

NBER WORKING PAPER SERIES

WHAT'S THE BIG IDEA? MULTI-FUNCTION PRODUCTS, FIRM SCOPE AND
FIRM BOUNDARIES

Mengxiao Liu
Daniel Trefler

Working Paper 26320
<http://www.nber.org/papers/w26320>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2019

We thank Professor Xiaoji Xu at Lehigh University for his help. We are grateful to Ross Jestrab and Yimin Yi at Syracuse University, and Poli Natama at the University of Toronto for their research assistance. We thank seminar participants at the University of Toronto, the Canadian Economic Association annual conference, the Midwest conference, and the Workshop on Technical Change, U.S.-China Trade and Labor Market Outcomes (Peking University) for their helpful comments. We also thank Devashish Mitra, Mary Lovely, and Yoonseok Lee at Syracuse University for their help and suggestions. All errors are our own. Trefler gratefully acknowledges funding from the Social Sciences and Humanities Research Council of Canada (SSHRC) and from the Canadian Institute for Advanced Research (CIFAR). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Mengxiao Liu and Daniel Trefler. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What's the Big Idea? Multi-Function Products, Firm Scope and Firm Boundaries
Mengxiao Liu and Daniel Trefler
NBER Working Paper No. 26320
September 2019
JEL No. F1,F12,F14,L14,L15,L23,L24

ABSTRACT

Products often bundle together many functions e.g., smartphones. The firm develops the big idea (which functions to bundle) and then chooses one supplier per function. We develop a model featuring holdup in which the firm's bargaining power declines in the number of suppliers. Greater scope as measured by the number of suppliers exacerbates holdup, but this is partially offset by the appropriate choice of vertical integration or outsourcing. Our main result flows from the empirical observation that the number of functions varies across products within an industry (firm heterogeneity). We introduce the notion of an 'ideas-oriented' industry in which more productive firms have higher marginal returns to introducing a new function. We show that more productive firms will (1) have more suppliers and (2) be more likely to integrate those suppliers. We take this to the data using a neural network to predict whether or not each of 29 million PATSTAT patent applications involves new/improved functions. We merge these patents with Capital IQ data on 55,000 companies and their supplier networks. We show that in industries where patents are skewed towards new or improved functions, more productive firms have more suppliers and are more likely to integrate these suppliers.

Mengxiao Liu
Department of Economics
Maxwell School of Citizenship and Public Affairs
Syracuse University
110 Eggers Hall
Syracuse, NY 13244
mengxiaoliu@gmail.com

Daniel Trefler
Rotman School of Management
University of Toronto
105 St. George Street
Toronto, ON M5S 3E6
CANADA
and NBER
dtrefler@rotman.utoronto.ca

1. Introduction

A single product often bundles together many functions. Smartphones and computers are extreme examples that combine communications (text, audio and visual) with photography, efficiency tools and other functions. Automobiles combine a basic driving experience with features ranging from a simple heated seat to consumer electronics through to AI-controlled brakes and steering. Even low-tech products can have many functions. A refrigerator may come with a defroster, an ice maker, an LED display, and a smart temperature control, none of which overlap technologically with the core compressor technology. Products often have multiple functions that are technologically distinct but are nevertheless bundled together to raise product demand. The firm provides the Big Idea by identifying and bundling clusters of functions that most interest consumers. For each function the firm then pairs with a supplier who helps develop and produce it. Multiple functions require a network of suppliers.

What implications flow from the fact that some firms are so much better than others at the Big Idea, that is, at identifying and bundling clusters of functions? And given the holdup problems associated with bringing ideas to market and coordinating multiple suppliers, how do firms decide whether to outsource functions or produce them in-house? We introduce the notion of an ‘ideas-oriented’ industry as one in which more productive firms have high marginal returns to adding a function relative to less productive firms. That is, productivity is the ability to identify and bundle valuable clusters of functions. It follows immediately that a more productive firm will imbed more functions into its product, which is what we mean by greater scope. However, if more functions require more suppliers then greater scope comes with greater potential for holdup by suppliers. Central to our paper is this trade-off between scope and holdup, and our main conclusion is that in ideas-oriented industries more productive firms can partially relax this trade-off and increase their scope by vertically integrating their suppliers.

This paper is about how multi-functionality connects two famous questions: What explains firm scope and what explains the boundaries of the firm? We assume that each function requires a unique supplier and non-contractible, relationship-specific investments from both the firm and the supplier. As a result, there is a bilateral holdup problem (Grossman and Hart, 1986, Hart and Moore, 1990). The firm engages in multilateral bargaining with its suppliers and the firm’s bargaining power is given by its Shapley value as share of total revenue. We assume that each function is essential for the final good, which makes the firm’s bargaining power a declining function of the number of suppliers. This sets up the scope-holdup trade-off: Multi-functionality increases the firm’s scope, but reduces the firm’s incentives to invest.

To be more concrete, we suppose as in Antràs (2003) that vertical integration raises the firm’s incentives by raising its bargaining power while outsourcing raises suppliers’ incentives by lowering the firm’s bargaining power. Start with a situation in which a firm has so few suppliers that its bargaining power is very high, so high that supplier investments are inefficiently discouraged. On ex ante efficiency grounds the firm should relinquish some bargaining power by outsourcing. As the number of suppliers increases, the firm’s bargaining power erodes and if it erodes enough, efficient incentives require the firm to vertically integrate in order to partially

restore its bargaining power. This creates a natural link between firm scope and firm boundary decisions.

Our main point flows from the empirical observation documented below of tremendous heterogeneity across firms in the number of suppliers used and the extent to which these suppliers are vertically integrated. To explain this, we first introduce within-industry productivity dispersion as in Melitz (2003) and especially Antràs and Helpman (2004). We then introduce the concept of an ‘ideas-oriented’ industry. In such an industry, productivity means the ability to squeeze out more demand from any level of functionality. More precisely, let θ be a firm’s productivity, let N be the firm’s number of functions or suppliers and let $D(N, \theta)$ be the firm’s demand shifter. An ideas-oriented industry is one for which D is log-supermodular. That is, D_N is increasing in θ .

Our model makes two predictions about heterogeneous firms in ideas-oriented industries.

1. *Firm scope with heterogeneity:* In ideas-oriented industries, more productive firms will have more suppliers.
2. *Firm boundaries with heterogeneity:* In ideas-oriented industries, more productive firms will be more likely to integrate these suppliers.

The logic for both is simple. A more productive firm has a higher marginal return to an additional function (a higher D_N) and so has more functions and, correspondingly, more suppliers. This reduces the firm’s bargaining power and to partially rebalance incentives the firm vertically integrates those suppliers.¹

This paper is about these two theoretical predictions and their empirical validity. The remainder of this introduction reviews the related theoretical literature and describes the empirics.

1.1. Related Theoretical Literature

The notion of endogenous limits to scope appears most famously in Kremer (1993), which takes a purely technological approach (O-ring technology). Incentives appear in Acemoglu, Antràs and Helpman (2007), henceforth AAH, a paper which provides the starting point for our own work. AAH discuss how the multilateral holdup problem constrains the number of suppliers. We depart from AAH in two ways. First and most importantly, in that paper a firm’s bargaining power is independent of the number of suppliers. Thus, our key mechanism is killed off. Second, in their paper all investment is done by the supplier and, since the firm does not need to be incentivized, there is always outsourcing.²

¹Each of the two predictions involves a double difference: (a) Within an industry it compares low-productivity firms with high-productivity firms and (b) across industries it compares ideas-oriented industries with non-ideas-oriented industries. A non-ideas-oriented industry is better described as a ‘cost-oriented’ industry. This is an industry in which more productive firms are better able to control costs as the number of functions increases. In such an industry the cost reduction from eliminating one extra function (i.e., from simplifying the production process) is larger for more productive firms than less productive firms. Letting $C(N, \theta)$ be the unit cost of producing a product with N functions, a cost-oriented industry is one for which C is log-submodular i.e., $C_N > 0$ is decreasing in θ . In cost-oriented industries (1) more productive firms will have fewer suppliers and (2) more productive firms will be more likely to outsource to these suppliers.

²AAH briefly discuss vertical integration, but vertical integration is preferred to outsourcing only if suppliers face a binding limited liability constraint.

Another closely related strand of the theoretical literature starts with Antràs and Helpman (2004) who introduce productivity heterogeneity into the Antràs (2003) model. To get at within-industry heterogeneity in firms' boundary choices, Antràs and Helpman (2004) introduce fixed costs of integrating and outsourcing. If the two fixed costs are equal then within an industry all firms either integrate or outsource while when fixed costs differ integration and outsourcing coexist within an industry. In contrast, we get coexistence without fixed costs. This is because in ideas-oriented industries more productive firms have more functions, more suppliers, and hence endogenously lower bargaining power. The lower bargaining power of more productive firms leads them to integrate while the higher bargaining power of less productive firms leads them to outsource.

More tangentially related to our work is Antràs and Chor (2013) and Alfaro, Antràs, Chor and Conconi (2019) who consider chains of suppliers and the decision about which suppliers on the chain to integrate.

1.2. Empirics

There is a long and established empirical literature on holdup that tests the international trade models of Antràs (2003) and Antràs and Helpman (2004). See for example Yeaple (2006), Nunn and Trefler (2008, 2013, 2014) and Alfaro and Charlton (2009). For a survey see Antràs (2015). There is also a related industrial organization literature of which Acemoglu, Johnson and Mitton (2009), Acemoglu, Aghion, Griffith and Zilibotti (2010) and Liu (2019) are most relevant.

We build on this literature by considering firms with multiple suppliers and by introducing the notion of ideas-oriented industries. An empirical assessment of our two hypotheses requires (1) data on whether or not an industry is ideas-oriented, (2) data on each firm's productivity and industry of affiliation, and (3) data on each firm's network of suppliers. To construct such a database we start with the S&P Capital IQ database, which contains firm-level data on industry of affiliation, sales and networks of suppliers. We then use an unsupervised machine learning algorithm (n-gram) to merge these data with the PATSTAT database on patent applications.

We use patent applications to define whether or not an industry is ideas-oriented. This is implemented as follows. We start with a random subset of 6,000 patent applications and, using their texts, we hand code a training set that assigns each patent a binary classifier that equals 1 if it "improves the performance of an existing function/product or introduces a new function/product" and equals 0 if it "improves production efficiency or reduces production costs." We use this to train a neural network model called a multi-layer perceptron (MLP) with 4 layers, 16 neurons per layer and a 20% dropout rate. Our trained MLP has an accuracy rate of over 85%. The model is then applied to the 29,666,609 PATSTAT patent applications taken out by firms that have been matched to the S&P Capital IQ database. Finally, an industry is classified as ideas-oriented if a large fraction of its patents improve functionality. The resulting classification is sensible. For example, cell phones and autos are classified as ideas-oriented whereas energy and materials are classified as cost-oriented.

The final database contains 251,484 companies that hold 29,666,609 patent applications, have on average 5.30 suppliers, and integrate on average 55% of these suppliers.

Armed with these data we examine our two predictions about heterogeneous firms in ideas-oriented industries. Both are supported. The rest of the paper is organized as follows. Section 2 develops the theory, Section 3 describes the data, Section 4 reports our empirical findings and Section 5 concludes the paper.

2. Theory

2.1. Setup

2.1.1. Preferences and Production

Consider a final good sector with a continuum of varieties. The representative consumer's preference is:

$$U = \left\{ \int_{\omega \in \Omega} [\varphi(\omega)^\nu y(\omega)]^{(\sigma-1)/\sigma} d\omega \right\}^{\sigma/(\sigma-1)},$$

where ω is a variety index. Ω is the set of varieties available to this consumer. $y(\cdot)$ is the consumer's consumption level of a variety. $\varphi(\cdot)^\nu$ is a demand shifter (ν is a parameter and φ is explained in detail below). σ is the elasticity of substitution. We assume $\sigma > 1$ and $\nu(\sigma - 1) \in (0,1)$.

Production of a variety has three stages. The firm first decides on a level of multi-functionality N , that is, on the number of functions the product will have. Second, the firm identifies N suppliers, each of which will help the firm develop one of the functions. This blueprint or 'ideas' stage involves non-contractible, relationship-specific inputs from both the firm and the supplier. Third, in the 'production' stage the final good is produced in a complete-contracting environment. The ideas stage is the key stage and we discuss it in detail next.

In the ideas stage, each function is developed using the shared inputs of the firm and the supplier. For simplicity, we assume that each function is developed by the firm with the help of a single supplier.³ A function can be of variable quality. For example, facial recognition is better in some cell phones than in others and compressors are better in some refrigerators than in others. Let q_j be the quality of function $j = 1, \dots, N$. It depends on the firm's input h_j and the seller's input m_j :

$$q_j = h_j^\eta m_j^{1-\eta} / \hat{\eta},$$

where $\hat{\eta} \equiv \eta^\eta (1 - \eta)^{1-\eta}$. $0 < \eta < 1$. Quality q_j and inputs (h_j, m_j) are non-contractible.

Consumer valuation of functionality and function quality is captured by the demand shifter

$$\varphi = D(N, \theta) \min\{q_1, q_2, \dots, q_N\}, \tag{1}$$

where $\theta \in [0,1]$ is a firm index that replaces ω ; it plays no role yet, but we will later interpret it as the firm's productivity as in Melitz (2003).

³It is possible to allow for multiple suppliers of a single function.

The particular functional form in equation (1) is not all that important to our argument. Similar results hold with an O-Ring production function. They also hold for a CES production function provided that one function is not too substitutable for another.⁴

The marginal cost of input $j \in \{h, m\}$ is $C_j(N, \theta)$. For simplicity, we assume that $C_j(N, \theta) = w_j C(N, \theta)$, where the constant w_j ($j = h, m$) captures the price of inputs and other things that are log-separable from N and θ . Note that both D and C depend on θ . Not surprisingly, we will find (roughly) that only D/C matters. This is the usual point that demand shifters and productivity are isomorphic.

The inverse demand for a final product is

$$y = A\varphi^\alpha p^{-\sigma},$$

where $\alpha \equiv \nu(\sigma - 1) \in (0, 1)$, and A is a collection of industry and country characteristics.

The firm is a monopolistic competitor and sets price equal to $[\sigma/(\sigma - 1)]c$, where c is the marginal cost for producing the final product. This generates the following revenue function:

$$R = \hat{A}\varphi^\alpha = \hat{A}[D(N, \theta) \min\{q_1, q_2, \dots, q_N\}]^\alpha \quad (2)$$

where $\hat{A} \equiv \sigma^{-\sigma}[(\sigma - 1)/c]^{\sigma-1}A$.

2.1.2. Timing

The production process is as follows. First, the firm and all the potential suppliers observe θ . The firm then chooses organizational form $k \in \{O, V\}$ (O is outsourcing and V is vertical integration), adopts technology N , and offers contract $\{\tau_j\}_{j=1}^N$, where τ_j is an upfront payment to supplier j . $\tau_j \in \mathbb{R}, \forall j$. A continuum of potential suppliers apply for the contract and the firm chooses N suppliers from them. The firm and the suppliers then simultaneously choose their investment levels $\{(h_j, m_j)\}_{j=1}^N$. After investments are made, the firm and the suppliers bargain over the division of future revenue. At this stage, the firm and the suppliers can decide to withdraw their investments. After the firm and the suppliers reach an agreement, ideas are generated (φ is determined). Output is produced and sold. Revenue is divided according to the bargaining agreement.

2.1.3. Holdup

We assume that in the bargaining stage, if supplier j decides to withdraw from the production process, the firm cannot use the input as efficiently. We model this by assuming that the quality of the input drops from q_j to $\Delta^k q_j$, where $k \in \{O, V\}$ and $0 \leq \Delta^O < \Delta^V < 1$. On the other hand, if the firm withdraws its investment for function j , q_j drops to 0 regardless of the organizational form k .

⁴To see this, note that under the symmetry that we impose below, $\min\{q_1, \dots, q_N\} = q$. Under CES, $\{\sum_{j=1}^N q_j^\beta\}^{1/\beta}$ becomes $N^{1/\beta}q$. Under O-Ring, $B(N)\Pi_{j=1}^N q_j$ becomes $B(N)q^N$. These different specifications affect the functional form of the optimal inputs (h_j, m_j) , but otherwise do not matter. See the online appendices A and B for proof.

2.2. Equilibrium

2.2.1. SSPE

We define the symmetric sub-game perfect equilibrium (henceforth SSPE) as a tuple $\{N, \tau, h, m\}$, where N is the firm's choice of functionality. In SSPE, τ is the firm's upfront payment to every supplier, that is, $\tau_j = \tau$ for $j = 1, \dots, N$. Similarly, h is the firm's investment for each function, and m is each supplier's investment. That is $(h_j, m_j) = (h, m)$, for $j = 1, \dots, N$.

SSPE can be characterized by backward induction as in AAH. Since this is familiar (and notationally difficult) territory, we jump immediately to the revenue in any SSPE.⁵ This is given by

$$R = \hat{A} \left\{ D(N, \theta) h^\eta m^{1-\eta} / \hat{\eta} \right\}^\alpha, \quad (3)$$

where \hat{A} and $\hat{\eta}$ are as previously defined.

Lemma 1 *In every SSPE, the firm's Shapley value under organizational form $k \in \{O, V\}$ is $\gamma^k(N)R$, where*

$$\gamma^k(N) = \frac{\delta^k N + 1}{N + 1},$$

and $\delta^k \equiv (\Delta^k)^\alpha$. Each seller's Shapley value is $(1 - \gamma^k(N))R/N$.

Proof Appendix B.⁶ ■

In AAH, the firm's share of revenue γ^k is independent of N . Here, organizations with more suppliers face larger holdup problems, as reflected in the fact that γ^k is decreasing in N . This has an important implication. If in our model γ^k were independent of N , then the choice of number of suppliers and the choice of organizational form would not interact. Specifically, the choice of organizational form would be determined as in Antràs (2003) or as in Antràs and Helpman (2004) with $f_V = f_O$, i.e., if η is large then all firms integrate and if η is small then all firms outsource. Here, a productive firm may want to have a large N that will lead to a smaller share of revenue (a small γ^k); the firm may find it optimal to offset this loss of revenue by moving from the O form to the V form, which has the effect of increasing the firm's revenue share from γ^O to γ^V .

2.2.2. Optimal Choice of Inputs

The firm and the suppliers' problems are familiar from Antràs (2003) and Antràs and Helpman (2004), henceforth AH. They simultaneously choose their investment levels taking the others' investment levels as given.

The firm's problem can be written as:

$$\max_{(h_1, h_2, \dots, h_N)} \gamma^k(N) \frac{\hat{A}}{\hat{\eta}^\alpha} \left[D(N, \theta) \min_{j=1, \dots, N} \{h_j^\eta m_j^{1-\eta}\} \right]^\alpha - w_h C(N, \theta) \sum_{j=1}^N h_j. \quad (\text{FP1})$$

⁵Appendix A proves the existence and uniqueness of SSPE.

⁶Appendix B derives the Shapley value under a Leontief production function. See the online appendix for the Shapley values under O-Ring and CES production functions.

Supplier j 's problem can be written as:

$$\max_{m_j} \frac{1 - \gamma^k(N)}{N} \frac{\hat{A}}{\hat{\eta}^\alpha} \left[D(N, \theta) \min_{j=1, \dots, N} \{h_j^\eta m_j^{1-\eta}\} \right]^\alpha - w_m C(N, \theta) m_j. \quad (\text{SP1})$$

We assume $\alpha \in (0, 1)$ so that the firm's problem (FP1) and the supplier's problem (SP1) are concave.

Lemma 2 *In every SSPE, the unique solution to (FP1) and (SP1) under organizational form k is*

$$\begin{aligned} h^k(N, \theta, \eta) &= \left\{ \frac{\alpha \hat{A}}{\hat{\eta}} \frac{D(N, \theta)^\alpha}{NC(N, \theta)} \left[\frac{\eta \gamma^k(N)}{w_h} \right]^{1-\alpha+\alpha\eta} \left[\frac{(1-\eta)(1-\gamma^k(N))}{w_m} \right]^{\alpha-\alpha\eta} \right\}^{1/(1-\alpha)} \\ m^k(N, \theta, \eta) &= \left\{ \frac{\alpha \hat{A}}{\hat{\eta}} \frac{D(N, \theta)^\alpha}{NC(N, \theta)} \left[\frac{\eta \gamma^k(N)}{w_h} \right]^{\alpha\eta} \left[\frac{(1-\eta)(1-\gamma^k(N))}{w_m} \right]^{1-\alpha\eta} \right\}^{1/(1-\alpha)} \end{aligned} \quad (4)$$

with $h^k(N, \theta, \eta)$ and $m^k(N, \theta, \eta)$ satisfying the following relationship:

$$\frac{h^k(N, \theta, \eta)}{m^k(N, \theta, \eta)} = \frac{\gamma^k(N)}{1 - \gamma^k(N)} \frac{\eta/w_h}{(1-\eta)/w_m}. \quad (5)$$

Proof Appendix C. ■

These are messy expressions, but ones that are not fundamentally new. The only new insight comes from equation (5): h/m will vary *within an industry* not only because different firms choose different organizational forms k , but also because they choose different-sized organizations which affect h^k/m^k via the effects of N on γ^k . Thus, our framework offers a natural explanation of the enormous within-industry heterogeneity in relationship-specific investments that we see in the data. There are two main (old) insights from equation (4). First and obviously, the optimal input levels are both less than the first-best (contractible) input levels, as summarized by the product of the exponents of γ^k and $(1 - \gamma^k)$. Second, h^k/m^k equals the first-best input ratio if and only if $\gamma^k = 1/2$. This points to how the Grossman-Hart logic plays out in this model. When η is large so that the firm's investment is most important, the firm wants to choose a form that will raise h^k/m^k . This is the form with the larger γ^k and, since $\gamma^V > \gamma^O$, vertical integration is preferred.

2.2.3. Optimal Choice of Scope and Organizational Form

Rewriting the firm's problem in equation (FP1) with the optimal inputs from Lemma 2 generates the firm's surplus $\Pi^k(N, \theta, \eta)$. The firm designs a contract (or blueprint) to maximize its surplus:

$$\max_{k \in \{O, V\}, N \in [1, \infty)} \Pi^k(N, \theta, \eta) = \tilde{A} G(N, \theta) \Psi(\gamma^k(N), \eta), \quad (\text{FP2})$$

where $\tilde{A} \equiv \hat{A}^{1/(1-\alpha)}$,

$$G(N, \theta) \equiv \left[\frac{D(N, \theta)}{NC(N, \theta)} \right]^{\frac{\alpha}{1-\alpha}},$$

and

$$\Psi(\gamma, \eta) \equiv \frac{1 - \alpha[\gamma\eta + (1 - \gamma)(1 - \eta)]}{[\gamma^\eta(1 - \gamma)^{1-\eta}]^{-\frac{\alpha}{1-\alpha}}}.$$

It is now apparent that only $D/(NC)$ matters, not D or NC separately.⁷ Note that up to this point we have not said anything about θ . It is now clear that the appropriate assumption is that G is increasing in θ .

Assumption 1 $G(N, \theta)$ is strictly increasing in θ .

This is a good spot to compare our model with that of Antràs and Helpman (2004), equation (10). Their model has an almost identical profit function: In our notation it is basically $\Pi^k(1, \theta, \eta) = \theta^{\sigma-1} \Psi(\gamma^k, \eta)$ where, as is standard in Melitz-like models, $G(1, \theta) = \theta^{\sigma-1}$. However, there are three differences to note:

1. N is a choice variable.
2. There are no fixed costs of organizations (f_V and f_O in their notation). Recall that in their model, when there are no fixed costs as is the case here (or even when there are fixed costs and $f_V = f_O$) then their model reduces to Antràs (2003). That is, when η is small all firms outsource and when η is large all firms vertically integrate.
3. The most important difference is that $\Psi(\gamma^k(N), \eta)$ depends on N . In Antràs (2003) or Antràs and Helpman (2004) with $f_V = f_O$, the firm chooses the organizational form k that maximizes $\Psi(\gamma^k(1), \eta)$ where $\gamma^k(1)$ and η are parameters. In our setting, the larger is the organization (N), the smaller is $\gamma^k(N)$. This creates a tension: the firm might want to increase the number of functions N in order to increase demand, but this weakens the firm's bargaining power $\gamma^k(N)$. In the next section, we show how this leads to within-industry heterogeneity of organizational forms even though there are no fixed costs.

This is also a good spot to compare our profit function to that in AAH. First, in AAH only the supplier makes a relationship-specific investment ($\eta = 0$) so that the firm always outsources. Second and more importantly, in AAH the Shapley value is completely determined by exogenous parameters so that there is no trade-off between size and holdup, i.e., $\gamma^k(N)$ is independent of N in AAH but decreasing in N in our model.

We now make assumptions that make it easier to solve for the optimal N . We will use first-order conditions and so ignore the integer constraint on N . The following assumption ensures that for each choice of k , there is a unique number of suppliers that is bounded away from 1 and infinity.

Assumption 2 $G(N, \theta)$ satisfies the following conditions:

1. $G(N, \theta)$ is strictly log-concave in N : $\frac{\partial^2 \ln G(N, \theta)}{\partial N^2} < 0$.
2. $\lim_{N \rightarrow 1} \frac{\partial \ln G(N, \theta)}{\partial \ln N} > \frac{1}{2}$.
3. $\lim_{N \rightarrow \infty} \frac{\partial \ln G(N, \theta)}{\partial \ln N} < 0$.

Note that some of our main results rely on monotone comparative static arguments and thus do not require convexity or uniqueness.

⁷Note that in the expressions for h^k and m^k in equation (4), what matters is D^α/NC , so D and NC matter separately. However, they only matter for the levels of h^k , m^k and hence for quality q_j . They do not matter separately for anything else whatsoever. See Appendix C for the expressions for Qual, demand shifter, and revenue.

2.3. Two Types of Industries

There are two types of industries, ideas-oriented and cost-oriented. In ideas-oriented industries, consumers highly value multi-functionality N so that $D_N > 0$ is salient. Further, high-productivity firms develop the best functions in the sense that each function generates a high marginal revenue conditional on the same N . Mathematically, D_N is increasing in θ or $D(N, \theta)$ is log-supermodular in (N, θ) .⁸ One can get at this same notion of ideas-oriented industries from the cost side by noting that in these industries, high-productivity firms are really good at managing the integration of complex designs. With complex designs, more functions raise the marginal costs for each supplier because each firm-supplier pair must ensure its design is compatible with all the other suppliers' designs. That is $C_N > 0$. Moreover, this problem is less salient for more productive firms. That is $C(N, \theta)$ is log-submodular in (N, θ) . Thus in ideas-oriented industries, more productive firms hire managers who are better at keeping cost low. Whether tackled from the demand side or the supply side, both imply the following:

Definition 1 *Ideas-oriented industries are industries where $G(N, \theta)$ is log-supermodular in (N, θ) .*

In cost-oriented industries, high multi-functionality comes with a complicated production chain that involves many steps. A productive firm does not get a big bang for its multi-functionality, rather, a complex production network comes at great management cost. Mathematically, D_N is non-increasing in θ and $C_N > 0$ is salient. A better manager is able to reduce production cost by reducing the size of the production network. That is $C(N, \theta)$ is log-supermodular in (N, θ) .

Definition 2 *Cost-oriented industries are industries where $G(N, \theta)$ is log-submodular in (N, θ) .*

2.3.1. Ideas-Oriented Industry

Taking the log-transformation of the firm's problem in (FP2) yields

$$\max_{k \in \{O, V\}, N \in [1, \infty)} \pi^k(N, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi(\gamma^k(N), \eta), \quad (\text{fp1})$$

where $\pi^k(N, \theta, \eta) \equiv \ln \Pi^k(N, \theta, \eta)$, $\tilde{a} \equiv \ln \tilde{A}$, $g(N, \theta) \equiv \ln G(N, \theta)$, and $\psi(\gamma, \eta) \equiv \ln \Psi(\gamma, \eta)$. Since the transformation from (FP2) to (fp1) is monotone, the optimal k and N that solve (FP2) also solve (fp1).

By choosing $k \in \{O, V\}$, the firm is indirectly choosing the value of $\delta^k \in \{\delta^O, \delta^V\}$. To find the optimal δ^k we adopt the methodology used in Antràs and Helpman (2004), where we begin by allowing the firm to treat δ as a continuous variable on the interval $(0, 1)$. Then (fp1) generalizes to

$$\max_{\delta \in (0, 1), N \in [1, \infty]} \pi(N, \delta, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi(\gamma(N, \delta), \eta), \quad (\text{fp2})$$

⁸We assume log-supermodularity instead of supermodularity for $D(N, \theta)$ (and log-submodularity instead of submodularity for $C(N, \theta)$) because the former is more convenient for our derivation. Mathematically, log-supermodularity implies supermodularity when $D_\theta D_N / D^2 \geq 0$, and log-submodularity implies submodularity when $C_\theta C_N / C^2 \leq 0$. For these two conditions to hold, we need only assume that $D_\theta \geq 0$ and $C_\theta \leq 0$, which are consistent with the definition of θ as firm productivity.

where $\gamma(N, \delta) \equiv \frac{\delta^\alpha N + 1}{N + 1}$. By Assumption 1, $G(N, \theta)$ is log-supermodular in (N, θ) , so $g(N, \theta)$ is supermodular in (N, θ) . In equation (fp2), N and θ jointly appear in $g(N, \theta)$ only, so the log-profit function $\pi(N, \delta, \theta, \eta)$ is supermodular in (N, θ) .

In ideas-oriented industries, the profit function is log-supermodular in (N, θ) , meaning a more productive firm has a higher profit margin from a larger N . However, the firm's revenue share $\gamma^k(N)$ is decreasing in N . Therefore, a more productive firm is more likely to choose $k = V$ because integration helps mitigate the firm's loss in revenue share from a larger N . This tension only works in certain industries. In industries with extremely low η , suppliers' inputs are extremely important. The firm will always find it optimal to incentivize its suppliers by outsourcing. In industries with extremely high η , the firm's inputs are much more important than the suppliers. The firm always chooses $k = V$ to incentivize itself. Therefore, there are two threshold values of η , η_{io}^L and η_{io}^H with $\eta_{io}^L < \eta_{io}^H$, such that in industries with $\eta < \eta_{io}^L$, firms always choose $k = O$ regardless of their productivity levels. In industries with $\eta > \eta_{io}^H$, firms always choose $k = V$ regardless of its productivity. In those industries in between, a more productive firm has a larger N and so chooses $k = V$ to compensate its lower revenue share. A less productive firm has a smaller N and so has a higher revenue share even with $k = O$. This intuition is formalized in Theorem 1.

Theorem 1 *In an ideas-oriented industry, there exist two threshold values of η , η_{io}^L and η_{io}^H , with $0 < \eta_{io}^L < \eta_{io}^H < 1$, such that:*

1. *In industries with $\eta < \eta_{io}^L$, all firms choose outsourcing;*
2. *In industries with $\eta > \eta_{io}^H$, all firms choose vertical integration;*
3. *In industries with $\eta \in (\eta_{io}^L, \eta_{io}^H)$, there exists a threshold $\tilde{\theta}_{io}(\eta)$, such that*
 - (a) *firms with $\theta < \tilde{\theta}_{io}(\eta)$ choose outsourcing,*
 - (b) *firms with $\theta > \tilde{\theta}_{io}(\eta)$ choose vertical integration,*
 - (c) *$\tilde{\theta}_{io}(\eta)$ is strictly decreasing in η .*

Proof Appendix D. ■

Compared to Antràs (2003), Antràs and Helpman (2004) and AAH, the within-industry heterogeneity in organizational forms in our model does not rely on the assumptions on fixed costs of production. In Antràs (2003), all firms outsource in low η industries and integrate in high η industries. In Antràs and Helpman (2004), productive firms integrate because integration brings higher variable profit that outweighs the high fixed costs. In Acemoglu et al. (2007), a firm never chooses integration because all relationship-specific investments are made by the suppliers, not the firm.

We now focus on the ideas-oriented industries with heterogeneous organizational forms, i.e., industries with $\eta \in (\eta_{io}^L, \eta_{io}^H)$. The firm's problem in (fp1) can be broken down into two steps. First, the firm chooses an optimal N for each organizational form $k \in \{O, V\}$. Denote this choice by $N_{io}^k(\theta, \eta)$, and the resulting profit by $\pi_{io}^k(N_{io}^k(\theta, \eta), \theta, \eta)$. The firm then compares its profits under

$k = O, V$, and chooses the k that brings it the higher profit. Denote this optimal solution to (fp1) by $N_{io}^*(\theta, \eta)$. Denote the corresponding revenue shares under the three scenarios (outsourcing, integration and optimal) by $\gamma_{io}^O(\theta, \eta)$, $\gamma_{io}^V(\theta, \eta)$ and $\gamma_{io}^*(\theta, \eta)$.

Theorem 2 *In an ideas-oriented industry with $\eta \in (\eta_{io}^L, \eta_{io}^H)$, the following results are true:*

1. $N_{io}^O(\theta, \eta)$, $N_{io}^V(\theta, \eta)$ and $N_{io}^*(\theta, \eta)$ are strictly increasing in θ .
2. $\gamma_{io}^O(\theta, \eta)$, $\gamma_{io}^V(\theta, \eta)$ and $\gamma_{io}^*(\theta, \eta)$ are strictly decreasing in θ .
3. $N_{io}^O(\tilde{\theta}_{io}(\eta), \eta) < N_{co}^V(\tilde{\theta}_{io}(\eta), \eta)$ and $\gamma_{io}^O(\tilde{\theta}_{io}(\eta), \eta) > \gamma_{io}^V(\tilde{\theta}_{io}(\eta), \eta)$.

Proof Appendix D. ■

Parts 1 and 2 of Theorem 2 capture the key tradeoff of the paper: A more productive firm chooses a larger scope (a larger N), but also faces a more severe holdup problem (a lower γ). Part 3 deals with a firm that is just indifferent between the two organizational forms. By Theorem 1, this firm has productivity $\theta = \tilde{\theta}_{io}(\eta)$. As the firm moves from O to V , two offsetting things happen to its revenue share. The direct effect is the improved outside option ($\delta^O < \delta^V$), which raises its share of revenue. The indirect effect is that the firm expands its organization ($N^O < N^V$) which lowers the firm's share of revenue. Part 3 states that the indirect effect dominates, meaning the revenue share is lower under V . Part 3 is ancillary to parts 1 and 2.

2.3.2. Cost-Oriented Industries

In cost-oriented industries, a high-productivity firm features a manager that is good at cutting costs. The profit function is log-submodular in N and θ , meaning a high-productivity (high θ) firm benefits more from a smaller organization (N). A smaller N increases $\gamma^k(N)$. The firm finds it less compelling to choose $k = V$ to compensate for its loss of bargaining power. Therefore, high-productivity firms choose smaller, more outsourced production networks. Again, this is only true in industries where η is neither too high nor too low. When η is close to 1, the firm may always find it optimal to choose $k = V$; when η is close to 0, the firm may always find it optimal to choose $k = O$. This intuition is formally stated in the following theorem.

Theorem 3 *In a cost-oriented industry, i.e., when $G(N, \theta)$ is log-submodular in (N, θ) , there exist two threshold values of η , η_{co}^L and η_{co}^H , with $0 < \eta_{co}^L < \eta_{co}^H < 1$, such that:*

1. In industries with $\eta < \eta_{co}^L$, all firms choose outsourcing;
2. In industries with $\eta > \eta_{co}^H$, all firms choose vertical integration;
3. In industries with $\eta \in (\eta_{co}^L, \eta_{co}^H)$, there exists a threshold $\tilde{\theta}_{co}(\eta)$, such that
 - (a) firms with $\theta > \tilde{\theta}_{co}(\eta)$ choose outsourcing,
 - (b) firms with $\theta < \tilde{\theta}_{co}(\eta)$ choose vertical integration,
 - (c) $\tilde{\theta}_{co}(\eta)$ is strictly increasing in η .

Now focus on the cost-oriented industries with heterogeneous organizational forms, i.e., industries with $\eta \in (\eta_{co}^L, \eta_{co}^H)$. Denote firm θ 's optimal choice of N under $k \in \{O, V\}$ by $N_{co}^k(\theta, \eta)$, and the resulting profit by $\pi_{co}^k(N_{co}^k(\theta, \eta), \theta, \eta)$. Similar to the ideas-oriented industries, the firm compares its profit under $k = O, V$, and chooses the k that brings it a higher profit. Denote this optimal solution by $N_{co}^*(\theta, \eta)$ and the revenue shares under the three scenarios (outsourcing, integration and optimal) by $\gamma_{co}^O(\theta, \eta)$, $\gamma_{co}^V(\theta, \eta)$ and $\gamma_{co}^*(\theta, \eta)$.

Theorem 4 *In a cost-oriented industry with $\eta \in (\eta_{co}^L, \eta_{co}^H)$, the following results are true:*

1. $N_{co}^O(\theta, \eta)$, $N_{co}^V(\theta, \eta)$ and $N_{co}^*(\theta, \eta)$ are strictly decreasing in θ .
2. $\gamma_{co}^O(\theta, \eta)$, $\gamma_{co}^V(\theta, \eta)$ and $\gamma_{co}^*(\theta, \eta)$ are strictly increasing in θ .
3. $N_{co}^O(\tilde{\theta}_{co}(\eta), \eta) > N_{co}^V(\tilde{\theta}_{co}(\eta), \eta)$ and $\gamma_{co}^O(\tilde{\theta}_{co}(\eta), \eta) < \gamma_{co}^V(\tilde{\theta}_{co}(\eta), \eta)$.

Parts 1 and 2 of Theorem 4 is a description of *lean-and-mean* production. High-productivity firms feature managers that can consolidate the production process (reducing N), which mitigates the holdup problem (a higher γ). Part 3 deals with a firm that is indifferent between the two organizational forms (the firm with productivity $\theta = \tilde{\theta}_{co}(\eta)$). As the firm moves from O to V , the direct effect is improved outside option ($\delta^O < \delta^V$), which raises the firm's revenue share. The indirect effect is a smaller organization ($N^O > N^V$), which also raises the firm's revenue share. Both effects imply that the firm's revenue share increases.

3. Data and Descriptive Statistics

We compile a novel dataset from two sources. (1) PATSTAT – a patent database with patent applications from 194 patenting authorities around the world.⁹ (2) S&P Capital IQ – a global database with information on companies' industry affiliation, financial statement variables, and relationship information.¹⁰ We use the first database to define ideas-oriented and cost-oriented industries, and the second database to gather firm-level financial and production-network characteristics.

The dataset is constructed in two steps. First, we use an approximate string matching algorithm that links 29,966,609 PATSTAT patent applications to 251,484 S&P Capital IQ companies,¹¹ which is 55% of all PATSTAT patents owned by companies.¹² We then train a neural network model called multilayer perceptron (MLP) to classify the matched patents. To this end we start

⁹The top patenting authorities in PATSTAT include the Japanese Patent Office (19,779,900 patent applications), the USPTO (15,161,843 patent applications), the Chinese National Intellectual Property Administration (14,535,117 patent applications), the German Patent and Trade Mark Office (7,424,621 patent applications), the Korean Intellectual Property Office (3,810,155 patent applications), the UK Patent Office (3,440,561 patent applications), and the European Patent Office (3,397,668 patent applications). The period of coverage begins at different times for different patenting authorities, and ends in January, 2018.

¹⁰The period of coverage for the financial data is 2009-2016. The period of coverage for the industry affiliation and relationship data is 2010-2017.

¹¹The majority of these companies come from the U.S., China, Japan, and Europe. Table A2 in Appendix F summarizes the geographic location of these companies.

¹²Other patent holders include individuals, governments, and universities.

with a random sample of 6,000 PATSTAT patent applications and, using their texts, we hand code a training set that assigns each patent a binary classifier that equals 1 if it “improves the performance of an existing function/product or introduces a new function/product,” and equals 0 if it “improves production efficiency or reduces production costs.” We use this hand-coded sample to train an MLP with 4 layers, 16 neurons per layer and a dropout rate of 20%.¹³ Our trained model has an accuracy rate of over 85%. The model is then applied to the 29,966,609 PATSTAT patent applications taken out by firms that have been matched to the S&P Capital IQ database. The matching and classification procedures are further explained in Appendix F and Appendix G, and elaborated in the online appendix.

Since the S&P Capital IQ database reports each company’s industry, we can compute the fraction of patents in each industry that improve the performance of an existing function/product or introduce a new function/product. To help readers understand this, we start with S&P Capital IQ aggregation of industries into 10 sectors.¹⁴ In column 1 of Table 1, the sectors are ranked by sectors’ fractions of such patents. At the 10-sector level, we refer to a sector as ideas-oriented if its fraction is above 0.60 and cost-oriented if the fraction is below 0.50. We define two binary variables at the sector level: IdeaDummy equals 1 if the fraction is above 0.60 and equals 0 if the fraction is below 0.60; CostDummy equals 1 if the fraction is below 0.50 and equals 0 if the fraction is above 0.50.

Table 1: Firm Innovation and Production Network Characteristics

Sector	Fraction of Ideas-oriented Patents	Patents/Revenue	R&D/Revenue	Partners	Fraction of Integrated Partners
	(1)	(2)	(3)	(4)	(5)
Ideas-oriented					
Telecom. Services	0.68	0.09	0.54%	9.90	0.45
Consumer Discretionary	0.63	0.22	1.53%	4.87	0.47
Information Technology	0.60	0.30	5.57%	7.74	0.43
Healthcare	0.60	0.20	5.15%	5.02	0.40
Intermediate					
Financials	0.58	0.12	0.10%	3.80	0.76
Industrials	0.57	0.19	1.01%	5.13	0.55
Consumer Staples	0.55	0.06	0.34%	4.83	0.54
Cost-oriented					
Energy	0.50	0.04	0.22%	7.33	0.52
Materials	0.49	0.17	0.80%	4.86	0.58
Utilities	0.48	0.04	0.13%	6.50	0.25
Total	0.57	0.15	1.25%	5.30	0.55

Notes: This table reports statistics on the 251,484 S&P Capital IQ companies that are matched with PATSTAT companies. R&D and revenue are averaged over the period 2009-2016 at historical rates and measured in millions of U.S. dollars.

¹³These hyperparameters are chosen based on training and validation accuracy rates. See the online appendix for more information.

¹⁴The 10 sectors are listed in Table 1 and come from the Global Industry Classification Standard (GICS), which is the generic industry classification in S&P Capital IQ. GICS (2017 version) contains 10 sectors, 24 industry groups, 67 industries, and 156 subindustries.

Table 1 provides several indications that our classification is sensible. First, the highest fractions are for Telecommunications Services (e.g., AT&T in mobile phones and Comcast in cable TV), Information Technologies (e.g., Samsung and Apple) and Consumer Discretionary (e.g., automobiles, auto components, multiline retail, and household durables). These are sectors that are consumer-facing and whose firms are sensitive to designing the right bundles for consumers. The lowest fractions are for Energy (e.g., oil, gas, and consumable fuels), Materials (e.g., chemicals, metals, and mining) and Utilities (e.g., electric and gas utilities), sectors that tend to be less sensitive to consumer demand for multi-function products.

A second indication that our classification is sensible is given by columns (2) and (3) in Table 1. Compared to companies in the cost-oriented group, companies in the ideas-oriented group tend to have higher innovation intensities as measured by patent-to-revenue and R&D-to-revenue ratios. The average patent-to-revenue ratio is 0.20 in the ideas-oriented group and only 0.08 in the cost-oriented group. The average R&D-to-revenue ratio is 3.20% in the ideas-oriented group and only 0.38% in the cost-oriented group. Thus, ideas-oriented sectors tend to be more innovation-intensive relative to cost-oriented sectors.

IdeaDummy and CostDummy will be central to our regression specification below. We also define them at finer levels of industry aggregations. At the level of 67 industries, we compute the total number of patents held by companies in each industry, and the fraction of these patents that improve the quality of an existing function/product or introduce a new function/product. IdeaDummy equals 1 if an industry's fraction of such patents is above the mean for all industries, and equals 0 otherwise. CostDummy equals 1 if an industry's fraction of such patents is below the mean for all industries, and equals 0 otherwise. We also consider using the median of all industries rather than the mean. In this case, IdeaDummy equals 1 if an industry's fraction of such patents is above the median for all industries, and equals 0 otherwise. CostDummy equals 1 if an industry's fraction of such patents is below the median for all industries, and equals 0 otherwise.¹⁵

Having constructed key variables IdeaDummy and CostDummy, we turn to data on production networks and vertical integration. We construct production network information for the 251,484 S&P Capital IQ companies that are matched with patent information.¹⁶ Each company's production network is composed of the focal company and its representative customers and suppliers over 2010-2017.¹⁷ In the theory section we defined a company's production network as consisting of itself and its suppliers; however, the empirical counterpart to a supplier is not immediate because a company's downstream customer could also be its "supplier." For example, consider the relationship between Apple headquarters and its retail arm, the Apple Store. The Store is an (upstream) customer in that it receives iPhones and it is a (downstream) supplier in

¹⁵We also define IdeaDummy and CostDummy at the levels of 24 industry groups and 156 subindustries. They generate similar empirical results as the sector- and industry-level variables.

¹⁶The production networks are constructed using S&P Capital IQ's business relationships data. Appendix H explains how we compile the business relationships data from S&P Capital IQ. Appendix I explains how we use the business relationships data to construct production networks.

¹⁷A customer is a company that purchases products or services from the focal company. A supplier is a company that sells its products or services to the focal company. The customer and supplier relationships are either reported by S&P Capital IQ or imputed from parent-subsidiary relationships. See Appendix I for more information.

that it supplies retail services. For this reason, we include both the company’s customers and suppliers in its production network.¹⁸ We refer to any company in a focal company’s production network other than itself as its partner. We refer to a firm’s number of partners as its scope. A partner is integrated if the focal company owns more than a 50% stake in the partner. Since we assume that the focal company is the one deciding whether to integrate its partners, the focal company’s owner (a company that owns more than 50% of the focal company) is excluded from its production network.¹⁹ The dataset contains 2,611,861 firm-partner relationships for 615,405 companies. 842,774 of these relationships are integrated. The average company has 4.24 partners, with 1.37 of them integrated. 73,914 of these companies are matched with patent information. Their production network characteristics are reported in columns (4) and (5) of Table 1.

The dataset does not contain information on a firm’s entire production network. Rather, it is a collection of the firm’s representative customer and supplier relationships.²⁰ To proceed, we need two minimal assumptions: (i) a company’s measured number of partners is proportional to its actual number of partners, and (ii) a company’s measured integration decision is driven by the same factors as all of its integration decisions.

4. Empirical Results

We examine two testable predictions from Theorems 1-4, one on integration decisions and the other on scope decisions, where scope is measured by a company’s number of partners.

Theorems 2 (part 2) and 4 (part 2) make predictions about scope decisions. Combined, they imply the following:

Hypothesis 1 (Firm scope with heterogeneity) *In ideas-oriented industries, a high-productivity firm is likely to have more partners than a low productivity firm. In cost-oriented industries, a high-productivity firm is likely to have fewer partners than a low productivity firm.*

Theorems 1 (part 3) and 3 (part 3) make predictions about integration decisions. Combined, they imply the following:

Hypothesis 2 (Firm boundaries with heterogeneity) *In ideas-oriented industries, a high-productivity firm is more likely to integrate its partner than a low productivity firm. In cost-oriented industries, a high-productivity firm is less likely to integrate its partner than a low productivity firm.*

¹⁸This is not important empirically. We can define a company’s downstream production network as consisting of itself and its customers, and its upstream production network as consisting of itself and its suppliers. In the empirical section, we focus on the production network defined in the main text. Repeating the empirical exercise for only the upstream or downstream production networks does not qualitatively change our empirical results.

¹⁹In actual execution, we also have to deal with those “partially owned” relationships where one company is the other’s limited partner, investor, or pending parent/investor. Appendix H elaborates on our treatment of such relationships.

²⁰2,227,879 customer and supplier relationships are used to construct production network characteristics. 954,410 of these relationships are collected by S&P Capital IQ from sources including companies’ 10K and annual reports, and newswires such as Bloomberg and Reuters. 1,286,076 relationships are imputed from parent-subsidiary relationships using S&P Capital IQ’s ownership data and BEA 2002 Input-Output Table (Acemoglu et al., 2010, Antràs and Chor, 2013). There is an overlap of 12,617 relationships between these two sources. See Appendix I for more information.

4.1. The Scope Decision

Hypothesis 1 predicts that a firm's scope decision depends on its productivity and the type of industry (ideas-oriented or not, cost-oriented or not) that it operates in. We test this hypothesis using the following system of equations:

$$\text{Partners}_{fi} = \alpha_1^I \ln(\text{Sales}_{fi}) + \alpha_2^I \text{IdeaDummy}_i + \alpha_3^I \ln(\text{Sales}_{fi}) * \text{IdeaDummy}_i + \beta^I X_{fi} + \gamma_{i'}^I + \varepsilon_{fi}^I \quad (6)$$

$$\text{Partners}_{fi} = \alpha_1^C \ln(\text{Sales}_{fi}) + \alpha_2^C \text{CostDummy}_i + \alpha_3^C \ln(\text{Sales}_{fi}) * \text{CostDummy}_i + \beta^C X_{fi} + \gamma_{i'}^C + \varepsilon_{fi}^C. \quad (7)$$

f is a firm index and i is either a sector index (10 sectors) or an industry index (67 industries). Note that each firm appears in only one sector or industry. Partners_{fi} is the number of partners firm f had over the period 2010-2017.²¹ It measures firm f 's scope decision. $\ln(\text{Sales}_{fi})$ is the log of firm f 's average sales over 2009-2016. Since state-of-the-art productivity measures (e.g., Orr, Trebler and Yu, 2019) require data that are very often missing, we use sales as a proxy for firm productivity.²² IdeaDummy_i and CostDummy_i are indicators for ideas-oriented and cost-oriented sectors or industries as defined in Section 3. X_{fi} is a vector of control variables including the log of firm f 's partners' average sales and the log of the average number of firm f 's partners' partners.²³ We include these variables in order to control for other production network characteristics that may be correlated with firm productivity. ε_{fi}^I and ε_{fi}^C are error terms. Finally, we include subindustry dummies (156 subindustries) $\gamma_{i'}^I$ and $\gamma_{i'}^C$ where i' indexes subindustries. Notice that IdeaDummy_i and CostDummy_i are subsumed in these subindustry dummies.

These specifications are difference-in-difference specifications. In the first equation we have $\ln(\text{Sales}_{fi})$, IdeaDummy_i (subsumed by the subindustry dummies) and their interaction $\ln(\text{Sales}_{fi}) * \text{IdeaDummy}_i$. The focus is on the interaction terms. That is, the coefficients of interests are α_3^I and α_3^C . According to Hypothesis 1, $\alpha_3^I > 0$ and $\alpha_3^C < 0$. Note that when they are defined using the industry mean, IdeaDummy_i and CostDummy_i always sum up to one, so that $\alpha_3^I + \alpha_3^C = 0$.

Table 2 reports the regression results. Panels (a) and (b) respectively correspond to equations (6) and (7). The positive coefficients on $\ln(\text{Sales}_{fi}) * \text{IdeaDummy}_i$ suggest that in ideas-oriented industries, high-productivity firms tend to have more partners (larger scope) than low-productivity ones. The negative coefficients on $\ln(\text{Sales}_{fi}) * \text{CostDummy}_i$ suggest that in cost-oriented industries, high-productivity firms tend to have fewer partners (smaller scope) than low-productivity firms. The results confirm Hypothesis 1.

The table has other results that are unrelated to Hypothesis 1. The coefficients on $\ln(\text{Sales}_f)$ are positive in both panels, suggesting that in both ideas-oriented and cost-oriented industries, high-productivity firms manage larger production networks. In addition, the positive coefficients

²¹More specifically, Partners_{fi} is the number of unique companies that have appeared as a customer or a supplier of the focal company during 2010-2017. We do not know the beginning and ending time of a customer or supplier relationship so an annual partner count does not have much advantage.

²²The financial variables are obtained from S&P Capital IQ, where non-revenue variables contain many missing values.

²³For example, if a firm has two partners, the first partner has 3 partners and the second partner has 5 partners, then the firm's average partner has 4 partners.

Table 2: Hypothesis 1: The Scope Decision (Number of Partners)

(a) Ideas-oriented Industries

IdeaDummy _i defined at:	sector	industry mean	industry median
	(1)	(2)	(3)
ln(Sales _{fi})	5.875*** (0.189)	6.662*** (0.241)	6.911*** (0.252)
ln(Sales _{fi})*IdeaDummy _i	4.695*** (0.576)	1.597*** (0.437)	1.591** (0.502)
ln(Average Partner Sales _{fi})	3.291*** (0.0705)	3.294*** (0.0700)	3.305*** (0.0695)
ln(Average Partners' Partners _{fi})	-2.096*** (0.120)	-2.036*** (0.120)	-2.042*** (0.120)
constant	-27.68*** (0.913)	-27.43*** (0.902)	-27.24*** (0.884)
Subindustry fixed effects	Y	Y	Y
Observations	55,353	55,353	55,353
R-squared	0.171	0.165	0.165

(b) Cost-oriented Industries

CostDummy _i defined at:	sector	industry mean	industry median
	(1)	(2)	(3)
ln(Sales _{fi})	8.223*** (0.290)	8.259*** (0.369)	8.321*** (0.422)
ln(Sales _{fi})*CostDummy _i	-3.124*** (0.379)	-1.597*** (0.437)	-1.319** (0.492)
ln(Average Partner Sales _{fi})	3.277*** (0.0704)	3.294*** (0.0700)	3.300*** (0.0697)
ln(Average Partners' Partners _{fi})	-2.001*** (0.121)	-2.036*** (0.120)	-2.039*** (0.120)
constant	-27.71*** (0.917)	-27.43*** (0.902)	-27.29*** (0.888)
Subindustry fixed effects	Y	Y	Y
Observations	55,353	55,353	55,353
R-squared	0.166	0.165	0.164

Notes: The dependent variable is a firm's number of partners. Panels (a) and (b) respectively report the results for equations (6) and (7). There are 156 subindustry fixed effects. The complete list of the GICS subindustries can be found at https://en.wikipedia.org/wiki/Global_Industry_Classification_Standard. Numbers in parentheses report robust standard errors. *p<0.05, ** p<0.01, *** p<0.001.

on ln(Average Partner Sales_f) and the negative coefficients on ln(Average Partner's Partners_f) in both panels suggest that connected firms (i.e., firms with more partners) tend to work with firms that are also connected. This is consistent with Bernard, Moxnes and Saito (2019) using Japanese production networks.²⁴

²⁴See also Bernard and Moxnes (2018) for a review of the literature on production networks.

4.2. Firm Boundaries/Integration Decision

Hypothesis 2 predicts that a firm's decision about integrating its partner depends on (1) the firm's productivity and (2) the industry's type (ideas-oriented or not, cost-oriented or not). Since the dependent variable is a firm's choice of whether or not to integrate its partner, we need to work at the firm-partner level. This raises a new issue because the theory assumes that all partners are symmetric, while in the data they definitely are not. So we need to deal with partner heterogeneity.

The most direct way is to allow partners to differ in productivity. Conditional on firm characteristics, a high-productivity partner is less likely to be integrated than a low-productivity one. To understand why, recall that a key parameter driving firm integration decisions is η , which measures the relative importance of the firm's non-contractible, relationship-specific investments in a firm-partner relationship. $1 - \eta$ is therefore the relative importance of the partner's non-contractible, relationship-specific investments. Introducing partner heterogeneity is similar to allowing η to vary across partners for a given firm. If a high-productivity partner also makes more non-contractible, relationship-specific investments, it is also more important in a firm-partner relationship. In other words, a high-productivity partner corresponds to a lower η . According to Theorems 2 and 4, the firm is less likely to integrate a partner with low η because integration hurts this important partner's investment incentive.

We specify the following system of equations to examine Hypothesis 2:

$$\begin{aligned} \text{Integration}_{fi,pi'} &= \alpha_1^I \ln(\text{Sales}_{pi'}) * \text{IdeaDummy}_i * \text{HighProd}_{fi} + \alpha_2^I \ln(\text{Sales}_{pi'}) * \text{HighProd}_{fi} \\ &+ \alpha_3^I \ln(\text{Sales}_{pi'}) * \text{IdeaDummy}_i + \alpha_4^I \text{IdeaDummy}_i * \text{HighProd}_{fi} + \alpha_5^I \text{IdeaDummy}_i + \alpha_6^I \text{HighProd}_{fi} \quad (8) \\ &+ \alpha_7^I \ln(\text{Sales}_{pi'}) + \beta^I X_{pi'} + \gamma_{fi}^I + \gamma_{i'}^I + \varepsilon_{fi,pi'}^I; \end{aligned}$$

$$\begin{aligned} \text{Integration}_{fi,pi'} &= \alpha_1^C \ln(\text{Sales}_{pi'}) * \text{CostDummy}_i * \text{HighProd}_{fi} + \alpha_2^C \ln(\text{Sales}_{pi'}) * \text{HighProd}_{fi} \\ &+ \alpha_3^C \ln(\text{Sales}_{pi'}) * \text{CostDummy}_i + \alpha_4^C \text{CostDummy}_i * \text{HighProd}_{fi} + \alpha_5^C \text{CostDummy}_i + \alpha_6^C \text{HighProd}_{fi} \\ &+ \alpha_7^C \ln(\text{Sales}_{pi'}) + \beta^C X_{pi'} + \gamma_{fi}^C + \gamma_{i'}^C + \varepsilon_{fi,pi'}^C. \quad (9) \end{aligned}$$

f and p are firm and partner indexes. i is an index for the firm's sector or industry. i' is an index for the partner's subindustry. $\text{Integration}_{fi,pi'}$ is a binary variable that equals 100 if partner p is integrated by firm f , and 0 otherwise. The dependent variable allows us to interpret the coefficients in terms of percentages. $\ln(\text{Sales}_{pi'})$ is the log of partner p 's average sales over the period 2009-2016. It is a proxy for the partner's productivity. IdeaDummy_i and CostDummy_i are respectively indicators for ideas-oriented and cost-oriented industries as defined in Section 3. HighProd_{fi} is a dummy variable that equals 1 if firm f 's sales is above the industry mean, and 0 otherwise. $X_{pi'}$ is a vector of control variables including the log of the average sales of all the firms that work with partner p , and the log of the average number of partners of these firms. They control for the partner's production network characteristics. $\varepsilon_{fi,pi'}^I$ and $\varepsilon_{fi,pi'}^C$ are error terms. Note that these are relationship-level specifications so a firm can appear in multiple regressions. We use γ_{fi}^I and γ_{fi}^C to control for firm fixed effects. We also include the partner's subindustry dummies $\gamma_{i'}^I$ and $\gamma_{i'}^C$.

The specification in equation (8) includes all possible interactions between $\ln(\text{Sales}_{pi'})$, IdeaDummy_i , and CostDummy_i . However, the double-interaction term $\text{IdeaDummy}_{fi} * \text{HighProd}_{fi}$ and the individual terms IdeaDummy_i^I and HighProd_{fi}^I are absorbed by the firm fixed effect γ_{fi}^I . Similarly, in equation (9) the double-interaction term $\text{CostDummy}_{fi} * \text{HighProd}_{fi}$ and the individual terms CostDummy_i^C and HighProd_{fi}^C are absorbed by the firm fixed effect γ_{fi}^C .

Of all these interactions it is the triple-interaction terms that are the key variables of interest. According to Hypothesis 2, we expect to see two patterns: (i) In an ideas-oriented industry, conditional on partner characteristics, a high-productivity firm is more likely to integrate its partner than a low-productivity firm; (ii) In a cost-oriented industry, conditional on partner characteristics, a high-productivity firm is less likely to integrate its partner than a low-productivity firm. In other words, we expect to see $\alpha_1^I > 0$ and $\alpha_1^C < 0$.

Table 3 reports the regression results. Panels (a) and (b) respectively correspond to equations (8) and (9). Consistent with Hypothesis 2, $\alpha_1^I > 0$ and $\alpha_1^C < 0$. This is the main result in the table.

In addition, the coefficients on the $\ln(\text{Average Firm Sales}_{pi'})$ are negative in both panels, and the coefficients on $\ln(\text{Average Firm Partners}_{pi'})$ are positive in both panels. The results suggest that a partner that works with large firms is less likely to be integrated, but a partner that works with more connected firms is more likely to be integrated. These findings have not been previously documented.

5. Conclusions

One of the firm's most important tasks is to design a product and bring it to market, that is, to identify a cluster of functions which consumers value and then develop and bundle the functions into a final product. Examples abound, from high-tech smartphones to lowly refrigerators. Since these functions are often technologically distinct, we assumed that the firm needs a separate supplier for each function. We also assumed that for each function both the firm and supplier make non-contractible, relationship-specific investments that create a two-sided holdup problem. Finally, we assumed that once a product is designed each function (supplier) is essential. As a result the firm's Shapley value as share of total revenue or bargaining power is decreasing in the number of suppliers. This sets up a trade-off between the number of functions and bargaining power.

Following Antràs (2003), we modeled vertical integration as the firm's way of increasing its bargaining power relative to outsourcing. While greater scope as measured by the number of functions/suppliers reduces the firm's bargaining power, this can be partially offset by the appropriate choice of vertical integration or outsourcing. The starting point of this paper is this link between firm scope (number of functions/suppliers) and firm boundaries (choice between vertical integration and outsourcing).²⁵

²⁵Our approach built on Acemoglu et al. (2007), but there the firm's bargaining power is independent of the number of suppliers. They thus do not consider our core mechanism. Indeed, they are more interested in technology adoption (what we called the adoption of a function) than in the choice between vertical integration and outsourcing.

Table 3: Hypothesis 2: The Firm Boundaries Decision (Integration vs. Outsourcing)

(a) Ideas-oriented Industries

IdeaDummy _i defined at:	sector-level (1)	industry mean (2)	industry median (3)
$\ln(\text{Sales}_{pi}) * \text{IdeaDummy}_i * \text{HighProd}_{fi}$	0.622*** (0.0641)	0.311*** (0.0669)	0.378*** (0.0635)
$\ln(\text{Sales}_{pi}) * \text{HighProd}_{fi}$	-1.314*** (0.0505)	-1.242*** (0.0544)	-1.223*** (0.0455)
$\ln(\text{Sales}_{pi}) * \text{IdeaDummy}_i$	0.835*** (0.0474)	0.271*** (0.0491)	0.271*** (0.0465)
$\ln(\text{Sales}_{pi})$	-1.763*** (0.0375)	-1.456*** (0.0402)	-1.418*** (0.0336)
$\ln(\text{Average Firm Sales}_{pi})$	-3.588*** (0.0423)	-3.569*** (0.0423)	-3.555*** (0.0423)
$\ln(\text{Average Firm Partners}_{pi})$	0.944*** (0.0784)	0.774*** (0.0783)	0.762*** (0.0783)
constant	48.71*** (0.376)	49.22*** (0.377)	49.19*** (0.376)
Firm fixed effects	Y	Y	Y
Partner subindustry fixed effects	Y	Y	Y
Observations	443,636	443,636	443,636
R-squared	0.461	0.459	0.459

(b) Cost-oriented Industries

CostDummy _i defined at:	sector-level (1)	industry mean (2)	industry median (3)
$\ln(\text{Sales}_{pi}) * \text{CostDummy}_i * \text{HighProd}_{fi}$	-0.101 (0.0964)	-0.311*** (0.0669)	-0.378*** (0.0635)
$\ln(\text{Sales}_{pi}) * \text{HighProd}_{fi}$	-1.039*** (0.0346)	-0.931*** (0.0397)	-0.845*** (0.0451)
$\ln(\text{Sales}_{pi}) * \text{IdeaDummy}_i$	-0.249*** (0.0712)	-0.271*** (0.0491)	-0.271*** (0.0465)
$\ln(\text{Sales}_{pi})$	-1.251*** (0.0257)	-1.186*** (0.0296)	-1.147*** (0.0334)
$\ln(\text{Average Firm Sales}_{pi})$	-3.579*** (0.0423)	-3.569*** (0.0423)	-3.555*** (0.0423)
$\ln(\text{Average Firm Partners}_{pi})$	0.782*** (0.0785)	0.774*** (0.0783)	0.762*** (0.0783)
constant	49.23*** (0.377)	49.22*** (0.377)	49.19*** (0.376)
Firm fixed effects	Y	Y	Y
Partner subindustry fixed effects	Y	Y	Y
Observations	443,636	443,636	443,636
R-squared	0.459	0.459	0.459

Notes: The dependent variable is a binary variable that equals 100 if a firm-partner relationship is integrated, and 0 other wise. Panels (a) and (b) respectively report the results for equations (8) and (9). 25,834 singleton observations were dropped. Numbers in parentheses report robust standard errors. *p<0.05, ** p<0.01, *** p<0.001.

Our main interest flowed from the empirical observation that the number of functions varies across products within an industry (firm heterogeneity). We introduced the notion of an ‘ideas-oriented’ industry as one in which more productive firms have higher marginal returns to introducing a new function. This leads to two testable hypotheses.

1. Firm scope with heterogeneity: In ideas-oriented industries, more productive firms will have more suppliers.
2. Firm boundaries with heterogeneity: In ideas-oriented industries, more productive firms will be more likely to integrate its suppliers.

In contrast, in cost-oriented industries, more productive firms will have fewer suppliers and will be less likely to integrate its suppliers.

We took these predictions to the data by training a neural network model called a multilayer perceptron to predict whether or not each of 29 million PATSTAT patent applications involves new/improved functions. Industries where patents are skewed towards new/improved functions were deemed ideas-oriented ($IdeaDummy = 1$), while industries where patents are skewed towards cost reductions were deemed cost-oriented ($CostDummy = 1$).

We then merged these patents with S&P Capital IQ data on about 55,000 companies and their supplier networks. We considered two regressions. At the firm level, we regressed the number of a firm’s partners on an interaction between $IdeaDummy$ and firm productivity. We found that more productive firms have more partners and, importantly, they have more partners in ideas-oriented industries relative to other industries. This difference-in-difference (DiD) confirmed our firm-scope hypothesis. In cost-oriented industries this DiD pattern was reversed, as expected.

At the firm-partner level we regressed a binary indicator of whether or not the partner is vertically integrated on a triple interaction between $IdeaDummy$, firm productivity, and partner productivity. We found that more productive firm-partner pairs are more likely to integrate and, importantly, they are more likely to integrate in ideas-oriented industries relative to other industries. This triple difference confirmed our firm-boundaries hypothesis. Further, in cost-oriented industries the triple difference pattern was reversed, as expected.

Summarizing, we presented a theory of the Big Idea, namely, a firm’s choice of how many functions to bundle in a product. We showed that this leads naturally to a link between firm scope and firm boundaries. Finally and most importantly, we derived testable implications for how more productive firms use vertical integration to increase scope while limiting the damage from holdup.

References

- Acemoglu, Daron, Philippe Aghion, Rachel Griffith, and Fabrizio Zilibotti**, "Vertical Integration and Technology: Theory and Evidence," *Journal of European Economic Association*, December 2010, 8 (5), 989–1033.
- , **Pol Antràs, and Elhanan Helpman**, "Contracts and Technology Adoption," *American Economic Review*, June 2007, 97 (3), 916–943.
- , **Simon Johnson, and Todd Mitton**, "Determinants of Vertical Integration: Financial Development and Contracting Costs," *Journal of Finance*, June 2009, 64 (3), 1251–1290.
- Alfaro, Laura and Andrew Charlton**, "Intra-Industry Foreign Direct Investment," *American Economic Review*, December 2009, 99 (5), 2096–2119.
- , **Pol Antràs, Davin Chor, and Paola Conconi**, "Internalizing Global Value Chains: A Firm-Level Analysis," *Journal of Political Economy*, April 2019, 127 (2), 508–559.
- Antràs, Pol**, "Firms, Contracts, and Trade Structure," *Quarterly Journal of Economics*, November 2003, 118 (4), 1375–1418.
- , *Global Production: Firms, Contracts, and Trade Structure*, Princeton, NJ: Princeton University Press, 2015.
- Antràs, Pol and Davin Chor**, "Organizing the Global Value Chain," *Econometrica*, November 2013, 81 (6), 2127–2204.
- Antràs, Pol and Elhanan Helpman**, "Global Sourcing," *Journal of Political Economy*, June 2004, 112, 552–580.
- Bernard, Andrew B. and Andreas Moxnes**, "Networks and Trade," *Annual Review of Economics*, March 2018, 10 (1), 65–85.
- , —, and **Yukiko U. Saito**, "Production Networks, Geography and Firm Performance," *Journal of Political Economy*, April 2019, 127 (2), 639–688.
- Grossman, Sanford J. and Oliver D. Hart**, "Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, August 1986, 94 (4), 691–719.
- Hart, Oliver D. and John Moore**, "Property Rights and the Nature of the Firm Property Rights and the Nature of the Firm," *Journal of Political Economy*, December 1990, 98 (6), 1119–1158.
- Kremer, Michael**, "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics*, August 1993, 108 (3), 551–75.
- Liu, Mengxiao**, "The Missing Option in Firm Boundary Decisions," 2019. Mimeo, Syracuse University.
- Melitz, Marc J.**, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, 71 (6), 1695–1725.
- Nunn, Nathan and Daniel Trefler**, "The Boundaries of the Multinational Firm: An Empirical Analysis," in Elhanan Helpman, Dalia Marin, and Thierry Verdier, eds., *The Organization of Firms in the Global Economy*, Cambridge, MA: Harvard University Press, 2008, pp. 55–83.
- and —, "Incomplete Contracts and the Boundaries of the Multinational Firm," *Journal of Economic Behavior & Organization*, October 2013, 94, 330–344.

- **and** —, “Domestic Institutions as a Source of Comparative Advantage,” in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Trade*, Amsterdam: Elsevier, 2014, pp. 263–315.
- Orr, Scott, Daniel Trefler, and Miaojie Yu**, “Estimating Productivity Using Chinese Data: Methods, Challenges and Results,” in Lili Yan Ing and Miaojie Yu, eds., *World Trade Evolution: Growth, Productivity and Employment*, Routledge, 2019, pp. 229–260.
- Yeaple, Stephen R.**, “Offshoring, Foreign Direct Investment, and the Structure of U.S. Trade,” *Journal of the European Economic Association*, April-May 2006, 4, 602–611.

Appendix

Appendix A. Existence and Uniqueness of SSPE

We show that in any SSPE, the firm chooses $(h_1, h_2, \dots, h_N) = (h, h, \dots, h)$, where h solves the following problem:

$$\max_h \gamma^k(N) \frac{\hat{A}}{\hat{\eta}^\alpha} D(N, \theta)^\alpha h^{\alpha\eta} m^{\alpha(1-\eta)} \} - w_h C(N, \theta) N h.$$

Each supplier chooses $m_j = m$, where m solves the following problem:

$$\max_m \gamma^k(N) \frac{\hat{A}}{\hat{\eta}^\alpha} D(N, \theta)^\alpha h^{\alpha\eta} m^{\alpha(1-\eta)} \} - w_m C(N, \theta) m.$$

The solutions to h and m simultaneously and uniquely solve the above two problems.

First, consider the firm's problem in (FP1):

$$\max_{(h_1, h_2, \dots, h_N)} \gamma^k(N) \frac{\hat{A}}{\hat{\eta}^\alpha} D(N, \theta)^\alpha \min_{j=1, \dots, N} \{h_j^{\alpha\eta} m_j^{\alpha(1-\eta)}\} - w_h C(N, \theta) \sum_{j=1}^N h_j.$$

Suppose all suppliers stick to their equilibrium strategies. The firm's problem can be simplified to

$$\max_{(h_1, h_2, \dots, h_N)} \gamma^k(N) \frac{\hat{A}}{\hat{\eta}^\alpha} D(N, \theta)^\alpha \min_{j=1, \dots, N} \{h_j^{\alpha\eta}\} m^{\alpha(1-\eta)} - w_h C(N, \theta) \sum_{j=1}^N h_j.$$

The firm always chooses $(h_1, h_2, \dots, h_N) = (h, h, \dots, h)$ to maximize its surplus because it is never optimal for the firm to deviate from this strategy. If the firm deviates by choosing $(h_1, h_2, \dots, h_N) \neq (h, h, \dots, h)$:

1. The firm always chooses (h_1, h_2, \dots, h_N) such that $h_1 = h_2 = \dots = h_N \equiv h'$. Because if not, the firm can do strictly better by lowering the levels of all $h_i > \min_j \{h_j\}$ to $h_i = \min_j \{h_j\}$. The firm's problem can therefore be further simplified to:

$$\max_{h'} \gamma^k(N) \frac{\hat{A}}{\hat{\eta}^\alpha} D(N, \theta)^\alpha (h')^{\alpha\eta} m^{\alpha(1-\eta)} - w_h C(N, \theta) N h'.$$

2. It is never optimal for the firm to choose $h' \neq h$ because the objective function is strictly concave in h' , so $h' = h$ is, by definition, the unique maximizer of the objective function.

Therefore, as long as the suppliers stick to their equilibrium strategies, the firm always chooses an h that maximizes the firm's surplus.

Now consider supplier j 's problem:

$$\max_{m_j} \frac{1 - \gamma^k(N)}{N} \frac{\hat{A}}{\hat{\eta}^\alpha} D(N, \theta)^\alpha \min_{-j \in \{1, \dots, N\} \setminus j} \{h_{-j}^{\alpha\eta} m_{-j}^{\alpha(1-\eta)}, h_j^{\alpha\eta} m_j^{\alpha(1-\eta)}\} - w_m C(N, \theta) m_j.$$

Suppose the firm and all the other suppliers stick to their equilibrium strategies. Supplier j 's problem can be written as:

$$\max_{m_j} \frac{1 - \gamma^k(N)}{N} \frac{\hat{A}}{\hat{\eta}^\alpha} D(N, \theta)^\alpha h^{\alpha\eta} \min\{m^{\alpha(1-\eta)}, m_j^{\alpha(1-\eta)}\} - w_m C(N, \theta) m_j.$$

Supplier j is strictly worse off if it deviates by choosing $m_j \neq m$ because its objective function is strictly concave in m_j , which means $m_j = m$ is the unique maximizer of the supplier's objective function.

Appendix B. Multilateral Bargaining Problem with Leontief Production Function

We use Shapley value to solve for the multilateral bargaining problem between the firm and its N suppliers. Each player's Shapley value is the average of her contributions to all coalitions that consist of players ordered below her in all permutations of the order. A coalition generates one of three possible values.

1. In a coalition without the firm, the value is $V_1 = 0$.
2. In a coalition with the firm and all the suppliers, the value is revenue $V_2 = R = \hat{A}D(N, \theta)^\alpha q^\alpha$, where $q = h^\eta m^{1-\eta} / \hat{\eta}$ as in the statement of the Lemma.
3. In a coalition with the firm, but not all the suppliers, the minimum quality is $\Delta^k q$ so that the value is a fraction of total revenue: $V_3 = \delta^k R = \hat{A}D(N, \theta)^\alpha (\Delta^k q)^\alpha$, where $\delta^k \equiv (\Delta^k)^\alpha$.

Consider the firm's contribution. Pick a permutation (a ranking of each player from 0 to N) and let $g(B)$ be the firm's rank in this permutation. If $g(B) < N$ then there is at least one supplier not in the coalition and the firm's contribution is $V_3 - V_1 = \delta^k R$. If $g(B) = N$ then all suppliers are in the coalition and the firm's contribution is $V_2 - V_1 = R$. The share of permutations with $g(B) = N$ is $1/(N+1)$. The share of permutations with $g(B) < N$ is $N/(N+1)$. Therefore, the firm's Shapley value is

$$R \frac{1}{N+1} + \delta^k R \frac{N}{N+1} = \frac{\delta^k N + 1}{N+1} R.$$

The value generated by a coalition of all players is R (case 2). Since the Shapley value is efficient, suppliers must receive

$$R - \frac{\delta^k N + 1}{N+1} R = \frac{1 - \delta^k}{N+1} NR.$$

The Shapley value is symmetric so that all suppliers have the same Shapley value. Dividing the above expression by the N suppliers gives each supplier's Shapley value: $[(1 - \delta^k)/(N+1)]R$.

The online appendix considers multilateral bargaining problems with CES and O-Ring production functions.

Appendix C. Firm and Suppliers' Levels of Investments

Solving for the firm and supplier's problems in SSPE gives the following expressions:

$$h^k(N, \theta, \eta) = \left\{ \frac{\alpha \hat{A}}{\hat{\eta}} \frac{D(N, \theta)^\alpha}{NC(N, \theta)} \left[\frac{\eta \gamma^k(N)}{w_h} \right]^{1-\alpha+\alpha\eta} \left[\frac{(1-\eta)(1-\gamma^k(N))}{w_m} \right]^{\alpha-\alpha\eta} \right\}^{1/(1-\alpha)};$$

$$m^k(N, \theta, \eta) = \left\{ \frac{\alpha \hat{A}}{\hat{\eta}} \frac{D(N, \theta)^\alpha}{NC(N, \theta)} \left[\frac{\eta \gamma^k(N)}{w_h} \right]^{\alpha\eta} \left[\frac{(1-\eta)(1-\gamma^k(N))}{w_m} \right]^{1-\alpha\eta} \right\}^{1/(1-\alpha)}.$$

Substituting $h^k(N, \theta, \eta)$ and $m^k(N, \theta, \eta)$ into the definitions of q , φ and R gives the following expressions:

$$q^k(N, \theta, \eta) = \left\{ \alpha \hat{A} \frac{D(N, \theta)^\alpha}{NC(N, \theta)} \left(\frac{\gamma^k(N)}{w_h} \right)^\eta \left(\frac{1-\gamma^k(N)}{w_m} \right)^{1-\eta} \right\}^{1/(1-\alpha)};$$

$$\varphi^k(N, \theta, \eta) = \left\{ \alpha \hat{A} \frac{D(N, \theta)}{NC(N, \theta)} \left(\frac{\gamma^k(N)}{w_h} \right)^\eta \left(\frac{1-\gamma^k(N)}{w_m} \right)^{1-\eta} \right\}^{1/(1-\alpha)};$$

$$R^k(N, \theta, \eta) = \left\{ \alpha \hat{A}^{1/\alpha} \frac{D(N, \theta)}{NC(N, \theta)} \left(\frac{\gamma^k(N)}{w_h} \right)^\eta \left(\frac{1-\gamma^k(N)}{w_m} \right)^{1-\eta} \right\}^{\alpha/(1-\alpha)}.$$

Appendix D. Firm Decisions in the Ideas-oriented Industry

The log-transformation of the firm's problem in FP2 is

$$\max_{k \in \{O, V\}, N \in [1, \infty)} \pi^k(N, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi(\gamma^k(N), \eta), \quad (\text{fp1})$$

which can be indirectly solved by solving

$$\max_{\delta \in (0, 1), N \in [1, \infty)} \pi(N, \delta, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi(\gamma(N, \delta), \eta). \quad (\text{fp2})$$

Appendix D.1. $\pi(N, \delta)$ is strictly concave in (N, δ)

Since the firm takes θ and η as given, the choice variables in (fp2) are N and δ . Let us write the log-profit function as $\pi(N, \delta)$. $\pi(N, \delta)$ is strictly concave in (N, δ) if and only if its Hessian matrix is negative definite. Assume that $\pi(N, \delta)$ is twice continuously differentiable, The Hessian matrix can be written as

$$\begin{pmatrix} \pi_{NN} & \pi_{N\delta} \\ \pi_{\delta N} & \pi_{\delta\delta} \end{pmatrix} = \begin{pmatrix} g_{NN} + \psi_{\gamma\gamma}\gamma_N^2 & \psi_{\gamma\gamma}\gamma_N\gamma_\delta \\ \psi_{\gamma\gamma}\gamma_N\gamma_\delta & \psi_{\gamma\gamma}\gamma_\delta^2 \end{pmatrix}, \quad (10)$$

The above matrix is negative definite if and only if g_{NN} and $\psi_{\gamma\gamma}$ are both negative.²⁶ $\psi_{\gamma\gamma}$ is always negative because

$$\psi_{\gamma\gamma} = - \left\{ \frac{\alpha(2\eta - 1)}{1 - \alpha[\gamma\eta + (1 - \gamma)(1 - \eta)]} \right\}^2 - \frac{\alpha}{1 - \alpha} \left[\frac{\eta}{\gamma^2} + \frac{1 - \eta}{(1 - \gamma)^2} \right] < 0,$$

so $\pi(N, \delta)$ is strictly concave if and only if

$$g_{NN} = \frac{\alpha}{1 - \alpha} \left\{ \frac{\partial^2 \ln D(N, \theta)}{\partial N^2} - \frac{\partial^2 \ln C(N, \theta, \eta)}{\partial N^2} + 1 \right\} < 0,$$

or

$$\frac{\partial^2 \ln C(N, \theta, \eta)}{\partial N^2} > \frac{\partial^2 \ln D(N, \theta)}{\partial N^2} + \frac{1}{N^2}.$$

By Assumption 2, $G(N, \theta)$ is strictly log-concave in N , so $g_{NN} < 0$. The Hessian matrix is thus negative definite so that $\pi(N, \delta, \theta, \eta)$ is strictly concave in (N, δ) . The firm's choice of N solves

$$\pi_N(\delta(N), N) = 0.$$

Taking the derivative of the above equation w.r.t. N generates the following equation:

$$\frac{\partial \delta(N)}{\partial N} = - \frac{\pi_{NN}}{\pi_{N\delta}} = - \frac{g_{NN} + \psi_{\gamma\gamma}\gamma_N^2}{\psi_{\gamma\gamma}\gamma_N\gamma_\delta}.$$

Since $\gamma(N, \delta) \equiv (\delta N + 1)/(N + 1)$, $\gamma_N < 0$ and $\gamma_\delta > 0$. We have shown that $\psi_{\gamma\gamma} < 0$ and $g_{NN} < 0$. Therefore, $\partial \delta(N)/\partial N > 0$. $\delta(N)$ is strictly increasing in N . Figure A1 shows a simulation of $\delta(N)$. We use this figure and the strict concavity of $\pi(N, \delta)$ to illustrate our later proofs.

²⁶A 2×2 matrix is negative definite if and only if its first determinant is negative and its second determinant is positive. These conditions translate to $g_{NN} + \psi_{\gamma\gamma}\gamma_N^2 < 0$ and $g_{NN}\psi_{\gamma\gamma}\gamma_\delta^2 > 0$. These two inequalityities hold if and only if $g_{NN} < 0$ and $\psi_{\gamma\gamma} < 0$.

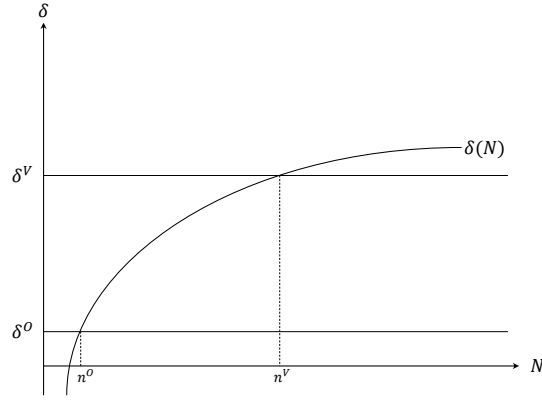


Figure A1: Simulation of $\delta(N)$

Appendix D.2. The marginal firm's organizational behavior

Now refer to the firm's actual problem in equation (fp1). The firm cannot choose any combination of (N, δ) on $\delta(N)$. Instead, the firm can only choose N from two horizontal lines $\delta = \delta^O$ and $\delta = \delta^V$. Assume for now that $\delta(N)$ crosses $\delta = \delta^O$ at (n^O, δ^O) and $\delta = \delta^V$ at (n^V, δ^V) , as shown in Figure A1. Define the marginal firm as the firm that is indifferent between $k = O$ and V . Denote the marginal firm's productivity by $\tilde{\theta}(\eta)$ and its choice under $k = O$ and V by N^O and N^V , respectively. Depending on the values of N^O and N^V relative to the interval (n^O, n^V) , there are 9 cases, as shown in the table below:

Table A1: Relationship between (N^O, N^V) and (n^O, n^V)

	$N^V < n^O$	$n^O < N^V < n^V$	$n^V < N^V$
$N^O < n^O$	N/A	N/A	N/A
$n^O < N^O < n^V$	N/A	$N^O < N^V, \gamma^O < \gamma^V$	N/A
$n^V < N^O$	N/A	N/A	N/A

We show by exclusion that for $\tilde{\theta}(\eta)$ to exist, only the middle cell in Table A1 is possible, i.e., when $N^O, N^V \in (n^O, n^V)$. We also show that in this scenario, the marginal firm's choice satisfies $N^O < N^V$ and $\gamma^O < \gamma^V$, as stated in Theorem 2, part 3. We then prove the existence and monotonicity of $\tilde{\theta}(\eta)$ under this scenario.

1. $N^V < n^O$ contradicts the definition of $\tilde{\theta}(\eta)$.

Refer to Figure A2. Recall that we assumed $\delta(N)$ crosses $\delta = \delta^O$ and $\delta = \delta^V$ at (n^O, δ^O) and (n^V, δ^V) . Since $\pi(N, \delta)$ is strictly concave in (N, δ) , moving from (N^V, δ^V) to (N^V, δ^O) increases firm's profit because keeping $N = N^V$ constant, we are approaching the optimal δ at $\delta(N^V)$. Denote the profits at (N^V, δ^V) and (N^V, δ^O) by $\pi(N^V, \delta^V, \tilde{\theta}(\eta), \eta)$ and $\pi(N^V, \delta^O, \tilde{\theta}(\eta), \eta)$. Based on our argument,

$$\pi(N^V, \delta^V, \tilde{\theta}(\eta), \eta) < \pi(N^V, \delta^O, \tilde{\theta}(\eta), \eta).$$

N^V may or may not be the optimal N that maximizes firm's profit at $\delta = \delta^O$, so

$$\pi(N^V, \delta^O, \tilde{\theta}(\eta), \eta) \leq \max_{N \in (1, \infty)} \pi(N, \delta^O, \tilde{\theta}(\eta), \eta).$$

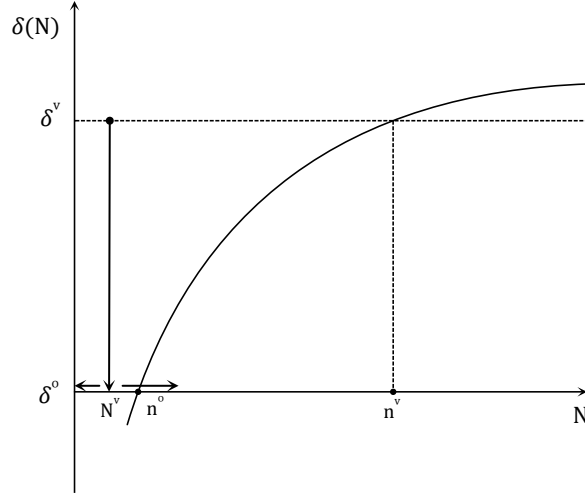


Figure A2: $N^V < n^O$

The above two inequalityities imply

$$\pi(N^V, \delta^V, \tilde{\theta}(\eta), \eta) < \pi(N^V, \delta^O, \tilde{\theta}(\eta), \eta) \leq \max_{N \in (1, \infty)} \pi(N, \delta^O, \tilde{\theta}(\eta), \eta).$$

The left and right ends of the above inequalityity are respectively the firm's optimal profits at $k = V$ and $k = O$. This inequalityity contradicts the definition of the marginal firm because it implies that the marginal firm's profit under integration is lower than its profit under outsourcing. The first column of Table A1 is ruled out.

2. $N^O > n^V$ contradicts the definition of $\tilde{\theta}(\eta)$.

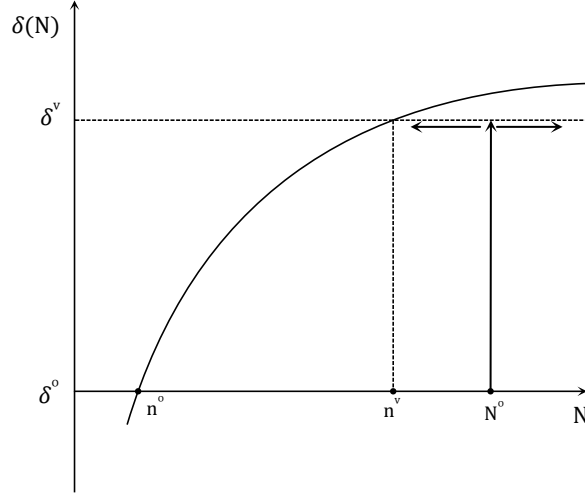


Figure A3: $N^O > n^V$

Refer to Figure A3. Since $\pi(N, \delta)$ is strictly concave in (N, δ) , moving from (N^O, δ^O) to (N^O, δ^V) increases the firm's profit because keeping $N = N^O$ constant, we are approaching the optimal $\delta(N^O)$. Denote the firm's profit at (N^O, δ^O) and (N^O, δ^V) by $\pi(N^O, \delta^O, \tilde{\theta}(\eta), \eta)$ and $\pi(N^O, \delta^V, \tilde{\theta}(\eta), \eta)$. Our argument implies

$$\pi(N^O, \delta^O, \tilde{\theta}(\eta), \eta) < \pi(N^O, \delta^V, \tilde{\theta}(\eta), \eta).$$

N^O may or may not be the optimal N that maximizes the marginal firm's profit at $k = V$, so

$$\pi(N^O, \delta^V, \tilde{\theta}(\eta), \eta) \leq \max_{N \in [1, \infty)} \pi(N, \delta^V, \tilde{\theta}(\eta), \eta).$$

The above two inequalityities imply

$$\pi(N^O, \delta^O, \tilde{\theta}(\eta), \eta) < \pi(N^O, \delta^V, \tilde{\theta}(\eta), \eta) \leq \max_{N \in [1, \infty)} \pi(N, \delta^V, \tilde{\theta}(\eta), \eta).$$

The left and right ends of the above inequality are respectively the firm's optimal profits under $k = O$ and $k = V$. The inequality contradicts the definition of the marginal firm because it implies that the marginal firm's profit under outsourcing is lower than its profit under integration. The third row of Table A1 is ruled out.

3. $N^O < n^O < N^V < n^V$ contradicts the definition of $\tilde{\theta}(\eta)$.

We now show that the upper-middle cell of Table A1 is impossible. To see this, draw an iso- γ line through (N^V, δ^V) . Since $\gamma \equiv (\delta N + 1)/(N + 1)$ is increasing in δ and decreasing in N , the iso- γ line is upward-sloping. Suppose this iso- γ line crosses $\delta = \delta^O$ at N' . Based on the relationship between N' and (N^O, N^V) , there are two cases: $N^O \leq N' < N^V$ and $N' < N^O < N^V$.²⁷

3(a). $N^O \leq N' < N^V$

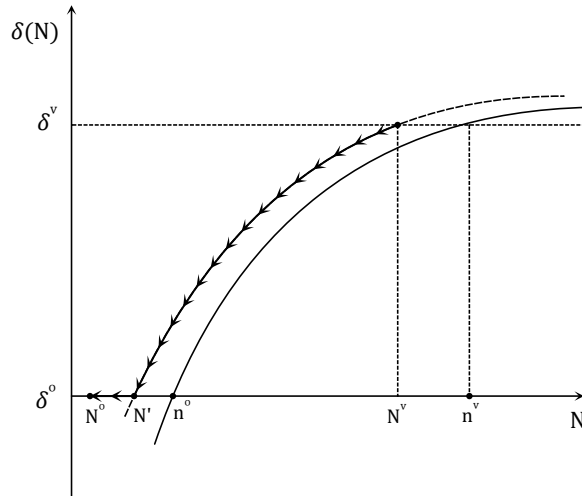


Figure A4: $N^O \leq N' < N^V$

See Figure A4. In this case, both (N^O, δ^O) and (N^V, δ^V) are above $\delta(N)$, which is the “ridge” of the profit function. It implies that δ^k ($k = O, V$) is “too small”—if the firm were allowed to choose δ^k from $[0, 1]$ (conditional on the same N^k), the firm would have chosen a bigger δ^k , which translates to a bigger γ^k . Since the profit function is strictly concave in γ , $\pi_\gamma > 0$ when γ^k is too small.²⁸ To generalize this argument, $\pi_\gamma = \psi_\gamma > 0$ whenever (N^k, δ^k) is above $\delta(N)$, and $\pi_\gamma = \psi_\gamma < 0$ whenever (N^k, δ^k) is below $\delta(N)$. Since the profit function is strictly concave in N ,

²⁷ $N' > N^V$ is not possible because of the monotonicity of the iso- γ line.

²⁸ The profit function can be written as $\pi(\gamma, N, \tilde{\theta}(\eta), \eta) = \psi(\gamma, \eta) + g(N, \tilde{\theta}(\eta))$. γ appears only in $\psi(\gamma, \eta)$ so $\pi_{\gamma\gamma} = \psi_{\gamma\gamma}$. We have previously shown that $\psi_{\gamma\gamma} < 0$, which implies $\pi_{\gamma\gamma} < 0$. The profit function is strictly concave in γ , so that when γ is smaller (bigger) than the optimal value, π_γ is strictly positive (negative).

N^k ($k = O, V$) solves

$$\pi_N(N^k, \delta^k, \tilde{\theta}(\eta), \eta) = g_N(N^k, \tilde{\theta}(\eta), \eta) + \psi_\gamma(\gamma^k, \eta) \gamma_N(N^k, \delta^k) = 0.$$

We have shown that $\psi_\gamma < 0$. $\gamma_N(N^k, \delta^k) = (\delta^k - 1)/(N + 1)^2 < 0$, so that $\psi_\gamma \gamma_N^k > 0$. It must be that $g_N(N^k, \tilde{\theta}(\eta), \eta) < 0$ for $k = O, V$. By Assumption 2, $G(N, \theta)$ is log-concave in N so $g_{NN} < 0$. Since $N^V > N' \geq N^O$, $g_N(N^V, \tilde{\theta}) < g_N(N', \tilde{\theta}) \leq g_N(N^O, \tilde{\theta}) < 0$. If we move from (δ^V, N^V) to (δ^O, N') along the iso- γ line (the arrowed path), $\psi(\gamma, \eta)$ remains constant. But $g(N, \theta)$ increases because $g_N(N, \theta)$ remains negative as we decrease the value of N . It follows that profit increases from (N^V, δ^V) to (N', δ^O) . If we then move from (N', δ^O) to (N^O, δ^O) , profit continues to increase because $g_N(N, \tilde{\theta})$ remains negative and $\psi_\gamma(\gamma, \eta)$ remains positive as we decrease N and increase γ , so both $g(N, \tilde{\theta}, \eta)$ and $\psi(\gamma)$ increase. This argument implies the following inequalityities:

$$\pi(N^V, \delta^V, \tilde{\theta}(\eta), \eta) < \pi(N', \delta^O, \tilde{\theta}(\eta), \eta) < \pi(N^O, \delta^O, \tilde{\theta}(\eta), \eta)$$

The above inequality implies that the marginal firm's profit under integration is lower than its profit under outsourcing, hence contradicting the definition of a marginal firm. $N^O \leq N' < N^V$ is impossible when $N^O < n^O < N^V < n^V$.

3(b). $N' < N^O < N^V$

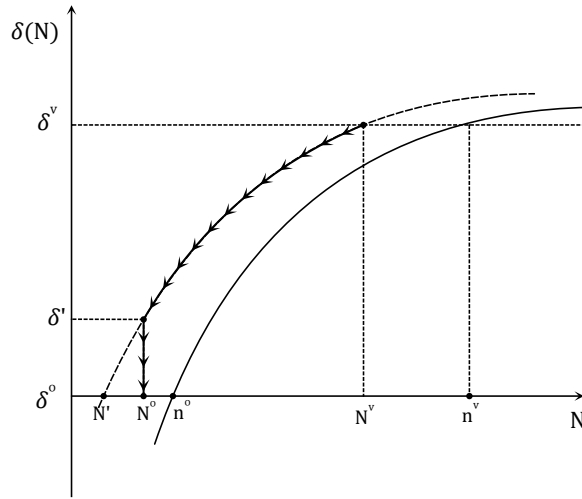


Figure A5: $N' < N^O < N^V$

See Figure A5. We can use similar logic to show that $g_N(N', \tilde{\theta}) < g_N(N^V, \tilde{\theta}) < g_N(N^O, \tilde{\theta}) < 0$ holds. From (N^V, δ^V) to (N^O, δ') along the iso- γ line, $\psi(\gamma, \eta)$ remains constant, $g(N, \tilde{\theta})$ increases because $g_N(N, \tilde{\theta}) < 0$ and N decreases, so profit increases. From (N^O, δ') to (N^O, δ^O) , $g(N, \tilde{\theta})$ remains constant because N does not change. $\psi(\gamma, \eta)$ increases because $\psi_\gamma < 0$ and γ decreases as δ shrinks and N remains constant. So profit increases along the arrowed path, which again implies that the marginal firm's profit under integration is lower than its profit under outsourcing, contradicting the definition of the marginal firm. Combined with the previous part, we have shown that $N^O < n^O < N^V < n^V$ is impossible, ruling out the upper-middle cell of Table A1.

4. $N^O < n^O < n^V < N^V$ is impossible.

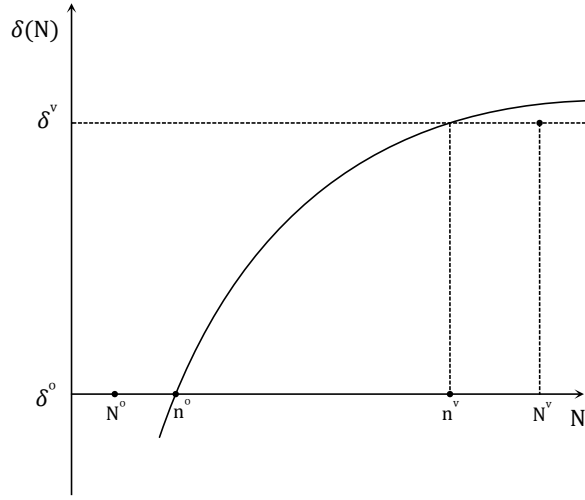


Figure A6: $N^O < n^O, N^V > n^V$

See Figure A6. In this case, (N^O, δ^O) is above $\delta(N)$ and (N^V, δ^V) is below $\delta(N)$, so $\psi_\gamma(\gamma^O, \eta) < 0 < \psi_\gamma(\gamma^V, \eta)$. Recall from previous analyses that $g_N(N^k, \tilde{\theta})$ is of the same sign as $\psi_\gamma(\gamma, \eta)$, so $g_N(N^O, \theta) < 0 < g_N(N^V, \theta)$. This implies $N^O > N^V$ because $g_{NN}(N, \theta) < 0$, which contradicts the assumption of $N^O < N^V$. The upper-right cell of Table A1 is ruled out.

5. $n^O < N^O < n^V < N^V$ contradicts the definition of $\tilde{\theta}(\eta)$.

We now rule out the middle-right cell of Table A1. In this case, both (N^O, δ^O) and (N^V, δ^V) are below $\delta(N)$ so $\psi_\gamma(\gamma^k, \eta) > 0$ for $k \in \{O, V\}$. We can further deduce that $g_N(N^k, \tilde{\theta}) > 0$ for $k \in \{O, V\}$. Let's draw an iso- γ line through (N^O, δ^O) . Suppose it crosses $\delta = \delta^V$ at (N', δ^V) . There are two possible cases: $N^O < N' \leq N^V$ and $N^O < N^V < N'$.

5(a). $N^O < N' \leq N^V$

See Figure A7. $g_{NN} < 0$, $g_N > 0$, and $N^O < N' < N^V$ imply that $g_N(N^O, \tilde{\theta}) > g_N(N', \tilde{\theta}) > g_N(N^V, \tilde{\theta}) > 0$. From (N^O, δ^O) to (N', δ^V) along the iso- γ line, γ remains constant while N increases so $\psi(\gamma, \eta)$ remains constant and $g(N, \tilde{\theta})$ increases. From (N', δ^V) to (N^V, δ^V) , δ remains constant and N increases so $\psi(\gamma, \eta)$ increases because $\psi_\gamma(\gamma, \eta) > 0$ and γ increases. $g(N, \theta)$ increases because $g_N > 0$ and N increases, so profit increases along the arrowed path, which implies that the marginal firm's profit is higher under integration than outsourcing, contradicting the definition of a marginal firm.

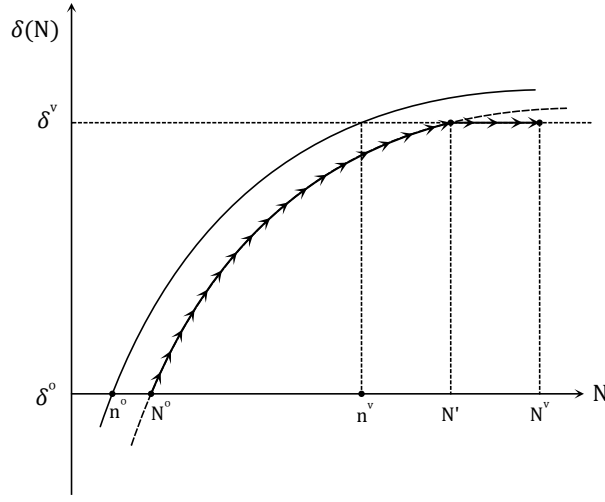


Figure A7: $N^O < N' \leq N^V$

5(b). $N^O < N^V < N'$

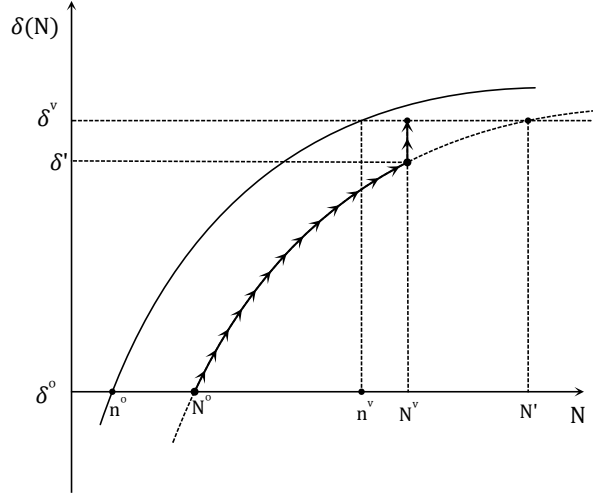


Figure A8: $N^O < N^V < N'$

See Figure A8. From (N^O, δ^O) to (N^V, δ') along the iso- γ line, γ remains constant while N increases, so $\psi(\gamma, \eta)$ remains constant while $g(N, \tilde{\theta})$ increases and profit increases. From (N^V, δ') to (N^V, δ^V) , δ increases and N^V remains constant, so $g(N, \tilde{\theta})$ remains constant and $\psi(\gamma, \eta)$ increases because $\psi_\gamma > 0$ and γ increases. Profit increases along the arrowed path, contradicting the definition of the marginal firm. Combined with the previous part, we can rule out the middle-right cell of Table A1.

6. $N^O, N^V \in (n^O, n^V)$.

We have excluded all the other possibilities in Table A1 except for the middle cell. If $\tilde{\theta}(\eta)$ does exist, it must be that $N^O, N^V \in (n^O, n^V)$. We now prove that if this condition holds, $N^O < N^V$ and $\gamma^O < \gamma^V$.

6(a). $N^O < N^V$

We have shown that $\pi(N, \delta; \theta(\eta))$ is strictly concave in (N, δ) , which implies that $\pi(N, \delta, \tilde{\theta}(\eta), \eta)$ is supermodular in (N, δ) . Since $\delta^V > \delta^O$, by Topkis's Theorem, it must be that $N^V > N^O$.

6(b). $\gamma^O < \gamma^V$

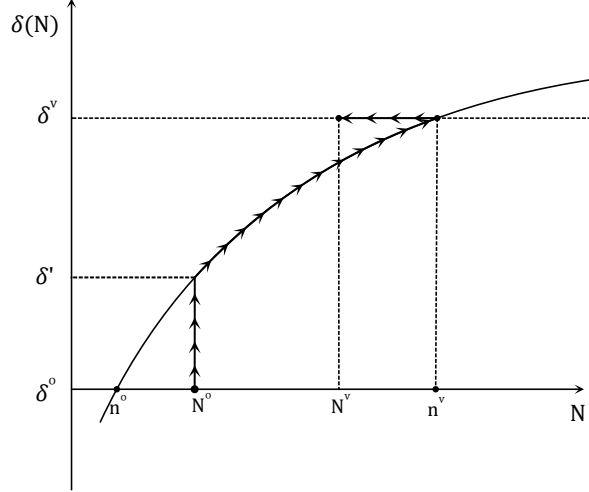


Figure A9: $n^O < N^O < N^V < n^V$

See Figure A9. (N^O, δ^O) and (N^V, δ^V) are the firm's optimal choice under $k = O, V$. From (N^O, δ^O) to (N^O, δ') , N is constant while δ increases, so γ increases. From (N^O, δ') to (n^V, δ^V) along the iso- γ line, γ remains constant. From (n^V, δ^V) to (N^V, δ^V) , δ is constant while N decreases. γ increases because $\gamma_N < 0$. Along the arrowed path, γ increases so $\gamma^O < \gamma^V$.

We have now proved that if $\delta(N)$ crosses both $\delta = \delta^O$ and $\delta = \delta^V$, and if the marginal firm exists, then the marginal firm's choice satisfies $N^O < N^V$ and $\gamma^O < \gamma^V$. We will then prove that the marginal firm does exist for $\underline{\eta} < \eta < \bar{\eta}$, and that $\delta(N)$ has to cross $\delta = \delta^O$ and $\delta = \delta^V$ when $\eta \in (\underline{\eta}, \bar{\eta})$.

Appendix D.3. Uniqueness of $\tilde{\theta}(\eta)$

Denote by $N^k(\theta, \eta)$ the scope decision of a firm with productivity θ in industry η under organizational form $k = O, V$. Express this firm's profit under organizational form k as

$$\pi(N^k, \delta^k, \theta, \eta) = \psi(\gamma^k, \eta) + g(N^k, \theta).$$

By Envelope Theorem, $\pi_\theta(N^k, \delta^k, \theta, \eta) = g_\theta(N^k, \theta) > 0$ iff

$$g_\theta(N^k, \theta) = \frac{\alpha}{1 - \alpha} \left\{ \frac{\partial \ln D(N, \theta)}{\partial \theta} - \frac{\partial \ln C(N, \theta, \eta)}{\partial \theta} \right\} > 0,$$

or

$$\frac{\partial \ln D(N, \theta)}{\partial \theta} > \frac{\partial \ln C(N, \theta, \eta)}{\partial \theta}.$$

We have shown in Appendix D.2 that for the marginal firm, $N^V(\tilde{\theta}, \eta) > N^O(\tilde{\theta}, \eta)$. By Assumption 1, $g(N, \theta)$ is supermodular in (N, θ) in the ideas-oriented industry, meaning that $g_\theta(N, \theta)$ is increasing in N . Thus

$$g_\theta(N^V, \tilde{\theta}) > g_\theta(N^O, \tilde{\theta}),$$

which is equivalent to

$$\pi_\theta(N^V, \delta^V, \theta, \eta) > \pi_\theta(N^O, \delta^O, \theta, \eta).$$

As θ increases, the difference between π^V and π^O increases. This means that if π^V and π^O cross, they can cross only once. This crossing point is $\tilde{\theta}(\eta)$. Therefore, if $\tilde{\theta}(\eta)$ exists, it is unique. Moreover, firms with $\theta < \tilde{\theta}(\eta)$ choose outsourcing and firms with $\theta > \tilde{\theta}(\eta)$ choose vertical integration.

Appendix D.4. Monotonicity of $\tilde{\theta}(\eta)$

By definition of $\tilde{\theta}$,

$$\pi(N^V, \delta^V, \tilde{\theta}, \eta) = \pi(N^O, \delta^O, \tilde{\theta}, \eta).$$

By Implicit Function Theorem,

$$\begin{aligned} \frac{d\tilde{\theta}}{d\eta} &= - \frac{\pi_\eta(N^V, \delta^V, \tilde{\theta}, \eta) - \pi_\eta(N^O, \delta^O, \tilde{\theta}, \eta)}{\pi_\theta(N^V, \delta^V, \tilde{\theta}, \eta) - \pi_\theta(N^O, \delta^O, \tilde{\theta}, \eta)} \\ &= - \frac{\psi_\eta(\gamma^V, \eta) - \psi_\eta(\gamma^O, \eta)}{g_\theta(N^V, \tilde{\theta}) - g_\theta(N^O, \tilde{\theta})} \end{aligned}$$

We have shown that $\gamma^V > \gamma^O$ and $N^V > N^O$. $\psi(\gamma, \eta)$ is supermodular in (γ, η) . By Assumption 1, $g(N, \theta)$ is supermodular in (N, θ) . Therefore, $\psi_\eta(\gamma^V, \eta) > \psi_\eta(\gamma^O, \eta)$ and $g_\theta(N^V, \tilde{\theta}) > g_\theta(N^O, \tilde{\theta})$. $d\tilde{\theta}/d\eta < 0$. If it exists, $\tilde{\theta}$ is decreasing in η .

Appendix D.5. Existence of $\tilde{\theta}(\eta)$, $\underline{\eta}$, and $\bar{\eta}$

1. Firms' choice of scope (N) is bounded between 1 and ∞ .

Since $\pi(N, \delta, \theta, \eta)$ is strictly concave in N , the sufficient conditions for $1 < N^k(\theta, \eta) < \infty$ are $\lim_{N \rightarrow 1} \pi_N(N, \delta, \theta, \eta) > 0$ and $\lim_{N \rightarrow \infty} \pi_N(N, \delta, \theta, \eta) < 0$.

$$\pi_N(N, \delta, \theta, \eta) = \frac{\alpha}{1 - \alpha} g_N(N, \theta) - \frac{1 - \delta}{(N + 1)^2} \cdot \psi_\gamma(\gamma(N, \delta), \eta)$$

$$\begin{aligned} \lim_{N \rightarrow 1} \pi_N(N, \delta, \theta, \eta) &= \frac{\alpha}{1 - \alpha} \cdot \lim_{N \rightarrow 1} g_N(N, \theta) - \frac{1 - \delta^\alpha}{4} \cdot \psi_\gamma\left(\frac{\delta^\alpha + 1}{2}, \eta\right) \\ &> \frac{\alpha}{1 - \alpha} \cdot \lim_{N \rightarrow 1} g_N(N, \theta) - \frac{1 - \delta^\alpha}{4} \cdot \psi_\gamma\left(\frac{\delta^\alpha + 1}{2}, 1\right) \\ &= \frac{\alpha}{1 - \alpha} \left\{ \lim_{N \rightarrow 1} g_N(N, \theta) - \frac{(1 - \delta^\alpha)^2}{2(1 + \delta^\alpha)(2 - \alpha - \alpha\delta^\alpha)} \right\} \\ &\geq \frac{\alpha}{1 - \alpha} \left\{ \lim_{N \rightarrow 1} g_N(N, \theta) - \lim_{\delta \rightarrow 0} \frac{(1 - \delta^\alpha)^2}{2(1 + \delta^\alpha)(2 - \alpha - \alpha\delta^\alpha)} \right\} \\ &= \frac{\alpha}{1 - \alpha} \left\{ \lim_{N \rightarrow 1} g_N(N, \theta) - \frac{1}{4 - 2\alpha} \right\} \\ &> \frac{\alpha}{1 - \alpha} \left\{ \lim_{N \rightarrow 1} g_N(N, \theta) - \frac{1}{2} \right\}. \end{aligned}$$

By Assumption 2, $g_N(N, \theta) > 1/2$, so $\lim_{N \rightarrow 1} \pi_N(N, \delta, \theta, \eta) > 0$

$$\begin{aligned}\lim_{N \rightarrow \infty} \pi_N(N, \delta, \theta, \eta) &= \frac{\alpha}{1 - \alpha} \cdot \lim_{N \rightarrow \infty} g_N(N, \theta) - 0 \cdot \psi_\gamma(\delta^\alpha) \\ &= \frac{\alpha}{1 - \alpha} \cdot \lim_{N \rightarrow \infty} g_N(N, \theta).\end{aligned}$$

By Assumption 2, $\lim_{N \rightarrow \infty} g_N(N, \theta) < 0$, so $\lim_{N \rightarrow \infty} \pi_N(N, \delta, \theta, \eta) < 0$. Since $\lim_{N \rightarrow 1} \pi_N(N, \delta, \theta, \eta) > 0$ and $\lim_{N \rightarrow \infty} \pi_N(N, \delta, \theta, \eta) < 0$, it must be that $1 < N^k(\theta, \eta) < \infty$ for $k \in \{O, V\}$. Denote the optimal scope decision by $N(\theta, \eta)$. $N(\theta, \eta)$ must also be bounded between 1 and ∞ .

2. Existence of the threshold industries

$\psi(\gamma, \eta)$ ranges from zero to infinity on $\eta \in (0, 1)$ for any given $\gamma \in (0, 1)$. Given that there is increasing difference between π^V and π^O , for each value of θ , there must be at least one η , such that $\pi^V(\theta, \eta) = \pi^O(\theta, \eta)$. There cannot be more than one η that satisfies this condition because this would violate the monotonicity of $\theta(\eta)$. Therefore, there is a one-to-one mapping from θ to η . Since $\theta(\eta)$ is strictly decreasing, $\eta(\theta)$ is also strictly decreasing over the interval $\theta \in [0, 1]$, so $\eta(1) \leq \eta(\theta) \leq \eta(0)$. Define $\underline{\eta} \equiv \eta(1)$ and $\bar{\eta} \equiv \eta(0)$. For $\eta < \underline{\eta}$, $\pi^V(\theta, \eta) - \pi^O(\theta, \eta) < 0$ for all θ , all firms choose outsourcing. For $\eta > \bar{\eta}$, $\pi^V(\theta, \eta) - \pi^O(\theta, \eta) > 0$, all firms choose vertical integration. In other words, $\underline{\eta}$ and $\bar{\eta}$ exist and $0 < \underline{\eta} < \bar{\eta} < 1$.

Appendix E. Firms' scope decisions in the ideas-oriented industry

We have proved that the marginal firm's behavior satisfies $N^V > N^O$ and $\gamma^V > \gamma^O$. We now prove the first two statements in Theorem 2.

Appendix E.1. Firm's scope decision in (fp2) is monotone in θ .

Since $\pi(N, \delta, \theta, \eta)$ is strictly concave in (N, δ) , the optimal (N, δ) is determined by the two first order conditions, $\pi_N = 0$ and $\pi_\delta = 0$. Differentiating these two equations with respect to θ and rearranging,

$$\begin{pmatrix} dN/d\theta \\ d\delta/d\theta \end{pmatrix} = \frac{1}{\det} \begin{pmatrix} \pi_{\delta\delta} & -\pi_{N\delta} \\ -\pi_{N\delta} & \pi_{NN} \end{pmatrix} \begin{pmatrix} \pi_{N\theta} \\ \pi_{\delta\theta} \end{pmatrix}$$

where \det is the determinant of the Hessian matrix. The above equation can be simplified to

$$\begin{pmatrix} dN/d\theta \\ d\delta/d\theta \end{pmatrix} = \frac{g_{N\theta}\psi_{\gamma\gamma}}{\det} \begin{pmatrix} \gamma_\delta^2 \\ -\gamma_\delta\gamma_N \end{pmatrix}$$

It can be easily shown that $\psi_{\gamma\gamma} < 0$, $\det > 0$. $\gamma_\delta = \frac{N}{N+1} > 0$ and $\gamma_N = \frac{\delta-1}{(N+1)^2} < 0$. By Assumption 1, $g_{N\theta} > 0$, so $dN/d\theta > 0$, and $d\delta/d\theta > 0$.

Appendix E.2. Firm's scope decision in (fp1) is monotone in θ .

We know that $\pi^k(N, \theta, \eta) = \tilde{a} + g(N, \theta, \eta) + \psi(\gamma^k(N), \eta)$, and $\pi_{N\theta}^k = g_{N\theta}$, so $g_{N\theta} > 0$ implies $\pi_{N\theta}^k > 0$, and that $N^k(\theta, \eta)$ is strictly increasing in θ .

Appendix F. Merging S&P Capital IQ and PATSAT

We use an unsupervised machine learning algorithm (an n-gram model) to match 3,165,143 PATSTAT companies with 33,783,284 S&P Capital IQ companies based on their names and countries of location. The matching is implemented in three stages:

1. Parsing: company names are standardized through a string cleaning procedure including converting company names to unicode letters in lower cases, removing stock exchange abbreviations and legal suffixes such as ltd, gmbh, and llc.
2. Matching: the company names across the two sources are combined and transformed into a sparse matrix, with each row corresponding to a company name and each column a numeric value indicating whether a three-letter gram exists in the company name, and how important this gram is (its TF-IDF score).
3. Filtering: similarity scores are computed for each pairwise combination of company names from the two sources (PATSTAT and S&P Capital IQ). Mutual top matches, i.e., those matches where the two companies are each other's top match are kept.

Using this method, we are able to match 251,484 companies across the two sources. Table A2 reports the top locations of these companies. See the online appendix for a more detailed description of the matching process.

Table A2: Geographic Location of the Matched Companies

Country/region	Companies	Percentage
United States	73,708	29%
China	31,266	12%
Japan	22,082	9%
Germany	15,131	6%
United Kingdom	12,681	5%
South Korea	9,583	4%
France	8,959	4%
Canada	5,803	2%
Italy	5,015	2%
Spain	4,085	2%
Taiwan	3,361	1%
Australia	3,340	1%
Sweden	3,073	1%
Switzerland	2,915	1%
Netherlands	2,896	1%
Russia	2,625	1%
Other	44,961	18%
Total	251,484	100%

Appendix G. Patent Classification

We use a supervised learning algorithm (a multilayer perceptron, or MLP model) to classify the 29,666,609 patents held by the 251,484 PATSTAT companies that are matched with S&P Capital IQ companies.

First, the research assistants are asked to read the titles and abstracts of 6,000 randomly selected patents, and assign a binary classifier for each patents. The classifier equals 1 if the patent improves the quality of an existing function/product, and 0 if the patent improves production efficiency or reduces production cost.

Second, we use the classified sample as a training set to train a multilayer perceptron (MLP) with different combinations of hyperparameters. The best MLP model has 4 layers with 16 neurons per layer, and generates training and validation accuracies of over 85%. This model is then used to classify the patents.

A detailed description of the classification procedure can be found in the online appendix.

Appendix H. Production Network Construction

S&P Capital IQ collects firms' customer and supplier relationships from various sources including company regulatory and annual reports, and newswires such as Thomson Reuters and Bloomberg. Each firm-customer and firm-supplier relationship can be considered as a buyer-seller relationship. The data platform reported 954,420 buyer-seller relationships during 2010-2017. We match each firm in a buyer-seller relationship with its ownership information to define the ownership structure variable. Table A3 reports the ownership structure of the 954,420 buyer-seller relationships.

Table A3: Ownership Structure in the Original Dataset

Ownership Structure	Relationships	Percentage
Buyer owns seller (level two)	13,167	1.38%
Buyer owns seller (level one)	3,485	0.37%
Seller owns buyer (level two)	3,714	0.39%
Seller owns buyer (level one)	16,587	1.74%
Neither	917,467	96.13%
Total	954,420	100.00%

A level-one ownership is a relationship where one firm is the other firm's investor, limited partner, or pending parent/investor; a level-two ownership is a relationship where one firm is the other firm's parent, merged entity, holding company, or ultimate parent. In cases where there are more than one type of relationship between two firms, level-two ownership dominates level-one ownership. There are no relationships where mutual ownership exist at the same level between the buyer and seller.

See the online appendix for more information on the sources for collecting the buyer-seller relationships data and the definition of the various owners.

Appendix I. Imputing Buyer-seller Relationships from S&P Capital IQ

Table A4: Ownership Structure in the Extended Dataset

Ownership Structure	Relationships	Percentage
Buyer owns seller (level two)	674,904	30.29%
Buyer owns seller (level one)	3,485	0.16%
Seller owns buyer (level two)	626,918	28.14%
Seller owns buyer (level one)	3,714	0.17%
Neither	918,858	41.24%
Total	2,227,879	100.00%

The percentage of integrated relationships (either defined at level-one or level-two) Table A3 is only 3.87. We have reasons to believe that there are integrated relationships not captured by S&P Capital IQ's sources. Therefore, we use the now standardized imputation method to uncover 1,286,076 buyer-seller relationships from parent-subsidiary relationships (also from the S&P Capital IQ data platform). 12,617 of these relationships overlap with the buyer-seller relationships originally collected from S&P Capital IQ. Table A4 reports the ownership structure in the extended dataset. This is the sample used to construct production network information. Both level one and level two ownerships are considered integrated.

The imputation method is elaborated in the online appendix.

Online Appendix to

**“What’s the Big Idea:
Multi-Function Products, Firm Scope and Firm Boundaries”**

by Mengxiao Liu and Daniel Trebler

Appendix A. Multilateral Bargaining with CES Production Function

Suppose the demand-shifter is:

$$\varphi = D(N, \theta) N Q,$$

where $Q = \left(\sum_{j=1}^N q_j^\beta \right)^{1/\beta}$ is the overall quality of the firm's product. In a symmetric equilibrium, $q_j = q$ for all j . The product quality is

$$Q = (N q^\beta)^{1/\beta} = N^{1/\beta} q.$$

The demand-shifter is

$$\varphi = D(N, \theta) N^{-1/\beta} Q = D(N, \theta) N^{-1/\beta} N^{1/\beta} q = D(N, \theta) q.$$

The revenue is

$$R = \hat{A} \varphi^\alpha = \hat{A} D(N, \theta)^\alpha n^{\alpha/\beta} q^\alpha.$$

Similar to the previous section, a coalition generates one of three possible values:

1. In a coalition without the firm, the value is 0.
2. In a coalition with the firm and all the suppliers, the value is revenue $R = \hat{A} D(N, \theta)^\alpha q^\alpha$.
3. In a coalition with the firm, but not all the suppliers, the overall quality is

$$\left[n q^\beta + (N - n) (\Delta^k q)^\beta \right]^{1/\beta} = [n + (N - n) (\Delta^k)^\beta]^{1/\beta} q,$$

where n is the number of suppliers who are in the coalition. The demand-shifter generated by this coalition is

$$D(N, \theta) N^{-1/\beta} Q = D(N, \theta) [n/N + (1 - n/N) (\Delta^k)^\beta]^{1/\beta} q.$$

The value of this coalition is

$$\hat{A} \{ D(N, \theta) [n/N + (1 - n/N) (\Delta^k)^\beta]^{1/\beta} q \}^\alpha = [n/N + (1 - n/N) (\Delta^k)^\beta]^{\alpha/\beta} R.$$

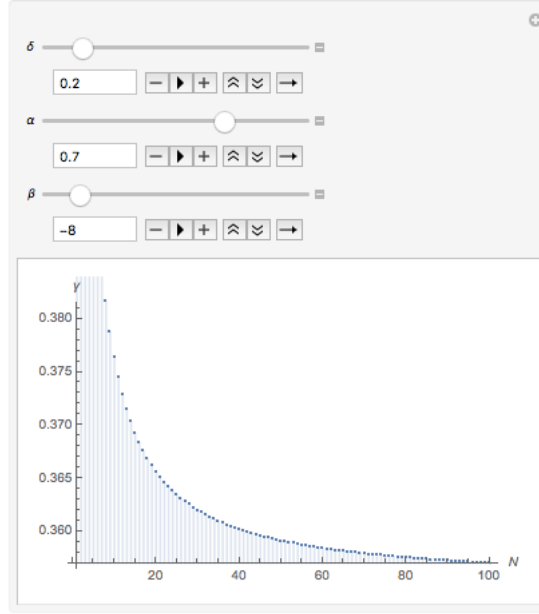
The firm's Shapley value can be calculated as:

$$\frac{1}{N+1} \sum_{n=0}^N \{ [1 - (\Delta^k)^\beta] n/N + (\Delta^k)^\beta \}^{\alpha/\beta} R \equiv \gamma(N, \Delta^k) R,$$

where $\gamma(N, \Delta^k)$ is the firm's revenue share, and

$$\gamma(N, \Delta^k) \equiv \frac{\sum_{n=0}^N \{ [1 - (\Delta^k)^\beta] n/N + (\Delta^k)^\beta \}^{\alpha/\beta}}{N+1}.$$

A mathematical proof for when $\gamma(N, \Delta^k)$ is strictly decreasing in N is impossible, but simulation shows that this holds when Δ^k and β are not too big and α is not too small. The figure below shows one scenario:



Appendix B. Multilateral Bargaining with O-Ring Production Function

Suppose the overall quality Q is an o-ring function of the individual functions:

$$Q = \prod_{j=1}^N q_j.$$

The firm's demand shifter is $\varphi = D(N, \theta)Q$, and its revenue is $R = \hat{A}D(N, \theta)^\alpha Q^\alpha$. In a symmetric equilibrium, $q_j = q$ for all j . The firm's quality is $Q = q^N$, and its revenue is $R = \hat{A}D(N, \theta)^\alpha q^{\alpha N}$. Again, there are three values that can be generated by a coalition:

1. In a coalition without the firm, the value is 0.
2. In a coalition with the firm and all the suppliers, the value is revenue $R = \hat{A}D(N, \theta)^\alpha q^{\alpha N}$.
3. In a coalition with the firm, but not all the suppliers, the overall quality is $Q = q^n (\Delta^k q)^{N-n} = (\Delta^k)^{N-n} q^N$, where n is the number of suppliers who are in the coalition. The corresponding value is $\hat{A}D(N, \theta)^\alpha Q^\alpha = (\Delta^k)^{\alpha(N-n)} R$. Note again, case 2 is a special case of case 3 when $n = N$.

The firm's Shapley value is

$$\frac{1}{N+1} \sum_{n=0}^N (\delta^k)^{N-n} R = \gamma(N, \delta^k) R,$$

where $\delta^k \equiv (\Delta^k)^\alpha$, and $\gamma(N, \delta^k) \equiv \frac{1 - (\delta^k)^{N+1}}{(1 - \delta^k)(N+1)}$ is increasing in δ^k and decreasing in N .

Appendix C. Merging S&P Capital IQ and PATSTAT

We use n-grams matching algorithm to match 3,165,143 PATSTAT companies with 33,783,284 S&P Capital IQ companies. We match these companies based on their names and countries of locations. The matching contains three stages: parsing, matching, and filtering.

In the parsing stage, we perform the following standard string cleaning procedures:

1. Convert company names to Unicode letters.
2. Remove the stock exchange abbreviations from the company names.
3. Convert all letters to lower cases.
4. Replace “,” by space and “..” by “.”.
5. Remove leading and trailing blanks.
6. Collapse multiple, consecutive internal blanks.
7. Randomly split all companies into chunks (10 for S&P Capital IQ and 20 for PATSTAT).
8. Use Python module called “cleanco” to remove company suffixes.
9. Drop duplicates and store to .csv files.

We then match the parsed company names using n-grams matching algorithm. Each company name is broken into a corpus of three-letter grams.²⁹ A dictionary is built by collecting all the grams of all the company names across the two databases. We then use the dictionary to transform each corpus into a vector of 0s and 1s. The list of company names is then transformed into a sparse matrix where each row corresponds to a company name and each column corresponds to a gram in the dictionary.³⁰ Each positive entry is then multiplied by a TF-IDF weight, where TF stands for term frequency and IDF stands for inverted dictionary frequency. Term frequency refers to the frequency that a gram appears in a corpus. Inverted document frequency is the inverse of the frequency that a gram appears in the entire document. TF allocates a higher weight to a gram that appear multiple times in a company name, and IDF reduces the weight of a gram that appears frequently in the dictionary.

Now that the company names are digitized, we use cosine similarity scores to select the top matches across the two databases. Technically, the dictionary can be considered as a multi-dimensional space, where each company name corresponds to vector in this space. Cosine similarity score is the cosine of the angle between two vectors. It ranges from 0 to 1. The higher the score, the more similar the company names.

The matching is implemented in two steps. We first match those company names for which we have country information. For each country, we calculate the similarity score between each and every pair of company names, and keep the top 20 matches with similarity score higher than or equal to 0.8. We then combine all the companies that do not have a match from this round of matching with those companies for which we do not have country information, and calculate pairwise similarity scores. After the second round of matching is done, we end up with 1,650,624 matches. 487,160 matches from the first step and 1,163,464 matches from the second step. After eye inspection, we decide to set the similarity score threshold for the two matched sets to be 0.95 and 0.96, respectively, which leaves us with 253,565 matches. We then keep the top match for each S&P Capital IQ company and the top match for each PATSTAT company. The resulting database contains 251,484 one-to-one matches. The location of these companies are summarized in Table A2.

²⁹For example, Panasonic is broken into the following grams: *pan, ana, nas, aso, son, oni, nic*.

³⁰The ordering of these grams is ignored, meaning that we discard information on the position of a gram in a corpus.

Table Appendix C.1: Geographic Location of the Matched Companies

Country/region	Companies	Percentage
United States	73,708	29%
China	31,266	12%
Japan	22,082	9%
Germany	15,131	6%
United Kingdom	12,681	5%
South Korea	9,583	4%
France	8,959	4%
Canada	5,803	2%
Italy	5,015	2%
Spain	4,085	2%
Taiwan	3,361	1%
Australia	3,340	1%
Sweden	3,073	1%
Switzerland	2,915	1%
Netherlands	2,896	1%
Russia	2,625	1%
Other	44,961	18%
Total	251,484	100%

Appendix D. Patent Classification

The 251,484 PATSTAT companies matched with S&P Capital IQ companies hold a total of 29,666,609 patent applications with English abstracts, which is 64% of all 46,470,475 English patent applications.³¹ We use supervised learning to train a multi-layer perceptron (MLP) model to classify these patent applications. For brevity, we will refer to these patent applications as patents.

We hire two research assistants, referred to as RA1 and RA2, to manually classify 6,000 randomly selected patents based on their titles and abstracts. RA1 has a science background and RA2 has an economics background. The patents are classified into two categories: ideas-oriented and cost-oriented, shortened as quality and cost. Quality patents include those patents that improve the quality of an existing function, or creates new function of an existing product, or creates a new product. Cost patents are those patents that improves production efficiency or reduces production costs. The research assistants' classification results are listed in the table below:

Table Appendix D.1: Patent Classification: 6000 Samples

		RA2's Classification		
		Cost	Quality	Total
RA1's Classification	Cost	920	351	1,271
	Quality	2,932	1,797	4,729
	Total	3,852	2,148	6,000

The two RAs' classifications are not quite consistent and we trust RA1's classification over RA2's classification because of two reasons. First, we hired two additional RAs to classify 1000 of

³¹Over 90% of the PATSTAT patent applications have English abstracts.

the 6000 patent sample. RA1's classification has higher overlaps with the additional classification results than RA2. Second, RA1's classification achieves much higher accuracy rates than RA2's classifications in the training process, as shown later in this section.

We then train a neural network model based on the classified results. We use three samples to train the model: RA1's classifications, RA2's classifications, and the common classifications, which include only those patents for which the two research assistants have assigned the same category. A multi-layer perceptron model can be considered as a feed-forward neural network that contains a number of neurons per layer, where the first layer transforms the input data (the sparse matrix created from patent texts) to logistic models and the last layer transforms logistic probabilities to categories. The first layer is called the input layer and the last layer is called the output layer. Those layers in-between are called the hidden layers. We choose three hyperparameters to optimize model performance: number of layers, units per layer, and dropout rate. The dropout rate is the fraction of samples that are randomly thrown out of the training sample at each training session. It is designed to reduce the overfitting problem. The MLP models are assessed by their performance in the training set and the validation set. In each set, we observe how training and validation accuracies improve over the training sessions (or epoch), and how the training and validation losses decline over training sessions (or epochs). Figures Appendix D.1 and Appendix D.2 respectively show the training results using patent abstracts and patent titles (with the best hyperparameters). Based on the model performances, we use the common classifications on patent abstracts as the main training sample and set the MLP model to have 4 layers with 16 neurons per layer with a dropout rate of 0.2.

Appendix E. Production Networks Construction

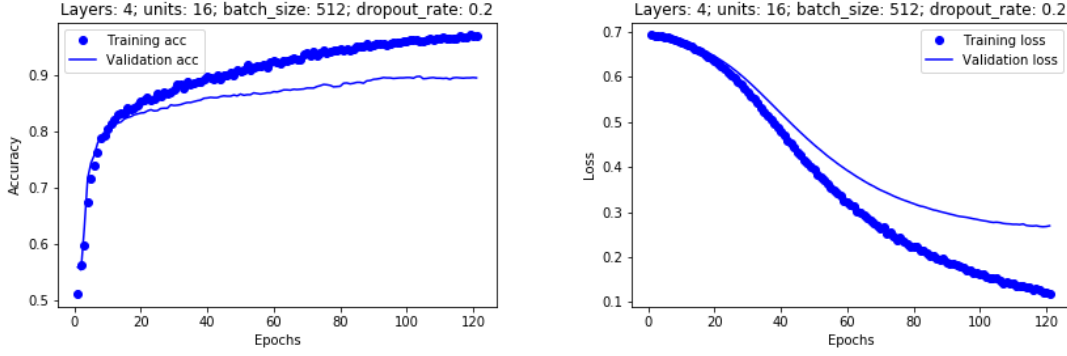
1. S&P Capital IQ Business Relationships and Ownership Structure

Sources of Information

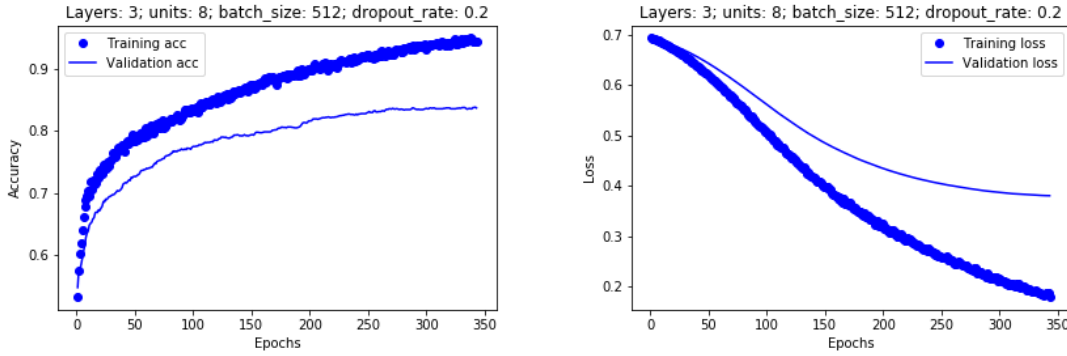
The S&P Capital IQ database contains 954,420 buyer-seller relationships between 429,776 companies over the period 2010-2017, where a buyer-seller relationship features the transaction of goods or services from the seller to the buyer. We use these buyer-seller relationships to construct information on a firm's production network, which is defined as all firms that engage in transaction relationships with the focal firm (including its customers and suppliers).

Definition of Firm Boundary and Ownership

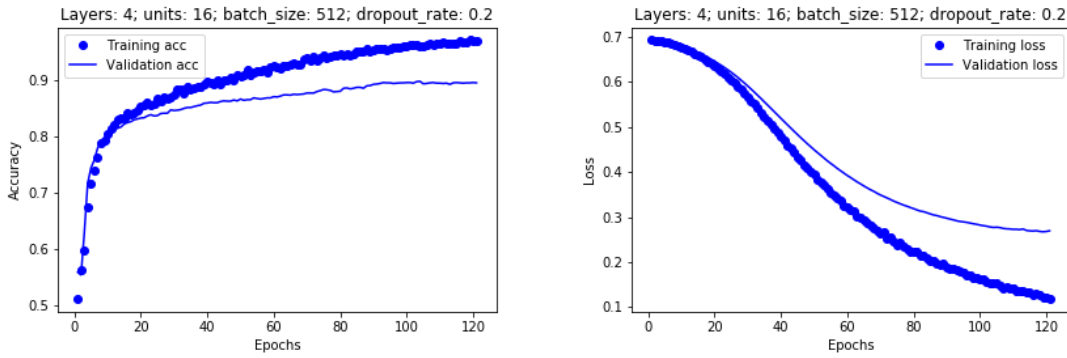
To identify the ownership structure in a buyer-seller relationship, we utilize the ownership information in S&P Capital IQ. The main ownership types include: parent, ultimate parent, limited partner, merged entity, investor, and pending parent/investor. We define two levels of ownership based on the degree of control. Level one includes investor, limited partner, and pending parent/investor; level two includes parent, merged entity, holding company, and ultimate parent. Obviously, level two is stronger than level one. Accordingly, there are two levels of firm boundary. Level one refers to those buyer-seller relationships where one firm is the other firm's investor, limited partner, or pending parent/investor; level two refers to those buyer-seller relationships where one firm is the other firm's parent, merged entity, holding company, or ultimate parent. If there are two levels of ownership between two firms, for example, if company A is both company B's parent and investor, we consider the higher-level ownership. If a seller S



(a) Training Accuracy and Loss for RA1's Classification

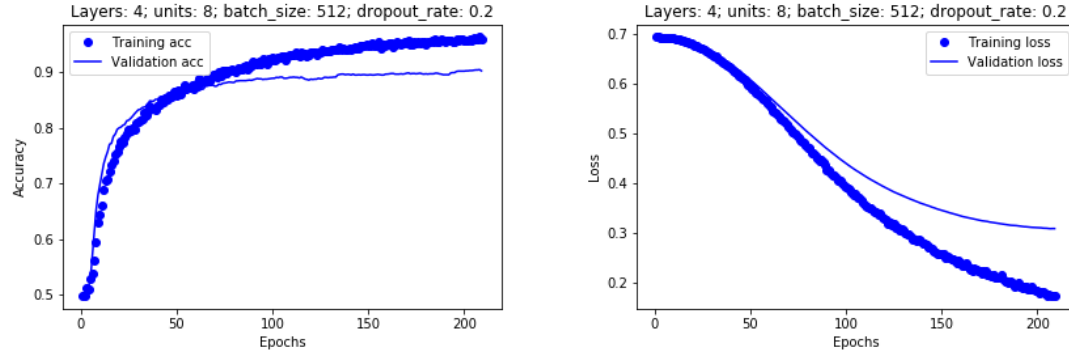


(b) Training Accuracy and Loss for RA2's Classification

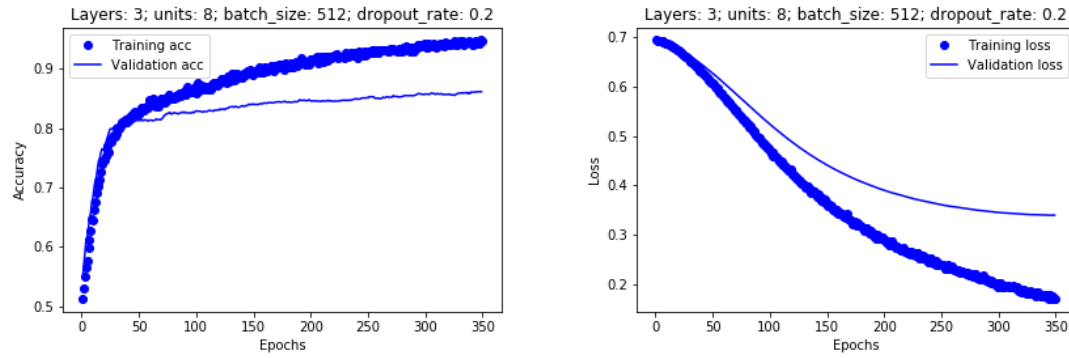


(c) Training Accuracy and Loss for Common Classification

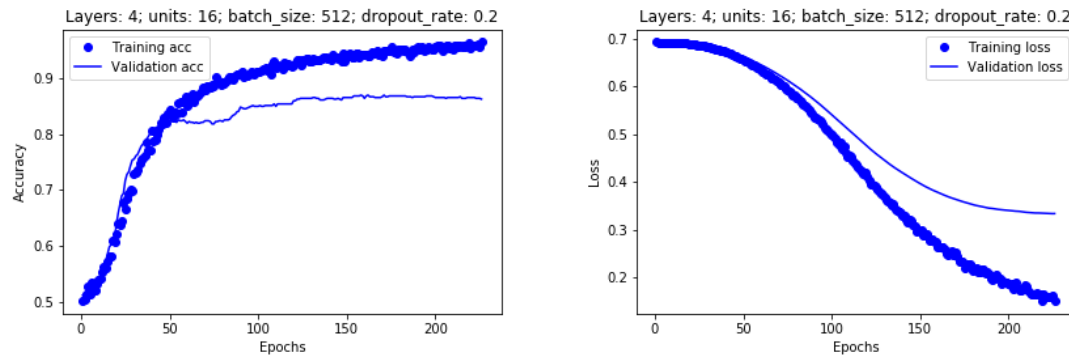
Figure Appendix D.1: Training Results Using Patent Abstracts



(a) Training Accuracy and Loss for RA1's Classification



(b) Training Accuracy and Loss for RA2's Classification



(c) Training Accuracy and Loss for Common Classification

Figure Appendix D.2: Training Results Using Patent Titles

invests in buyer B and buyer B is seller S's parent, we consider the second relationship to dominate the first one.³² There are cases where the two companies share the same parent or ultimate parent, but are not otherwise related, i.e., there is no direct ownership relationship between these two companies. In such cases, we consider the two companies to be integrated (thus have level two firm boundary), but neither firm owns the other. However, these two firms will not be used to construct production network information, as the relationship does not have a clearly-defined "owner".

Under the above definition, each buyer-seller relationship is characterized by five possible ownership structures: buyer owns seller at level two, buyer owns seller at level one, seller owns buyer at level two, seller owns buyer at level one, and neither firm owns the other. If there are bilateral transactions between two firms, the ownership structure should be mirrored between the two relationships. Table A3 shows the percentage of each of these relationships.

Table Appendix E.1: Ownership Structure in the Original Dataset

Ownership Structure	Relationships	Percentage
Buyer owns seller (level two)	13,167	1.38%
Buyer owns seller (level one)	3,485	0.37%
Seller owns buyer (level two)	3,714	0.39%
Seller owns buyer (level one)	16,587	1.74%
Neither	917,467	96.13%
Total	954,420	100.00%

2. Imputing Firm-Supplier Relationships from Parent-Subsidiary Information.

Let us refer to those buyer-seller relationships compiled from S&P Capital IQ as the original dataset. According to Table A3, integrated relationships (including both level one and level two ownership) are merely 3.87% of all 954,420 relationships, which indicates that the sample ignores many within-firm transaction relationships. We therefore adopt an imputation method that has been formalized in previous literature to uncover those within-firm transaction relationships.³³

Table Appendix E.2: Ownership Structure in the Extended Dataset

Ownership Structure	Relationships	Percentage
Buyer owns seller (level two)	674,904	30.29%
Buyer owns seller (level one)	3,485	0.16%
Seller owns buyer (level two)	626,918	28.14%
Seller owns buyer (level one)	3,714	0.17%
Neither	918,858	41.24%
Total	2,227,879	100.00%

We use U.S. BEA 2002 Input-Output Table to impute buyer-seller relationships from parent-subsidiary relationships in the S&P Capital IQ database. More specifically, denote the industries of the parent and the subsidiary by $i(p)$ and $i(s)$, respectively. Denote the input-output linkage between upstream industry i and downstream industry j by $l(i,j)$, where linkage is the

³²There is no incidence where the two firms have the same degree of control over each other.

³³See for example Acemoglu et al. (2010) and Antràs and Chor (2013).

amount of input that the downstream industry directly purchases from the upstream industry to produce one dollar of output. A transaction relationship between p and s exists as long as $\max\{l(i(p),i(s)),l(i(s),i(p))\} > 0.01$. If the condition is satisfied, the transaction goes from the parent to the subsidiary if $l(i(p),i(s)) > l(i(s),i(p))$, and from the subsidiary to the parent if $l(i(s),i(p)) > l(i(p),i(s))$. If $l(i(s),i(p)) = l(i(p),i(s))$, we consider the transaction to be bilateral. The direction of transaction does not matter for our exercise because in the empirical section, we define a firm's production network as composing of itself, its customers, *and* its suppliers. A firm's partner in its production network could refer to either its customer or its supplier. The direction of transaction is unimportant. Therefore bilateral transaction relationships do not result in over-counting in the firm production networks. In addition, we do not include those partners that own the focal firm,³⁴ so the parent will not be counted as the subsidiary's partner.

We impute 1,286,076 buyer-seller relationships from 1,157,184 parent-subsidiary relationships using the above methodology. 12,617 of them overlap with the original dataset. Table A4 summarizes the ownership structure in the extended dataset.

³⁴Because the integration decision is not made by the focal firm in this case.