Come Together: Firm Boundaries and Delegation*

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

November 2022

Abstract

We jointly study firm boundaries and the allocation of decision rights within them by confronting an incomplete-contracts model with data on vertical integration and delegation for thousands of firms around the world. Integration has an option value: it confers authority to delegate or centralize decision rights, depending on who can best solve problems that arise in the course of an uncertain production process. In line with the model's predictions, we find that firms are more likely to integrate suppliers that produce more valuable inputs and operate in industries with more dispersed productivity, and that firms delegate more decisions to integrated suppliers that produce more valuable inputs and operate in more productive industries.
JEL classifications: D2, L2.
Keywords: Integration, delegation, uncertainty, real options, supply assurance.

*We are extremely grateful for comments from Philippe Aghion, Ricardo Alonso, Jacques Crémer, Wouter Dessein, Luis Garicano, Bob Gibbons, Oliver Hart, Louis Kaplow, Francine Lafontaine, Rocco Macchiavello, Ameet Morjaria, Marian Moszoro, Jawwad Noor, Andrea Prat, Mike Powell, Veronica Rappoport, Kathy Spier, 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, Universitat Autònoma de Barcelona, LSE, King's College London, Southampton, HSE Moscow, and the Harvard Law, Economics and Organization Research Seminar. 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 for Legros and ERC Grant Agreement N. 834253 for Conconi).

1 Introduction

Why do firms integrate suppliers? One benefit, of course, is direct control of production: in a world replete with contracting frictions, ownership can facilitate top management's ability to impose productivity-enhancing decisions, such as re-tooling, human capital investment, or standards conformity, that its suppliers otherwise would not make.

Equally important, 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. Over the course of a lengthy and uncertain production process, different types of problems are bound to arise, and top managers may wish to re-allocate decision rights among themselves and their suppliers to solve them according to their relative expertise. Within a firm's boundaries, management can do this relatively seamlessly, choosing to delegate production decisions to its integrated suppliers or to centralize those decisions, depending on which problem arises. This option is hardly available outside the firm, where suppliers retain those control-allocation rights among their prerogatives of ownership, and are less likely to exercise them in management's interests. 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.

Despite their evident connection, the interplay between vertical integration and delegation has scarcely been explored. This paper brings these organizational design decisions together, both theoretically and empirically. It first develops a model to jointly study vertical integration and delegation. It then assesses the evidence in light of the model, combining information on vertical integration and delegation for firms in multiple countries and industries.

Our analysis is founded on a well-known conceptual distinction between outsourcing (non-integration) and delegation. While both are instruments for re-assigning control rights, the first is defined by formal titles of ownership and requires legal intervention to reverse, via an asset purchase. By contrast, delegation is a non-contractible act of relinquishing control that can in principle be revoked at will by owner fiat.[1] One contribution of this paper is to make this distinction operational by generating testable predictions about the differential

[1] See for instance Williamson (1991) and Aghion and Tirole (1997). Indeed, non-integration has legal force, while delegation does not. The law not only regulates and registers asset sales, it frequently adjudicates disputes between parties who hold separate titles. Once integrated, the parties largely forego appeal to the law in many of their disputes.

responses of delegation and outsourcing to changes in the firm's environment.

In our theoretical model, the primary organizational design problem is the allocation of control over production decisions. We consider a world in which such allocations are not contractible, but are among the prerogatives of ownership. A headquarters (HQ) and one of its input suppliers decide whether to integrate before starting production and learning which of many potential "problems" — new designs, quality control or compatibility issues, supply chain disruptions, new hiring or capital equipment requirements — will have to be tackled. In the course of addressing a problem, a non-contractible production decision, modeled as a choice among horizontally differentiated standards, will also have to be made. All decisions work equally well for generating the input, but preferences over them diverge between the parties, possibly due to differences in training, background, corporate culture, or managerial vision.[2] These private costs are the source of interest conflict in the model. Solving problems and choosing standards cannot be separated: both are in the hands of whoever controls the production process.

HQ and the supplier differ along two dimensions beyond their conflicting preferences over production standards. First, distinct problems play differently to their comparative advantages: some will be easier to solve for the supplier, others for HQ. Second, contracting frictions imply that the supplier has smaller residual stakes in the enterprise profit than HQ.

Under integration, HQ owns the primary asset the supplier uses. After a problem arises, she can therefore decide whether to retain control (centralize), solving the problem and implementing the production decision herself; or to offer control to the supplier (delegate), who then chooses the production standard and addresses the problem. HQ's decision to delegate comes down to trading off a supplier's sometimes superior problem solving ability against the private costs he will impose on her through his own choice of production standard. In so doing, she does not internalize the supplier's private cost incurred because of her own choice of production standard. By contrast, with non-integration, the supplier owns the asset, and therefore the option to delegate to HQ. However, given his small residual stakes, and his own neglect of HQ's private costs, he never has an interest in doing so. When the integration decision is made before production begins, the parties are deciding between two imperfect control allocation mechanisms: one that eventually offers the supplier too little control (integration), and one that offers him too much (non-integration).

The model delivers several testable predictions about input value and uncertainty as ex-

[2]For instance, HQ favors redundant features and user-friendliness in the software bundled with the game consoles it sells, while the software engineers prefer elegant design and low maintenance. Or each party bristles at the personality types of the otherwise capable candidates it finds to fill a post. This "standards" approach to modeling tensions about the way operations should be carried out within firms has been employed by Van den Steen (2005, 2010) and Hart and Holmström (2010), among others.

ogenous determinants of integration and delegation decisions. First, the greater the value contribution of an input to the enterprise, the more likely HQ should be to delegate to the integrated supplier who provides it: with greater value, there is a larger opportunity cost for HQ to indulge in the private benefit of keeping control and providing a mediocre problem solution rather than letting the supplier provide a superior one. Second, a supplier of a more valuable input should be more likely to be integrated. This result derives from the difference in the parties' delegation behavior: integration becomes relatively more efficient as value increases, because HQ is more likely to allocate control to the party with the better problem solution; by contrast, since a non-integrated supplier never delegates, his relative efficiency never increases. Third, first-order shifts in the supplier's problem-solving ability should lead to more delegation. Finally, under some empirically mild conditions, HQ should be more likely to integrate inputs whose production carries more problem risk. The logic for this result is similar to that concerning the role of uncertainty in option theory (Dixit and Pindyck, 1994), since the flexibility to reassign control in the face of new information turns integration into a real option.

To assess the evidence, we combine information from the WorldBase dataset by Dun & Bradstreet and the World Management Survey (WMS), which allows us to measure integration and delegation choices for firms in 20 countries. A first look at these data reveals a positive correlation between delegation and integration: more vertically integrated firms tend to grant more autonomy to their plant managers.[3] This result underscores the conceptual distinction between delegation and outsourcing (non-integration): while both involve shifting control from an HQ to a supplier, they are clearly not interchangeable, since they are negatively correlated in the data.

In our empirical analysis, we assess the validity of the predictions generated by the model regarding the effects of input value and uncertainty. We measure vertical integration at the firm-input level, coding which inputs the firm integrates within its boundaries. This reflects the choices faced by firms in our theoretical model, in which HQ decides whether or not to integrate each input supplier. Our delegation measure captures the degree of autonomy granted by a firm to the managers of its integrated suppliers when faced with various production decisions (e.g., hiring a new employee, introducing a new product, making a capital investment).

To test the model's predictions about the role of input value, we exploit exogenous vari-

[3] See Section A-2.2 of the Empirical Appendix. A simple application of our theoretical model's results offer a rationale for the positive correlation between vertical integration and autonomy: both could be driven by a common (unobserved) driver in the form of enterprise value (e.g., HQ productivity or final good price). Higher value enterprises should integrate more inputs and grant more authority to their integrated suppliers.

ation in input-output (IO) coefficients using data from the US Bureau of Economic Analysis (BEA). IO coefficients capture exactly the role of input value in our model, since they measure the dollar value of each input used in the production of an output. To capture exogenous variation in uncertainty, we use the mean and dispersion of the distribution of productivity of independent suppliers in an industry.

The evidence confirms the role of input value for integration and delegation, in line with the predictions of our model. We find that final good producers are indeed more likely to integrate suppliers of more valuable inputs, as proxied by input-output coefficients. Among integrated suppliers, more autonomy is granted to those producing more valuable inputs.

Also consistent with the theoretical model, we find that delegation increases with the mean productivity in the input industry, while riskiness of the input industry has no significant effect on delegation. Moreover, the probability that firms integrate a particular input increases with the riskiness of the input industry. This finding is reminiscent of the literature on supply assurance motives for integration, as discussed in the next section.

The empirical results of our baseline regressions hold up in a battery of robustness checks, such as restricting the analysis to different samples (e.g., only multi-plant firms), including input-industry and output-industry fixed effects, and additional controls (e.g., firm and plant characteristics, as well as measures of contracting frictions). In the integration regressions, we can also include firm fixed effects, exploiting variation within firms across input industries (in the riskiness of the industry or the value of the input) to identify the role of input value and uncertainty.

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 data and variables used in the empirical analysis. Section 5 presents the empirical results. Section 6 discusses alternative theoretical explanations for these results. Section 7 concludes.

2 Related Literature

Our work is mainly related to two streams of literature. First, we build on the vast literature on the determinants of firm boundaries. Many theoretical studies have looked the technological/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); another strand has focused on market determinants (e.g., McLaren, 2000; Grossman and Helpman, 2002; Legros and Newman, 2013; Conconi *et al.*, 2012). A number of papers study legal/institutional determinants of integration (e.g., Acemoglu *et al.*, 2009, Macchiavello, 2012). 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 *et al.*, 2002).

The empirical literature on firm boundaries is also very large, and growing; Lafontaine and Slade (2007, 2013) provide excellent overviews. Our empirical analysis of ownership is closest to some recent papers emphasizing supply assurance (e.g., Macchiavello and Miquel-Florensa, 2017) and the relative technological importance of inputs (e.g., Berlingieri *et al.*, 2021; Hansman *et al.*, 2020) as motives for integration.

Second, our paper contributes to to the literature on delegation within firms. Theoretical studies in this literature include, among others, Holmström (1984), Aghion and Tirole (1997), Dessein (2002), Hart and Moore (2005), Alonso *et al.* (2008), Alonso and Matouschek (2008), and Marin and Verdier (2008).[4] Our approach to modeling delegation is related to work on the design of knowledge hierarchies (Garicano, 2000) and referrals (Garicano and Santos, 2004) insofar as we are concerned with allocation of decision making among the organization's members according to the expertise at solving particular production problems.[5] On the empirical side, contributions include Acemoglu *et al.* (2007), Guadalupe and Wulf (2010), Bloom *et al.* (2012), McElheran (2014), Graham *et al.* (2015), Wu (2017), and Katayama *et al.* (2018).

In this paper, we bring these organizational choices together. 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), and Friebel and Raith (2010). Although some formal studies have emphasized the conceptual difference between integration and delegation (Baker *et al.*, 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.

Our paper also contributes to the literature on supply assurance motives for integration (e.g., Carlton, 1979; Bolton and Whinston, 1993). Those papers tend to focus on demand uncertainty and the ability of non-integrated suppliers to sell their inputs to other buyers. The

[4]Much (but by no means all) of this literature views delegation as a means of achieving better outcomes by assigning decision rights to (ex-ante) better informed parties; often this helps to incentivize delegates to become more informed in the first place. In our simplified model of delegation, the assignment of control is instead a *response* to (symmetric) information: the (ex-post) sufficiently more capable (or possibly less time constrained) party gets it. The two approaches are complementary — the production decisions could involve the acquisition of further information — and our approach is mainly for tractability.

[5]In those papers the allocation of control is decided contingently through contracts rather than managerial authority, and they abstract from incentive problems, which play a key role in the comparative statics of our model. Moreover, they endogenize knowledge acquisition, while we take it as given. A more minor distinction is that those papers treat the quality of problem solutions as discrete rather than continuous.

supply assurance in our model derives from uncertainty about problems that may arise during the production process and the ability of the firm's members to solve them.

Our model refines the "value principle" that emphasizes the role of pecuniary variables such as profitability and prices in organizational design, and extends it beyond the integration context where it has already been applied (Legros and Newman, 2013, 2017; Alfaro *et al.*, 2016). While option value has been considered in some studies of integration (e.g., Bradley *et al.*, 1988; Grullon *et al.*, 2012; Bena and Li, 2014), our analysis provides a novel mechanism, based on the notion of asset ownership as a bundle of control rights, rather than just a claim on return streams. Moreover, input value affects not only firm boundaries, but also the internal control structure, through its effect on centralization/delegation decisions.

3 The Model

We present a simple theoretical model in which some modeling assumptions are aimed at delivering predictions that we can test with our data rather than at theoretical generality. The interested reader is referred to Legros and Newman (2022) for a broader treatment.

3.1 Environment

Consider a production process in which a final good j is produced with a I inputs indexed by i. An enterprise is composed of an HQ, who produces the final good, and I suppliers S_i, each comprising a manager (who will also be referred to as a supplier) and an indivisible productive asset that can be owned by the supplying manager or by HQ. For empirical purposes, we shall think of each supplier S_i as representative of its industry; there may be differences across industries. All players are risk neutral.

The expected value of a good produced by the enterprise can be written as

$$A_f P_j \sum_{i=1}^{I} \pi_{i,j}\, \mathbb{E}V_{i,j}. \tag{1}$$

A_f represent the exogenous productivity of the HQ of the enterprise, capturing her entrepreneurial competence or the profitability of her product. P_j is the price of the final good and $\pi_{i,j}\, \mathbb{E}V_{i,j}$ is the contribution of each supplier i. The latter is decomposable into an exogenous, technologically dictated value share $\pi_{i,j}$ and an endogenous quantity $\mathbb{E}V_{i,j}$ that will be shaped by production and organizational decisions made by HQ and supplier S_i. For now, we focus on the relationship between HQ and a typical supplier, suppressing the index notation and normalizing $A_f = P_j = 1$. Discussion of the effects of these parameters will be taken up in Section 3.2.2.

There are two types of uncertainty in this model. The first is the familiar quantity or quality risk, and related information asymmetries, that cloud inference about underlying decision variables, making them non-contractible. The second is more specific to the organizational design concerns in this paper, which is uncertainty over which specific problems will arise in the course of production — and therefore who is best suited to to tackle them — that drives the delegation decision and is the source of the option value of integration.

3.1.1 Technology, Timing and Information

We model the production process in the simplest possible way: there is a binary decision $d \in \{h, s\}$ over how to use the asset that can be carried out by either party, depending on the assignment of control. The choices h, s are equally effective from the viewpoint of production efficiency, but impose different private costs on the two parties, e.g., resulting from different corporate cultures or different preferences over design standards. Choice h is HQ's preferred option, costing her zero, but imposing a positive cost c^S per unit of produced input on the supplier; similarly, option s costs the supplier nothing but imposes cost $c^H > 0$ per unit on HQ. To ensure gains from trade, the cost parameters satisfy $\pi > \max\{c^S, c^H\}$ (since control is ultimately allocated to one party or the other, only one will bear a cost).

The sequence of events is as follows. First S and HQ agree on who owns the primary productive asset S uses, and on a (positive or negative) side-payment T from HQ to S to ensure participation. At this point, as is standard in models of ownership, the two parties are "locked in" to each other for the duration, because some sunk investment (physical, human capital, enduring the legal process of registering title, or simply finding each other) has to be made to get production going. After production commences, a "problem" (or opportunity) arises; this is represented by the random variable $y \geq 0$ with c.d.f $G(y)$ and density $g(y)$. The problem can be solved by either party, but to do so, the party needs to be given control of the production decision d. At this point, the parties agree on a payment $t(y)$ from HQ to S, due on receipt of the input batch, which may depend on the problem that arose. Next, the owner of the asset decides who will solve the problem/make the production decision, and the assignee then chooses d. Finally, the inputs are delivered, HQ privately observes the number of usable units in the batch, and appropriates the realized value less the payment.

How well the problem is solved depends not only on the problem itself, but also on the party that solves it: the efficacy of the solution under HQ-control is a constant $z > 0$, while under S-control it is y (thus, we label problems by the competence with which the supplier would solve them). Efficacy here is measured in terms of the *expected* number of usable inputs that will ultimately be produced; the actual number depends partly on random factors outside the parties' control, and only HQ ultimately observes how many units are usable.[6]

HQ is a generalist, equally adept at addressing all problems (hence the constant efficacy z; this can be relaxed) The supplier is a specialist, better than HQ at solving some of problems, worse at others ($0 < G(z) < 1$). It is convenient, though not strictly necessary, that HQ is on average better at solving problems than the supplier:

$$\mathbb{E}y \leq z. \tag{A1}$$

HQ's competence z, as well as the supplier's effectiveness distribution $G(y)$ are common knowledge: each party knows exactly how well the other would perform in any given situation y; there is only uncertainty as to which situation will arise.

3.1.2 Contracting

Contracting is limited to the transfers of ownership and side payments to facilitate those, and to further monetary transfers payable upon delivery of a batch of inputs. What is not contractible are the decision d (making control relevant), and the number of inputs that HQ finds usable (thereby eliminating the supplier's stake in the continuation value of the relationship). The identity of the actual decision makers/problem solvers is also not contractible. This ensures that allocating control must be informal, determined by the owner, who could always overturn a decision that was supposed to be under the control of the other party, yet deny that in court. Even if the court knows the state y (or infers it from the payment $t(y)$), that is not enough to establish who made a decision. Finally, as HQ's aggregate profit depends on many different inputs, profit sharing would add little in terms of meaningful incentives, so we abstract from those. For empirical purposes, we take the payment to be equal to the market price of the input p multiplied by the expected quantity. Thus, $t(y)$ is equal to py if the anticipated party in control is the supplier, and pz if it is HQ.

When negotiating ownership, the supplier and HQ take into account the private costs that will ensue. We assume that both parties have sufficient cash at the time of ownership contracting to make any side payments needed to settle the distribution of surplus required to strike a deal. Thus we will be considering "efficient" ownership choices in what follows.

[6] The private observation prevents the payment from depending on realized rather than expected quantity; thus only HQ has an interest in the continuation value once $t(y)$ is negotiated. Private observation could be interpreted as HQ's subjective assessment of quality, or the compatibility of the input with HQ's production technology or with the inputs of her other suppliers; none of these are likely to be contractible – if a part looks wrong, or doesn't work, it would be too costly for a court to determine fault and compensation.

Note that in contrast to ownership, the availability of side payments does not imply that an efficient control allocation can be achieved, because of the non-verifiability of the decision maker's identity.

By way of summary, the sequence of events is illustrated in Figure 1.

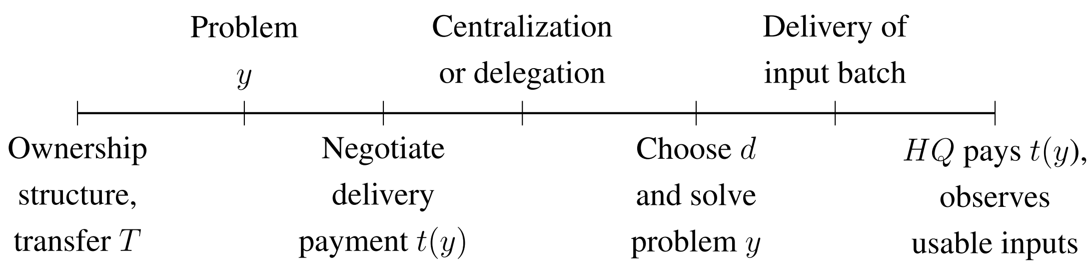


Figure 1: Timing of Events

3.1.3 A Benchmark: Efficient Allocation of Control

A surplus maximizing control allocation would have S deciding whenever the total surplus is greater under S control than under HQ control:

$$y(\pi - c^H) \geq z(\pi - c^S),$$

or

$$y \geq y^*(\pi) := z\frac{\pi - c^S}{\pi - c^H}. \tag{2}$$

In a world where the identity of the deciding party is contractible, ownership and delegation are irrelevant, at least as far as the control allocation is concerned: the optimal allocation can be achieved by contract, and need not be relegated to those blunt mechanisms. A contract could directly implement a control allocation by specifying who decides as a function of y, independent of the compensation structure or of who owns the asset.[7]

3.2 Input Value and Organizational Design

We now return to the world in which rather than being contractible, control is instead assigned by the owner of the asset. First we consider delegation under each ownership structure, and then, using results of that analysis, discuss the determination of ownership.

[7]Having ownership determined after y is realized would also yield efficient control: if $y \geq y^*$, let S own; if $y \leq y^*$, let HQ own. We have already ruled this out, as it may forego many of the (here un-modeled) benefits of ownership that pervade the literature, specifically providing incentives for (or direct implementation of) sunk relationship-specific investments.

3.2.1 Delegation Under Integration and Non-integration

Because HQ privately observes how many inputs actually are usable (generate the positive value π), she is the residual claimant of value. By contrast, S receives a transfer that cannot depend on realized value, only on its expectation y.

It follows that S will never delegate when he owns the asset (i.e. under non-integration): centralizing pays him $t(y)$, while delegating yields $t(y) - zc^S$. On the other hand, if HQ owns the asset (i.e. under integration), she prefers to delegate when $y(\pi - c^H) \geq z\pi$, in other words when

$$y \geq y^H(\pi) := z\frac{\pi}{\pi - c^H}.$$

Note that like S, HQ delegates less often than in the first-best, since $y^H(\pi) > y^*(\pi)$. But $y^H(\pi)$ is decreasing in π, so the probability $1 - G(y^H(\pi))$ that HQ delegates increases with π. As the value of trade increases, its weight in HQ's calculations grows vis-à-vis her private costs, and she is more willing to delegate, approaching the efficient control allocation as π grows large.[8] The literature has frequently cited consonance in parties' preferences as a critical determinant of delegation (e.g., Aghion and Tirole, 1997; Dessein, 2002). Here, higher transaction values endogenously increase consonance. We can thus state the following:

Result 1. *(Value principle for delegation): HQ's propensity to delegate increases with* π*.*

Observe that in contrast to this strong prediction about the relationship between supplier control and input value under integration, *efficient* control allocations need not display such a property. Instead, the response would be sensitive to the private cost parameters: from (2), efficient supplier control increases with π if $c^S < c^H$, and decreases with π if $c^S > c^H$.

3.2.2 Comparing Ownership Structures

Since S always retains control when he owns the asset, the total expected surplus under non-integration (i.e. supplier ownership) is simply

$$W^S(\pi) := \mathbb{E}y(\pi - c^H).$$

The total expected surplus under integration (HQ ownership) is

$$W^H(\pi) := G(y^H(\pi))z(\pi - c^S) + (\pi - c^H)\int_{y^H(\pi)}^{\infty} yg(y)dy. \quad (3)$$

[8]This result does not depend on the constancy of z, which could instead be random variable $z(y)$, where $z(y) > y$ for some y and $z(y) < y$ for others. Then whenever $y > \frac{\pi}{\pi - c^H}z(y)$, $y > \frac{\pi'}{\pi' - c^H}z(y)$ for $\pi' > \pi$, so the set of problems delegated to the supplier is non-decreasing in π.

That is, $G(y^H(\pi))$ of the time HQ centralizes, choosing $d = h$, and generating the expected usable output z, which has net per-unit value $(\pi - c^S)$. The rest of the time, she bears the cost c^H in proportion to the number of usable units expected under S-control, where $d = s$.

Assuming efficient allocation of ownership, integration occurs whenever $W^H(\pi) - W^S(\pi)$ is non-negative, which can be written

$$\int_0^{y^H(\pi)} [z(\pi - c^S) - y(\pi - c^H)]g(y)dy \geq 0. \tag{4}$$

As π is embedded in this condition, it is natural to ask what the value principle implies for the propensity of HQ to integrate her suppliers (Legros and Newman, 2013; Alfaro *et al.*, 2016).[9] Supposing then that a supplier whose input is worth π is integrated, what can be said about a supplier with a more valuable input worth π'? Notice that the left hand side of expression (4) is increasing in π.[10] Thus, the value difference $W^H(\pi) - W^S(\pi)$ has the single-crossing property, so we can state the following monotonicity result:

Result 2. *(Value principle for integration): If a supplier who generates value π is integrated, so is a supplier who generates value $\pi' > \pi$.*

The result stems from the differences in HQ's and S's delegation behavior under their respective ownership. Compared to S-ownership, HQ-ownership becomes relatively more efficient at allocating control as π increases: $y^H(\pi) - y^*(\pi) = \frac{zc^S}{\pi - c^H}$ decreases with π; meanwhile S always chooses to retain control, so the corresponding efficiency gain is smaller or negative.[11]

Looking back at expression (1) for the expected value of an enterprise, Results 1 and 2 imply that the suppliers with higher input value are the ones most likely to be integrated and, among those that are, to have higher degrees of delegation.

Other sources of variation in enterprise value would have similar effects as input value. In particular, we could allow value to vary across firms within an industry, e.g., if their HQs have different productivity (A_f in equation (1), which we have so far set equal to 1). Then $A_f\pi$ would replace π in all the formal expressions above. The effects of variation in A_f on integration and delegation choices would be identical to those of input value: both the propensity

[9]We cannot simply lean on those papers' results because the mechanisms that allow parties to benefit from integration there differ from the control-over-control mechanism at work here.

[10]Its π-derivative is $g(y^H)y^H_\pi z(\pi - c^S) + G(y^H)z - y^H g(y^H)y^H_\pi(\pi - c^H) - \int_0^{y^H} yg(y)dy$, which using $y^H = z\pi/(\pi - c^H)$, reduces to $[zG(y^H) - \int_0^{y^H} yg(y)dy] - g(y^H)y^H_\pi zc^S$. From (A1), the bracketed term is non-negative ($z \geq \mathbb{E}y \geq \int_0^{y^H} yg(y)dy/G(y^H)$), and the last term is positive because the derivative $y^H_\pi < 0$.

[11]S ownership may become absolutely more efficient with increasing π through its effect on $y^*(\pi)$, but only if $c^S < c^H$, in which case, under (A1), HQ integrates for all $\pi > c^S$ anyway. Dispensing with (A1) may instead generate a U-shaped relation between value and integration, but only if c_H is larger than c_S and the supplier's average competence is sufficiently higher than HQ's.

to integrate suppliers within the firm boundaries and the propensity to delegate decisions to integrated suppliers would increase in HQ's productivity. Similarly, if the output price P_j varies across industries, firms in high-value industries would have greater propensities than firms in low-value industries both to integrate their suppliers and to delegate to them.[12]

These simple extensions of the model imply that there should be *positive* covariation between the degree to which firms are vertically integrated and the extent to which they delegate, in line with what we observe in our data (see Figure A-7 in the Appendix): more valuable firms (e.g., those that have more productive HQs) should integrate more suppliers and grant more autonomy to them.

Taken together, Results 1 and 2 help to underscore and operationalize the conceptual distinction between the informal organizational decision to delegate and the formal one of outsourcing (non-integration). Although both move decision-making away from the "center," the degree of delegation increases, while the propensity to outsource decreases, in response to changes in the same variable, namely input (or enterprise) value.

3.3 Uncertainty and Organizational Design

As has been noted elsewhere (Hart and Holmström, 2010), and as the present model illustrates, uncertainty plays an essential role for delegation as well as for ownership. In particular, if the parties could perfectly anticipate y at the time of the integration decision, there would never be a strict incentive to integrate and subsequently delegate: either $y \leq y^*(\pi)$, in which case there will be integration with centralization, or $y > y^*(\pi)$, in which case non-integration performs at least as well as integration (strictly so if $y < y^H$, for then HQ would inefficiently centralize, or if the acquisition of the supplier's asset has even infinitesimal cost). Contrary to the data, no variation in autonomy among integrated suppliers would be observed, as no HQ would delegate at all.

As we have already discussed, the the value of owning assets is affected by the option to centralize, which provides a hedge against risk in y. Without that uncertainty, the sole role of ownership would be to enforce the efficient assignment of control, as in footnote 7, by integrating if and only if $y \leq y^*(\pi)$. In that case, against the strong prediction in Result 2, integration would increase with π only if $c^S > c^H$.

To be more precise about how uncertainty affects delegation and integration, we consider the comparative statics of (first-order) stochastic dominance and of (Rothschild-Stiglitz) risk in the distribution $G(y)$.[13]

[12] Evidence on the integration effect can be found in Alfaro *et al.* (2016), and McGowan (2017).

[13] A distribution $\hat{G}$ is stochastically larger than G if $\hat{G}(y) \leq G(y)$ everywhere, with strict inequality on a

3.3.1 Delegation

The degree of delegation is $1-G(y^H(\pi))$, the probability that y exceeds the threshold $y^H(\pi)$. From the definition of stochastic dominance the following is immediate as well as intuitive:

Result 3. *Stochastic increases in the supplier productivity distribution $G(y)$ lead to a higher probability of delegation.*

When problems are more likely to land in the supplier's sphere of competence, it is more likely that he will be the one asked to solve them.

However, because the cumulative probability at any given interior y value — $y^H(\pi)$ in particular — can increase or decrease with increases in risk (risk-comparable c.d.f.'s cross at least once), there is no general prediction about the effects of risk on delegation: increasing risk tends to fatten both upper and lower tails, meaning only that there is a reduced likelihood that problems that crop up will be the ones in the supplier's "domain of mediocrity." Likely as not, they will play more to his strengths as to his weaknesses, yielding an ambiguous change in HQ's willingness to delegate.

3.3.2 Integration

It is illuminating to rewrite inequality (4) as the following "put-option" condition:

$$\mathbb{E}\max\{\pi z-(\pi-c^H)y,0\}\geq c^S zG(y^H(\pi)).$$

The integrand $\max\{\pi z-(\pi-c^H)y,0\}$ is decreasing and convex in y. So the relative benefit of integration over non-integration (value of the option) decreases with stochastic increases in $G(y)$: the more likely the supplier will be the one better able to solve the problem that arises, the less valuable is the option to centralize, and with that, the lower is the value of integration. Unlike in standard real-options theory, however, the cost of acquiring the asset, which is the supplier's expected private cost, is systematically tied to the underlying uncertainty, for it depends on the likelihood of centralization $G(y^H(\pi))$. This decreases with $G(y)$ as well, so the net effect on integration appears to be difficult to determine without more specific information about the distributions.

As would be expected from the theory of options, where more uncertainty increases option value, more risk in the supplier's problem-solving capacity raises the value of integration.

---

non-null set. $\hat{G}$ is riskier than G if the two distributions have the same mean and $\int_0^x \hat{G}(y)dy\geq\int_0^x G(y)dy$ for all x, with strict inequality on a non-null set; equivalently, $\hat{G}$ can be derived from G via a sequence of mean-preserving spreads; or $\mathbb{E}_G u(y)\geq\mathbb{E}_{\hat{G}}u(y)$ for all concave functions $u(\cdot)$.

But again, the cost of integration $c^S zG(y^H(\pi))$ could rise or fall with increases in risk. Despite this ambiguity, the effects on integration of changes in the riskiness of $G(y)$ can be signed for certain classes of distributions, in particular the lognormal family, which is salient for our empirical analysis. Within that class, the relative benefit effect dominates the cost effect, even if that happens to be countervailing, so we have:

Result 4. *If there is integration at lognormal $G(y)$, there is integration at a riskier lognormal $\hat{G}(y)$.*

For the proof of this result, see Section A-1.1 of the Theoretical Appendix.

Results 3 and 4 imply that delegation and integration co-vary with different moments of the productivity distribution: roughly speaking, delegation co-varies with the mean, but not with risk; integration co-varies with risk, but not with the mean. These results provide another illustration of how the conceptual distinction between delegation and outsourcing is manifested in terms of measurable quantities.

3.4 From Theory to Data

According to the theoretical model (Results 1 and 2), both integration and delegation choices should depend on input value. Suppliers of more valuable inputs should be more likely to be integrated with firms; among the integrated suppliers, those producing more valuable inputs should be delegated more decisions. These results lead to the following testable predictions, which we will bring to the data in Section 5:

> **P.1**: Firms should be more likely to delegate production decisions to integrated suppliers that produce more valuable inputs.
>
> **P.2**: Firms should be more likely to integrate suppliers that produce more valuable inputs.

In Section 4.3, we describe how we measure delegation, integration, and input value, to assess the validity of the above predictions.

The theory also implies that delegation and integration should depend on $G(y)$, the distribution of problems that can arise during production. In our empirical analysis, we will bring to the data Results 3 and 4 by constructing a proxy for $G(y)$ for each input industry. To this end, in what follows, we will assume that the distribution of problems is specific to each input industry and that each supplier has one customer, at least under the period of observation.[14]

[14]Alternatively, we could allow for multiple buyers for each supplier and assume that the problem drawn is perfectly correlated across his customers.

According to the model, a supplier who is not integrated will receive a payment py, no matter how large or small y is. Thus, interpreting y in the model as expected output per unit of resources (e.g., employment), the observable labor productivity of a non-integrated supplier should reflect the problem encountered producing his input. Under the assumption that the problems encountered are i.i.d. among suppliers in the same industry i, the empirical distribution of sales per worker of non-integrated suppliers in industry i can thus be used as a proxy for the distribution $G_i(y)$.[15]

In particular, if there is a large enough number of suppliers in an industry and the empirical distribution of labor productivity is approximately lognormal, as is the case in our data (see Figure A-3 in the Empirical Appendix), then we can take $G_i(y)$ to be lognormal, validating one of the hypotheses of Result 4. Lognormality also allows for simple parametric measurement of stochastic dominance and risk in terms of the mean and coefficient of variation (see Levy, 1973), which will be employed in the empirical analysis. In particular, controlling for the coefficient of variation, differences in the mean productivity of suppliers in input industries proxy for stochastic differences in $G_i(y)$; controlling for the mean, differences in the coefficient of variation proxy for differences in Rothschild-Stiglitz risk.[16]

In Section 5, we will assess the validity of the following predictions, which correspond to theoretical Results 3 and 4:

> **P.3**: Firms should be more likely to delegate to their integrated suppliers that operate in stochastically more productive input industries.
>
> **P.4**: Firms should be more likely to integrate suppliers that operate in riskier input industries.

[15]By contrast, the model implies that the observed productivity distribution of the integrated suppliers is not a good proxy for $G_i(y)$. This is because the contribution of an integrated supplier depends on whether or not HQ centralizes, which she does whenever the supplier has low competence. In that case, it is HQ's own competence z that determines productivity. The observed productivity of integrated suppliers is thus left-censored, clouding the relationship between various orderings of observed distributions from two industries $i, \hat{i}$ and orderings of the underlying true distributions $G_i(y), G_{\hat{i}}(y)$. In the case of non-integrated suppliers, who endogenously do not delegate, there is no censoring problem. Another reason for excluding integrated suppliers when constructing the uncertainty measures is that transfer pricing effects may distort their measured labor productivity.

[16]Variation in the distribution of labor productivity across industries reflects not only differences in output, but also in the prices prevailing in those industries. Our proxy for first-order stochastic differences in $G_i(y)$ will thus suffer from measurement error, if we do not control for industry prices. The coefficient of variation of labor productivity is instead independent of prices.

4 Data and Variables

4.1 Datasets

In what follows, we describe the datasets used in our empirical analysis to construct firm-level integration and delegation measures and to assess the role of input value and uncertainty in shaping these decisions.

4.1.1 World Management Survey

The World Management Survey (WMS) 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 help to make the data of high quality. 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.[17] 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, day-of-the-week), the manager (seniority, job tenure) and the location of the CEO of the firm was collected. These survey measures 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. The survey achieved a 45% response rate, which is very high for company surveys.[18] 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 (see row "WMS dataset" in Table A-1).

For each firm, one (randomly selected) plant is surveyed in the WMS. Thus, we do not typically observe variation in delegation across plants belonging to the same firm. In some cases, the same plant was interviewed again in the later waves.

[17]The interviewers were given the name and contact details of the firm, but no financial details.

[18]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.

4.1.2 WorldBase

WorldBase provides coverage of public and private firms in more than 200 countries and territories.[19] This dataset has been used extensively in the empirical literature on firm boundaries (e.g., Acemoglu *et al.*, 2009; Alfaro *et al.*, 2019). The unit of observation 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 information on its primary industry and up to five secondary industries in which each establishment operates. These are classified based on the 1987 US Standard Industrial Classification (SIC). The activities are recorded at the SIC4 level (935 industries, of which 459 are in manufacturing). Worldbase also provides additional information of the plants (e.g., location, sales, employment).

WorldBase allows us to trace ownership linkages between establishments. In particular, we can use DUNS numbers to link plants that have the same domestic or 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.[20]

We use the 2005 WorldBase dataset and focus 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 row "WorldBase dataset" in Table A-1).

4.2 Samples

In the empirical analysis, we use two samples constructed from the datasets described above.

[19]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.

[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. In the case of multinational corporations, we follow Alfaro *et al.* (2016) and split them into several corporations, considering the domestic ultimate in each country as the relevant headquarters.

4.2.1 Delegation Sample

This is the sample that is used to test the predictions related to delegation. It is constructed by combining information from WorldBase and the WMS. Notice that we are not able to use the full WMS to test the model's predictions concerning the role of input value and uncertainty on delegation choices. This is because testing predictions P.1 and P.3 requires information on ownership and input-output linkages between the plant surveyed in the WMS and its central headquarters, which we can only get for plants that are also in the WorldBase dataset.

For the United States 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 firms in WorldBase. We have 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 manufacturing plants, while WorldBase contains lots of very small plants. The matched sample includes 3,444 plant-year observations located in 20 countries.[21] The final sample used in the delegation regressions, which we henceforth call "delegation sample", is a bit smaller due to missing information on some control variables for some of the plants. It consists of 2,889 plant-year observations, for plants operating in 574 sectors, corresponding to 2,253 firms. In this sample, 1,663 observations correspond to plants that are part of multi-plant firms, while the remaining observations correspond to single-plant firms. Note that, even though in principle only one plant per firms was interviewed in the WMS, some plants were interviewed more than once. This explains why the number of firms is smaller than the number of plant-year observations.[22]

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 delegation sample. Using information for all the countries in the sample is key to having a sufficient number of observations for our econometric analysis. Even the country with the maximum number of observations (the United States) accounts for only 20 percent of the sample.

[21]The remaining plants in WMS could not be matched to a firm in WorldBase, either because they are not in the WorldBase sample, or because they could not be uniquely assigned to a specific firm based on the string matching algorithm. Since the success of the algorithm is random, the matched sample corresponds to a random subsample of the WMS.

[22]The WMS includes only three cases in which two plants belonging to the same firm are surveyed.

4.2.2 Integration Sample

This sample is constructed exclusively from the WorldBase dataset. It is used to test the predictions related to integration. We exclude firms that have less than 20 employees, to correct for differences in the coverage of small firms across countries (see also Klapper *et al.*, 2006). Since the WMS contains exclusively manufacturing plants, we restrict the attention to firms that have a primary SIC code in manufacturing (between SIC 2000 and 3999). To be able to exploit within-firm variation across inputs, we select firms that integrate at least one input different from their primary output j.

The final sample used in the integration regressions, which we henceforth call "integration sample", includes 67,111 plants, corresponding to 66,102 firms, operating in 459 sectors, located in 20 countries. As shown in Table A-1, this sample features more variation in plant size compared to the delegation sample: median plant employment is 42, the mean is 147, and the standard deviation is 3,187. Table A-3 reports the number of observations (at the firm-input level) by country.

4.3 Key Variables

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

4.3.1 Delegation

Our measure of delegation comes from the WMS survey of Bloom *et al.* (2012). They conducted in-depth interviews with the plant managers of medium-sized manufacturing firms, excluding those where the CEO and the plant manager was the same person (4.9% of their interviews).

Plant managers were asked to state the degree of autonomy they have when hiring a new full-time permanent shop floor employee, introducing a new product, or in sales and marketing decisions. These 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). They were also 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.[23] Since the scaling may vary across questions, we have standardized the scores from the four autonomy

[23] In Appendix Figure A-6, we detail the individual questions in the same order as they appear in the survey.

questions to z-scores, by normalizing each question to mean zero and standard deviation one.[24]

The variable *Delegation*$_{f(j,c,)p(i)}$ measures the overall degree of autonomy that the CEO of firm f (with primary activity j, located in country c) grants to the senior manager of plant p (with primary activity i). It is the average across the four z-scores for plant p belonging to firm f.[25] We use information on ownership linkages from the WorldBase dataset to link a plant to its parent firm. Notice that the design of the survey implies that delegation is only measured for integrated suppliers. This is in line with our theoretical model, in which the control of the production of the input can be delegated to a supplier only within firm boundaries.[26] Figure A-1 in the Empirical Appendix shows the distribution of *Delegation*$_{f(j,c,)p(i)}$ in the delegation sample.

As pointed out by Bloom *et al.* (2012), their survey includes different types of firms: many have only one production plant, others are multi-plant firms;[27] in most cases, the CEO is located in a different location than the plant manager interviewed; in others cases, it is located on site.

In Section 5.2, we examine the effects of input value and supplier uncertainty on delegation choices. As mentioned before, the WMS typically contains information on one plant per firm. This implies that we can only observe the degree of autonomy granted by the CEO to one plant manager. We use the primary SIC4 code of the parent firm to identify the output industry j and the primary SIC4 code of the plant to identify the input industry i and rely on cross-firm variation in *Delegation*$_{f(j,c,)p(i)}$ to identify the role of input value.

4.3.2 Integration

To distinguish between integrated and non-integrated inputs, we build on the methodology developed by Fan and Lang (2000), combining information on firms' reported activities with input-output (IO) tables. As explained below, in our empirical analysis, we measure vertical

[24]The continuous measure of delegation using in the empirical analysis stands in apparent contrast to the binary delegation choice in the theoretical model. However, the two can be reconciled by supposing that the production process is subdivided into a number of tasks, each of which is subject to a problem shock, and can be delegated or centralized. HQ delegates a task whenever the supplier's productivity on it exceeds the threshold $y^*(\pi)$, and centralizes otherwise. Interpreting the number or fraction of tasks delegated as the degree of delegation yields a measure that has the same properties to the probability of delegation in the baseline model.

[25]All subcomponents are highly correlated with overall delegation score. The highest correlation (0.72) is with the z-score for marketing autonomy.

[26]In the case of non-integration, there is de-facto decentralization of production of the input: HQ does not have the authority to centralize production decisions, and the supplier has no incentive to give control to HQ.

[27]In the case of firms with multiple plants, the random plants surveyed are the best guess at the average degree of decentralization in the firm as a whole.

integration at the firm-input level, in line with our theoretical model: for each firm, we identify the set of inputs needed to produce its final good and estimate the probability that the firm produces each input within its boundaries.

To measure vertical integration, we need standardized and disaggregated data on input requirements for each output sector. It should be stressed that disaggregated IO tables are only available for a few countries. Moreover, when available, they are usually based on different sector classifications. For these reasons, we follow previous studies (e.g.,Acemoglu *et al.*, 2009; Alfaro *et al.*, 2016, 2019) in using US IO tables to capture technological linkages between sectors. These tables should be informative about input flows across industries to the extent that these are determined by technology. As pointed out by Acemoglu *et al.* (2009), assuming that the US IO structure carries over to other countries also mitigates concerns about the endogeneity of technology.[28]

The data are from the Benchmark IO tables of the Bureau of Economic Analysis (BEA). 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 is based on the SIC 1987 classification. To convert the input-output data at the 4-digit SIC level, we use the concordance guide provided by the BEA.[29] This allows us to measure vertical linkages between 935 manufacturing and non-manufacturing SIC4 industries. For each output industry j, IO tables report the dollar value of i used as an input in the production of $1 of j, also known as the direct requirements coefficient, $IO_{i,j}$.

To measure vertical integration, we combine the IO tables with information from WorldBase on the primary and secondary activities of each firm f.[30] We proceed in four steps. First, we use the primary SIC4 code of each firm f to identify its output sector j. Second, we use IO tables to identify the set of inputs required to produced j, which we denote by $S(j) = \{i : IO_{i,j} > 0\}$. Third, we identify which inputs firm f integrates, using the primary and secondary SIC codes reported by the firm and all its subsidiaries (if any). We define $I(f) \subseteq S(j)$ to be the set of integrated inputs, which the firm can in principle obtain within

[28]The results are robust to restricting the analysis to high-income members of the Organisation for Economic Co-operation and Development (OECD), which are closer to the United States in terms of technology.

[29]The concordance table can be found at `http://www.bea.gov/industry/exe/ndn0017.exe`. It assigns a unique 6-digit IO industry to each 4-digit SIC code in manufacturing (i.e., between 2000 and 3999). Outside these sectors, each SIC code may be matched to multiple IO codes. When this is the case, we consider all possible matches. In robustness checks, we show that our results are robust to restricting the analysis to manufacturing inputs, for which the concordance between the SIC and IO industry codes is one to-one.

[30]As mentioned above, each establishment in WorldBase reports a primary activity and up to five secondary SIC codes that are produced at the location of the production facility. A firm with a headquarters and k establishments can thus in principle report up to $5+6*k$ integrated secondary activities. Most firms in our integration sample have a single establishment. However, they still report secondary activities. Right censoring is not a problem: only 0.01 percent of plants in the WorldBase sample report 5 secondary SIC codes.

its ownership boundaries. The complement set $NI(f) = S(j) \setminus I(f)$ identifies the non-integrated inputs, i.e., the inputs required in the production of the firm's output that are not included in $I(f)$. Finally, having identified the set of integrated and non-integrated inputs for each firm f, we can construct the variable *Integration*$_{f(j,c),i}$. This is a dummy variable equal to 1 if firm f (producing primary output j, located in country c) integrates a supplier in input industry i within its boundaries.

In our empirical analysis, we exploit variation in *Integration*$_{f(j,c),i}$ within and across firms to study how input value and input risk shape integration choices. To keep the analysis tractable and exclude cases of mechanical integration (almost all sectors use their own output as an input), we limit the sample to firms that integrate at least one input different from their output j and to the top 100 (manufacturing and non-manufacturing) inputs i used by j, as ranked by the IO coefficients (see also Alfaro *et al.*, 2019).[31]

As mentioned before, US IO tables are highly disaggregated, providing information on vertical linkages between 935 manufacturing and non-manufacturing industries.[32] 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).

4.3.3 Input Value

To examine how input value affects delegation and integration choices, we use the variable $IO_{i,j}$ described above. This is the direct requirement coefficient for the pair of sectors i and j, which captures the dollar value of input i used in the production of one dollar of j. Thus IO coefficients IO_{ij} are a close empirical counterpart to π_{ij} in our theoretical model. In robustness checks, we use different quartiles of the $IO_{i,j}$ variable to check for non-linear effects of input value on organizational choices.

Figure A-2 in the Empirical Appendix shows the distribution of $IO_{i,j}$ in the WorldBase sample. Not surprisingly, given that the BEA input-output tables are highly disaggregated, the average $IO_{i,j}$ is only 4 cents in the delegation sample (see Table A-4) and 5 cents in the integration sample (see Table A-5).

4.3.4 Uncertainty in Input Industries

To test prediction P.3, we use information from the full WorldBase dataset to measure labor productivity (sales per employee) of non-integrated suppliers in an input industry. As

[31] The results are robust to restricting the analysis to the top 10 or top 20 inputs.

[32] Outside manufacturing, the correspondence between BEA and SIC industry codes is not always unique and this may induce measurement error in the IO coefficients. In fact, the effect of input value on integration is even stronger for manufacturing inputs, where the correspondence is one to one.

discussed in Section 3.4, this is can be used as a proxy for the distribution $G_i(y)$.[33]

In particular, we construct the variable *Mean Productivity*$_{i,c}$ as the arithmetic average of labor productivity of non-integrated suppliers of input i in country c. To minimize measurement error, we consider all plants that report SIC4 code i as their only production activity. Overall, labor productivity is computed for a sample of 12,063,180 non-integrated plants in WorldBase. The distribution of labor productivity of independent input suppliers approximates a lognormal distribution (see Figure A-3 in the Empirical Appendix).[34] Figure A-4 in the Empirical Appendix shows the distribution of *Mean Productivity*$_{i,c}$ in the integration sample.

To assess the validity of prediction P.4, we use *CV Productivity*$_{i,c}$, the coefficient of variation of productivity of suppliers in the same input industry.[35] Controlling for the mean of supplier productivity, this variable can be used as a proxy for riskiness of the input industry in the Rothschild-Stiglitz sense (see Levy, 1973). Figure A-5 in the Empirical Appendix shows the distribution of *CV Productivity*$_{i,c}$ in the integration sample.

In robustness checks, we construct *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ after winsorizing labor productivity at the 5th and 95th percentile. We also construct alternative measures of uncertainty in input industries, using data on stock market returns of US firms from Bloom *et al.* (2018).

5 Empirical Results

5.1 The Effects of Input Value and Uncertainty on Delegation

In this section, we assess our model's predictions concerning the effects of input value and of uncertainty in input industries on delegation and choices.

To this purpose, we use the delegation sample to estimate:

$$\begin{aligned} \textit{Delegation}_{f(j,c),p(i)} \quad = \quad & \beta_1 \textit{IO}_{i,j} + \beta_2 \textit{ Mean Productivity}_{i,c} + \beta_3 \textit{ CV Productivity}_{i,c} + \\ & + \beta_4 \mathbf{X}_p + \beta_5 \mathbf{X}_f + \delta_i + \delta_j + \delta_c + \epsilon_{f(j,c),p(i)}. \end{aligned} \tag{5}$$

The dependent variable is the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). The variable $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij.

[33]As noted in footnote 16, this proxy will suffer from measurement error if we do not control for prices.

[34]The figure is constructed by regressing log labor productivity of all independent suppliers in WorldBase on 4-digit-industry $\times$ country dummies. It thus shows within-industry-country variation in log labor productivity.

[35]Notice that, unlike *Mean Productivity*$_{i,c}$, this variable is independent of industry prices.

Mean Productivity$_{i,c}$ and *CV Productivity*$_{i,c}$ measure respectively the mean and coefficient of variation of labor productivity of independent suppliers in industry i in country c. 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). Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. We cluster standard errors at the input-industry i level.[36]

Prediction P.1 of our theoretical model suggests that a producer of good j should delegate more decisions to integrated suppliers that produce more valuable inputs, implying that the coefficient β_1 should be positive and significant. According to prediction P.3 of our model, firms should delegate more decisions to suppliers that operate in more productive industries, so we also expect the coefficient β_2 to be positive and significant.

The results of estimating (5) are reported in Table 1. We present first a specification that includes the key control variable with input-industry fixed effects (column 1), and then further include country fixed effects (column 2), output-industry fixed effects (column 3), and the plant and firm controls (column 4).

Consider first the effects of input value. As mentioned in Section 4, the WMS does not allows us to observe variation in delegation across plants belonging to the same firm. As a result, β_1 cannot be identified by comparing the degree of autonomy granted by firm f producing good j (e.g., automobiles) to the managers of two of its plants, one producing input i (e.g., plastics materials and resins), the other producing input i' (e.g., engines). We can, however, exploit cross-firm variation in the degree of delegation to identify the coefficient of *IO*$_{i,j}$. In particular, we can compare the degree of autonomy granted by two firms producing the same output (e.g., automobiles) to the managers of one of their plants, one producing input i (e.g., plastics materials and resins), the other producing input i' (e.g., engines). We can also compare the degree of autonomy granted to two plants producing the same input i (e.g., plastics materials and resins) who belong to firms making different final goods (e.g., automobiles and fabricated pipe and fittings).

As expected, the coefficient of *IO*$_{i,j}$ is positive and significant across all specifications, indicating that suppliers of more "important" inputs (i.e., those with larger value contributions) are granted more authority (prediction P.1). In terms of magnitude, based on the estimates reported in column 4 of Table 1, increasing the input-output coefficient by 1 standard deviation increases delegation by around 0.072 standard deviations.[37]

[36]The results are robust to clustering standard errors at the input-output level (the level of variation of *IO*$_{i,j}$) and at the industry-country level (the level of variation of *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$).

[37]The standard deviation of *IO*$_{i,j}$ in the delegation sample is 0.06, so 1.2*0.06 = 0.072. This effect might

Table 1

The effects of input value and uncertainty on delegation

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.9863***	0.9729***	1.1332**	1.2088**
	(0.3497)	(0.3445)	(0.5105)	(0.5008)
Mean $Productivity_{i,c}$	0.0266**	0.0276**	0.0332**	0.0352*
	(0.0131)	(0.0136)	(0.0164)	(0.0184)
CV $Productivity_{i,c}$	0.0025	0.0020	0.0026	0.0037
	(0.0029)	(0.0027)	(0.0027)	(0.0024)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	Yes
WMS noise controls	Yes	Yes	Yes	Yes
Observations	2,889	2,889	2,889	2,889

Notes: The dependent variable is $Delegation_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean* $Productivity_{i,c}$ and *CV* $Productivity_{i,c}$ are respectively the mean and coefficient of variation of labor productivity of independent suppliers in industry i in country c. Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. The noise controls are related to the WMS survey (e.g., duration and day-of-the-week of the interview process, seniority, and job tenure of the plant manager interviewed, and whether the CEO is located in plant p). 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 input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

The coefficient of *Mean* $Productivity_{i,c}$ is also positive and significant in all specifications. We thus find strong support for prediction P.3 of our model.[38] Based on the estimates

appear to be a small, but can be compared to the effect of other covariates. For example, the coefficient of log firm employment, one of the firm controls included in column 4 of Table 1, is 0.101, which implies that an increase in firm employment of 40% is associated with a 0.101*0.4=0.040 standard deviation change in the delegation index.

[38]This is notwithstanding the fact that our proxy for first-order stochastic differences in $G(y)$ suffers from measurement error, since we cannot fully control for industry prices. As mentioned before, variation in the distribution of labor productivity across industries reflects not only differences in output, but also in the prices prevailing in those industries. To the extent that prices vary at the industry level, they are accounted by the fixed effect δ_i in (5). Sector-country fixed effects cannot be included, since they would be collinear with the uncertainty measures.

reported in column 4 of Table 1, increasing mean supplier productivity by 1 standard deviation increases delegation by around 0.034 standard deviations.[39]

Recall that the theory has no implications for the effect of *CV Productivity*$_{i,c}$ on delegation choice. As it happens, the coefficient of this variable is not statistically significant. Concerning the auxiliary controls included in Table 1 (coefficient estimates not reported), we find that larger and older firms are more likely to delegate decisions to their suppliers and that plants that are larger and have a more educated workforce are granted more autonomy.

It is interesting to compare our results with those of Bloom *et al.* (2012), who use the same measure of within-firm delegation (which they refer to as decentralization). They find that trust is a key driver of the internal allocation of decision rights and has larger effects than any of the other covariates in their study: when including the full set of these covariates, they find that a 1 standard deviation increase in trust is associated with a 0.07 standard deviation increase in delegation. Our results indicate that input value is also an important driver of delegation, as its effect is as large as that of trust. Notice that, in our analysis, the role of national trust is absorbed by the country fixed effects.

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 Section A-2.3 of the Empirical Appendix. A first set of robustness checks addresses is related to the delegation variable, which measures the degree of autonomy granted by the firm's CEO to the plant's top manager. Recall that firms in which the plant manager is also the CEO are excluded from the dataset of Bloom *et al.* (2012) used in our analysis. Moreover, Table 1 includes "noise controls" (e.g., seniority and job tenure of the plant manager interviewed, and an indicator variable for whether the CEO is on site) to reduce measurement error in the delegation variable. Table A-7 restricts the analysis to multi-plant firms. Although this drastically reduces sample size to 1,663 observations, the coefficient of $IO_{i,j}$ is always positive and significant and the coefficient of *Mean Productivity*$_{i,c}$ is positive and significant in three of the four specifications.

A second set of robustness checks concerns the uncertainty measures. In Table A-8, we verify that the results of Table 1 are robust to controlling for the number of suppliers used to construct the variables *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$. Table A-9 shows that the delegation results continue to hold if we construct these variables after winsorizing labor productivity at the 5th and 95th percentile. In Table A-10, we use US stock market data from Bloom *et al.* (2018) to construct alternative uncertainty measures. The variable *Mean Stock Returns*$_i$ is the mean of stock market returns in SIC4 industry i, while *SD Stock Returns*$_i$ captures the cross-section dispersion in stock market returns across firms in that

[39]The standard deviation of *Mean Productivity*$_{i,c}$ in the matched sample is 0.97, so 0.035*0.97=0.034.

industry. Unlike *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$, these uncertainty measures vary only at the sector level. Moreover, they are only available for listed firms in some (manufacturing) industries, which reduces the size of the sample. Notwithstanding these limitations, the results are broadly in line with our model's predictions: the coefficient of *IO*$_{i,j}$ remains positive and significant, while the coefficient of *Mean Stock Returns*$_i$ is positive but not significant.

In Table A-11, we control for contracting frictions at the industry, country, and industry-country level, as in Nunn (2007). Following Nunn (2007), we interact a given sector's contract intensity with *Rule of Law*$_c$, a measure of each country's judicial quality. Our coefficients of interest remain unaffected while the interaction between a sector's contract intensity and legal quality is mostly insignificantly associated with more delegation. The coefficients of the additional controls suggest that contracting frictions increase the incentives to delegate decisions to integrated suppliers. We have also explored non-linearities in the effects of input value on delegation choices, by including different quartiles of the variable *IO*$_{i,j}$ in equation (5). The results of Table A-12 show that the effects are concentrated in the top quartile, which captures the most important inputs.

Overall, the results of Table 1 and the robustness checks in Section A-2.3 of the Empirical Appendix confirm the role of input value and supplier uncertainty on delegation choices, in line with predictions P.1 and P.3 of our model.

5.2 The Effects of Input Value and Uncertainty on Integration

We next assess validity of our model predictions about the effects of input value and uncertainty in supplier markets on integration choices. To this purpose, we use the integration sample and estimate the following linear probability model:

$$\begin{aligned} \mathit{Integration}_{f(j,c),i} \quad = \quad & +\alpha_1 \, \mathit{IO}_{i,j} + \alpha_2 \, \mathit{Mean\ Productivity}_{i,c} + \alpha_3 \, \mathit{CV\ Productivity}_{i,c} \\ & +\beta_4 \, \mathbf{X}_f + \delta_i + \delta_j + \delta_c + \epsilon_{f(j,c),i}, \end{aligned} \tag{6}$$

The dependent variable is the probability that firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. *IO*$_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and coefficient of variation of labor productivity of independent suppliers in input industry i in country c. $\mathbf{X}_f$ is a vector of firm-level controls, and δ_i, δ_j, and δ_c denote input-industry, output-industry, and country fixed effects. In the most demanding specifications, we exploit only within-firm variation to identify the role of input value and supplier uncertainty. In these specifications, we replace output-sector and country fixed effects with firm fixed effects (δ_f), which allow us to account for the role of unobservable firm characteristics. We cluster standard errors at the input-industry (i) level.[40]

According to prediction P.2 of our theoretical model, producers of good j should be more likely to integrate suppliers of more valuable inputs, implying that the coefficient of $IO_{i,j}$ should be positive and significant. According to prediction P.4, when controlling for *Mean Productivity*$_{i,c}$, the coefficient of *CV Productivity*$_{i,c}$ should be positive and significant.

Table 2

The effects of input value and uncertainty on integration

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.1499***	0.1791***	0.1789***	0.2030***
	(0.0134)	(0.0145)	(0.0145)	(0.0161)
Mean Productivity$_{i,c}$	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CV Productivity$_{i,c}$	0.0008***	0.0007***	0.0007***	0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	6,644,884	6,644,884	6,644,884	6,644,884

Notes: The dependent variable is *Integration*$_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. Firm-level controls include the employment and age of the parent firm. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

The results of estimating (7) are reported in Table 2.[41] We first regress *Integration*$_{f(j,c),i}$ against the key controls of interest and input-industry fixed effects (column 1).[42] We then

[40] The results of are unaffected if we cluster at the input-output (i, j) or at the industry-country (i, c) level.

[41] Recall that the WordlBase sample focuses on manufacturing firms that integrate at least one input different from their primary output. In these regressions, we do not include firms from Greece. This is because establishments in Greece only report their primary SIC codes. As a result, we cannot use within-firm variation to study the determinants of integration choices.

[42] Recall that, for each industry j, we focus on its top 100 inputs in terms of (strictly positive) IO coefficients.

add country fixed effects (column 2), output-industry 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). In this specification, firm controls as well as 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 coefficient of $IO_{i,j}$ is positive and significant at the 1 percent level across all specifications, confirming that final good producers are more likely to produce in house more valuable inputs, in line with prediction P.2 of our model. In terms of magnitude, based on the specification of column 3, moving the input-output coefficient by 1 standard deviation increases the probability of vertical integration by 0.64 percentage points — a 64 percent increase compared to the baseline probability of one percentage point.[43]

The estimated coefficient for *CV Productivity*$_{i,c}$ is also positive and highly significant. This finding is robust to including additional controls and different sets of fixed effects. In particular, it continues to hold when including firm fixed effects. In this specification, the coefficient α_3 in (7) is identified by exploiting only within-firm variation in the dispersion of supplier productivity across input industries. 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.[44] This corresponds to a 34% increase relative to the baseline integration probability of one percentage point.

These results provide strong support for prediction P.4. Increases in risk increase the option value of integration: when there is more uncertainty about the problems that will arise during production HQ attaches greater value to the option of being able to centralize or delegate production decisions. When productivity is distributed lognormally, as it is the case in our data, this effect dominates the (ambiguous) effect on the cost of integration, so firms should be more likely to integrate suppliers in riskier input industries.

Concerning the additional controls, the negative coefficient of *Mean Productivity*$_{i,c}$, although insignificant, is consistent with our theoretical model: recall from Section 3.3.2 that stochastic increases in the supplier's ability to solve problems reduce the option value of integration, but also reduce its cost, and it is an empirical question whether one effect dominates the other. In our data, the benefit effect appears to dominate the opposing cost effect. The coefficients of the firm controls (not reported), indicate that larger and older firms are more likely to integrate inputs within their boundaries.

In Section A-2.4 in the Empirical Appendix, we report the results of series of additional

[43] The standard deviation of $IO_{i,j}$ is 0.036. Thus, 0.179*0.036*100 = 0.644.

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

estimations to verify the robustness of the results of Table 2. We first consider different samples of firms. Some concerns have been raised in the literature about the measurement of vertical integration, particularly for multi-plant firms (Atalay, Hortaçsu, and Syverson, 2014). Measurement error should work against us, making it harder to find a significant effect of input value and supplier uncertainty on vertical integration. Nevertheless, we have verified that the results of Table 2 continue to hold when we restrict the analysis to single-plant firms (see Table A-13). The coefficients of input value and input risk also remain positive and significant when restricting the analysis to multi-plant firms (see Table A-14),[45] or to the firms included in the delegation sample (see Table A-15).

Another set of estimations is related to the uncertainty variables. Table A-16 shows that the integration results are robust to controlling for the number of suppliers in each sector-country i, c used to construct the variables *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$. In Table A-17 these variables are constructed after winsorizing labor productivity at the 5th and 95th percentile. In Table A-18, we restrict the analysis to input industries in which there are at least 50 suppliers in sector-country i, c, to rule out that the results of Table 2 are due to demand-driven supply assurance motives for integration (see discussion in Section 6). In Table A-19, we use data from Bloom *et al.* (2018) on stock market returns of US firms to measure uncertainty in input industries. As mentioned before, these alternative uncertainty measures vary only at the sector level and can only be constructed for some (manufacturing) industries. The results confirm that firms are more likely to integrate suppliers that produce more valuable inputs and operate in riskier industries.

In Table A-20, we control for contracting frictions at the industry, country, and industry-country level, as in Nunn (2007). In these regressions, the sample is restricted to manufacturing inputs, for which the variable *Contract Intensity*$_i$, measuring the fraction of inputs that are differentiated, can be constructed. Following Nunn (2007), we interact this variable with, *Rule of Law*$_c$, a measure of each country's judicial quality. The coefficient of $IO_{i,j}$ remains highly significant and increases substantially in magnitude. This is because measurement error in IO coefficients is substantially smaller for manufacturing inputs compared to others (see Section 4.3) and thus the attenuation bias is less severe. Note that the coefficient on the interaction of *Contract Intensity*$_i$ with *Rule of Law*$_c$ is negative and highly significant, in line with previous studies (e.g., Acemoglu *al.*, 2009).

A final set of robustness checks concern the input-output coefficients. We have explored non-linearities in the effects of input value on integration choices, by including different quartiles of the variable $IO_{i,j}$ in equation (7). The results of Table A-21 confirm that firms

[45] Table A-14 includes multi-plant firms that have establishments in different countries. The results are robust to excluding multinational firms.

are more likely to integrate suppliers of more valuable inputs.

Overall, the robustness checks of Table 2 confirm the role of input value and supplier uncertainty on integration choices, in line with our model's predictions.

6 Discussion: Alternative Mechanisms

Our empirical analysis establishes the following regularities: (i) firms delegate more decisions to integrated suppliers that produce more valuable inputs; (ii) firms are more likely to integrate suppliers of more valuable inputs; (iii) firms delegate more decisions to integrated suppliers in more productive input industries; (iv) 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 creates a real option for HQ to centralize or delegate decisions according to comparative advantage, and the value of this option increases with the degree of uncertainty about the severity of problems that may arise during the production process.

Existing alternative models could provide a rationale for some — but not all — of our empirical findings. For instance, as suggested earlier, the model in Legros and Newman (2013) could explain the finding that the likelihood that a supplier is integrated grows with the input's contribution to total firm value. However, that model does not consider the possibility of delegation and thus cannot rationalize the findings about its determinants; nor does it account for the empirical effects of risk on integration.

The finding that the likelihood that a supplier is integrated grows with the riskiness of the input industry is related to the literature on supply assurance motives for integration (e.g., Carlton, 1979; Bolton and Whinston, 1993). In these models, the assurance motive 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. These might also account for the positive coefficient of *CV Productivity*$_{i,c}$ in Table 2.

However, we should expect these demand-driven mechanisms, particularly the variants in which suppliers opportunistically sell to other buyers, to be less relevant when firms can source inputs from many suppliers. Against this hypothesis, the coefficient of *CV Productivity*$_{i,c}$ remains positive and highly significant when focusing on input industries with large numbers of suppliers (see Table A-18). Moreover, the result holds when we include output industry fixed effects, which account for product market uncertainty (see columns 2 and 3 of Table 2 and the corresponding specifications in Section A-2.4 of the Empirical Appendix), and firm fixed effects, which account for demand for inputs by other firms in the same country-output sector (see column 5 of Table 2 and the corresponding specifications in Section A-2.4 of the Empirical Appendix).

Moreover, while existing models of supply assurance could in principle explain why suppliers in riskier industries are more likely to be integrated, they do not rationalize the other empirical findings, since they neither have anything to say about delegation, nor address the role of input value in integration decisions.

In a similar vein, transaction-cost reasoning might account for the risk-integration finding.[46] Along these lines, one may argue that more uncertainty could lead to more contractual incompleteness, which in turn could lead to more integration.[47] As discussed above, our results are robust to including controls for the degree of contractibility (as in Nunn, 2007) to account for this alternative mechanism (see Table A-20). Finally, like classical supply assurance, this approach appears to be silent on the drivers of delegation and the organizational effects of input value.

As mentioned in the Introduction and documented in Section A-2.2 of the Empirical Appendix, our data also show that more vertically integrated firms tend to delegate more. Following the discussion in Section 3.2, this correlation is consistent with our theoretical model: higher value enterprises (e.g., that have more productive HQs or can sell their goods at higher prices) should be more vertically integrated while simultaneously granting more autonomy to their subordinates.

The covariation of delegation and integration could potentially be rationalized by managerial capacity models in which an HQ'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. However, there are grounds for skepticism. Among them, the positive correlation between delegation and integration is robust to controlling for measures of firm size and workforce education that could proxy for, or at least correlate with, managerial capacity.[48] Those measures are themselves positively correlated with delegation (see Table 1), whereas a managerial capacity model would tend to predict that firms with more capacity should delegate less. Finally, capacity-constrained HQs would tend to keep control of the decisions regarding more im-

[46] Lafontaine and Slade (2007) summarize the evidence from a few single-industry studies.

[47] Some of the steps in this argument are rather tenuous. One can think of mechanisms through which uncertainty could enhance rather than reduce contractibility, e.g., if greater risk encourages more investment in verification technologies. In our data, input risk (as measured by $\textit{CV Productivity}_{i,c}$) is uncorrelated with contracting frictions (as measured by the interaction between $\textit{Contract intensity}_i$ and $\textit{Rule of Law}_c$).

[48] We can also directly control for the plant's management practices, using data from Bloom *et al.* (2016). The coefficient of $\textit{Vertical Integration}_f$ in Table A-6 remains positive and significant.

portant inputs and delegate decisions concerning less important ones, which goes against the evidence in result (i) above: more autonomy is granted to suppliers who contribute more, rather than less value to the enterprise.

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 of the organization can mislead when embedded in two dimensions. For example, based on models that study separately firm boundaries and the allocation or control within firms, one may expect integration and centralization to covary positively at the firm level, while the opposite is true in the data.

In this paper, we have brought integration and delegation together, both theoretically and empirically. A number of insights emerge from the exercise. First, the analysis reveals a novel mechanism by which supply assurance motivates integration: the ability to redeploy control to the party most competent to solve problems that may arise during production ensures a minimal level of productivity. Formally, this places the evaluation of the costs and benefits of ownership in the rubric of real options. Our findings suggest that integration may increase managerial flexibility, because it allows re-allocating decision-making across different parts of the organization. Outsourcing can therefore be very costly for firms, which lose the ability to redeploy control in light of new information.[49] Thus, in a world in which input risk is rising, we would expect waves of vertical mergers, as the option value of being able to re-allocate decision-making within firm boundaries increases.

Our framework also helps to uncover unifying themes in the analysis of organizational design. In particular, it extends the value principle, already applied to integration in previous work, to delegation: the value of an input or the profitability of a firm affect how much autonomy will be granted to the input's supplier.

An important consequence of our findings concerns the connection between delegation (or decentralization within firms) and productivity. Delegation is often associated with better performance, i.e., higher generated value, though the direction of causality remains an open

[49] This is illustrated by Boeing's infamous 787 Dreamliner fiasco. In a departure from its practices with earlier models of aircraft, Boeing outsourced the design and manufacture of major components of the 787 (e.g., fuselage, wings, stabilizers) to independent suppliers. As it turned out, some of these suppliers were not competent to solve various design and procurement problems that arose during the course of production. The need for remedial fixes led to years of delivery delays and tens of billions of dollars in cost overruns. 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 (e.g., Tang *et al.*, 2009; McDonald and Kotha, 2015).

question. A frequent claim in the business press is that delegation increases value, for instance by improving incentives and morale, or freeing up top managerial resources. A few studies find evidence for links in this direction (e.g., Bresnahan *et al.*, 2002; Aghion *et al.*, 2021). The present model offers two selection mechanisms through which productivity could be driving delegation, rather than the other way around: exogenously more productive firms should be more willing to delegate; and subordinates with more autonomy may be the ones who happen to have encountered problems they are good at solving.

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. Studying firm boundaries together with other aspects of the firm's internal organization helps to illuminate interdependencies that are crucial for understanding the functioning and guiding the design and regulation of organizations.

References

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., N. Bloom, B. Lucking, R. Sadun, and J. Van Reenen (2021). "Turbulence, Firm Decentralization and Growth in Bad Times," *American Economic Journal: Applied* 13, 133-169.

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.

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.

Bena, J., and K. Li, (2014). "Corporate Innovations and Mergers and Acquisitions," *The Journal of Finance* 69, 1923-1960.

Berlingieri, G., F. Pisch, and C. Steinwender (2021). "Organizing Supply Chains: Input Cost Shares and Vertical Integration," *Journal of the European Economic Association* 19, 1816-1852.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta, and S. Terry (2018). "Really Uncertain Business Cycles," *Econometrica* 86, 1031-1065.

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.

Bradley, M., A. Desai, and E.H. Kiim (1988). "Synergistic Gains from Corporate Acquisitions and their Division between the Stockholders of Target and Acquiring Firms," *Journal of Financial Economics*, 21, 3–40.

Bresnahan, T., E. Brynjolfsson, and L. M. Hitt (2002). "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *Quarterly Journal of Economics* 117, 339-376.

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

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

Dixit, A. K., and R. S. Pindyck (1994). *Investment and Uncertainty*. Princeton, NJ: Princeton University Press.

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.

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.

Graham, J. R., C. R. Harvey, M. Puri (2015). "Capital Allocation and Delegation of Decision-Making Authority within Firms," *Journal of Financial Economics* 115, 449-470.

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.

Grullon, G., E. Lyandres, and A. Zhdanov (2012). "Real Options, Volatility, and Stock Returns," *The Journal of Finance* 67, 1499–1537.

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.

Hansman, C., J. Hjort, G. León-Ciliotta, and M. Teachout (2020). "Vertical Integration, Supplier Behavior, and Quality Upgrading Among Exporters," *Journal of Political Economy* 128, 3570-3625.

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.

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

______(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.

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

______(2022). "Controlling Control," mimeo, BU and ECARES.

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

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

Macchiavello, R., and J. Miquel-Florensa (2017). "Vertical Integration and Relational Contracts: Evidence from the Costa Rica Coffee Chain," CEPR Discussion Paper No. 11874.

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.

McGowan, D. (2017). "Digging Deep to Compete: Vertical Integration, Product Market Competition and Prices," *Journal of Industrial Economics* 65, 683-718.

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.

Nunn, N. (2007). "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade," *Quarterly Journal of Economics* 122, 569-600.

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.

______(2010) "Interpersonal Authority in a Theory of the Firm," *American Economic Review* 100, 466-490.

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

______(1991). "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives," *Administrative Science Quarterly* 36, 269-296.

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

Appendices

A-1 Theoretical Appendix

A-1.1 Integration, Risk, and the Log-normal Distribution

Rewrite the integration (HQ ownership) condition (4) as

$$z > \frac{\pi - c^H}{\pi - c^S}\mathbb{E}[y|y < y^H(\pi)].$$

The probability of integration increases (in the sense that it occurs for a larger set of parameters c^H, c^S, z, π) in response to changes in risk if $\mathbb{E}[y|y < y^H(\pi)]$ decreases.

If y is drawn from the family of log-normal distributions with parameters μ, σ, then the c.d.f. is $\Phi(\frac{\ln y - \mu}{\sigma})$, where $\Phi(\cdot)$ is the standard normal c.d.f, which is log-concave. The conditional expectation can be written (notationally suppressing dependence of y^H on π)

$$e^{\mu + \frac{\sigma^2}{2}} \cdot \frac{\Phi(\frac{\ln y^H - \mu}{\sigma} - \sigma)}{\Phi(\frac{\ln y^H - \mu}{\sigma})}.$$

One lognormal (μ', σ') is riskier than another (μ, σ) if their means $e^{\mu + \frac{\sigma^2}{2}}$ and $e^{\mu' + \frac{\sigma'^2}{2}}$ are equal and $\sigma' > \sigma$. Denoting $m = \mu + \frac{\sigma^2}{2}$, and rewriting $\mathbb{E}[y|y < y^H]$ in terms of m and σ, higher risk lowers the conditional expectation if and only if

$$\frac{\partial}{\partial \sigma} e^m \cdot \frac{\Phi(\frac{\ln y^H - m}{\sigma} - \frac{\sigma}{2})}{\Phi(\frac{\ln y^H - m}{\sigma} + \frac{\sigma}{2})} < 0. \tag{7}$$

Straightforward computation reveals that this condition is equivalent to

$$\frac{\phi(\frac{\ln y^H - m}{\sigma} + \frac{\sigma}{2})}{\Phi(\frac{\ln y^H - m}{\sigma} + \frac{\sigma}{2})}(\frac{\ln y^H - m}{\sigma^2} - \frac{1}{2}) - \frac{\phi(\frac{\ln y^H - m}{\sigma} - \frac{\sigma}{2})}{\Phi(\frac{\ln y^H - m}{\sigma} - \frac{\sigma}{2})}(\frac{\ln y^H - m}{\sigma^2} + \frac{1}{2}) < 0,$$

where $\phi(\cdot)$ is the standard normal density.

Now, logconcavity implies $\frac{\phi(\frac{\ln y^H - m}{\sigma} + \frac{\sigma}{2})}{\Phi(\frac{\ln y^H - m}{\sigma} + \frac{\sigma}{2})} < \frac{\phi(\frac{\ln y^H - m}{\sigma} - \frac{\sigma}{2})}{\Phi(\frac{\ln y^H - m}{\sigma} - \frac{\sigma}{2})}$. And because $y^H = \frac{z\pi}{\pi - c^H} > z$, under the maintained hypothesis (A1) that $z \geq \mathbb{E}y$, $\ln y^H > m$, so $\frac{\ln y^H - m}{\sigma^2} + \frac{1}{2}$ is positive and exceeds $\frac{\ln y^H - m}{\sigma^2} - \frac{1}{2}$. Thus, condition (7) is satisfied, and we conclude that increasing riskiness of lognormal distributions implies more integration.[50]

[50] A somewhat lengthier proof shows that the statement is true without (A1). See Legros-Newman (2022).

A-2 Empirical Appendix

A-2.1 Descriptive Statistics

Table A-1

Plant-level statistics

	Employment			Age			
	Mean	Med.	S.D.	Mean	Med.	S.D.	N. plants
WMS dataset	277	150	405	40	51	45	11,691
WorldBase dataset	12	2	606	17	13	60	17,152,559
Delegation sample	254	150	367	30	40	35	2,256
Integration sample	147	42	3,187	26	33	29	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, delegation sample

Country	Number of Observations	Percentage
Argentina	88	3.05
Australia	93	3.22
Brazil	194	6.72
Canada	195	6.75
Chile	25	0.87
China	58	2.01
France	159	5.50
Germany	219	7.58
Greece	101	3.50
India	75	2.60
Italy	89	3.08
Ireland	17	0.59
Japan	100	3.46
Mexico	82	2.84
New Zealand	88	3.05
Poland	13	0.45
Portugal	74	2.56
Sweden	218	7.55
United Kingdom	411	14.23
United States	590	20.42
Total	2,889	100.00

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

Table A-3

Observations by country, integration 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 integration 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 variables used in delegation regressions

	Mean	Median	S.D.	N. firms	N. observations
Delegation$_{f(j,c),p(i)}$	0.13	0.07	0.99	2,253	2,889
IO$_{i,j}$	0.04	0.01	0.06	2,253	2,889
Mean Productivity$_{i,c}$	0.26	0.16	0.97	2,253	2,889
CV Productivity$_{i,c}$	3.18	1.45	6.41	2,253	2,889
Share Employment$_p$	0.61	0.60	0.89	2,253	2,889
% Workers with College Degree$_p$	15.21	9.27	16.56	2,253	2,889
Employment$_f$	672.27	300.00	958.93	2,253	2,889
Age$_f$	40.08	31.00	35.02	2,253	2,889

Notes: The table reports descriptive statistics of the main variables used in the regressions of Table 1. *Delegation*$_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i), constructed combining information from WMS and WorldBase. *IO*$_{i,j}$ 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 based on the US IO table. *Mean Productivity*$_{i,c}$ is the mean of labor productivity of the independent suppliers in input industry i located in country c (in millions of US Dollars), while *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity. Both variables have been computed based on all non-integrated plants in WorldBase. *Share Employment*$_p$ is the plant's share of the firm's employment using information from Worldbase. *% Workforce with College Degree*$_p$ is the percentage of the plant's employees with a bachelor's degree or higher using information from WMS. *Employment*$_f$ and *Age*$_f$ respectively measure the number of employees of firm f and the number of years since the firm was established, based on information from WorldBase.

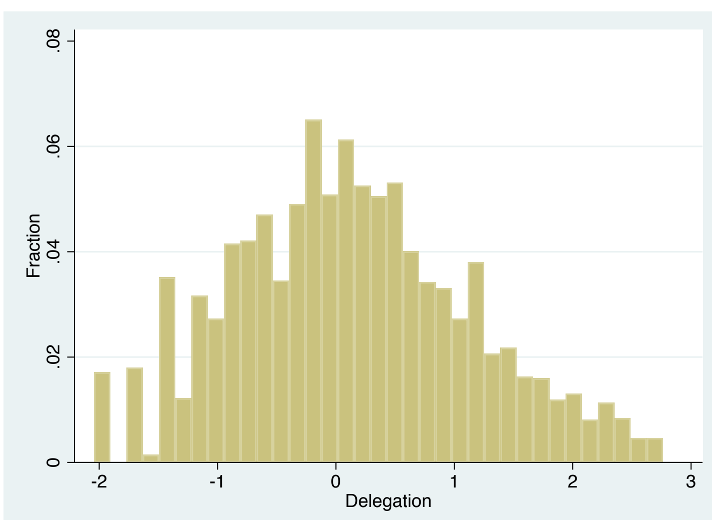
Table A-5

Descriptive statistics of variables used in integration regressions

	Mean	Median	Standard deviation	N. firms	N. observations
Integration$_{f(j,c),i}$	0.01	0.00	0.11	66,102	6,644,884
IO$_{i,j}$	0.05	0.05	0.036	66,102	6,644.884
Mean Productivity$_{i,c}$	0.50	0.30	10.50	66,102	6,644,884
CV Productivity$_{i,c}$	3.04	1.94	4.63	66,102	6,644,884
Employment$_f$	206.38	45.00	4,903.87	66,102	6,644,884
Age$_f$	33.56	26.00	28.98	66,102	6,644,884

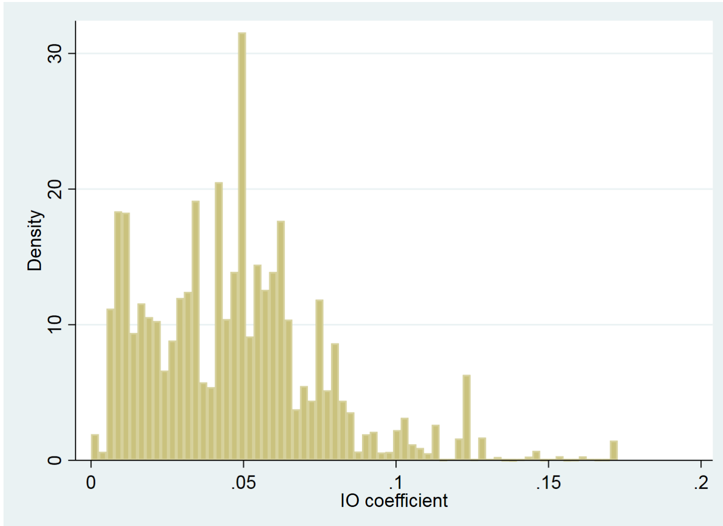
Notes: The table reports descriptive statistics of the variables used in Table 2 (and robustness checks), based on the integration sample. *Integration*$_{f(j,c),i}$ is a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries, constructed using information from WorldBase. *IO*$_{i,j}$ 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, based on US input-output tables. *Mean Productivity*$_{i,c}$ is the mean of labor productivity of the independent suppliers in input industry i located in country c (in millions of US Dollars), while *CV Productivity*$_{i,c}$ is the coefficient of variation of labor productivity. Both variables are constructed using information on non-integrated plants in WorldBase. *Employment*$_f$ and *Age*$_f$ respectively measure the number of employees of firm f and the number of years since the firm was established, based on information from WorldBase.

Figure A-1: Delegation



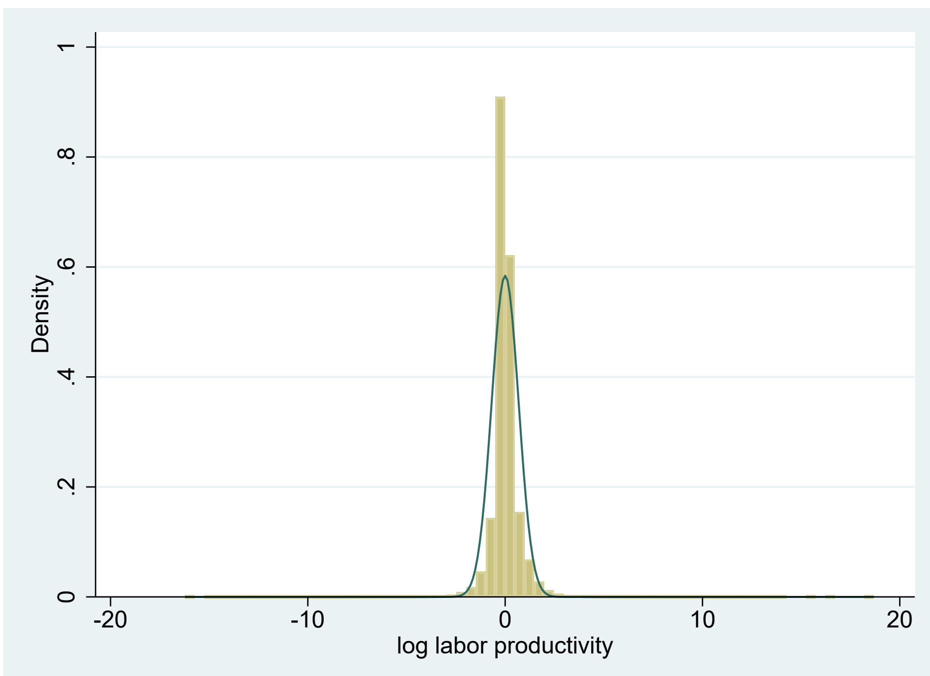
Notes: The figure shows the distribution of *Delegation*$_{f,p}$ in the delegation sample.

Figure A-2: Input Value



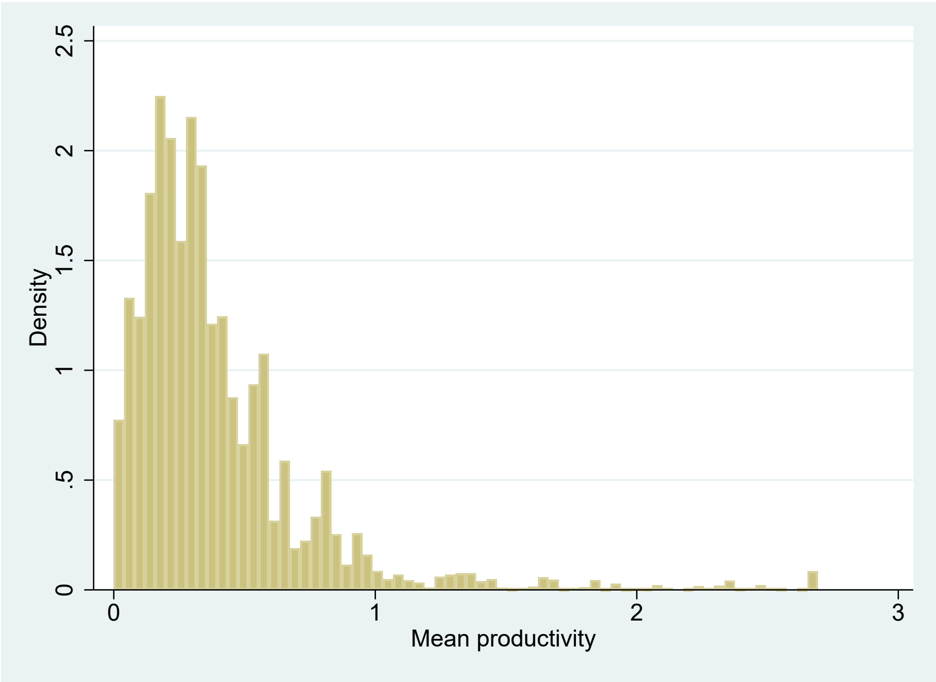
Notes: The figure shows the distribution of $IO_{i,j}$ in the integration sample.

Figure A-3: Productivity Distribution of Suppliers



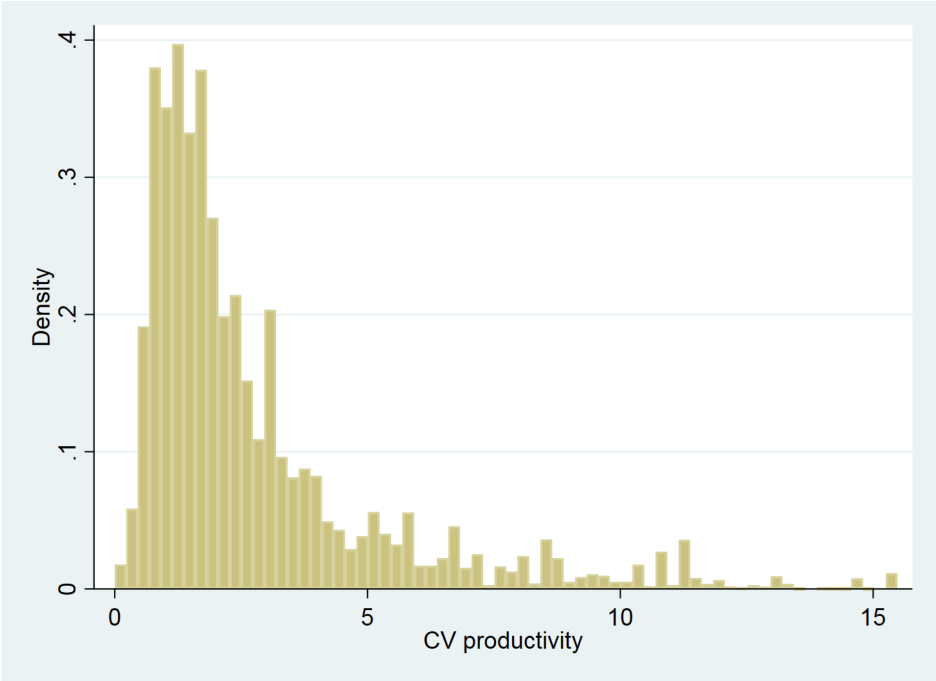
Notes: The figure shows that the distribution of labor productivity of independent suppliers approximates a lognormal distribution (in line with Result 4). It is constructed by regressing log labor productivity of all independent suppliers in WorldBase on 4-digit-industry × country dummies, and thus captures within-industry-country variation in log labor productivity.

Figure A-4: Mean Productivity



Notes: The figure shows the distribution of *Mean Productivity*$_{ic}$ in the integration sample.

Figure A-5: Input Risk



Notes: The figure shows the distribution of *CV Productivity*$_{ic}$ in the integration sample.

Figure A-6: 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 The Co-Variation of Integration and Delegation

We document the existence of a positive firm-level correlation between delegation and vertical integration. To measure delegation, we use the variable $Delegation_{f,p}$, which captures the degree of autonomy granted by the CEO of firm f to the manager of plant p. As discussed in Section 4.3, this is constructed using survey data from Bloom *et al.* (2012).

To measure firm-level vertical integration, we follow the methodology of Alfaro *et al.* (2016). For each firm f located in country c with primary activity j, we define $IO_{ij}^f \equiv IO_{ij} * \mathbb{I}_i^f$, where $\mathbb{I}_i^f \in \{0,1\}$ is an indicator variable that equals one if and only if firm f owns plants that are active in sector i. The firm's vertical integration index measures the fraction of inputs used in the production of a firm's final good that can be produced in house and is the sum of the IO coefficients for each input industry in which firm f is active: $\textit{Vertical Integration}_{f(j,c)} = \sum_i IO_{ij}^f$.

To examine the relationship between delegation and integration, we estimate

$$\textit{Delegation}_{f(j,c),p(i)} = \beta_1 \textit{ Vertical Integration}_{f(j,c)} + \beta_2 \mathbf{X}_p + \beta_3 \mathbf{X}_f + \delta_i + \delta_j + \delta_c + \epsilon_{f(j,c),p(i)}. \tag{A-1}$$

The dependent variable measures the extent to which firm f (with primary activity j, located in country c) delegates decisions to the manager of plant p (with primary activity i). The main control variable of interest is $\textit{Vertical Integration}_{f(j,c)}$, the vertical integration index of firm f. $\mathbf{X}_p$ and $\mathbf{X}_f$ are vectors of plant- and of firm-level controls, while δ_i, δ_j and δ_c are respectively input-sector, output-sector (at the 3-digit SIC level), and country fixed effects.[51]

The results are reported in Table A-6. In column 1, we regress the degree of delegation within firm f against the vertical integration index of the firm, including only input-industry fixed effects. The estimated coefficient of $\textit{Vertical Integration}_{f,j,c}$ is positive and significant (at the one-percent level).[52] Figure A-7 is a binned scatterplot that illustrates the results of this specification. Each of the 40 bins represents around 85 observations.

The remaining specifications of Table A-6 show that the positive correlation between delegation and vertical integration 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).[53]

[51] 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.

[52] The coefficient of $\textit{Vertical Integration}_{f,j,c}$ is also significant (at the five-percent level) in an even more parsimonious specification, in which we do not include any fixed effects.

[53] The variables $\%\ \textit{Workforce with College Degree}_p$ and $\textit{Employment}_p$ 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.

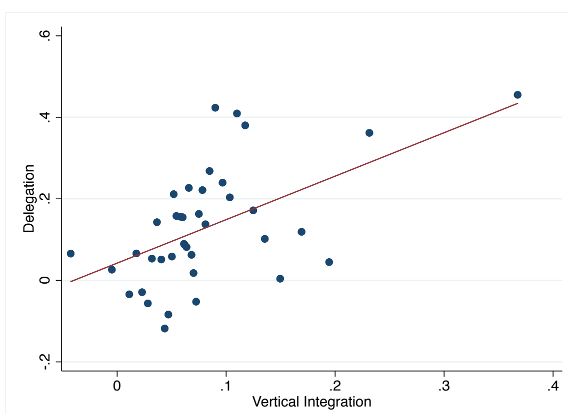
Table A-6

Delegation and integration

	(1)	(2)	(3)	(4)
Vertical Integration$_{f(j,c)}$	0.685***	0.794***	0.691***	0.577**
	(0.246)	(0.244)	(0.250)	(0.250)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	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 the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). *Vertical Integration*$_{f(j,c)}$ is the vertical integration index of firm f. Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. 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.

Figure A-7: Firm-Level Delegation and Vertical Integration



Notes: The figure is a binned scatterplot, constructed based on the regression in column 1 of Table A-6. We group the residuals into 40 bins, compute the mean for each bin and represent the resulting data in a scatterplot.

We have carried out a series of additional estimations to verify the robustness of the results of Table A-6. The results of these estimations are available upon request. 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. Second, we reproduce Table A-6 after winsorizing the vertical integration index at the 5th and 95th percentile. Finally, we restrict the analysis to single-plant firms, for which our measures of integration and delegation are less likely to suffer from measurement error. All specifications confirm that, in more vertically-related firms, HQ delegates more decisions to plant managers.

In principle, any relationship between delegation and integration is *a priori* possible: the number and value of suppliers an HQ owns has nothing to do with whether she grants one of them complete autonomy (fully delegates) or none, so the degree of vertical integration places no restriction on the degree of delegation we might observe.

As discussed at the end of Section 3.2.2, the positive covariation between integration and delegation documented Table A-6 can be rationalized by simple extensions of our baseline model: more valuable firms (e.g., those that have more productive HQs) should integrate more suppliers and grant more autonomy to them.

A-2.3 Delegation Results: Robustness Checks

Table A-7

The effects of input value and uncertainty on delegation
(only multi-plant firms)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.979**	0.952*	1.482**	1.593**
	(0.450)	(0.485)	(0.662)	(0.642)
$\textit{Mean Productivity}_{i,c}$	0.175	0.307**	0.274*	0.260*
	(0.147)	(0.150)	(0.143)	(0.133)
$\textit{CV Productivity}_{i,c}$	-0.086	-0.119	-0.076	-0.072
	(0.085)	(0.083)	(0.086)	(0.084)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	Yes
WMS noise controls	Yes	Yes	Yes	Yes
Observations	1,663	1,663	1,663	1,663

Notes: The dependent variable is $\textit{Delegation}_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. $\textit{Mean Productivity}_{i,c}$ and $\textit{CV Productivity}_{i,c}$ are respectively the mean and coefficient of variation of labor productivity of independent suppliers in industry i in country c. Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment, age of the parent firm, and an indicator variable identifying single-plant firms. The noise controls are related to the WMS survey (e.g., duration and day-of-the-week of the interview process, seniority, and job tenure of the plant manager interviewed, indicator for whether the CEO is located in plant p). 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 input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-8

The effects of input value and uncertainty on delegation
(controlling for the number of suppliers in each sector-country)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.9850***	0.9695***	1.1211**	1.1967**
	(0.3485)	(0.3442)	(0.5098)	(0.5005)
Mean Productivity$_{i,c}$	0.0267**	0.0279**	0.0339**	0.0359*
	(0.0132)	(0.0137)	(0.0167)	(0.0186)
CV Productivity$_{i,c}$	0.0024	0.0018	0.0019	0.0030
	(0.0030)	(0.0027)	(0.0026)	(0.0023)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	Yes
WMS noise controls	Yes	Yes	Yes	Yes
Observations	2,889	2,889	2,889	2,889

Notes: The dependent variable is *Delegation*$_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and coefficient of variation of labor productivity of independent suppliers in industry i in country c. We also control for the number of suppliers in input industry i located in country c (coefficient not reported). Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. The noise controls are related to the WMS survey (e.g., duration and day-of-the-week of the interview process, seniority, and job tenure of the plant manager interviewed, indicator for whether the CEO is located in plant p). 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 input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-9

The effects of input value and uncertainty on delegation
(winsorizing supplier productivity)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	1.0175***	0.9951***	1.1585**	1.2442**
	(0.3458)	(0.3393)	(0.4983)	(0.4893)
$Mean\ Productivity_{i,c}$	0.0475***	0.0473***	0.0601***	0.0686***
	(0.0110)	(0.0116)	(0.0141)	(0.0139)
$CV\ Productivity_{i,c}$	-0.0048	0.0270	0.0187	0.0120
	(0.0470)	(0.0488)	(0.0532)	(0.0530)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	Yes
WMS noise controls	Yes	Yes	Yes	Yes
Observations	2,889	2,889	2,889	2,889

Notes: The dependent variable is $Delegation_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. $Mean\ Productivity_{i,c}$ and $CV\ Productivity_{i,c}$ are respectively the mean and coefficient of variation of labor productivity of independent suppliers in industry i in country c. Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. The noise controls are related to the WMS survey (e.g., duration and day-of-the-week of the interview process, seniority, and job tenure of the plant manager interviewed, indicator for whether the CEO is located in plant p). 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 input (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-10

The effects of input value and uncertainty on delegation
(uncertainty measures based on US stock market data)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	1.150***	1.094***	1.386**	1.445**
	(0.401)	(0.398)	(0.599)	(0.553)
Mean Stock Returns$_i$	0.210	0.215	0.189	0.199
	(0.127)	(0.137)	(0.391)	(0.368)
SD Stock Returns$_i$	-1.236	-2.039	0.724	1.144
	(2.038)	(2.003)	(3.248)	(3.042)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	Yes
WMS noise controls	Yes	Yes	Yes	Yes
Observations	2,045	2,045	2,045	2,045

Notes: The dependent variable is *Delegation*$_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Stock Returns*$_i$ and *SD Stock Returns*$_i$ are respectively the mean and standard deviation of stock market returns of US firms operating in sector i. Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. The noise controls are related to the WMS survey (e.g., duration and day-of-the-week of the interview process, seniority, and job tenure of the plant manager interviewed, indicator for whether the CEO is located in plant p). 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 input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-11

The effects of input value and uncertainty on delegation
(controlling for contracting frictions)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.9585***	0.9539***	1.1442**	1.2206**
	(0.3573)	(0.3555)	(0.5102)	(0.5020)
Mean Productivity$_{i,c}$	0.0317***	0.0289**	0.0353**	0.0376**
	(0.0117)	(0.0133)	(0.0161)	(0.0178)
CV Productivity$_{i,c}$	0.0028	0.0020	0.0025	0.0035
	(0.0031)	(0.0027)	(0.0027)	(0.0024)
Contract Intensity$_i$	-0.4657	-0.5216	-0.3316	-0.4316
	(0.3630)	(0.3693)	(0.4288)	(0.4201)
Contract Intensity$_i$ × *Rule of Law*$_c$	1.4431***	3.4424	3.2988	5.8446
	(0.4506)	(7.6895)	(7.8708)	(7.6455)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	Yes
WMS noise controls	Yes	Yes	Yes	Yes
Observations	2,889	2,889	2,889	2,889

Notes: The dependent variable is *Delegation*$_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and coefficient of variation of labor productivity of independent suppliers in industry i in country c. The variables *Contract intensity*$_i$ and *Rule of Law*$_c$ are constructed as in Nunn (2007). *Contract intensity*$_i$ is defined at 4-digit SIC. The sample is restricted to manufacturing inputs, for which *Contract intensity*$_i$ can be constructed. Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. The noise controls are related to the WMS survey (e.g., duration and day-of-the-week of the interview process, seniority, and job tenure of the plant manager interviewed, indicator for whether the CEO is located in plant p). 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 input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-12

The effects of input value and uncertainty on delegation
(quartiles of $IO_{i,j}$)

	(1)	(2)	(3)	(4)
$IO_{i,j}^{4th}$	0.1436**	0.1372**	0.1436**	0.1320*
	(0.0637)	(0.0640)	(0.0720)	(0.0709)
$IO_{i,j}^{3rd}$	0.0223	0.0181	0.0383	0.0191
	(0.0606)	(0.0592)	(0.0629)	(0.0612)
$IO_{i,j}^{2nd}$	0.0671	0.0667	0.0464	0.0493
	(0.0563)	(0.0563)	(0.0578)	(0.0565)
Mean Productivity$_{i,c}$	0.0258*	0.0264*	0.0308*	0.0328*
	(0.0135)	(0.0138)	(0.0161)	(0.0181)
CV Productivity$_{i,c}$	0.0026	0.0021	0.0027	0.0037
	(0.0031)	(0.0029)	(0.0028)	(0.0025)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	No	Yes	Yes
Country FE	No	Yes	Yes	Yes
Plant controls	No	No	No	Yes
Firm controls	No	No	No	Yes
WMS noise controls	Yes	Yes	Yes	Yes
Observations	2,889	2,889	2,889	2,889

Notes: The dependent variable is *Delegation*$_{f(j,c),p(i)}$, the degree of autonomy granted by the CEO of the parent firm f (with primary activity j, located in country c) to the senior manager of plant p (with primary activity i). $IO_{i,j}^{n}$ are dummies for the n^{th} quartile of the the direct requirement coefficient for the sector pair ij (omitted category is the 1^{st} quartile). *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and coefficient of variation of labor productivity of independent suppliers in industry i in country c. Plant-level controls include the percentage of the plant's employees with a bachelor's degree or higher and the plant's share of the firm's employment. Firm-level controls include the employment and age of the parent firm. The noise controls are related to the WMS survey (e.g., duration and day-of-the-week of the interview process, seniority, and job tenure of the plant manager interviewed, indicator for whether the CEO is located in plant p). 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 input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

A-2.4 Integration Results: Robustness Checks

Table A-13

The effects of input value and uncertainty on integration (only single-plant firms)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.1439***	0.1715***	0.1715***	0.1943***
	(0.0210)	(0.0210)	(0.0210)	(0.0231)
$Mean\ Productivity_{i,c}$	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$CV\ Productivity_{i,c}$	0.0007***	0.0007***	0.0007***	0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	6,027,632	6,027,632	6,027,632	6,027,632

Notes: The dependent variable is $Integration_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. $Mean\ Productivity_{i,c}$ and $CV\ Productivity_{i,c}$ are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-14

The effects of input value and uncertainty on integration
(only multi-plant firms)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.1953***	0.2471***	0.2460***	0.2871***
	(0.0310)	(0.0316)	(0.0317)	(0.0348)
$Mean\ Productivity_{i,c}$	-0.0002**	-0.0002**	-0.0002**	-0.0002**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$CV\ Productivity_{i,c}$	0.0012***	0.0012***	0.0012***	0.0012***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	617,218	617,218	617,218	617,218

Notes: The dependent variable is $Integration_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. $Mean\ Productivity_{i,c}$ and $CV\ Productivity_{i,c}$ are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-15

The effects of input value and uncertainty on integration
(only firms in delegation sample)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.0770***	0.1263***	0.1283***	0.1375***
	(0.0143)	(0.0166)	(0.0165)	(0.0179)
Mean Productivity$_{i,c}$	-0.0000**	-0.0000*	-0.0000**	-0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CV Productivity$_{i,c}$	0.0006***	0.0006***	0.0006***	0.0006***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	249,471	249,471	249,471	249,471

Notes: The dependent variable is *Integration*$_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-16

The effects of input value and uncertainty on integration
(controlling for the number of suppliers in each sector-country)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.1490***	0.1776***	0.1775***	0.2012***
	(0.0214)	(0.0214)	(0.0214)	(0.0236)
Mean Productivity$_{i,c}$	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CV Productivity$_{i,c}$	0.0004***	0.0004***	0.0004***	0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	6,644,884	6,644,884	6,644,884	6,644,884

Notes: The dependent variable is *Integration*$_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. *$IO_{i,j}$* is the direct requirement coefficient for the sector pair ij. *Mean Productivity$_{i,c}$* and *CV Productivity$_{i,c}$* are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. We also control for the number of suppliers in input industry i in country c (coefficient not reported). Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-17

The effects of input value and uncertainty on integration
(winsorizing supplier productivity)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.1512***	0.1817***	0.1815***	0.2068***
	(0.0136)	(0.0217)	(0.0217)	(0.0239)
Mean Productivity$_{i,c}$	-0.0136***	-0.0137***	-0.0137***	-0.0139***
	(0.0040)	(0.0052)	(0.0052)	(0.0053)
CV Productivity$_{i,c}$	0.0066***	0.0066***	0.0066***	0.0066***
	(0.0018)	(0.0018)	(0.0018)	(0.0019)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	6,644,884	6,644,884	6,644,884	6,644,884

Notes: The dependent variable is *Integration*$_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-18

The effects of input value and uncertainty on integration
(only input sectors with 50+ suppliers)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.1602***	0.1978***	0.1977***	0.2293***
	(0.0248)	(0.0249)	(0.0249)	(0.0281)
Mean Productivity$_{i,c}$	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CV Productivity$_{i,c}$	0.0007***	0.0007***	0.0007***	0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	5,484,936	5,484,936	5,484,936	5,484,936

Notes: The dependent variable is *Integration*$_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Productivity*$_{i,c}$ and *CV Productivity*$_{i,c}$ are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-19

The effects of input value and uncertainty on integration
(uncertainty measures based on US stock market data)

	(1)	(2)	(3)	(4)
$IO_{i,j}$	0.2937***	0.2960***	0.2950***	0.3111***
	(0.0362)	(0.0343)	(0.0343)	(0.0374)
Mean Stock Returns$_i$	-0.0190	-0.0220	-0.0217	-0.0223
	(0.0396)	(0.0382)	(0.0382)	(0.0402)
SD Stock Returns$_i$	0.5420**	0.5986***	0.5989***	0.6482***
	(0.2098)	(0.2099)	(0.2236)	(0.2236)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	533,075	533,075	533,075	531,726

Notes: The dependent variable is *Integration*$_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. *Mean Stock Returns*$_i$ and *SD Stock Returns*$_i$ are respectively the mean and standard deviation of stock market returns of US firms operating in sector i. Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Table A-20

The effects of input value and uncertainty on integration
(controlling for contracting frictions)

	(1)	(2)	(3)	(4)
$\text{IO}_{i,j}$	0.28348***	0.30247***	0.30210***	0.30783***
	(0.02779)	(0.02840)	(0.02840)	(0.03004)
Mean Productivity$_{i,c}$	0.00003***	0.00003***	0.00003***	0.00003***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)
CV Productivity$_{i,c}$	0.00058***	0.00055***	0.00056***	0.00053***
	(0.00017)	(0.00016)	(0.00016)	(0.00016)
Contract Intensity$_i$	0.06370	0.07560*	0.07476*	0.15163***
	(0.05013)	(0.04410)	(0.04404)	(0.04679)
Contract Intensity$_i$ $\times$ *Rule of Law*$_c$	-0.10310**	-0.10457***	-0.10379***	-0.19025***
	(0.04476)	(0.03983)	(0.03976)	(0.04186)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	-
Observations	1,080,628	1,080,628	1,080,628	1,080,628

Notes: The dependent variable is *Integration*$_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO_{i,j}$ is the direct requirement coefficient for the sector pair ij. Firm-level controls include the employment and age of the parent firm. The variables *Contract intensity*$_i$ and *Rule of Law*$_c$ are constructed as in Nunn (2007). *Contract intensity*$_i$ is defined at 4-digit SIC. The sample is restricted to manufacturing inputs, for which *Contract intensity*$_i$ can be constructed. Output and input fixed effects defined at 3-digit SIC. In the specification in column 4, firm controls and 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-21

The effects of input value and uncertainty on integration
(quartiles of $IO_{i,j}$)

	(1)	(2)	(3)	(4)
$IO^{4th}_{i,j}$	0.0109***	0.0139***	0.0139***	0.0172***
	(0.0020)	(0.0020)	(0.0020)	(0.0023)
$IO^{3rd}_{i,j}$	0.0078***	0.0088***	0.0088***	0.0092***
	(0.0015)	(0.0018)	(0.0018)	(0.0020)
$IO^{2nd}_{i,j}$	0.0064***	0.0059***	0.0059***	0.0051***
	(0.0013)	(0.0015)	(0.0015)	(0.0017)
$\textit{Mean Productivity}_{i,c}$	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\textit{CV Productivity}_{i,c}$	0.0008***	0.0007***	0.0007***	0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Input FE	Yes	Yes	Yes	Yes
Output FE	No	Yes	Yes	-
Country FE	Yes	Yes	Yes	-
Firm FE	No	No	No	Yes
Firm controls	No	No	Yes	No
Observations	6,644,884	6,644,884	6,644,884	6,644,884

Notes: The dependent variable is $\textit{Integration}_{f(j,c),i}$, a dummy variable equal to 1 if firm f (with primary activity in sector j, located in country c) integrates input i within its boundaries. $IO^{n}_{i,j}$ are dummies for the n^{th} quartile of the the direct requirement coefficient for the sector pair ij (omitted category is the 1^{st} quartile. $\textit{Mean Productivity}_{i,c}$ and $\textit{CV Productivity}_{i,c}$ are respectively the mean and the coefficient of variation of labor productivity of the independent suppliers in input industry i in country c. Firm-level controls include the employment and age of the parent firm, and a dummy variable equal to 1 for single-plant firms. Output and input fixed effects defined at 3-digit SIC. Standard errors clustered at the input-industry (i) level in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.