will be adapted, and reverting to its generic form is no longer possible (little would change apart from some notational complexity if reversal were permitted). After the investment, adaptation involves further actions, such as design and process modifications in response to problems, that are performed by the supplier ($s \in [0,1]$) and by HQ ($h \in [0,1]$). These need to be coordinated for adaptation to be successful. In particular, adaptation succeeds (yields a return) with probability $p(s,h) = 1-(s-h)^2$, and fails (yields zero) otherwise.[8] HQ and the supplier have opposing preferences about how to carry out adaptation and find it costly to accommodate the other's approach (this could be due to differences in background, technologies, or "vision," possibly arising from the fact that they are in different industries). Specifically, HQ has private cost $(1-h)^2$, while the supplier has private cost cs^2 ($c > 0$). Hence, HQ prefers the decision to be close to $h = 1$, while the supplier likes the adaptation decision to be close to $s = 0$. Typically, c would be small, as HQ's practices or brand identity would matter more than that of a small component of her product.

The value of the adapted input depends on who decides which actions to perform. HQ has a capability at producing the input that is normalized to 1. If HQ centralizes decisions, choosing the action s as well as h, the expected contribution from the adapted input is $A\pi p(s,h)$. However, if the supplier chooses s, the expected value is $A\pi y p(s,h)$, reflecting his capability y.[9]

Summarizing, let $\mathbb{D}$ (for delegation) be the indicator function taking value 1 if the supplier chooses s for the enterprise, and value 0 if HQ does; let $\mathbb{I}$ be the indicator of whether the initial fixed investment is made. Then the expected supplier contribution is

$$\mathbb{E}\pi v = \mathbb{I}\pi p(s,h)(\mathbb{D}y + 1 - \mathbb{D}) + (1-\mathbb{I})\pi\lambda y.$$

[8]This specification provides a simple way to capture coordination problems within firms and is common in the organizations literature (e.g., Dessein and Santos, 2006; Alonso, Dessein, and Matouschek, 2008; Legros and Newman, 2013).

[9]The case in which the supplier is permitted to choose h is ruled out; this could be supposed to be technologically infeasible, but it can also be shown that HQ would never choose to delegate the h decision to S under the model's payoff and contractibility assumptions.

3.2 Contracting and Timing

Contracting is limited to fixed monetary payments and transfers of ownership. In particular, payments contingent on adaptation decisions or outcomes are not possible (e.g., because they are not observable or, if they are, they are not verifiable by third parties). The identity of the decision maker (hence the delegation decision) is also not contractible. Moreover, only aggregate output, and not the (relatively small) contribution of individual suppliers, is contractible, so that profit shares would provide no meaningful incentives.[10]

Ownership rights are contractible. If the supplier sells his asset to HQ, she gains the right to impose the initial adaptation investment on the supplier. However, she also has the (non-contractible) right to choose the control structure: she can choose whether to centralize (choose s for the supplier) or delegate (let the supplier choose s).

It is assumed that all parties have payoffs denominated in monetary terms, and that all have sufficient liquidity on hand to effectuate any side payments that might be needed to satisfy the distributional requirements among them. Thus, the enterprise will be choosing the organizational structure that maximizes the ex-ante total surplus.

Contracting occurs between HQ and all of the suppliers S_i simultaneously. First, HQ chooses the firm boundaries, by deciding which suppliers to integrate (the transfer can be interpreted as the asset purchase price in this case). Crucially, when making the integration choice, she does not yet know the capabilities of suppliers in an input industry, only the distribution $F(y)$. This assumption captures the fact that at least some aspects of supplier capability (e.g., their ability to solve new problems or carry out particular tasks) are revealed during the production process.[11]

HQ can then invoke the authority garnered from ownership to force all integrated suppliers to make the initial adaptation investment, but she has no such authority over the non-integrated ones.

After learning the capability of the suppliers, HQ can decide to which of the integrated ones she delegates the adaptation process.[12]

The timing for a single HQ-supplier relationship is summarized as follows:

1. Contracting: integration decision and monetary transfer.

[10] "Budget breaking" arrangements are also ruled out; since HQ makes production as well as control allocation decisions, she cannot profitably serve in this capacity, and third-party budget breaking arrangements are well-known to be vulnerable to collusion or sabotage.

[11] This does not mean that HQ has no idea about supplier capability, only that there is some residual uncertainty, represented by the distribution $F(y)$

[12] Learning about the capability of non-integrated suppliers may be possible, but as will become clear, is also useless, given that HQ cannot force the initial adaptation investment on them.

2. Adaptation investment choice at supplier cost ϕ.
3. Supplier capability y observed by HQ and S.
4. Delegation decision if the supplier is integrated.
5. s and h are chosen at costs cs^2 and $(1-h)^2$.
6. Output is realized.

3.3 Ownership Structures

Non-Integration. The supplier has ownership of his asset and will never make the initial adaptation investment, given that he bears the cost ϕ (which is non-contractible), while his continuation value cannot depend on the success of adaptation (also non contractible).

Hence under non-integration there is no adaptation. It follows that the expected value $\mathbb{E}v$ to HQ of a non-integrated supplier (which is also equal to the total surplus, since the private costs are zero) is given by

$$V^N = A\pi\lambda\mathbb{E}y. \tag{2}$$

Integration. If HQ has ownership of the supplier's asset, she can impose the initial adaptation investment on him. Notice that she will always choose to do so, given that she does not bear the cost ϕ and the investment has positive expected value (else she would not have integrated in the first place).

Under integration, HQ can also decide whether to centralize the adaptation decisions (s, h) or delegate them to her supplier. If HQ centralizes decision making, she will choose $s = h = 1$: this will maximize the probability that adaptation succeeds, while minimizing her private costs. The interim value to HQ of an integrated supplier under centralization is

$$v^C(A, \pi) = A\pi. \tag{3}$$

By contrast, if HQ decides to delegate the direction of the adaptation process to the supplier, by letting him choose s, she anticipates that he will set $s = 0$ (since this minimizes his private costs and he has no financial stake in the outcome of the process). HQ will then choose h to maximize $A\pi y(1-h^2) - (1-h)^2$, which yields $h = \frac{1}{1+A\pi y}$. It follows that the interim value to HQ of an integrated supplier under delegation is

$$v^D(A, \pi, y) = \frac{(A\pi y)^2}{1 + A\pi y}. \tag{4}$$

Notice that the function $v^D(A, \pi, y)$ is increasing and strictly convex in A, π and y.

To decide whether to delegate, HQ compares $v^C(A, \pi)$ with $v^D(A, \pi, y)$. She will thus delegate whenever the realized capability of the supplier exceeds a cutoff value $y^*(A\pi)$ defined by

$$v^C(A, \pi) = v^D(A, \pi, y^*(A\pi)). \tag{5}$$

From (3) and (4), $y^*(A\pi)$ is the unique positive solution to $A\pi(y^2 - y) = 1$. It is (i) greater than 1 (HQ's capability) and (ii) decreasing in A and π. The reason for property (i) is that delegation suffers from an incentive distortion, since HQ and the supplier make their decisions independently, while centralization, at least from HQ's point of view, does not suffers from such distortion. In order to compensate for the incentive loss, it takes a supplier capability strictly higher than HQ's to convince her to delegate. Property (ii) is a result of the relative rigidity of centralization: centralized decisions are independent of A (or π); by contrast, delegated decisions improve with firm value because of the incentive response (in this case HQ's). Thus, the value of delegation is more elastic with respect to A than is the value of centralization, implying that an HQ with a higher A is more willing to delegate.

Two comments are in order. First, it has long been argued (Coase, 1937) that managerial authority under integration can lead to various kinds of rigidities, while integration's flexibility, particularly with respect to organizational design, is being emphasized here. In this model, *centralized* decisions are indeed rigid, i.e. they do not depend on the parameters A, π and, most notably, information on y. However, this is not the case for delegated decisions, which do vary with all of these parameters, as does the decision whether to delegate in the first place. In light of the model, it may thus be more accurate to say that it is centralization, rather than integration *per se* that is rigid.

Second, observe that the value to HQ of an adapted input is not always greater than that of a generic one ($\max\{v^C(A, \pi), v^D(A, \pi, y)\}$ may be less than λAy), but will always exceed it for Ay sufficiently large: a capably adapted input is more valuable than a generic one, but a sloppily adapted one need not be.

The probability of delegation conditional on integration is $1 - F(y^*(A\pi))$. In turn, the cutoff value $y^*(A\pi)$ is decreasing; thus:

Lemma 1. *The probability that HQ delegates a decision to an integrated supplier is increasing in A and π.*

In the benchmark model, delegation is a binary variable: the integrated supplier either makes a single production decision involving adaptation of the input, or does not. The model can be extended to allow for multiple tasks to be performed by each supplier; in this manner, delegation becomes a continuous variable, in line with the measure of delegation used in the empirical analysis. See the Theoretical Appendix for details.

Firm Boundary Choices

At the contracting stage, HQ determines whether to integrate each supplier S. The total surplus of an integrated relationship is

$$V^I \equiv \mathbb{E}\max[v^C(A,\pi), v^D(A,\pi,y)] - cF(y^*(A\pi)) - \phi. \tag{6}$$

The first term $\mathbb{E}\max[v^C(A,\pi), v^D(A,\pi,y)]$ is the expected value accruing to HQ under integration. The remaining terms are the (expected) costs of integration. Both are borne directly by the supplier and include the centralization cost c, which is incurred with probability $F(y^*(A))$, and the investment cost ϕ. In order for the supplier to agree to sell his asset, he must be compensated for these costs via a monetary transfer at the time of contracting. Thus, HQ will choose to integrate the supplier whenever $V^I \geq V^N$. Combining (2) with (6), the condition for integration can be written as

$$\mathbb{E}\max[v^C(A,\pi), v^D(A,\pi,y)] - A\pi\lambda\mathbb{E}y \geq cF(y^*(A\pi)) + \phi. \tag{7}$$

The left-hand side is the option value of integration. While the value of both ownership structures increase with $A\pi$, the term $R^I \equiv \mathbb{E}\max[v^C(A,\pi), v^D(A,\pi,y)]$ (eventually) increases faster than that of non-integration because of the incentive response under delegation.[13]

Thus once $R^I > V^N$, which is necessary for integration, $R^I - V^N$ is increasing in A. Meanwhile, since $y^*(A\pi)$ is decreasing in A, the integration cost on the right-hand side decreases in A. It follows that the propensity to integrate a supplier is increasing in HQ's productivity A.[14]

[13]More precisely, this term is strictly convex in A, while V^N is linear: $R^I_{AA} = \pi f(y^*)y^*_A - v^D_A(A,\pi,y^*)f(y^*)y^*_A + \int_y^* v^D_{AA}(A,\pi,y)f(y)dy$.
The integral is positive because $v^D(A,\pi,y)$ is strictly convex in A; the sum of the first two terms is positive because $y^*_A < 0$ and $v^D_A(A,\pi,y^*) > \pi$, which is verified using (4) and $\frac{A^2\pi^2y^{*2}}{1+A\pi y^*} = A\pi$, the definition of y^*.

[14]Observe that HQ will integrate any supplier with positive π if her A is large enough, for then $R^I_A > V^N_A$. This is true because as A grows without bound, $y^* \to 1$, while $v^D_A(A,\pi,y) = \frac{2A\pi y + A^2\pi^2y^2}{(1+A\pi y)^2}\pi y \to \pi y$ for y bounded away from 0. Thus, $R^I_A = \pi F(y^*) + \int_{y^*}^{\overline{y}} v^D_A(A,\pi,y)f(y)dy \to \pi F(1) + \int_1^{\overline{y}} \pi y f(y)dy > \pi F(1) - \int_0^1 \pi F(y)dy + \int_1^{\overline{y}} \pi y f(y)dy = \pi\mathbb{E}y > \pi\lambda\mathbb{E}y = V^N_A$. Then since the cost is bounded by $c+\phi$, V^I exceeds V^N for A sufficiently large.

Lemma 2. *If an HQ in industry j with productivity A integrates a supplier in industry i, then an HQ in industry j with productivity $A' > A$ will also integrate a supplier in industry i.*

A corollary of this result is that the set of integrated suppliers will increase (in the set inclusion order) as A increases. That is, if an HQ with productivity A integrates a set $I(A) \subset \{1, 2, \dots, n\}$, then, all else equal, an HQ with productivity $A' > A$ integrates a superset $I(A') \supseteq I(A)$.

Applying the same reasoning to the technologically determined value share leads to:

Lemma 3. *Holding F, c, and ϕ fixed across input industries, if an HQ in industry j integrates a supplier from industry i, she also integrates suppliers from industries k for whom $\pi_{kj} > \pi_{ij}$.*

In the empirical analysis, the degree of vertical integration for a firm present in industry j is the sum $VI(A) \equiv \sum_{i \in I(A)} \pi_{ij}$. Lemma 2 implies that the degree of vertical integration is an increasing function of A. Since by Lemma 1 the degree of delegation is also an increasing function of A, it follows that firms with a more productive HQ will have stronger incentives both to integrate suppliers and to delegate the adaptation process to them. This leads to the first main result:

Proposition 1. *The degree of delegation and the degree of vertical integration covary positively across firms.*

This result generalizes straightforwardly to the case of multiple tasks (see the Theoretical Appendix for details). It can also be shown to hold true more broadly, including settings that allow for richer financial contracting possibilities, renegotiation, and strategic interactions among suppliers (Legros and Newman, 2015).

It should be stressed that, in this model, delegation is a contingent element of organizational design, one that is a response to the arrival of information. The comparative static implications of uncertainty are considered in Section 5.3.

3.4 Testable Predictions

Summarizing, the theoretical model generates three key predictions that will be brought to the data. The first follows directly from Proposition 1:

P.1: More vertically integrated firms should have a higher degree of delegation.

Moreover, Lemmas 1 and 3 predict how delegation and integration decisions should also vary across suppliers, depending on the technological importance of their inputs:

P.2: Final good producers should delegate more tasks to suppliers of more important inputs.

P.3: Final good producers should be more likely to integrate suppliers of more important inputs.

As has been noted, uncertainty about capability of suppliers creates an option value that enhances the benefits of integration. Characteristics of the distribution of supplier capability should therefore influence integration decisions: the value of the option is higher when the distribution of supplier capability is more dispersed. However, the integration cost can increase or decrease with risk, so the overall predicted effect on integration choices is ambiguous.

Section 5 will provide an assessment of predictions P.1-P.3, along with an examination of how input risk affects integration choices.

4 Dataset and Variables

This section describes the datasets, samples, and key variables used in the empirical analysis.

4.1 Datasets

World Management Survey

The first of the dataset used in the empirical analysis is the World Management Survey (WMS). This is a large scale project aimed at collecting high quality data on organizational design of firms around the world and has been used in many studies (e.g., Bloom, Sadun, and Van Reenen, 2012).

The survey is conducted through phone interviews with plant managers. Several features of the survey design are meant to guarantee the quality of the data. First, the survey is "double blind", i.e. managers do not know they are being scored and interviewers do not know the plant's performance.[15] This enables scoring to be based on the interviewer's evaluation of the firm's actual organizational practices, rather than their aspirations, the manager's perceptions, or the interviewer's impressions. Second, each interviewer ran 85 interviews on average. This allows to include interviewer fixed effects, which help to address concerns over the reliability and consistency of the answers. Third, information on the interview process itself (duration,

[15]This was achieved by providing only firm names and contact details to the interviewers (but no financial details).

day-of-the-week), and on the manager (seniority, job tenure and location) was collected. These survey metrics are used as "noise controls" to help reduce measurement error.

The main wave of interviews was run in the summer of 2006, followed by smaller waves in 2009 and 2010.[16] The survey achieved a 45% response rate, which is very high for company surveys.[17]

Overall, the WMS contains around 11,691 plants in 20 countries. The sampling frame was drawn to be representative of medium sized manufacturing firms in each country: median plant employment is 150, mean plant employment is 277, with a standard deviation of 405.

WorldBase

The second dataset used in the empirical analysis is WorldBase by Dun & Bradstreet, which provides coverage of public and private firms in more than 200 countries and territories.[18]

The WorldBase dataset has been used extensively in the empirical literature on firm boundaries (e.g., Acemoglu, Johnson, and Mitton, 2009; Alfaro *et al.*, 2019). The unit of observation in the dataset is the establishment/plant, namely a single physical location where industrial operations or services are performed or business is conducted.

Each establishment in WorldBase is identified by a unique nine-digit sequence called Data Universal Numbering System (DUNS) number. For each establishment, WorldBase provides:

1. Industry information: the 4-digit SIC code of the primary industry in which each establishment operates, and the SIC codes of up to five secondary industries.
2. Ownership information: information about the firms' family members (number of family members, domestic parent and global parent).[19]
3. Location information: country of each plant.
4. Additional information: sales, employment, age.

[16] A minority of the plants (668) have been interviewed in more than one wave of the WMS. The sample excludes plants where the CEO and the plant manager were the same person (only 4.9% of the interviews).

[17] The high success rate is due to the fact that (i) the interview did not discuss firm's finances, (ii) there were written endorsement of many institutions like the Bundesbank, Banque de France, UK Treasury, and World Bank, and (iii) high quality MBA-type students were hired to run the surveys.

[18] WorldBase is the core database with which D&B populates its commercial data products that provide information about the "activities, decision makers, finances, operations and markets" of the clients' potential customers, competitors, and suppliers. The dataset is not publicly available but was released to us by Dun and Bradstreet. The sample was restricted to plants for which primary SIC code information and employment were available (due to cost considerations). For more information see: http://www.dnb.com/us/about/db_database/dnbinfoquality.html.

[19] D&B also provides information about the firm's status (joint-venture, corporation, partnership) and its position in the hierarchy (branch, division, headquarters).

WorldBase allows us to trace ownership linkages between establishments. In particular, we can use DUNS numbers to link all plants that have the same domestic of global parent. D&B defines a parent as a corporation that owns more than 50 percent of another corporation. To construct firm-level variables, we link all plants that have the same domestic ultimate owner, as in Alfaro *et al.* (2016).[20]

We use the 2005 WorldBase dataset. When focusing on the 20 countries that are also included in the WMS, WorldBase contains 17,371,146 plants (corresponding to 16,718,199 parent firms). Median plant employment is 2, the mean is 288, and the standard deviation is 5,428 (see Table A-1).

4.2 Samples

To study delegation and vertical integration choices, we will use two samples constructed from the WMS and WorldBase datasets described above.

Matched Sample

To assess the validity of predictions P.1 and P.2 concerning the interplay between delegation and vertical integration and how delegation choices depend on input value, we use a sample constructed by combining information from the WMS and the WorldBase dataset.

For the US and Canada we have linked plants interviewed in the WMS to plants in WorldBase using a common plant identifier (the DUNS number). For the remaining countries, there is no common plant identifier, so have used a string matching algorithm based on company names and location information to link plants in the WMS to plants in WorldBase. We then manually checked the results of the matching process. To construct firm-level variables, we have used ownership information from Worldbase to identify the parent of any matched plant.

As mentioned above, the WMS is focused on medium-sized plants, while Worldbase contains lots of very small plants. The matched sample includes 3,444 plant observations located in 20 countries, operating in 574 sectors, and corresponding to 2,883 firms. As shown in Table A-1, this is a representative sample from the WMS: median plant employment is 150, the mean is 254, and the standard deviation is 367. Table A-2 reports the number of observations (at the plant level) by country in the matched sample.[21]

[20] A "Domestic Ultimate" is a subsidiary within the global family tree which is the highest ranking member within a specific country and is identified by a "domuduns"' code. A "Global Ultimate" is the top most responsible entity within the global family tree and is identified by "gluduns" code. The two codes only differ in the case of multinationals firms.

[21] Notice that the number of observations is larger than the number of plants, since a few plants were interviewed in more than one wave of the WMS.

WorldBase Sample

To assess the validity of prediction P.3 concerning the impact of input value of integration choices, we will use a larger sample from the WorldBase dataset (we will use the matched sample as a robustness check).

We have limited this sample to firms in WorldBase that have at least 20 employees. This helps to correct for differences in the coverage of small firms across countries (see also Klapper, Laeven, and Rajan, 2006). To keep the analysis tractable, we restrict the attention to firms that have a primary SIC code in manufacturing (between SIC 2000 and 3999) that integrate at least one input different from their primary output j.

The Worldbase sample includes 67,111 plants, corresponding to 66,102 firms, operating in 459 sectors, located in 19 countries. Table A-3 reports the number of observations (at the firm-input level) by country.[22] This sample features more variation in plant size compared to the matched sample: median plant employment is 42, the mean is 147, and the standard deviation is 3,187 (see Table A-5 in the Empirical Appendix).

4.3 Main Variables

In what follows, we define the main variables used in the empirical analysis. Tables A-4 and A-5 in the Empirical Appendix present summary statistics for these variables.

Delegation

In the WMS, plant managers were asked four questions on delegation from the central headquarters to the local plant manager.[23] First, they were asked how much capital investment they could undertake without prior authorization from the corporate headquarters. This is a continuous variable enumerated in national currency that is converted into dollars using PPPs. Plant managers had then to state the degree of autonomy they had in three other dimensions: (a) the introduction of a new product, (b) sales and marketing decisions, and (c) hiring a new full-time permanent shop floor employee. These more qualitative variables were scaled from a score of 1 (defined as all decisions taken at the corporate headquarters), to a score of 5 (defined as complete autonomy granted to the plant manager).

Since the scaling may vary across questions, we have standardized the scores from the four autonomy questions to z-scores, by normalizing each question to mean zero and stan-

[22] Comparing Table A-3 with A-2, notice that one country (Greece) is missing in the WorldBase sample. This is because establishments in Greece only report their primary SIC codes. As a result, we cannot use within-firm variation to study how integration choices depend on input value, to verify prediction P.3.

[23] In Appendix Figure A-5, we detail the individual questions in the same order as they appear in the survey.

dard deviation one. The variable $\textit{Delegation}_{f,p}$ is the average across the four z-scores for plant p belonging to firm f. We use information on ownership linkages from the WorldBase dataset to link a plant to its parent firm.[24] Figure A-1 in the Empirical Appendix shows the distribution of $\textit{Delegation}_{f,p}$ in the matched sample.

Vertical Integration

To measure vertical integration, we combine information from WorldBase on firms' production activities with data from Input-Output tables.[25]

Given the difficulty of finding highly disaggregated input-output matrices for all the countries in our dataset, we use U.S. input-output tables to provide a standardized measure of input requirements for each output sector.[26] The data are from the Benchmark Input-Output Tables of the Bureau of Economic Analysis (BEA), which include the make table, use table, and direct and total requirements coefficients tables. We employ the Use of Commodities by Industries after Redefinitions 1992 (Producers' Prices) tables. The BEA uses six-digit industry codes, while the classification of production activities in WorldBase follows the SIC classification. We convert the input-output data at the 4-digit SIC level, using the concordance guide provided by the BEA.[27]

For every pair of industries, ij, the input-output accounts provide the dollar value of i required to produce a dollar's worth of j. By combining information from WorldBase on firms' activities with U.S. input-output data, we construct the input-output coefficients for each firm f with primary activity j, IO_{ij}^{f}. Here, $IO_{ij}^{f} \equiv IO_{ij} * I_{i}^{f}$, where IO_{ij} is the direct requirement coefficient for the sector pair ij (i.e., the dollar value of i used as an input in the production of one dollar of j) at the 4-digit SIC level and $I_{i}^{f} \in \{0,1\}$ is an indicator variable that equals one if and only if firm f owns plants that are active in sector i. A firm with primary activity j that reports i as a secondary activity is assumed to supply itself with all the i it needs to produce j.

[24] In the case of multi-plant firms, $\textit{Delegation}_{f,p}$ captures the degree of autonomy granted to the manager of plant p by the owner of the parent firm f. In the case of single-plant firms, it captures the degree of autonomy that the owner/CEO gives to the plant manager.

[25] The methodology is based on Fan and Lang (2000) and has been used in several empirical studies on firm boundaries (e.g. Acemoglu, Johnson, and Mitton, 2009; Alfaro *et al.*, 2016, 2019).

[26] As pointed out by Acemoglu, Johnson and Mitton (2009), assuming that the U.S. input-output structure carries over to other countries mitigates concerns about the endogeneity of technology.

[27] The concordance table is based on the SIC 1987 classification. For codes for which the match is not one-to-one, we have randomized between possible matches.

To verify the first prediction of the model, we construct a firm's integration index:

$$\text{Vertical Integration}_{f,j} = \sum_i IO^f_{ij}, \tag{8}$$

which is the sum of the IO coefficients for each input industry in which firm f is active. This index measures the fraction of inputs used in the production of a firm's final good that can be produced in house.[28] In the case of multi-plant firms, we link the activities of all plants that report to the same domestic ultimate and consider the main activity of the domestic ultimate as the primary sector.

As an illustration of the procedure used to construct the vertical integration index, consider the example of a Japanese shipbuilder that reports two secondary activities, Fabricated Metal Structures (SIC 3441) and Sheet Metal Work (SIC 3444). The IO_{ij} coefficients for these sectors are:

		Output (j)
		Ships
Input (i)	*Ships*	0.0012
	Fab. Metal	0.0281
	Sheet Metal	0.0001

The table is just the economy-wide IO table's output column for the firm's primary industry, Ship Building and Repairing (3731/61.0100), restricted to the input rows for the industries in which it owns a plant (or reports a secondary activity). The IO_{ij} coefficient for fabricated metal structures to ships is 0.0281, indicating that 2.8 cents worth of metal structures are required to produce a dollar's worth of ships. The firm is treated as self-sufficient in the listed inputs but not any others, so its vertical integration index *Vertical integration*$_f$ is the sum of these coefficients, 0.0294: about 2.9 cents worth of the inputs required to make a dollar of primary output can be produced within the firm.[29]

Table A-4 provides summary statistics for the variable *Vertical integration*$_f$. Notice that the mean vertical integration index is 0.1, so the average firm in our sample produces 10 cents of each dollar of output within its boundaries. Given a total intermediate share in manufacturing of around 0.5, this corresponds roughly to 20 percent of the value of all inputs.

[28] Alternatively, we could normalize input-output coefficients by the total sector-specific intermediate share to make them sum to unity for each sector. Since we include output-sector fixed effects in our empirical specifications, such sector-specific normalization is absorbed by the fixed effects.

[29] Many industries, including Ship Building and Repairing, have positive IO_{jj} coefficients: some "ships" are used to ferry parts around a shipyard or are actually crew boats that are carried on board large ships; machine tools are used to make other machine tools; etc. As a result, firms will be measured as at least somewhat vertically integrated. To control for this, in the empirical analysis, we will include output industry fixed effects.

Figure A-2 in the Empirical Appendix shows the distribution of *Vertical integration*$_f$ in the matched sample.

To assess the validity of the second prediction of the model, we construct the dummy variable *Integration*$_{f,j,i,c}$, which is equal to 1 if firm f (producing primary output j and with a domestic ultimate located in country c) integrates a supplier in input industry i within its boundaries. To keep the analysis tractable, we limit the sample to firms that integrate at least one input different from their primary output j, and to the top 100 inputs i used by j, as ranked by the IO coefficients. It should be stressed that the BEA input-output table that we are using is highly disaggregated (based on 935 4-digit SIC industries). As a result, even when focusing on the top 100 inputs, the average probability that a firm integrates any input is only around 1 percent (see Table A-5).

To examine how input value affects delegation and integration choices, we will use the variable IO_{ij}. This is the direct requirement coefficient for the sector pair ij, which captures the dollar value of input i used in the production of one dollar of j. Not surprisingly, given that the BEA input-output tables are highly disaggregated, the average IO_{ij} is only 4 cents in the matched samples, and 5 cents in the WorldBase sample (see Tables A-4 and A-5). Figure A-3 in the Empirical Appendix shows the distribution of IO_{ij} in the Worldbase sample.

Firm-level and Plant-level Controls

Using information from WorldBase, we construct auxiliary firm-level controls. These include the variables *Employment*$_f$, the total number of employees of the firm, and *Age*$_f$, the number of years since its establishment.

The auxiliary plant-level controls are drawn from the WMS dataset. They include the plant's employment as a fraction of the firm's employment (*Employment Share*$_p$), and the education of the workforce, defined as the percentage of a plant's employees who have a bachelor's degree or higher (*% Workforce with College Degree*$_p$). In some specifications, we also include the variable *Management*$_p$ to control for the quality of a plant's management practices.[30]

Riskiness of Input Industries

In the model, the capability of a supplier is unknown ex-ante, and the characteristics of suppliers' capability distributions will therefore influence firm boundary choices. To empirically

[30]The WMS contains information on 18 management practices, measured on a scale from 1 to 5. The variable *Management*$_p$ is the average of the 18 individual management dimensions for plant p, after each has been normalized to a z-score (with a mean of zero and a standard-deviation of one).

explore the role of uncertainty on integration decisions, we need first to identify the relevant input industries for each firm, and then to construct proxies for the dispersion in supplier capability in each input industry.

To identify the relevant input industries, we combine information from WorldBase on firms' production activities with Input-Output data from the BEA. For each firm f producing good j in country c, we focus on the top 100 inputs i as ranked by the IO coefficients IO_{ij}.

In the theoretical model, the contribution of a non-integrated depends linearly on his capability y. Under the assumption that the capabilities are i.i.d. among suppliers in the same industry, the empirical productivity distribution of non-integrated suppliers in industry i approximates the distribution $F_i(y)$. By contrast, an integrated supplier's contribution depends nonlinearly on y.[31] Hence the model suggests that the observed productivity distribution of integrated suppliers in industry i is not a good proxy for $F_i(y)$.[32]

To construct our main measure of input risk, we thus focus on non-integrated suppliers of input i and use information from WorldBase on their labor productivity. To minimize measurement error, we consider all plants that report SIC4 code i as their only production activity.[33] The distribution of labor productivity of input suppliers approximates a lognormal distribution. Following Levy (1973), controlling for the mean, we can then use the coefficient of variation of the distribution to proxy for the riskiness of the input industry in the Rothschild-Stiglitz sense. Our preferred measure of input risk is *CV Productivity*$_{i,c}$, the coefficient of variation of labor productivity of suppliers of input i in country c.[34] Figure A-4 in the Empirical Appendix shows the distribution of *CV Productivity*$_{i,c}$ in the Worldbase sample.

As an alternative way to proxy for input risk, we use data from Bloom *et al.* (2018) on stock market returns of US firms in 2005 (the same year as our WorldBase dataset). Stock market returns are approximately normally distributed. The variable *SD Stock Returns*$_i$ captures the cross-section dispersion in stock market returns across firms in SIC4 industry i. Unlike *CV Productivity*$_{i,c}$ this risk measure varies only at the sector level (given that we only have disaggregated stock-market data for firms in the United States) and is only available for some (manufacturing) industries.[35]

[31] According to the model, the capability of an integrated supplier only affects output under delegation, but not under centralization. Since HQ will centralize whenever the capability is low enough, the observed productivity of integrated suppliers is left-censored, which can cloud the relationship between various orderings (most saliently, Rothschild-Stiglitz riskiness) of the observed distributions and orderings of the underlying ones.

[32] In addition to the theoretical reasons for not including integrated suppliers, transfer pricing effects may also distort their measured labor productivity.

[33] The results continue to hold if we consider all plants that report SIC code i as one of their production activities, including those integrated in larger firms.

[34] In robustness checks, we focus on input industries with at least 50 independent suppliers (in industry i country c), or construct *CV Productivity*$_{i,c}$ after winsorizing labor productivity at the 5th and 95th percentile.

[35] We construct the variable *SD Stock Returns*$_i$ at the SIC4 level, though in some cases the underlying stock

5 Empirical Results

We now assess the validity of the model's predictions. Section 5.1, studies the relationship between delegation and integration, to verify prediction P.1. Section 5.2, reports the impact of input value shares on delegation and integration choices, to verify predictions P.2 and P.3.

The model suggests that the supply assurance motive for integration should be higher in riskier input industries, in which supplier capability is more dispersed. However, the cost of integration can increase or decrease with risk, so the overall effect is generally ambiguous. Section 5.3, empirically examines the impact of supplier uncertainty on integration choices, and discusses how the findings can be rationalized in light of the theoretical model.

5.1 Delegation and Integration

We first consider prediction P.1 concerning the relationship between delegation and integration. According to the model, firms with a more productive HQ will have stronger incentives both to integrate suppliers and to delegate the adaptation process to them. As a result, the two organizational variables should be positively correlated.

To assess the validity of the first prediction of the model, we estimate the following:

$$\text{Delegation}_{f,p,i,j,c} = \beta_1 \text{ Vertical Integration}_{f,j,c} + \beta_2 \mathbf{X}_p + \beta_3 \mathbf{X}_f + \delta_i + \delta_j + \delta_c + \epsilon_{f,p,i,j,c}. \quad (9)$$

The dependent variable is the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j, located in country c). The main control of interest is *Vertical Integration*$_{f,i,c}$, the vertical integration index of firm f. According to prediction P.1, the estimated coefficient β_1 should be positive and significant. $\mathbf{X}_p$ and $\mathbf{X}_f$ are vectors of plant- and of firm-level controls, while δ_i, δ_j and δ_c are input-sector, output-sector (at the 3-digit SIC level), and country fixed effects.[36] We include input-sector (output-sector) fixed effects to control for the average amount of delegation to a given input industry (by a given output industry).[37] We cluster standard errors at the firm level.

The results are reported in Table 1. In column 1 we regress *Delegation*$_{f,p,i,j,c}$ against the key control of interest, *Vertical Integration*$_{f,j,c}$, and input-industry fixed effects. In line with prediction P.1 of the model, the estimated coefficient of *Vertical Integration*$_f$ is positive and

market data is available at a more aggregate level (SIC2 or SIC3).

[36]Given that the data on delegation were collected in different waves of surveys and by different interviewers, we also include in these regressions survey noise controls and fixed effects for the year in which the firm was surveyed to reduce measurement error in the dependent variable.

[37]For the vast majority of firms, the delegation survey provides information on one plant only. Thus, unlike in the integration regressions below, we cannot include firm fixed effects when estimating (9).

significant (at the one-percent level).[38] This result continues to hold when we further include country fixed effects (column 2), output-industry fixed effects (column 3), and control for the size and age of the parent firm, as well as the plants' size and level of education of the workforce (column 4).[39]

Table 1
Delegation and Integration

	(1)	(2)	(3)	(4)
Vertical Integration$_f$	0.685***	0.794***	0.691***	0.577**
	(0.246)	(0.244)	(0.250)	(0.250)
log(Employment$_f$)				0.087***
				(0.022)
log(Age$_f$)				0.035*
				(0.021)
Share Employment$_p$				0.323***
				(0.073)
log(% Workforce with College Degree$_p$)				0.056***
				(0.016)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
R-squared	0.182	0.198	0.206	0.216
N. observations	3,444	3,444	3,444	3,444

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 3-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

In terms of magnitude, the point estimates reported in column 3 of Table 1 indicate that increasing *Vertical Integration*$_f$ by one standard deviation is associated with an increase in

[38]The coefficient of *Vertical Integration*$_f$ is also significant (at the five-percent level) in an even more parsimonious specification, in which we do not include any fixed effects.

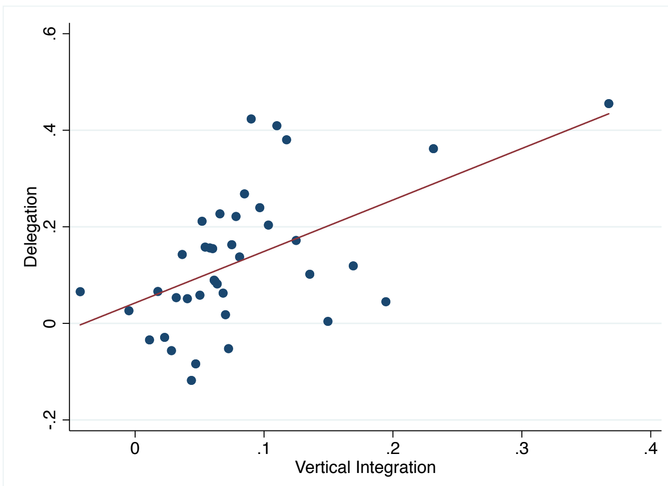
[39]The variables *% Workforce with College Degree*$_p$ and *Employment*$_p$ are missing for a few plants. To avoid dropping observations, in the specifications in which we include these variables, we replace missing values with -99 and use a dummy variable to control for these instances.

delegation of around 0.06 standard deviations.[40]

Concerning the auxiliary controls, the coefficients of *log(Employment)*$_f$ and *log(Age)*$_f$ are positive and significant, indicating that larger and older firms grant more authority to their plant managers. The coefficient of the variables *Share Employment*$_p$ and *log(% Workforce with College Degree)*$_f$ are also positive and significant, suggesting that firms delegate more to plants that are larger and have a more educated workforce.

The binned scatterplot in Figure 1 illustrates the key result of Table 1, based on the specification of column 1 of Table 1.[41]

Figure 1: Delegation and Vertical Integration



Notes: Binned scatterplot of the relationship between *Delegation*$_{f,p,i,j,c}$ and *Vertical Integration*$_f$ (based on 40 bins).

Recall that, according to the model, the reason why delegation and vertical integration should be correlated is that both should be increasing in A, which captures exogenous characteristics of the HQ that increase the profitability of the enterprise (e.g., product appeal, entrepreneurial ability of the CEO). The positive coefficient of *Vertical Integration*$_f$ should thus not be interpreted in a causal sense, i.e., more integration leading to more delegation. Rather, the model suggests that integration and delegation choices are endogenously correlated, because firms that have a more productive HQ (higher A) have stronger incentives to integrate suppliers and delegate production decisions to them. Looking at the results in column 4 of Table 1, notice that the coefficient of *Employment*$_f$ is positive and significant and

[40]The standard deviation of *Vertical Integration*$_f$ in the matched sample is 0.09 , so 0.691*0.09 = 0.062.

[41]Figure 1 is created by regressing *Delegation*$_{f,p,i,j,c}$ on *Vertical Integration*$_f$, input industry dummies, and noise controls. We then take the residuals of the delegation and integration variables and group them into 40 equal-sized bins, compute the mean of the variables within each bin, and create a scatterplot of these data points.

the coefficient of *Vertical Integration*$_f$ is significantly smaller than in column 3. To the extent that firm size is correlated with A, this is what one would expect based on the model.[42]

We have carried out a series of additional estimations to verify the robustness of the results of Table 1. The results of these estimations are reported in the Empirical Appendix.

First, we use more disaggregated industry fixed effects (defined at the SIC4 level instead of SIC3) to control for the primary activities of the plant and its parent firm (see Table A-6). The coefficients of the key variable of interest, *Vertical Integration*$_f$, remains positive and significant. The main drawback is that we lose some observations.

Second, we verify that the results of Table 1 are robust to restricting the analysis to the 10 largest countries in the sample, i.e. those with the highest number of firms (see Table A-7).

Third, we reproduce Table 1 after winsorizing *Vertical Integration*$_f$ at the 5th and 95th percentile. The results continue to hold (see Table A-8).

Finally, one may be concerned about measurement error in the vertical integration index. In an influential study, Atalay, Hortaçsu, and Syverson (2014) find little evidence of intra-firm shipments between related plants within the United States. This suggests that using Fan and Lang (2000)'s methodology to construct the variable *Vertical Integration*$_f$ may lead us to mis-classify some inputs as being integrated, when the firm is actually sourcing them from the market. Random measurement error in the vertical integration index should work against us, by attenuating the coefficient β_1, making it harder to find support for prediction P.1. Nevertheless, we have verified that the positive relationship between delegation and vertical integration holds even when we restrict the analysis to single-plant firms.[43] For these firms, measurement error in the vertical integration index should be less of a concern, since it is unlikely that a parent would not use the inputs produced in its own establishment. Delegation and vertical integration remain positively correlated (see Table A-9).

5.2 Delegation, Integration, and Input Value Shares

In this section, we assess the validity of Predictions P.2 and P.3 of the model concerning the impact of input value shares on delegation and integration choices.

[42] According to the model, if we could actually measure the exogenous productivity of the HQ and include it in regression (9), the coefficient of *Vertical Integration*$_f$ should become insignificant. The model also suggests that measures of firm performance (e.g., productivity, stock market returns) are endogenous to organizational choices, reflecting not only the exogenous productivity of the HQ (A), but also the capability of suppliers (y) and HQ's choice to delegate or centralize productive decisions. If we control for firm labor productivity in the specification in column 4 of Table 1, the coefficient of *Vertical Integration*$_f$ is unaffected, while the coefficient of the productivity measure is positive but not significant.

[43] In these regressions, we can only include one set of industry fixed effects (given that the primary SIC code of the parent firm coincides with the primary SIC code of the plant) and one employment variable (given that the number of employees of the plant and the firm are the same).

To verify how the technological importance of an input affects delegation choices, we estimate the following regression:

$$\text{Delegation}_{f,p,i,j,c} = \beta_1 \, \text{IO}_{ij} + \beta_2 \mathbf{X}_p + \beta_3 \mathbf{X}_f + \delta_i + \delta_j + \delta_c + \epsilon_{f,p,i,j,c}. \tag{10}$$

Table 2

Delegation and Input Value Shares

	(1)	(2)	(3)	(4)
IO_{ij}	0.986***	0.937**	0.980**	0.941**
	(0.379)	(0.380)	(0.452)	(0.441)
Vertical Integration$_f$				0.505**
				(0.253)
log(Employment$_f$)				0.084***
				(0.023)
log(Age$_f$)				0.048**
				(0.022)
Share Employment$_p$				0.321***
				(0.075)
log(% Workforce with College Degree$_p$)				0.057***
				(0.017)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
N. observations	3,179	3,179	3,179	3,179

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 3-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

As in model (9), the dependent variable in (10) is the degree of autonomy granted to plant p (with primary activity i, located in country c) by parent firm f (with primary activity j, located in country c). The main control of interest is IO_{ij}, the direct requirement coefficient for the sector pair ij. According to prediction P.2, the coefficient of this variable should be positive and significant. Some specifications include vectors of plant-level controls ($\mathbf{X}_p$), firm-level controls ($\mathbf{X}_f$), input-sector and output-sector fixed effects (δ_i and δ_j), and country fixed effects (δ_c). We cluster standard errors at the firm level. Results are practically identical if we cluster at the industry-pair level (the level of variation of the IO coefficient).

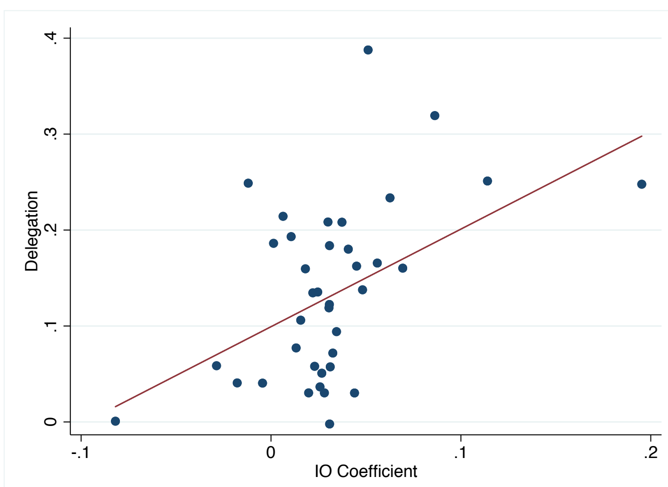
The results are reported in Table 2. Across all specifications, the coefficient of IO_{ij} is positive and significant. In line with prediction P.2 of the model, this result indicates that final good producers delegate more decisions to suppliers of more valuable inputs.

In terms of magnitude, based on the estimates reported in column 3 of Table 2, increasing the input-output coefficient by one standard deviation increases delegation by around 0.05 standard deviations.[44]

Notice that the coefficient of the overall vertical integration index of the firm remains positive and significant, confirming that more integrated firms give more autonomy to their suppliers, in line with prediction P.1. As expected, given the positive correlation with IO_{ij}, the coefficient of $\textit{Vertical Integration}_f$ in column 4 of Table 2 (0.505) is slightly lower than in column 4 of Table 1 (0.577).

The binned scatterplot of Figure 2 illustrates the positive relationship between input value and delegation: suppliers of more valuable inputs are granted more autonomy. The regression line is based on column 1 of Table 2.

Figure 2: Delegation and Input Value Shares



Notes: Binned scatterplot of the relationship between $\textit{Delegation}_{f,p,i,j,c}$ and IO_{ij} (based on 40 bins).

[44]The standard deviation of IO_{ij} in the matched sample is 0.055, so 0.980*0.055 = 0.054.

We have carried out a series of additional estimations to verify the robustness of the results of Table 2. The results of these estimations are reported in the Empirical Appendix. First, we have used more disaggregated industry fixed effects (defined at the SIC4 level) to control for the primary activities of the plant and its parent firm (see Table A-10). Second, we have verified that the results are robust to restricting the analysis to the 10 largest countries in the sample (see Table A-11).[45] In all specifications, the coefficient of IO_{ij} is positive and significant, confirming that final good producers are more likely to delegate decisions to suppliers of more valuable inputs, in line with prediction P.2 of the model.

We next assess the validity of prediction P.3, according to which final good producers should be more likely to integrate suppliers of more valuable inputs. To this purpose, we estimate the following linear probability model:

$$\text{Integration}_{f,j,i,c} = \beta_1 \, \text{IO}_{i,j} + \beta_2 \, \mathbf{X}_f + \delta_i + \delta_f + \epsilon_{f,j,c,i}. \tag{11}$$

The dependent variable is $\textit{Integration}_{f,j,i,c}$, which is equal to 1 if firm f (with primary activity in sector j and located in country c) integrates input i within its boundaries. The key control of interest is $IO_{i,j}$, the direct requirement coefficient for the sector pair ij. Our value theory of integration suggest that a producer of good j should be more likely to integrate suppliers of more valuable inputs, implying that the estimated coefficient of $IO_{i,j}$ should be positive and significant. $\mathbf{X}_f$ is a vector of firm-level controls, while δ_i denotes input-industry fixed effects at the 4-digit SIC level. In the most demanding specifications, we include firm fixed effects (δ_f), which allow us to account for the role of unobservable firm characteristics. In alternative specifications, we replace firm fixed effects with output-sector and country fixed effects (δ_j and δ_c). We cluster standard errors at the input-output level, the same as the main variable of interest, $IO_{i,j}$.[46]

The observations in (11) are at the firm-input level. We focus on the 67,105 firms in the WorldBase sample and consider the top 100 inputs (based on the IO coefficients) necessary to produce the firm's output.[47]

The results are reported in Table 3. in which we regress $\textit{Integration}_{f,j,i,c}$ against the key

[45] These robustness checks are similar to those performed for Table 1. The only difference if that, when verifying the robustness of the results of Table 2, we cannot restrict the analysis to the sub-sample of single-plant firms. Given that these firms tend to be engaged in fewer production activities, there is not enough variation in IO_{ij} to identify the effect of input value on delegation choices.

[46] The results of are unaffected if we use two-way clustering at the input and firm level: changes in the standard errors are minimal and $IO_{i,j}$ is always positive and significant at the 1 percent level.

[47] The number of observations is 7,042,966, which is more than 67,105 firms * 100 inputs. This is because the inputs included are in some cases more than 100, due to the fact that the matching between BEA and SIC codes is not always one-to-one (see footnote 27). For each output j, we have chosen the top 100 inputs based on the ranking of IO coefficients; in the case of a tie, we have included all inputs with the same IO_{ij}.

control of interest, IO_{ij}, and input-industry fixed effects. We then add country fixed effects (column 2), output fixed effects (column 3), and additional firm-level controls (column 4). In the last specification, we include firm fixed effects, exploiting only within-firm variation in integration choices (column 5). Notice that, in this specification, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). The results strongly support prediction P.3 of the theoretical model. Across all specifications, the coefficient of IO_{ij} is positive and significant (at the one-percent level), confirming that final good producers are more likely to produce in house more valuable inputs.

Table 3

Integration and Input Value Shares

	(1)	(2)	(3)	(4)	(5)
IO_{ij}	0.14430***	0.14446***	0.17037***	0.17019***	0.19218***
	(0.02054)	(0.02057)	(0.02465)	(0.02465)	(0.02782)
log($\mathrm{Employment}_f$)				0.00135***	
				(0.00011)	
log(Age_f)				0.00017**	
				(0.00008)	
Input FE	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes	No
Country FE	No	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes
N. observations	7,042,966	7,042,966	7,042,966	7,042,966	7,042,966

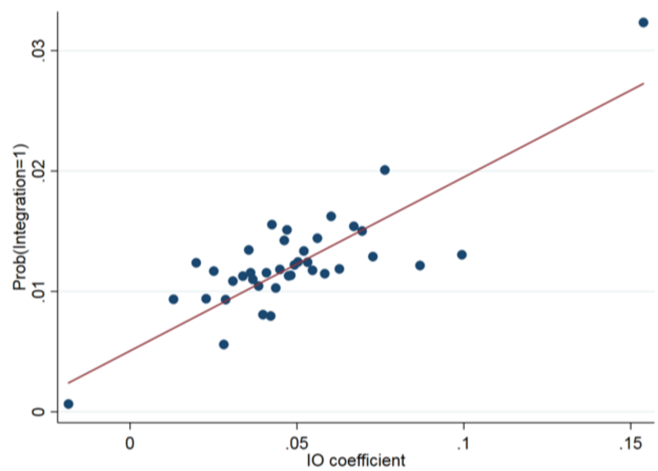
Notes: The dependent variable is $Integration_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. IO_{ij} is the direct requirement coefficient for the sector pair ij. $Employment_f$ measures firm employment, and Age_f is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-output level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

In terms of magnitude, based on the specification of column 3, moving the input-output coefficient by one standard deviation increases the probability of vertical integration by 0.6 percentage points – a 60 percent increase compared to the baseline probability of one percentage point.[48]

[48]The standard deviation of IO_{ij} is 0.036. Thus, 0.170*0.036*100 = 0.612.

The binned scatterplot in Figure 3 illustrates the positive relationship between input value and integration choices: suppliers of more valuable inputs are more likely to be integrated. The regression line is based on column 1 of Table 3.

Figure 3: Integration and Input Value Shares



Notes: Binned scatterplot of the relationship between $Integration_{f,j,i}$ and IO_{ij} (based on 40 bins).

The coefficients of the auxiliary firm controls, $Employment_f$ and Age_f, are positive and significant coefficient, indicating that larger and older firms are more likely to integrate inputs within their boundaries.

We have carried out a series of additional estimations to verify the robustness of the results reported in Table 3. The coefficients of the key variable of interest, IO_{ij}, remains positive and significant if we restrict the analysis to firms included in the matched sample (see Table A-12). The results are also robust to using more disaggregated industry fixed effects (see Table A-13) and restricting the analysis to the 10 largest countries in the sample (see Table A-14). This is also true when we restrict the analysis to single-plant firms (see Table A-15), for which measurement error in the dependent variable should not be a concern.[49]

Overall, the results of Tables 1-3 and in the corresponding robustness checks strongly support predictions P.1-P.3 of the theoretical model concerning the relationship between delegation and integration choices and how these choices are affected by the value of inputs.

[49] As mentioned before, the results of Atalay, Hortaçsu, and Syverson (2014) suggest that using the methodology of Fan and Lang (2000) may lead us to mistakenly classify some inputs as being sourced within firm boundaries, while they are actually bought from independent suppliers. In regression (11), this would imply a measurement error in the dependent variable $Integration_{f,j,i,c}$, which should make it harder to find support for the model's predictions.

5.3 Supplier Uncertainty and the Option Value of Integration

In the theoretical model, because HQ has the right to reassign decision making power to and from her integrated suppliers in the light of new information about their capabilities, integration has an option value. By contrast, non-integrated suppliers always retain control over their decisions, so there is no option value. From HQ's perspective, the payoff structures of the two modes of ownership are therefore very different: for non-integration the payoff is linear in capability, so only mean values matter, but for integration there is a convexity introduced by the option to centralize or delegate, so higher moments also count.

Following intuition from option theory, this observation suggests that characteristics of the distribution of supplier capability should influence integration decisions. In finance, the value of a classical call option increases with the Rothschild-Stiglitz risk of the underlying asset. According to the model, integration has costs that are also affected by the distribution of supplier productivity, so the net predicted impact of an increase in input risk on firm boundary choices is ambiguous. In what follows, we first examine the role of uncertainty empirically, and then rationalize them in light of the theoretical model.

To proxy for input risk, we use information on the dispersion of labor productivity and of stock market returns in input industries. To identify the relevant input industries, we combine information from WorldBase on firms' production activities with Input-Output data from the BEA. For each firm f producing good j in country c, we focus on the top 100 inputs i as ranked by the IO coefficients IO_{ij}.

Our main measure of input risk is *CV Productivity*$_{i,c}$, the coefficient of variation of labor productivity of suppliers of input i in country c. As discussed in Section 4, when constructing this measure, we focus on independent suppliers: according to the model, the empirical productivity distribution of non-integrated suppliers in industry i approximates the distribution $F_i(y)$.[50]

To assess how input risk affects integration choices, we estimate

$$\begin{aligned}\text{Integration}_{f,j,i,c} \quad = \quad & \gamma_1 \text{ CV Productivity}_{i,c} + \gamma_2 \text{ Mean Productivity}_{i,c} + \\ & \gamma_3 \text{ IO}_{i,j} + \gamma_4 \, \mathbf{X}_f + \delta_i + \delta_f + \epsilon_{f,j,c,i}.\end{aligned} \tag{12}$$

[50]As noted in footnotes 31 and 32, censoring introduced by HQs' exercise of the option to centralize, and transfer pricing may distort measured supplier productivity, so the observed productivity distribution of integrated suppliers is not a reliable representation of the underlying capability distribution. The results reported below nevertheless continue to hold if integrated suppliers are included in the construction of *CV Productivity*$_{i,c}$.

Table 4

Integration and Riskiness of Input industries

	(1)	(2)	(3)	(4)	(5)	(6)
CV Productivity$_{i,c}$	0.00064***	0.00075***	0.00075***	0.00074***	0.00074***	0.00074***
	(0.00009)	(0.00010)	(0.00010)	(0.00010)	(0.00010)	(0.00010)
Mean Productivity$_{i,c}$	-0.01305	-0.01677	-0.01652	-0.01609	-0.01607	-0.01598
	(0.01272)	(0.01544)	(0.01525)	(0.01523)	(0.01522)	(0.01510)
IO$_{ij}$			0.14985***	0.17906***	0.17888***	0.20304***
			(0.01342)	(0.01447)	(0.01446)	(0.01611)
log(Employment$_f$)					0.00144***	
					(0.00013)	
log(1+ Age$_f$)					0.00017	
					(0.00011)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	6,644,884	6,644,884	6,644,884	6,644,884	6,644,884	6,644,884

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of input supplier productivity. IO_{ij} is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

As in model (11), the dependent variable in (13) is an indicator equal to 1 if firm f (with primary activity in sector j and located in country c) integrates input i within its boundaries. The key control of interest is *CV Productivity*$_{i,c}$, capturing the riskiness of input industry. In all specifications, we also control for *Mean Productivity*$_{i,c}$, the mean of supplier productivity. We cluster standard errors at the input industry level (the results are unaffected if we use two-way clustering at the input and firm level).

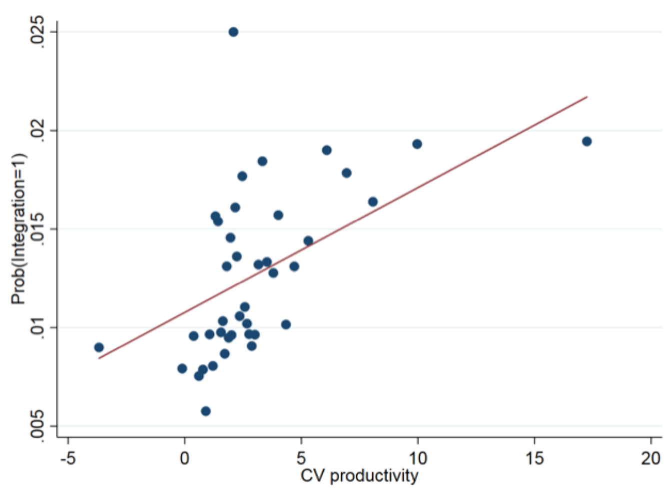
The results are reported in Table 4. Following the same structure as in Table 3, we present first a specification that includes the key control variables with input-industry fixed effects (column 1), and then further include country fixed effects (column 2), output-industry fixed effects (column 3), and the firm controls (column 4). In the last specification, we replace country and output-industry fixed effects with firm fixed effects (column 5).

In all specifications, the estimated coefficient for *CV Productivity*$_{i,c}$ is positive and significant (at the one percent level), indicating that firms are more likely to integrate suppliers that operate in industries with greater productivity dispersion.

As for the economic magnitude of the effects, based on the specification in column 3, a one-standard-deviation increase in CV Productivity$_{i,c}$ increases the probability of integrating a supplier by around 0.34 percentage points. This corresponds to a 34% increase relative to the baseline integration probability of one percentage point.[51]

The binned scatterplot in Figure 4 illustrates the positive relationship between *CV Productivity* and *Integration*$_{f,j,i,c}$. The regression line is based on column 1 of Table 4.

Figure 4: Integration Probability and Riskiness of Input Industry



Notes: Binned scatterplot of the relationship between *Integration*$_{f,j,i,c}$ and *CV Productivity*$_{i,c}$ (based on 40 bins).

[51]The standard deviation of CV Productivity$_{i,c}$ is 4.63, thus 0.0007*4.63*100 =0.342.

Concerning the other controls, the positive and significant coefficient of IO_{ij} confirms that producers are more likely to integrate more valuable inputs, in line with prediction P.3 of the theoretical model. The positive and significant coefficient of $Employment_f$ indicates that the propensity to integrate inputs is higher for larger firms.

We have performed a series of additional estimations to verify how the riskiness of input industries affects firms' decisions to integrate suppliers. First, we have carried out the analysis using the matched sample instead of the larger WorldBase sample (Table A-16). The results confirm that higher dispersion in the productivity of input suppliers increases the the probability of integration. The results are also robust to using more disaggregated industry fixed effects (Table A-17), restricting the analysis to the top-10 countries (Table A-18), and focusing on single-plant firms (Table A-19). We have also constructed the variable $CV\ Productivity_{i,c}$ after winsorizing labor productivity at the 5th and 95th percentile (Table A-20) and restricted the analysis to input industries in which there are at least 50 suppliers in each input industry-country (Table A-21). In all these regressions, the coefficient of $CV\ Productivity_{i,c}$ remains positive and significant.

We have also reproduced Table 4 using data on stock market returns to capture cross-industry variation in the distribution of supplier productivity (see Table A-22). As discussed in Section 4, the variable $SD\ Stock\ Returns_i$ is constructed using data on stock market returns of US firms. Using this alternative risk measure reduces the number of observations and forces us to assume that the distribution of supplier productivity within an industry is the same across all countries in the sample. Additional noise comes from the fact that $SD\ Stock\ Returns_i$ can only be constructed at the SIC4 level for some manufacturing industries, which further reduces the number of observations. Notwithstanding these limitations, the results confirm that firms' propensity to integrate an input increases with the dispersion in the productivity of suppliers in the input industry: the coefficient of $SD\ Stock\ Returns_i$ is always positive and significant.

Overall, the empirical results strongly indicate that the probability that firms integrate a particular input within their boundaries increases with the dispersion in supplier capability in that industry. To rationalize these results, note first that the option value of integration (ignoring the cost) can be written as

$$\mathbb{E}_F \max[v^C(A,\pi) - A\pi\lambda y, v^D(A,\pi,y) - A\pi\lambda y].$$

The integrand is a convex function of y on $[0,\overline{y})$, since it is the maximum of two convex functions. Consequently, the option value rises when the distribution F increases in Rothschild-Stiglitz riskiness. In other words, supply assurance is a more compelling motive to integrate when there is more uncertainty about supplier capability.

Of course Rothschild-Stiglitz riskiness is not always easily quantified by single parameters. But for lognormal families of distributions indexed by (μ, σ), wherein the mean is $e^{\mu+\frac{\sigma^2}{2}}$ and the coefficient of variation is $\sqrt{e^{\sigma^2}-1}$, holding the mean fixed while increasing CV results in greater Rothschild-Stiglitz risk.

In short, if only the value of integration matters, the empirical results can easily be understood: the positive response to increases in CV corresponds precisely to increases in Rothschild-Stiglitz risk, at least for the lognormal distributions our data approximate. The only confounding effects are on the cost side. The integration cost $cF(y^*(A)) + \phi$ responds ambiguously to changes in risk, so the net predicted effect is ambiguous. However, the empirical results are clearly consistent with the theoretical model. One interpretation is that the cost parameter c is fairly small: the value effects dominate the cost effects.[52]

6 Alternative Mechanisms

Our empirical analysis establishes the following regularities:

1. Firms that delegate more decisions tend to be more vertically integrated.
2. Firms delegate more decisions to integrated suppliers that produce more valuable inputs.
3. Firms are more likely to integrate suppliers of more valuable inputs.
4. Firms are more likely to integrate inputs in industries in which supplier productivity is more dispersed.

These results can be rationalized by our theoretical model, in which integration enhances efficiency and creates a real option for HQ to retain control or delegate according to comparative advantage. Below we discuss other possible explanations for our findings.

[52] While not as robust, the empirical finding that increasing the mean of supplier productivity lowers the propensity to integrate can also be rationalized by the model. For lognormal families, holding CV fixed while increasing the mean corresponds to a (first order) stochastic increase in the distribution. Intuitively, this reduces the concern about downside capability risk and thereby reduces the supply assurance motive to integrate; against this, such increases in the distribution enhance the relative value of adapted over generic inputs, but only if the shift is happening at very high values of y. Formally, writing the integrand above as $\max\{(1-\lambda y)A\pi, v^D(A,\pi,y) - A\pi\lambda y\}$, the first argument of $\max\{\cdot,\cdot\}$ is clearly decreasing, while for the second argument, note $v^D_y(A,\pi,y) = \frac{2A\pi y+(A\pi y)^2}{(1+A\pi y)^2}A\pi < A\pi\lambda$ unless y is sufficiently large. Stochastic increases in capability do reduce the integration cost, leading to net ambiguous effects. If the effect on costs are small, as our findings on input risk suggest, increases in the mean weaken integration incentives.

The covariation of delegation and integration might be rationalized by models in which headquarter's attention is a scarce corporate resource (e.g., Geanakoplos and Milgrom, 1991; Aghion and Tirole, 1995). If vertical integration increases the scope of decisions in a firm, HQ may simply need to cede control to lower-level managers.

We believe that theories of limited managerial capacity do not provide a rationale for our empirical findings. There are four reasons for this. First, the positive correlation between delegation and integration is robust to controlling for the size of the firm as captured by its total number of employees (see column 3 of Table 1).

Table 5

Delegation and Integration, Controlling for Management

	(1)	(2)	(3)	(4)
Vertical Integration$_f$	0.577**	0.554**	0.505**	0.484*
	(0.250)	(0.249)	(0.253)	(0.253)
IO$_{ij}$			0.941**	0.912**
			(0.441)	(0.439)
log(Employment$_f$)	0.087***	0.061***	0.084***	0.058**
	(0.022)	(0.022)	(0.023)	(0.023)
log(Age$_f$)	0.035*	0.033	0.048**	0.045**
	(0.021)	(0.021)	(0.022)	(0.022)
Share Employment$_p$	0.323***	0.293***	0.321***	0.291***
	(0.073)	(0.073)	(0.075)	(0.075)
log(% Workforce with College Degree$_p$)	0.056***	0.043***	0.057***	0.044***
	(0.016)	(0.016)	(0.017)	(0.017)
Management$_p$		0.096***		0.093***
		(0.021)		(0.022)
Input FE	Yes	Yes	Yes	Yes
Output FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
N. observations	3,444	3,444	3,179	3,179

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. *Management*$_p$ is the normalized z-score capturing the quality of the plant's management practices. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 3-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Second, these theories would view delegation and management as substitutes, to the extent that good management reduces headquarters' overload. To verify this, we have included the quality of a plant's management practices as a control in (9). The results in Table 5 suggest that delegation and management are complements rather than substitutes: the better the plant's management practices the higher is the degree of autonomy given to plant-level managers.[53] Also, if good management reduces headquarters' overload, the partial correlation between delegation and vertical integration should become *larger* once one controls for the quality of management. Comparing the coefficients of $Vertical\ Integration_f$ reported in Table 5, this clearly not the case: the coefficients in columns 2 and 4 are, if anything, smaller than those in columns 1 and 3.[54]

Third, in a model in which top managers are capacity constrained, one would expect them to keep control of the decisions related to the most important inputs and delegate decisions concerning less important ones. The empirical analysis shows instead that more decision-making autonomy is granted to suppliers of more important inputs (regularity 2 above), in line with prediction P.2 of our model.

The final reason for skepticism is that theories of limited managerial capacity have little to say about how input value shares affect integration decisions (regularity 3) and how the dispersion in supplier capability affects firm boundaries (regularity 4).

The model in this paper exhibits a novel kind of supply assurance motive for integration: the ability to centralize control under integration affords the HQ at least a moderate level of input value, even if her supplier turns out to be quite inept.[55] Integration helps guarantee the firm a minimum quality of inputs, because it allows redeployment of control in response to information that arrives during the course of production. Note that it is *interim* uncertainty (after production begins, but before the input is produced) that is hedged here, in line with the empirical finding that input risk increases integration propensities.

In "ex-post" forms of supply assurance (e.g., Carlton, 1979; Bolton and Whinston, 1993; Baker, Gibbons, and Murphy, 2002), firms also integrate in order to guarantee a stable supply of inputs. But the assurance motive for integration is driven by uncertainty resolved after input production (e.g., product demand), possibly augmented by the supplier's hold-up behavior. Broadly speaking, one would expect less integration when there is less of a risk of suppliers coming up short, for technological or behavioral reasons. This might then provide

[53]We Substituting the variable $Management_p$ with its four components (see footnote 30) reveals that only those management practices related to providing targets and incentives to personnel are significantly correlated with the degree of autonomy granted to the plant manager.

[54]The coefficients in columns 1 and 3 (2 and 4) are not statistically different.

[55]One Boeing 787 engineer complained that some of the outsourced electrical components were "like Radio Shack ... cheap, plastic and prone to failure" (Gates, 2013).

an explanation for the positive coefficient of *CV Productivity*$_{i,c}$ in Table 4 (regularity 4).

Typically, the ex-post assurance motives for integration would be mitigated when there are many suppliers in an input industry. Against this hypothesis, focusing on input industries in which there are many suppliers, leads to the finding that the coefficient of *CV Productivity*$_{i,c}$ remains positive and highly significant (see Table A-21), albeit with somewhat diminished magnitude (the difference in the coefficients is significant at the 5% level). This is also true in the specification that includes firm fixed effects, which account for demand for inputs by other firms in the same country-output sector (column 4), while output industry fixed effects in other columns control for product market uncertainty.

More generally, existing supply assurance theories have little to say about our other empirical findings, concerning the interplay between firm boundaries and the allocation of decision rights within firms and the role of input value shares in shaping delegation choices (regularities 1 and 2).

7 Conclusion

Organizations are complicated. Understanding them entails simplification, and a lot has been learned by isolating distinct organizational design elements. But there are costs to isolation. Formally similar models that focus only on one dimension or another of the organization (integration or centralization) can mislead when embedded in two dimensions (integration and centralization might be predicted to covary positively, unlike in our model or in the data).

A multidimensional approach helps to illuminate interdependencies that can be crucial for understanding organizational functioning and guiding organizational design. The present analysis is predicated on a well-known conceptual distinction between ownership (integration) and delegation choices: ownership is formal, delegation informal.[56] But the thrust of the analysis emphasizes their interrelatedness, and that generates some new insights.

For example, there is a complementarity between authority over production decisions (e.g., the fixed investment that starts the adaptation process in the model) and authority to allocate control (the delegation choice): without the first, the second has no bite. Moreover, scholarship on firm boundaries, which tends to ignore the potential interplay between integration and delegation, has often argued (at least since Coase, 1937) that managerial authority under integration leads to rigidities of various sorts. Here it has been suggested, on the con-

[56] Indeed, the law regulates and registers asset sales and adjudicates disputes between parties who hold separate titles. Once they are integrated, however, the parties largely forego appeal to the law in many of their disputes, and via the business judgment rule, are immune to its intervention in most matters, in particular who will make various business decisions.

trary, that integration may actually increase managerial flexibility, because it facilitates the re-allocation of decision-making across different parts of the organization. In the more nuanced view that emerges from a multidimensional approach, it is only a particular form of integrated ownership structure – centralization – that appears rigid. Finally, studying integration and delegation together reveals a novel mechanism by which supply assurance motivates integration: it facilitates redeployment of control over production decisions, thereby assuring a minimal level of competence across them.

The empirical results provide strong support for the model's predictions. More vertically integrated firms tend to delegate more decisions to lower management. Moreover, input value affects both integration and delegation choices: suppliers of more important inputs are more likely to be integrated; among integrated suppliers, those producing more valuable inputs are granted more autonomy from top management. Finally, firms are more likely to integrate suppliers in "riskier" input industries. This is consistent with the idea that the option value of integration is higher – and the supply assurance motive stronger – in input industries in which supplier ability is more dispersed.

The findings also raise new questions, both about how integration and delegation interact with each other, and about how they do so with other aspects of organization. An example of the first concerns the dynamics of this relationship. The present analysis – in large measure due to the cross-sectoral nature of the data – has been static, with an assumed one-off integration and delegation decision per producer-supplier relationship. Exploring both theoretically and empirically how changes in the capabilities of suppliers or market conditions affect the interplay between integration and delegation choices over time, and with them the option value of ownership, is an interesting avenue for future research.

With respect to other aspects of organization, the empirical results in Table 5 suggest that firms in which central headquarters give more autonomy to their subordinate suppliers tend to adopt better management practices; that is, some of these practices are complementary to delegation. Further theoretical and empirical investigation of the relationship among integration, delegation, and management is an important direction for future inquiry. This would contribute to the broader agenda of understanding how choices and efficacy of management practices depend on the organizational environment, and how these decisions affect firm performance.

As evidence mounts that organization matters for the performance of individual firms, industries, and aggregate economies, it is becoming ever more imperative to understand the functioning of organizations as a whole rather than just their parts. We hope this exercise is an encouraging illustration of what can be learned by bringing together disparate elements of organizational design within a single framework, as well as rich data to measure them.

References

Acemoglu, D., P. Aghion, R. Griffith, and F. Zilibotti (2010). "Vertical Integration and Technology: Theory and Evidence," *Journal of European Economic Association* 8, 989-1033.

Acemoglu, D., P. Aghion, C. Lelarge, J. Van Reenen, and F. Zilibotti (2007). "Technology, Information and the Decentralization of the Firm," *Quarterly Journal of Economics* 122, 1759-1799.

Acemoglu, D., S. Johnson, and T. Mitton (2009). "Determinants of Vertical Integration: Financial Development and Contracting Costs," *Journal of Finance* 63, 1251-1290.

Aghion, P., M. Dewatripont, and P. Rey (2002). "On Partial Contracting," *European Economic Review* 46, 745-753.

Aghion, P. and J. Tirole (1995). "Some Implications of Growth for Organizational Form and Ownership Structure," *European Economic Review* 39, 440-455.

______(1997). "Real and Formal Authority in Organizations," *Journal of Political Economy* 105, 1-29.

Akcigit, U., H. Alp, and M. Peters (2018). "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries," NBER Discussion Paper 21905.

Alfaro, L., P. Antràs, D. Chor, and P. Conconi (2019). "Internalizing Global Value Chains: A Firm-Level Analysis," *Journal of Political Economy* 127, 508-559.

Alfaro, L., P. Conconi, H. Fadinger, A. Newman (2016). "Do Prices Determine Vertical Integration?,"*Review of Economic Studies* 83, 1-35.

Alonso, R., W. Dessein, and N. Matouschek (2008). "When does Coordination require Centralization?" *American Economic Review* 98, 145-179.

Alonso, R. and N. Matouschek (2008). "Optimal Delegation," *Review of Economic Studies* 75, 259-93.

Atalay, E., A. Hortaçsu, and C. Syverson (2014), "Vertical Integration and Input Flows," *American Economic Review* 104, 1120-1148.

Baker, G., R. Gibbons, and K. Murphy (1999). "Informal Authority in Organizations," *Journal of Law Economics and Organization* 15, 56-73.

______(2002). "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics* 117, 39-84.

Baker, G., and T. Hubbard (2003). "Make Versus Buy in Trucking: Asset Ownership, Job Design and Information," *American Economic Review* 93, 551-572.

Berlingieri, G., F. Pisch, and C. Steinwender (2018). "Organizing Supply Chains: Input Cost Shares and Vertical Integration," NBER Working Paper No. 25286.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). "Does Management Matter? Evidence from India," *Quarterly Journal of Economics* 128, 1-51.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta, and S. Terry (2018). "Really Uncertain Business Cycles," forthcoming, *Econometrica*.

Bloom, N., R. Sadun, and J. Van Reenen (2010). "Does Product Market Competition Lead Firms to Decentralize?," *American Economic Review* 100, 434-438.

______(2012). "The Organization of Firms Across Countries," *Quarterly Journal of Economics* 127, 1663-1705.

______(2016). "Management as a Technology?," NBER Working Paper No. 22327.

Bloom, N., and J. Van Reenen (2007). "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics* 122, 1351-1408.

Bolton, P. and M. Whinston (1993). "Incomplete Contracts, Vertical Integration, and Supply Assurance," *Review of Economic Studies* 60, 121-148.

Caliendo, L., F. Monte, and E. Rossi-Hansberg (2015). "The Anatomy of French Production Hierarchies," *Journal of Political Economy* 123, 809-852.

Carlton, D. W. (1979). "Vertical Integration in Competitive Markets Under Uncertainty," *Journal of Industrial Economics* 27, 189-209.

Certo, S.C. and S.T. Certo (2012). *Modern Management* 12th ed. Prentice Hall.

Coase, R. H. (1937). "The Nature of the Firm," *Economica* 4, 386-405.

Conconi, P., P. Legros, and A. F. Newman (2012). "Trade Liberalization and Organizational Change," *Journal of International Economics* 86, 197-208.

Chipty, T. (2001). "Vertical Integration, Market Foreclosure, and Consumer Welfare in the Cable Television Industry," *American Economic Review* 91, 428-53.

Dessein, W. (2002). "Authority and Communication in Organizations," *Review of Economic Studies* 69, 811-838.

Dessein, W. (2014). "Incomplete Contracts and Firm Boundaries: New Directions" *Journal of Law Economics and Organizations* 30, i13-i36.

Dessein W., L. Garicano, and R.H. Gertner (2010), "Organizing for Synergies," *American Economic Journal: Microeconomics* 2, 77-114.

Dessein, W. and T. Santos (2006). "Adaptive Organizations," *Journal of Political Economy* 114, 956-995.

Fan, J. P. H., and L. H. Lang (2000). "The Measurement of Relatedness: An Application to Corporate Diversification," *Journal of Business* 73, 629-660.

Forbes, S. and M. Lederman (2010). "Does Vertical Integration Affect Firm Performance? Evidence from the Airline Industry," *Rand Journal of Economics* 41, 765-790.

Friebel, G., and M. Raith (2010). "Resource Allocation and Organizational Form," *American Economic Journal: Microeconomics* 2, 1-33.

Garicano, L. (2000). "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy* 108, 874-904.

Garicano, L. and T. Santos (2004). "Referrals," *American Economic Review* 94, 499-525.

Gates, D. (2013). "Boeing 787's Problems Blamed on Outsourcing, Lack of Oversight," *Seattle Times*, February 2, 2013.

Geanakoplos, J., and P. Milgrom (1991). "A Theory of Hierarchies Based on Limited Managerial Attention," *Journal of the Japanese and International Economies* 5, 205-225.

Gibbons, R., and J. Roberts (2013). *The Handbook of Organizational Economics*, Princeton: Princeton University Press.

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

Grossman, G. M., and E. Helpman (2002), "Integration Versus Outsourcing In Industry Equilibrium," *Quarterly Journal of Economics* 117, 85-120.

Guadalupe, M., and J. Wulf (2010). "The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization," *American Economic Journal: Applied Economics* 4, 105-127.

Hart, O. D., and B. R. Holmström (2010). "A Theory of Firm Scope," *Quarterly Journal of Economics* 125, 483-513.

Hart, O. D., and J. H. Moore (1990). "Property Rights and the Nature of the Firm," *Journal of Political Economy* 98, 1119-1158.

Hart, O. D., and J. H. Moore (2005). "On the Design of Hierarchies: Coordination versus Specialization," *Journal of Political Economy* 113, 675-702.

Holmström, B. (1984). "On the Theory of Delegation," in M. Boyer and R. Kihlstrom (eds.), *Bayesian Models in Economic Theory*, Elsevier.

Holmström, B. and P. Milgrom (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics and Organization* 7, 24-52.

Holmström, B., and J. Tirole (1991). "Transfer Pricing and Organizational Form," *Journal of Law, Economics, and Organization* 7, 201-28.

Hortaçsu, A., and C. Syverson (2007). "Cementing Relationships: Vertical Integration, Foreclosure, Productivity and Prices," *Journal of Political Economy* 115, 250-301.

Joskow, P. (1987). "Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets," *American Economic Review* 77, 168-185.

Katayama, H., Meagher, K., Wait, A. (2018). "Authority and Communication in Firms," *Journal of Economic Behavior and Organization* 155, 315-348.

Klapper, L., L. Laeven, and R. Rajan (2006). "Entry Regulation as a Barrier to Entrepreneurship," *Journal of Financial Economics* 82, 591-629.

Kogut, B., and U. Zander (1992). "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology," *Organization Science* 3, 383-397.

Lafontaine, F., and M. Slade (2007). "Vertical Integration and Firm Boundaries: The Evidence" *Journal of Economic Literature* 45, 629-685.

——— Lafontaine, F., and M. Slade (2013). "Inter-Firm Contracts," in R. Gibbons and J. Roberts, *Handbook of Organizational Economics*, 958-1013.

Legros, P., and A. F. Newman (2008). "Competing for Ownership," *Journal of the European Economic Association* 6, 1279-1308.

——— (2013). "A Price Theory of Vertical and Lateral Integration," *Quarterly Journal of Economics* 128, 725-770.

——— (2015). "Integration, Delegation and Management," mimeo.

——— (2017). "Demand-Driven Integration and Divorcement Policy," *International Journal of Industrial Organization* 53, 306-325.

Levy, H. (1973). "Stochastic Dominance Among Log-Normal Prospects," *International Economic Review* 14(3): 601-14.

Marin, D., and T. Verdier (2008). "Power Inside the Firm and the Market: A General Equilibrium Approach," *Journal of the European Economic Association* 6, 752-788.

McDonald, R., and S. Kotha (2015). "Boeing 787: Manufacturing a Dream," Harvard Business School Case 615-048.

McElheran, K. (2014). "Delegation in Multi-Establishment Firms: Adaptation vs. Coordination in I.T. Purchasing Authority," *Journal of Economics and Management Strategy* 23, 225-257.

McLaren, J. (2000). "Globalization and Vertical Structure," *American Economic Review* 90, 1239-1254.

Milgrom, P., and J. Roberts (1990). "The Economics of Modern Manufacturing: Technology, Strategy and Organization," *American Economic Review* 80, 511-528.

Nelson, R. R., and S. G. Winter (1982). *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge, MA.

Powell, M. (2015). "An Influence-Cost Model of Organizational Practices and Firm Boundaries,"*Journal of Law, Economics, and Organization* 31, i104-i140.

Rantakari, H. (2013), "Organizational Design and Environmental Volatility," *Journal of Law, Economics, and Organization* 29, 569-607.

Roberts, J. (2007). *The Modern Firm: Organizational Design for Performance and Growth*, Oxford University Press.

Tang, C.S., J. D. Zimmerman, and J. Nelson (2009). "Managing New Product Development and Supply Chain Risks: The Boeing 787 Case," *Supply Chain Forum: An International Journal* 10, 74-86.

Van den Steen, E. (2005). "Organizational Beliefs and Managerial Vision," *Journal of Law, Economics, and Organization* 21, 256-283.

Van den Steen, E. J. (2010) "Interpersonal Authority in a Theory of the Firm," *American Economic Review* 100, 466-490.

Wernerfelt, B. (1984). "A Resource-Based View of the Firm," *Strategic Management Journal* 5, 171-80.

Williamson, O. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*, Free Press.

Woodruff, C. (2002). "Non-contractible Investments and Vertical Integration in the Mexican Footwear Industry," *International Journal of Industrial Organization* 20, 1197-1224.

Woodward, J. (1958) *Management and Technology*. Cambridge: Cambridge University Press.

Wu, Y. (2017). "Authority, Incentives, and Performance: Evidence from a Chinese Newspaper," *Review of Economics and Statistics* 99, 16-31.

Zhao, Y., and X. Xu (2013). "Build-to-Performance: The Boeing 787 Dreamliner: A Case on Development Outsourcing and Supply Chain Management."

A-1 Theoretical Appendix

Multiple Tasks Extension

In the model described in Section 3, delegation is a binary variable: HQ either centralizes or delegates the production of an input to an integrated supplier. Below we extend the model to allow for multiple tasks to be performed by each integrated supplier. This allows us to generate a continuous delegation choice, in line with the measure of delegation used in our empirical analysis.

Consider a supplier industry i and suppose that there are many tasks $t \in \{1, \ldots, T\}$ that need to be performed in order to adapt the input. On each task the capability of the supplier is a random variable $y + \epsilon_t$, where y has distribution F and ϵ_t are i.i.d., and independent of y, with distribution $G(\epsilon)$ and mean zero.

The ϵ_t as well as the single draw of y are realized and observed before task assignment. We think of the distribution of task-specific capability G as independent of input i, while the overall capability F depends on i, as in the benchmark model described in Section 3. As before, HQ has capability 1 on all tasks and can separately delegate or retain control over each task. Each task contributes equally and additively to the overall supplier value, and costs of decision on each task are weighted by $1/T$. Then, the capability $x_t \equiv y + \epsilon_t$ has distribution given by the convolution

$$C(x_t) = \int_0^\infty G(x_t - y) f(y) dy.$$

Now, centralizing a task yields a payoff to HQ equal to $A\pi/T$. Delegation yields $(1/T)A\pi(y+\epsilon_t)(1-(s_t-h_t)^2)$ at cost $(1/T)(1-h_t)^2$ to HQ, $(1/T)cs_t^2$ to S. As before, $s_t = 0$, so now delegation of task t yields $(A\pi(y+\epsilon_t))^2/(1+A\pi(y+\epsilon_t))$, provided $y+\epsilon_t > 0$ (there is never delegation if $y+\epsilon_t \leq 0$). In other words, $y+\epsilon_t$ replaces y in the benchmark model, and delegation occurs when $y+\epsilon_t > y^*(A\pi)$. The probability of centralizing one task is $C(y^*(A\pi))$, increasing in $y^*(A\pi)$, therefore decreasing in $A\pi$. So the probability of delegating the task is $1 - C(y^*(A\pi))$, increasing in $A\pi$.

In this setting, the degree of delegation is simply the number (or fraction) of tasks delegated, which is very close to the delegation measure employed in our empirical analysis. It is a binomial random variable with parameters $(1 - C(y^*), T)$, which like the delegation probability of Lemma 1 is (stochastically) increasing in A and π.

Of course this formulation modifies the value of integration somewhat. For each task, HQ obtains value:

$$v_t(A, \pi, y + \epsilon_t) = \begin{cases} A\pi, & \text{if } y + \epsilon_t \leq y^*(A\pi) \\ v^D(A, \pi, y + \epsilon_t), & \text{if } y + \epsilon_t > y^*(A\pi). \end{cases}$$

Using the change of variable $x_t = y + \epsilon_t$, there is integration iff:

$$\frac{1}{T} \sum_{t=1}^{T} \left(\mathbb{E} v_t(A, \pi, x_t) - V^N \right) > C(y^*(A\pi))c + \phi.$$

Under non-integration, the adaptation tasks do not enter, since there is no adaptation; the non-integration value is still governed by the expectation of the random variable y, that is $V^N = A\pi\lambda\mathbb{E}y$.

We can then reformulate the main results about firm boundary choices as follows:

Proposition 2. *In the tasks model, (i) the propensity to integrate increases with A and π and (ii) the option value of integration increases in the riskiness of $F(y)$.*

Proof. (i) Integration increases with A and with π. As in the baseline model, because $C(y^*(A\pi)))$ is a decreasing function of A, it is enough to show that the left hand side is an increasing function of A and of π, a sufficient condition being that each term $\mathbb{E}v_t(A, \pi, x_t) - V^N$ is increasing in A and in π. The argument mimics those in footnotes 13 and 14.

(ii) The option value increases in the riskiness of $F(y)$. We can write

$$\mathbb{E} v_t(A, \pi, x_t) = \int \left(\int \max\{A\pi, v^D(A, \pi, y + \epsilon_t)\} dF(y) \right) dG(\epsilon_y).$$

The term $\max\{A\pi, v^D(A, \pi, y + \epsilon_t)\}$ is convex in y, so the integral in parentheses is increasing in the riskiness of $F(y)$; thus the expectation with respect to ϵ_t increases with risk. As $F(y)$ is common to all tasks, the total output option value $\frac{1}{T} \sum_{t=1}^{T} \mathbb{E} v_t(A, \pi, x_t)$ increases with the riskiness of $F(y)$. □

Since integration and delegation both continue to increase with A, the co-variation of integration and delegation is preserved.

As with the baseline model, for the propensity to integrate to increase with risk, the cost of integration $C(y^*(A\pi))c$ must not increase too quickly. Rather than bounding the cost parameter c, as discussed at the end of section 5.3, we can invoke an alternative condition that is independent of c: that the distribution of the noise $G(\epsilon_t)$ have a decreasing density. Indeed, in this case $G(x - y)$ is concave in y, and therefore riskier $F(y)$ distributions reduce the probability $C(y^*(A\pi)) = \int_0^\infty G(y^*(A\pi)) - y)dF(y)$ of centralizing a task.

A-2 Empirical Appendix

A-2.1 Descriptive Statistics

Table A-1

Size of plants

	Mean	Median	Standard deviation	N. plants
WMS dataset	277	150	405	11,691
WorldBase dataset	288	2	5,428	17,371,146
Matched sample	254	150	367	3,444
Worldbase sample	147	42	3,187	67,111

Notes: The table reports statistics on the plants included in the WMS and WorldBase datasets, and in the samples used in our empirical analysis.

Table A-2

Observations by Country, Matched Sample

Country	Number of Observations	Percentage
Argentina	100	2.90
Australia	133	3.86
Brazil	234	6.79
Canada	207	6.01
Chile	95	2.76
China	64	1.86
France	212	6.16
Germany	224	6.50
Greece	104	3.02
India	104	3.02
Italy	106	3.08
Ireland	75	2.18
Japan	102	2.96
Mexico	86	2.50
New Zealand	118	3.43
Poland	27	0.78
Portugal	78	2.26
Sweden	330	9.58
United Kingdom	432	12.54
United States	613	17.80
Total	3,444	100.00

Notes: The table reports the number of plant observations by country in the matched sample.

Table A-3

Observations by Country, WorldBase Sample

Country	Number of Observations	Percentage
Argentina	17,081	0.26
Australia	61,489	0.93
Brazil	3,857	0.06
Canada	149,022	2.24
Chile	4,570	0.07
China	558,337	8.40
France	35,617	0.54
Germany	1,985,864	29.89
India	101,107	1.52
Italy	412,315	6.20
Ireland	5,804	0.09
Japan	1,088,345	16.38
Mexico	30,865	0.46
New Zealand	44,824	0.67
Poland	28,116	0.42
Portugal	142,727	2.15
Sweden	17,319	0.26
United Kingdom	156, 962	2.36
United States	1,800,663	27.10
Total	6,664,884	100.00

Notes: The table reports the number of observations by country in the WorldBase sample. The observations are at the firm-input level. For each firm in the WorldBase sample, we consider the top 100 inputs (based on the IO coefficients) necessary to produce the firm's output.

Table A-4

Descriptive Statistics of Main Variables, Matched Sample

	Mean	Median	Standard deviation	N. observations	N. firms
Delegation_p	0.13	0.07	0.99	3,444	2,883
$\text{Share Employment}_p$	0.61	0.60	0.89	3,384	2,621
% $\text{Workers with College Degree}_p$	15.20	10.00	16.34	3,225	2,655
Management_p	3.05	3.06	0.65	3,444	2,883
$\text{Vertical Integration}_f$	0.10	0.08	0.08	3,444	2,883
Employment_f	674.89	300.00	1,043.32	3,444	2,883
Age_f	40.08	30.00	35.02	3,444	2,883
$\text{IO}_{i,j}$	0.04	0.04	0.035	3,179	2,428

Notes: The table reports descriptive statistics of the main variables used in the regressions of Tables 1, 2, and 5 (and robustness checks), based on the matched sample. *Delegation*$_p$, is the overall autonomy index of plant p. *Share Employment*$_p$ is the plant's share of the firm's employment. % *Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. *Management*$_p$ is the normalized z-score capturing the quality of the plant's management practices. *Employment*$_f$ measures the number of employees of firm f. *Age*$_f$ is the number of years since the firm was established. *Vertical integration*$_f$ is the vertical integration index of firm f. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij, measured at the 4-digit SIC level for the top 100 inputs of each industry j.

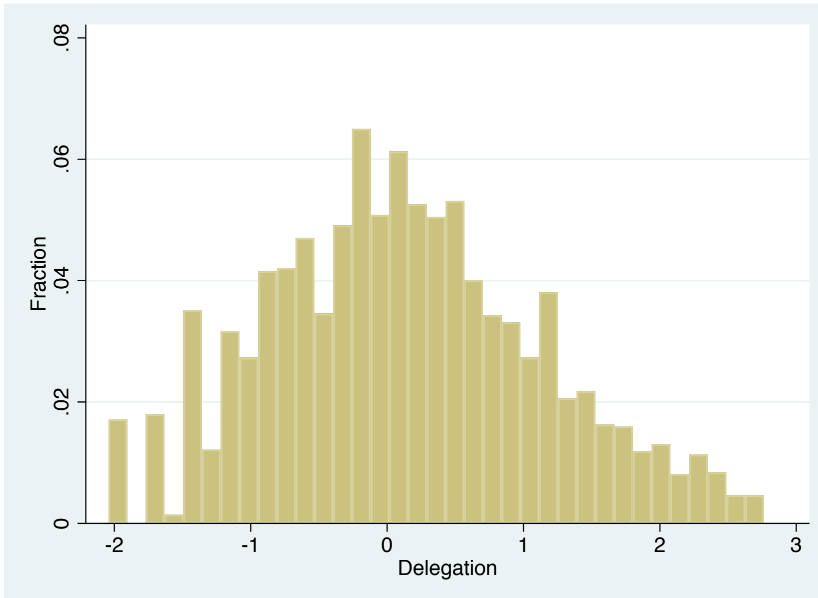
Table A-5

Descriptive Statistics of Main Variables, WorldBase Sample

	Mean	Median	Standard deviation	N. observations	N. firms
Integration$_{f,i}$	0.01	0.00	0.11	6,644,884	66,102
Employment$_f$	206.38	45.00	4,903.87	6,644,884	66,102
Age$_f$	33.56	26.00	28.98	6,644,884	66,102
IO$_{i,j}$	0.05	0.05	0.036	6,644,884	66,102
CV Productivity$_{i,c}$	3.04	1.94	4.63	6,644,884	66,102
Mean Productivity$_{i,c}$	0.0005	0.0003	0.0150	6,644,884	66,102
SD Stock Returns$_i$	0.03	0.02	0.02	533,075	65,714
Mean Stock Returns$_i$	0.04	0.02	0.25	533,075	65,714

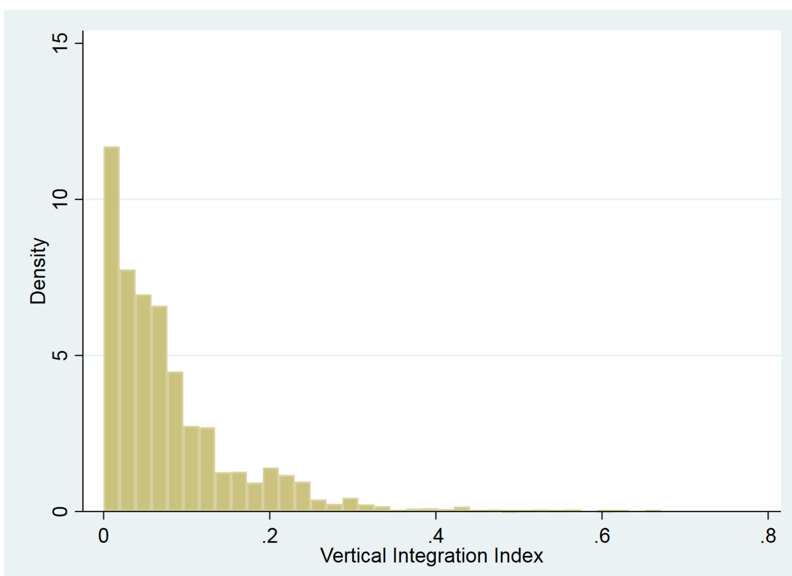
Notes: The table reports descriptive statistics of the variables used in Tables 3 and 4 (and robustness checks), based on the WorldBase sample. *Integration*$_{f,i}$ is a dummy equal to 1 if firm f integrates input i within its boundaries, measured at the 4-digit SIC level for the top 100 inputs of each industry j. *Employment*$_f$ measures the number of employees of firm f. *Age*$_f$ is the number of years since the firm was established. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij, measured at the 4-digit SIC level for the top 100 inputs of each industry j. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of supplier productivity (in billions of US Dollars). *SD Stock Returns*$_i$ is the standard deviation of stock market returns of firms in industry i, while *Mean Stock Returns*$_i$ is their mean stock market returns (in 2005).

Figure A-1: Delegation



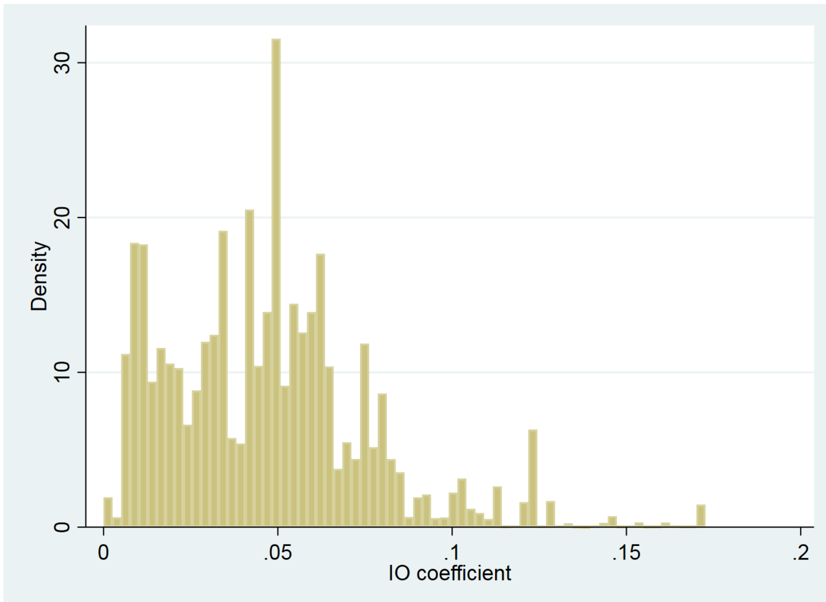
Notes: The figure shows the distribution of $Delegation_p$ in the matched sample.

Figure A-2: Vertical Integration



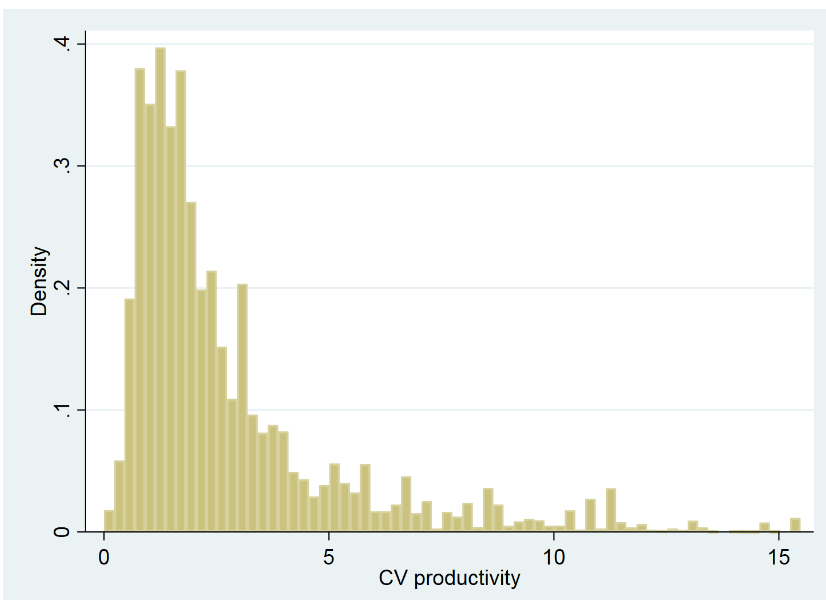
Notes: The figure shows the distribution of $Vertical\ integration_f$ in the matched sample.

Figure A-3: Input Value



Notes: The figure shows the distribution of IO_{ij} in the WorldBase sample.

Figure A-4: Input Risk



Notes: The figure shows the distribution of $CV\ Productivity_{icj}$ in the WorldBase sample.

Figure A-5: Survey on Delegation

For Questions D1, D3, and D4 any score can be given, but the scoring guide is only provided for scores of 1, 3, and 5.

Question D1: "*To hire a FULL-TIME PERMANENT SHOPFLOOR worker what agreement would your plant need from CHQ (Central Head Quarters)?*"

Probe until you can accurately score the question—for example if they say "*It is my decision, but I need sign-off from corporate HQ.*" ask "*How often would sign-off be given?*"

	Score 1	Score 3	Score 5
Scoring grid:	No authority—even for replacement hires	Requires sign-off from CHQ based on the business case. Typically agreed (i.e. about 80% or 90% of the time).	Complete authority—it is my decision entirely

Question D2: "*What is the largest CAPITAL INVESTMENT your plant could make without prior authorization from CHQ?*"

Notes: (a) Ignore form-filling

(b) Please cross check any zero response by asking "*What about buying a new computer—would that be possible?*" and then probe….

(c) Challenge any very large numbers (e.g. >$¼m in US) by asking "*To confirm your plant could spend $X on a new piece of equipment without prior clearance from CHQ?*"

(d) Use the national currency and do not omit zeros (i.e. for a U.S. firm twenty thousand dollars would be 20000).

Question D3: "*Where are decisions taken on new product introductions—at the plant, at the CHQ or both*"?

Probe until you can accurately score the question—for example if they say "*It is complex, we both play a role*," ask "*Could you talk me through the process for a recent product innovation?*"

	Score 1	Score 3	Score 5
Scoring grid:	All new product introduction decisions are taken at the CHQ	New product introductions are jointly determined by the plant and CHQ	All new product introduction decisions taken at the plant level

Question D4: "*How much of sales and marketing is carried out at the plant level (rather than at the CHQ)*"?

Probe until you can accurately score the question. Also take an average score for sales and marketing if they are taken at different levels.

	Score 1	Score 3	Score 5
Scoring grid:	None—sales and marketing is all run by CHQ	Sales and marketing decisions are split between the plant and CHQ	The plant runs all sales and marketing

Question D5: "*Is the CHQ on the site being interviewed*"?

Notes: The electronic survey, training materials and survey video footage are available on www.worldmanagementsurvey.com

A-2.2 Robustness Checks

Table A-6

Delegation and Integration (4-digits SIC Industry FE)

	(1)	(2)	(3)	(4)
Vertical Integration$_f$	0.671**	0.756***	0.792**	0.701**
	(0.276)	(0.276)	(0.314)	(0.314)
log(Employment$_f$)				0.082***
				(0.027)
log(Age$_f$)				0.037
				(0.026)
Share Employment$_p$				0.295***
				(0.088)
log(% Workforce with College Degree$_p$)				0.061***
				(0.020)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
N. observations	3,257	3,257	3,257	3,257

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 4-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A-7

Delegation and Integration (Largest 10 Countries)

	(1)	(2)	(3)	(4)
Vertical Integration$_f$	0.865***	0.867***	0.653**	0.545*
	(0.286)	(0.287)	(0.305)	(0.307)
log(Employment$_f$)				0.088***
				(0.026)
log(Age$_f$)				0.028
				(0.024)
Share Employment$_p$				0.323***
				(0.088)
log(% Workforce with College Degree$_p$)				0.055***
				(0.018)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
N. observations	2,512	2,512	2,512	2,512

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 3-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A-8

Delegation and Integration (Winsorizing Vertical Integration)

	(1)	(2)	(3)	(4)
Vertical Integration$_f$	0.899***	1.060***	0.950***	0.798**
	(0.323)	(0.316)	(0.328)	(0.330)
log(Employment$_f$)				0.086***
				(0.022)
log(Age$_f$)				0.036*
				(0.021)
Share Employment$_p$				0.320***
				(0.073)
log(% Workforce with College Degree$_p$)				0.056***
				(0.016)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
N. observations	3,444	3,444	3,444	3,444

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 3-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A-9

Delegation and Integration (Single-Plant Firms)

	(1)	(2)	(3)
Vertical Integration$_f$	0.871*	1.010**	0.851*
	(0.446)	(0.447)	(0.448)
log(Employment$_f$)			0.110***
			(0.034)
log(Age$_f$)			0.005
			(0.036)
Share Employment$_p$			0.287**
			(0.114)
log(% Workforce with College Degree$_p$)			0.072***
			(0.025)
Output FE	Yes	Yes	Yes
Country FE	No	Yes	Yes
Noise controls	Yes	Yes	Yes
N. observations	1,480	1,480	1,480

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output fixed effects are the primary activities of the single-plant firm (defined at 3-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A-10

Delegation and Input Value Shares (4-digits SIC Industry FE)

	(1)	(2)	(3)	(4)
IO_{ij}	0.932**	0.862*	1.189*	1.410**
	(0.458)	(0.464)	(0.704)	(0.683)
Vertical Integration_f				0.717**
				(0.314)
$\log(\text{Employment}_f)$				0.091***
				(0.027)
$\log(\text{Age}_f)$				0.040
				(0.026)
Share Employment_p				0.312***
				(0.088)
log(% Workforce with College Degree_p)				0.061***
				(0.020)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
N. observations	3,179	3,179	3,179	3,179

Notes: The dependent variable is $Delegation_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). IO_{ij} is the direct requirement coefficient for the sector pair ij. *Vertical* $Integration_f$ is the vertical integration index of firm f. $Employment_f$ measures the firm's employment, Age_f is the number of years since its establishment, *Share* $Employment_p$ is the plant's share of the firm's employment, and *% Workforce with College* $Degree_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 4-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A-11

Delegation and Input Value Shares (Largest 10 Countries)

	(1)	(2)	(3)	(4)
IO_{ij}	1.458***	1.463***	1.836***	1.781***
	(0.472)	(0.469)	(0.562)	(0.545)
Vertical Integration$_f$				0.379
				(0.307)
log(Employment$_f$)				0.091***
				(0.028)
log(Age$_f$)				0.047*
				(0.024)
Share Employment$_p$				0.313***
				(0.090)
log(% Workforce with College Degree$_p$)				0.056***
				(0.019)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes
N. observations	2,369	2,369	2,369	2,369

Notes: The dependent variable is *Delegation*$_{f,p,i,j,c}$, the degree of autonomy granted to plant p (with primary activity i, located in country c) by the parent firm f (with primary activity j). *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Vertical Integration*$_f$ is the vertical integration index of firm f. *Employment*$_f$ measures the firm's employment, *Age*$_f$ is the number of years since its establishment, *Share Employment*$_p$ is the plant's share of the firm's employment, and *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher. Output and input fixed effects are respectively the primary activities of the parent and of the plant (defined at 3-digit SIC). Standard errors clustered at the firm level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A-12

Integration and Input Value Shares (Matched Sample)

	(1)	(2)	(3)	(4)	(5)
IO_{ij}	0.04845*	0.06449**	0.10164***	0.10144***	0.10665***
	(0.02903)	(0.02999)	(0.01857)	(0.01842)	(0.01943)
log(Employment$_f$)				0.00570***	
				(0.00034)	
log(Age$_f$)				-0.00023	
				(0.00039)	
Input FE	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	-
Firm FE	No	No	No	No	Yes
N. observations	292,744	292,744	292,744	292,744	292,744

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 5, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input-output level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-13

Integration and Input Value Shares (4-digits SIC Industry FE)

	(1)	(2)	(3)	(4)	(5)
IO_{ij}	0.14712***	0.14719***	0.17564***	0.17545***	0.19912***
	(0.02050)	(0.02054)	(0.02484)	(0.02485)	(0.02827)
log(Employment$_f$)				0.00136***	
				(0.00011)	
log(Age$_f$)				0.00017**	
				(0.00008)	
Input FE	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	-
Firm FE	No	No	No	No	Yes
N. observations	7,042,966	7,042,966	7,042,966	7,042,966	7,042,966

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. IO_{ij} is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 4-digit SIC. In the specification in column 5, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input-output level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-14

Integration and Input Value Shares (Largest 10 Countries)

	(1)	(2)	(3)	(4)	(5)
IO_{ij}	0.14536***	0.14551***	0.17133***	0.17117***	0.19351***
	(0.02075)	(0.02077)	(0.02486)	(0.02486)	(0.02806)
$\log(\text{Employment}_f)$				0.00137***	
				(0.00011)	
$\log(\text{Age}_f)$				0.00016*	
				(0.00008)	
Input FE	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	-
Firm FE	No	No	No	No	Yes
N. observations	6,776,732	6,776,732	6,776,732	6,776,732	6,776,732

Notes: The dependent variable is $\textit{Integration}_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. IO_{ij} is the direct requirement coefficient for the sector pair ij. $\textit{Employment}_f$ measures firm employment, and $\textit{Age}_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 5, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input-output level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-15

Integration and Input Value Shares
(Single-Plant Firms)

	(1)	(2)	(3)	(4)	(5)
IO_{ij}	0.13929***	0.13960***	0.16427***	0.16426***	0.18487***
	(0.02037)	(0.02041)	(0.02468)	(0.02468)	(0.02780)
$\log(\text{Employment}_f)$				0.00008	
				(0.00007)	
$\log(\text{Age}_f)$				0.00015*	
				(0.00009)	
Input FE	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	-
Firm FE	No	No	No	No	Yes
N. observations	6,361,633	6,361,633	6,361,633	6,361,633	6,361,633

Notes: The dependent variable is $Integration_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. IO_{ij} is the direct requirement coefficient for the sector pair ij. $Employment_f$ measures firm employment, and Age_f is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 5, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input-output level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-16

Integration and Riskiness of Input industries (Matched Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
CV Productivity$_{i,c}$	0.00062***	0.00062***	0.00062***	0.00057***	0.00056***	0.00056***
	(0.00015)	(0.00015)	(0.00015)	(0.00014)	(0.00014)	(0.00015)
Mean Productivity$_{i,c}$	-0.00314**	-0.00314**	-0.00316**	-0.00246*	-0.00264**	-0.00245**
	(0.00145)	(0.00145)	(0.00145)	(0.00133)	(0.00124)	(0.00124)
IO$_{ij}$			0.07702***	0.12626***	0.12829***	0.13752***
			(0.01188)	(0.01495)	(0.01485)	(0.01592)
log(Employment$_f$)					0.00672***	
					(0.00040)	
log(1+ Age$_f$)					-0.00009	
					(0.00028)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	249,471	249,471	249,471	249,471	249,471	249,471

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of supplier productivity. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 5, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-17

Integration and Riskiness of Input industries (4-digits SIC Industry FE)

	(1)	(2)	(3)	(4)	(5)	(6)
CV Productivity$_{i,c}$	0.00038***	0.00048***	0.00047***	0.00047***	0.00047***	0.00047***
	(0.00008)	(0.00007)	(0.00007)	(0.00007)	(0.00007)	(0.00007)
Mean Productivity$_{i,c}$	-0.00760	-0.01099	-0.01078	-0.01021	-0.01018	-0.01017
	(0.00790)	(0.00991)	(0.00979)	(0.00968)	(0.00968)	(0.00965)
IO$_{ij}$			0.15325***	0.21249***	0.21252***	0.21173***
			(0.01299)	(0.01534)	(0.01534)	(0.01534)
log(Employment$_f$)					0.00144***	
					(0.00012)	
log(1+ Age$_f$)					0.00015	
					(0.00010)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	6,644,884	6,644,884	6,644,884	6,644,884	6,644,884	6,644,884

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of supplier productivity. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 4-digit SIC. In the specification in column 6, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-18

Integration and Riskiness of Input industries (Largest 10 Countries)

	(1)	(2)	(3)	(4)	(5)	(6)
CV Productivity$_{i,c}$	0.00064***	0.00075***	0.00074***	0.00074***	0.00074***	0.00073***
	(0.00009)	(0.00010)	(0.00010)	(0.00010)	(0.00010)	(0.00010)
Mean Productivity$_{i,c}$	-0.01285	-0.01648	-0.01623	-0.01583	-0.01582	-0.01573
	(0.01253)	(0.01516)	(0.01498)	(0.01496)	(0.01495)	(0.01482)
IO$_{ij}$			0.14988***	0.17844***	0.17829***	0.20249***
			(0.01396)	(0.01505)	(0.01505)	(0.01676)
log(Employment$_f$)					0.00146***	
					(0.00013)	
log(1+ Age$_f$)					0.00017	
					(0.00011)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	6,430,959	6,430,959	6,430,959	6,430,959	6,430,959	6,430,959

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of supplier productivity. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 6, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-19

Integration and Riskiness of Input industries (Single-Plant Firms)

	(1)	(2)	(3)	(4)	(5)	(6)
CV Productivity$_{i,c}$	0.00060***	0.00070***	0.00070***	0.00069***	0.00069***	0.00069***
	(0.00008)	(0.00010)	(0.00009)	(0.00009)	(0.00009)	(0.00009)
Mean Productivity$_{i,c}$	-0.01084	-0.01385	-0.01360	-0.01320	-0.01321	-0.01313
	(0.01139)	(0.01353)	(0.01336)	(0.01335)	(0.01335)	(0.01324)
IO$_{ij}$			0.14390***	0.17148***	0.17147***	0.19426***
			(0.01365)	(0.01475)	(0.01475)	(0.01646)
log(Employment$_f$)					0.00010	
					(0.00010)	
log(1+ Age$_f$)					0.00014	
					(0.00012)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	6,027,632	6,027,632	6,027,632	6,027,632	6,027,632	6,027,632

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of supplier productivity. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 6, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-20

Integration and Riskiness of Input industries (Winsorizing Supplier Productivity)

	(1)	(2)	(3)	(4)	(5)	(6)
CV Productivity$_{i,c}$	0.00646***	0.00656***	0.00640***	0.00639***	0.00640***	0.00645***
	(0.00178)	(0.00182)	(0.00178)	(0.00179)	(0.00178)	(0.00182)
Mean Productivity$_{i,c}$	0.00134	0.00173	0.00140	0.00172	0.00174	0.00173
	(0.00419)	(0.00421)	(0.00420)	(0.00424)	(0.00426)	(0.00424)
IO$_{ij}$			0.15107***	0.18158***	0.18141***	0.20665***
			(0.01363)	(0.01474)	(0.01474)	(0.01649)
log(Employment$_f$)					0.00147***	
					(0.00013)	
log(1+ Age$_f$)					0.00018	
					(0.00011)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	6,532,715	6,532,715	6,532,715	6,532,715	6,532,715	6,532,715

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of supplier productivity. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 6, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-21

Integration and Riskiness of Input industries (50+ Suppliers per Input Sector)

	(1)	(2)	(3)	(4)	(5)	(6)
CV Productivity$_{i,c}$	0.00054***	0.00069***	0.00069***	0.00068***	0.00068***	0.00067***
	(0.00009)	(0.00009)	(0.00009)	(0.00009)	(0.00009)	(0.00009)
Mean Productivity$_{i,c}$	-0.00964	-0.01414	-0.01387	-0.01326	-0.01324	-0.01304
	(0.01025)	(0.01340)	(0.01323)	(0.01317)	(0.01316)	(0.01292)
IO$_{ij}$			0.16020***	0.19779***	0.19767***	0.22927***
			(0.01628)	(0.01784)	(0.01784)	(0.02034)
log(Employment$_f$)					0.00164***	
					(0.00016)	
log(1+ Age$_f$)					0.00020	
					(0.00012)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	5,484,936	5,484,936	5,484,936	5,484,936	5,484,936	5,484,936

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity of the independent suppliers in input industry i located in country c, while *Mean Productivity*$_{i,c}$ is the mean of supplier productivity. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 6, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-22

Integration and Riskiness of Input industries (Alternative Risk Measure)

	(1)	(2)	(3)	(4)	(5)	(6)
SD Stock Returns$_i$	0.50514**	0.49893**	0.54178**	0.59859***	0.59893***	0.64823***
	(0.22636)	(0.22459)	(0.23294)	(0.20984)	(0.20992)	(0.22362)
Mean Stock Returns$_i$	-0.02182	-0.02330	-0.01907	-0.02199	-0.02166	-0.02232
	(0.04118)	(0.04103)	(0.03964)	(0.03817)	(0.03818)	(0.04017)
IO$_{ij}$			0.29371***	0.29569***	0.29536***	0.31112***
			(0.03619)	(0.03427)	(0.03428)	(0.03742)
log(Employment$_f$)					0.00540***	
					(0.00075)	
log(1+ Age$_f$)					0.00243***	
					(0.00055)	
Input FE	Yes	Yes	Yes	Yes	Yes	Yes
Output FE	No	No	No	Yes	Yes	-
Country FE	No	Yes	Yes	Yes	Yes	-
Firm FE	No	No	No	No	No	Yes
N. observations	533,075	533,075	533,075	533,075	533,075	531,726

Notes: The dependent variable is *Integration*$_{f,j,i,c}$, a dummy equal to 1 if firm f producing final product j (defined at 4-digit SIC) and located in country c integrates input i (defined at 4-digit SIC) within its boundaries. *SD Stock Returns*$_i$ is the standard deviation of stock market returns of firms in industry i. *Mean Stock Returns*$_i$ is the mean of stock market returns of firms in industry i. *IO*$_{ij}$ is the direct requirement coefficient for the sector pair ij. *Employment*$_f$ measures firm employment, and *Age*$_f$ is the number of years since the firm's establishment. Output and input fixed effects defined at 3-digit SIC. In the specification in column 56, country and output-industry fixed effects are absorbed by the firm fixed effects (each firm f is associated to one location and one primary activity). Standard errors clustered at the input level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.