A price theory of vertical and lateral integration under two-sided productivity heterogeneity*

Kaniṣka Dam† Konstantinos Serfes‡

Abstract

We analyze the interplay between product market prices and firm boundary decisions. Heterogeneous units are matched to form enterprises. Each enterprise can choose between two ownership structures – centralized ownership (integration) performs well in coordinating managerial actions but ignores private costs; dispersed ownership (non-integration), on the other hand, overvalues private costs but is conducive to poor coordination. A more productive unit on one side of the market matches with a more productive unit on the other side following a positive assortative matching pattern. The equilibrium ownership structure is monotone, i.e., high productivity units integrate while the low productivity ones stay separate. Product market price can be positively or negatively associated with the incidence of integration, depending on how a price change affects the endogenously determined distribution of surplus. As higher prices may imply less integration in the market, the industry supply may be backward-bending. Our model delivers new empirical and policy implications.

Key words: Two-sided heterogeneity; vertical and lateral integration; positive assortative matching.

JEL codes: D21, D86, L14, L22.

*We are grateful to Alessandro Bonatti (the associate editor) and two anonymous referees for insightful comments that helped improve the paper. We thank Eva Chalioti, Joseph Harrington, Matthew Nagler, David Pérez-Castrillo, Patrick Rey, Ramya Shankar, Jack Stecher, Chad Syverson, Galina Vereshchagina, Nikos Vettas, the seminar participants at CCNY, CIDE, SWUFE, the conference participants at CRETE 2013, SAET 2013 and EARIE 2014 for helpful suggestions, and Mitchell Valdes for excellent research assistance. Dam acknowledges the research grant, "Fondo de Apoyo a la Investigación" from CIDE.

†Corresponding author. Center for Research and Teaching in Economics (CIDE), 3655 Carretera Mexico-Toluca, Lomas de Santa Fe, 01210 Mexico City, Mexico. E-mail: kaniska.dam@cide.edu

‡School of Economics, LeBow College of Business, Drexel University, 3220 Market Street, Philadelphia PA 19104. E-mail: ks346@drexel.edu

1 Introduction

There is a plethora of evidence on heterogeneity of firm productivity within an industry, which is also associated with organizational variation at the firm level (e.g. Gibbons, 2010; Syverson, 2011) and endogenous sorting among firms (e.g. Hortaçsu and Syverson, 2007; Atalay, Hortaçsu, and Syverson, 2014; Braguinsky, Ohyama, Okazaki, and Syverson, 2015). In this paper, we analyze the interplay between product market competition and firm boundary decisions (choice of ownership structure) under ex-ante productivity heterogeneity. We intend to shed light on the following questions pertaining to the Organizational Industrial Organization (OIO) literature which lies at the intersection of organizational and industrial economics: (i) do firms establish relationships with other firms following a positive assortative matching with respect to their productivities?, (ii) are higher productivity enterprises more likely to integrate?, (iii) what is the effect of product market competition on the likelihood of integration? and (iv) is the industry supply curve always upward sloping, when we incorporate the possible reorganization of the firms a price change induces?

The building block for the analysis is the model by Legros and Newman (2013). Following their paper, we posit a model where output is produced by combining two complementary assets, where each asset is run by a (cash constrained) manager. In the airline industry, for example, regional airlines operate as "subcontractors" for major US network carriers on short- and medium-haul routes, often connecting smaller cities to the major carrier hubs, (see Forbes and Lederman, 2009, 2010). The output in this example is a traveler who travels on two complementary flight segments, one served by a regional carrier, and the other, by a major. In general, majors own some regional flights (integration), and outsource services to independent firms (non-integration). As another example, consider the healthcare industry where as of 2015 approximately 20 percent of all Medicare fee-for-service hospital admissions ended in skilled nursing facility (SNF) stays (e.g. Zhu, Patel, Shea, Neuman, and Werner, 2018). In this example, hospitals and SNFs are the two complementary units that should coordinate to guarantee the best health outcome for each patient. Traditionally, hospitals and SNFs receive separate payments for the care they provide. To reduce spending and improve quality of care, Medicare recently introduced bundled payment programs that link payments for multiples services related to a single episode of care.

In our model, the two units can integrate or stay separate and use revenue sharing contracts to govern their relationship. In order to study possible inefficiencies that emerge due to firm boundary decisions, we abstract from any other form of inefficiency that can arise from imperfections in the product market. In the decision of the two units about whether to integrate or not they face a trade-off between coordination benefits and private costs. Under integration, better coordination of non-contractible managerial efforts boosts firm revenue, but at the expense of higher private costs for the managers, whereas the opposite is true when they remain separate. Air-carriers, for instance, under integration, can adopt common practices that reduce delays and other costs, but at the expense of a change in the managers' daily routine these common practices will introduce.[1] In response to the bundled payment initiative by Medicare, coordination and communication between hospitals and SNFs have been improved either by integration or by formal contractual agreements (see Zhu et al., 2018).

Neither organization achieves full efficiency. Which one dominates the other depends crucially on the

[1]As discussed in Hart and Hölmstrom (2010), two firms may want to adopt a common standard, as in Cisco's acquisition of Stratacom. The benefits for the organization from a common standard can be enormous but the private costs within the firms may increase because of the change the new standard introduces. Moreover, there is no agreement between the firms about which "approach" should be adopted. However, agreeing on a common approach (coordination) boosts firm revenue.

market price and how the total surplus in the relationship is distributed between the two units. A higher market price favors integration because it increases the value of output, and hence, the benefits from coordination, while a more balanced distribution of the total surplus favors non-integration because it adequately incentivizes the two managers to coordinate their efforts even when the two units are separate (as is standard in models with moral hazard in teams).

We extend Legros and Newman (2013) by introducing two-sided productivity heterogeneity and endogenous sorting between the firms. One firm from the one side of the market, say side A, matches (forms a relationship) with one firm from the other (one-to-one matching), say side B, to create an enterprise. These two firms also decide whether to remain separate as production units or to integrate, facing the trade-off between coordination benefits and private costs we discussed above. Higher productivity units are more desirable, so there is competition among all units for the high productivity ones, which endogenizes the allocation of utilities. The utility of any firm is determined by the characteristics (e.g. marginal productivity) of all the inframarginal firms in the market (i.e., firms with lower productivities).

We find that, under reasonable conditions, matching is positive assortative (PAM) with respect to firm productivities irrespective of the choice of ownership structure.[2] Then, we establish conditions under which high productivity enterprises integrate, while low productivity ones choose to remain as separate units. Both of these predictions have empirical support, as we discuss in Section 8.

The aforementioned equilibrium outcome yields interesting predictions with respect to the competitiveness of the product market as measured by the market price. A higher market price can induce more or less integration in the market. The intuition for a negative relationship between market price and the incidence of integration is as follows. A higher market price, as we discussed above, is a force in favor of integration, but at the same time it can lead to more balanced equilibrium utility allocation among the units, which is a force in favor of non-integration. If the outside option of the units on one side of the market, say side B, is high then the equilibrium utility they will receive will also be high. High outside option may be because the assets of the B firms are highly valuable in other markets. Now consider the enterprise that is indifferent between integrating and staying non-integrated. The utility allocation at this enterprise is unbalanced in favor of the B firm (if the utilities were balanced then the enterprise would prefer non-integration). But if at the same time the marginal productivity of the inframarginal B firms is low or they are not heterogeneous enough, then competition among the firms on the other side of market, side A, for the relatively high productivity B firms is mitigated, and hence, most of the extra surplus due to a price increase goes to the A side. This makes the allocation of utilities more balanced, and hence, the indifferent enterprise may prefer non-integration at a higher price.[3] But if the inframarginal firms on the B side have a high marginal productivity, or productivity heterogeneity is high, then a higher price will make utilities even more unbalanced in favor of B and more enterprises will integrate.

Holding the organizational structures in the market fixed, higher market price implies higher aggregate output, because coordination under non-integration improves, and the aggregate output in the integrated enterprises is at its highest potential. But if a higher price implies less integration, then this

[2]In particular, we find conditions under which the generalized increasing difference (GID) condition, which is a necessary and sufficient condition for positive sorting, proposed by Legros and Newman (2007), holds. Proving GID in our context is not straightforward because the bargaining frontiers of some enterprises are non-differentiable and non-convex due to the choice between the two ownership structures.

[3]Consider, for example, Chakraborty and Citanna (2005) who examine an occupational choice model with wealth heterogeneity, two-sided moral hazard and matching. As in our model, the division of the gains from a match is determined by competitive forces. They show that matches are typically wealth heterogeneous, with richer individuals choosing the occupation for which incentives are more important.

introduces a countervailing force on aggregate output because non-integrated firms produce less output than the integrated ones. If this force is stronger, then industry output may decrease with price. In other words, the organizationally augmented supply curve (OAS) can be *backward-bending* due to the organizational restructuring a price change triggers.

Our results are significant because they help to show why "opening the black box" of the firm may have dramatic implications for understanding industry behavior and performance, even to the point of challenging some of the most unquestioned ideas in industrial economics, such as upward-sloping supply curves. We offer more discussion about the implications of our results in Sections 8 and 9.

2 Related literature and our contribution

The literature on OIO, which is concerned with how market structure affects firm boundaries decisions, is still in its early stages of development.[4] Our paper adds to the recent contributions in this strand of literature. As we have mentioned earlier, for any given enterprise the decision to integrate depends on how the enterprise surplus is divided among the units. When shares are unbalanced, non-integration performs poorly in coordinating managerial actions because incentives cannot be easily aligned, and hence, integration is the preferred choice. However, this basic trade-off is not new. It has been explored in Legros and Newman (2013) and other prior works (e.g. Grossman and Helpman, 2002). Legros and Newman (2013) derive two main results: (i) despite a continuum of firms being homogeneous, the two organizational modes may co-exist in the market, and (ii) higher product market price makes integration more likely, and hence, the organizationally augmented industry supply curve (OAS) is upward sloping.

Coordination motives as the main driver for vertical and lateral integration has earlier been analyzed by Hart and Hölmstrom (2010), although there is no scope for ex-ante revenue sharing contract as ex-post bargaining is efficient. The role of surplus sharing in determining the choice between vertical integration and outsourcing has also been analyzed in Grossman and Helpman (2002). But their mechanism is different in the sense that large governance costs in the vertically integrated firms are balanced against costs arising from a holdup problem (as in Grossman and Hart, 1986) and search for suitable partners under outsourcing. In the equilibrium of the market with identical participants, either all firms vertically integrate or there is pervasive outsourcing. Extreme revenue share makes an equilibrium with integration more likely to occur because it generates excess demand or excess supply for intermediate inputs.

We differ from the aforementioned works in the following aspects. In Legros and Newman (2013), all units are homogeneous, the revenue shares are endogenously determined, but the outside option of the firms on one side of the market is exogenously fixed. In Grossman and Helpman (2002), on the other hand, the share of intermediate input suppliers is exogenously given, which reflects the degree of input market competition, and it does not interact with the degree of product market competition. As we allow units to be ex-ante heterogeneous, endogenous sorting leads to an endogenous distribution of surplus. Even when Legros and Newman (2013) allow for ex-ante heterogeneity in firm productivities, because the matching is exogenous, all units who receive offers consume their fixed reservation payoff. In light of our framework, the same utility allocation or surplus sharing cannot be part of a stable equilibrium

[4]See Legros and Newman (2014) for an excellent survey. These authors argue: "Nascent efforts at developing an OIO already suggest that market conditions or industrial structure matter for organization design. At the same time, organizational design will affect the productivity of firms, hence eventually the total industry output, the quality of products and information about this quality for consumers. Organizational design matters for consumers, hence for IO."

as more productive units must receive higher utility. The endogenous distribution of surplus in our model has further implications for the association between product market price and integration. Due to endogenous sorting, the surplus division in the infra-marginal firms determines how balanced the surplus sharing is in the marginal firm, which in turn determines the choice of ownership structure of the marginal firm following a price increase. Higher surplus can generate a more even allocation of utilities in the marginal firm, which favors non-integration. Therefore in our model, a rise in price may lead to less integration, and consequently, the industry supply curve may be backward-bending.[5]

To summarize, our contribution is that we provide a particular mechanism, which relies on two-sided heterogeneity and sorting, to illustrate how the effect of surplus sharing on firm boundary decisions manifests itself in the market.[6] In doing so, we offer a more complete picture of the interaction of market price with integration decisions and output, given that productivity heterogeneity and sorting are ubiquitous in markets. On the technical side, ours is one of the first few applications of recent results in the theory of assortative matching under imperfectly transferable utility to a non-smooth environment.

Gibbons, Holden, and Powell (2012) obtain generic heterogeneity of ownership by analyzing a rational-expectations equilibrium of price formation and endogenously chosen governance structures. They show that the informativeness of the price mechanism can induce ex-ante homogeneous firms to choose heterogeneous governance structures. Aghion, Griffith, and Howitt (2006) provide evidence of a U-shaped relationship between product market competition and vertical integration. Integration is more likely when competition is either soft or intense. In addition to the role of product market competition, we emphasize the significance of competition for heterogeneous units in the input market in an endogenous matching framework.

Alfaro, Conconi, Fadinger, and Newman (2016) use changes in trade policy, e.g. tariffs, as an exogenous source of price variation. They find empirical evidence that the level of product prices do affect vertical integration decisions. Acemoglu, Griffith, Aghion, and Zilibotti (2010) show that technology intensity affects the likelihood of integration. They find that technology intensity of the downstream producers is positively associated with integration, while technology intensity of the upstream suppliers is negatively associated. Our model also disentangles the contributions of the two units to overall firm productivity.

In a model of managerial incentives under endogenous matching, Alonso-Paulí and Pérez-Castrillo (2012) analyze the choice between incentive and Codes of Best Practice (CBP) contracts. The presence of two different contracting modes gives rise to a non-concave bargaining frontier for each shareholder-manager pair, as it is the case in our model. They show that for considerable range of parameter values the equilibrium matching may be negative assortative with a high-revenue shareholder offering a CBP contract to a less efficient manager, whereas a low-revenue shareholder offers an incentive contract to lure a more efficient manager. Macho-Stadler, Pérez-Castrillo, and Porteiro (2014) shows the robust co-existence of two contracting modes – namely, short- and long-term contracts in a labor market where

[5] Legros and Newman (2013, pp.746-747) discuss how a technological shock that impacts only a fraction of the firms in the market affects the organization of all firms. It turns out that the shock leads to less integration by the firms unaffected by the shock, and hence, less output. This is an example of an organizational external effect: the organizational change is coming from outside the firm, transmitted by the market. The reorganization in our model, caused by a price change, is operating through different channels: the distributional as well as the price channels, but there are similarities. Most notably, the organizational external effect is also a force in our model and is operating through the endogenously determined outside options.

[6] Two-sided productivity heterogeneity and matching have also become increasingly important in the context of international trade, as there are easily available datasets with detailed information about the matching between exporters and importers (e.g. Bernard, Moxnes, and Ulltveit-Moe, 2014; Dragusanu, 2014; Sugita, Teshima, and Seira, 2015).

heterogeneous firms are endogenously matched with heterogeneous workers. There is positive sorting in the market for short-term contracts, whereas under long-term contracts, sorting does not emerge because the gains from incentive provision by the long-term contracts are superior to the benefits of sorting under the short-term contracts.[7]

3 The Model

3.1 Technology and matching

Consider a two-sided market where on each side there is a continuum of firms or (supplier) units of measure 1. Firms are vertically differentiated with respect to their productivity.[8] In particular, let $J_A = [0,\, 1]$ be the set of "A firms" on the one side of the market and $J_B = [0,\, 1]$, the set of "B firms" on the other side. Each unit $i \in J_A$ is assigned a type or 'productivity' $a = a(i) \in A$ and each $j \in J_B$ has an assigned type $b = b(j) \in B$ where the type spaces $A = [\underline{a},\, \bar{a}]$ and $B = [\underline{b},\, \bar{b}]$ are subintervals of $\mathbb{R}_{++}$. Let $G(a)$ be the fraction of A firms with productivity lower than a, i.e., $G(a)$ is the cumulative distribution function of a with the associated density function $g(a) > 0$ for all $a \in A$. Similarly, let $F(b)$ be the distribution function of b with the associated density function $f(b) > 0$ for all $b \in B$.

Production of a homogeneous consumer good requires one A unit and one B unit who are matched one to one to form an 'enterprise'. All decisions and payoffs of each enterprise will only depend on the types of the two participating units, and hence, a typical enterprise will be denoted by $(a,\, b)$. A matching is a one-to-one mapping $\alpha : B \to A$ which assigns to each $b \in B$ a type $a = \alpha(b) \in A$. Such enterprises may include lateral as well as vertical relationships. The stochastic output of an enterprise $(a,\, b)$ is given by:

$$\tilde{y}(a,\, b) = \begin{cases} z(a,\, b) & \text{with probability } \ \pi(e_A,\, e_B) \equiv 1 - (e_B - e_A), \\ 0 & \text{otherwise.} \end{cases}$$

The success output $z(a,\, b)$ can be thought of as the productivity of an enterprise $(a,\, b)$.[9] We assume that $z(a,\, b)$ is twice continuously differentiable, strictly increasing in a and b, and supermodular in $(a,\, b)$, i.e., $z_{ab}(a,\, b) \geq 0$.[10] Each unit must make a non-contractible production decision: $e_A \in [0,\, 1]$ by an A firm and $e_B \in [0,\, 1]$ by a B firm. These decisions can be made by the manager of the assets or by someone else. If the two managers coordinate their decisions and set $e_A = e_B$, then inefficiencies disappear and the enterprise reaches its full potential $z(a,\, b)$ with probability 1. The manager of each firm is risk neutral and incurs a private cost for the managerial action. The private cost of an A unit is ce_A^2 and that

[7]Chade, Eeckhout, and Smith (2017) offer an excellent survey of recent results in the theory of assortative matching.

[8]Bloom and van Reenen (2007), using a survey from medium-sized manufacturing firms from four countries, document that management practices are heterogeneous and affect firm performance. Gibbons (2010) offers a more detailed account of various empirical studies that document persistent performance differences (PPDs). In the computer industry, computer systems manufacturers rely on networks of independent component suppliers. These suppliers are of various 'qualities' and produce components that are used as inputs in the production of the final product (see Fallick, Fleischman, and Rebitzer, 2006).

[9]Legros and Newman (2013) assume a quadratic probability of success of the form $\pi(e_A,\, e_B) \equiv 1 - (e_A - e_B)^2$. We use a linear one instead for tractability. The linear probability of success is everywhere below the quadratic one implying that non-integration is more profitable under the quadratic probability. That is why in Legros and Newman (2013) integration is never a strictly dominant organizational mode. Nevertheless, the important qualitative features of the model are not affected by the choice of linear probability.

[10]We use subscripts to denote partial derivatives and when the subscript is a number it denotes the position of the variable with respect to which we differentiate, e.g. $\phi_2(a,\, b,\, u)$ is the partial derivative of ϕ with respect to b.

of a B unit is $c(1-e_B)^2$ with $c > 0$. Clearly, there is a disagreement about the direction of the decisions, what is easy for one is hard for the other. Also, managers with zero cash endowments are protected by limited liability, i.e., their state-contingent incomes must always be non-negative. The importance of this assumption is that the division of surplus between the managers will affect the organizational choice.

3.2 Ownership structures and contracts

The ownership structure can be contractually determined. We assume two different options, each of which implies a different allocation of decision rights. First, the production units can remain as separate firms (the *non-integration* regime, denoted by N). In this case, managers have full control over their decisions. Second, the two units can integrate, a regime denoted by I, into a single firm by selling their assets to a third party, called the headquarter (HQ), which gives HQ full control over managerial decisions, e_A and e_B, assuming that the third party possesses enough cash to finance the acquisition.[11] The headquarter is motivated entirely by revenue and incurs no costs from the managerial decisions. These costs are still borne by the managers. As argued by Hart and Hölmstrom (2010), integration results in an organization where less weight is placed on private costs than under non-integration. This, however, is offset by the fact that under integration total revenue, rather than individual unit profits, is maximized.

The revenue of each enterprise is publicly verifiable, and hence, ex-ante contractible. We assume that each A firm has all the bargaining power in an arbitrary enterprise $(a,\, b)$ and makes take-it-or-leave-it offers to the B firm.[12] A contract $(s,\, d) \in [0,\, 1] \times \{N,\, I\}$ specifies a revenue share s for the B unit and an ownership structure d. Consider an arbitrary enterprise $(a,\, b)$. If the members of this enterprise stay separate, then a revenue-sharing contract is simply a share s of the total revenue that accrues to the B unit. As we assume limited liability, the units get nothing in the case of failure.

When the two units integrate, HQ buys the assets of the A and B units at predetermined prices in exchange of a share contract $\mathbf{s} = (s_A,\, s_B,\, s_{HQ}) \in \mathbb{R}^3_+$ with $s_A + s_B + s_{HQ} = 1$. HQs are supplied perfectly elastically with an opportunity cost normalized to zero.

3.3 The product market

The product market is perfectly competitive where consumers and producers take the product price P as given. Identical consumers maximize a smooth quasi-linear utility which gives rise to a downward-sloping demand curve $D(P)$. Enterprises correctly anticipate price P when they sign contracts and make their production decisions. Define $R(a,\, b) \equiv Pz(a,\, b)$, the revenue of an enterprise $(a,\, b)$ in the event

[11] The two units supply complementary inputs to produce a single homogenous good. If we think of enterprises as vertical relationships, one unit, say, A may be named the "upstream" firm, and the other, the "downstream" firm. In our model, lateral and vertical relationships are somewhat equivalent because the sole motive for integration is to improve coordination among the units which is achieved by conferring the decision making rights on a third party. We do not consider vertical integration in a more traditional sense where the rights to make decision belong to the integrated entity, and in which there are the usual efficiency gains such as ameliorating the problem of double marginalization.

[12] In a model with a continuum of types, a particular bargaining protocol is irrelevant, and hence, assuming a take-it-or-leave-it bargaining protocol is innocuous. This is because, due to the continuity assumptions, the factor owners do not earn rents over their next best opportunity within the market, as the next competitor is arbitrarily close. However, in a model with discrete types there would be a match-specific rent left for bargaining.

of success.

3.4 Timing of events

The economy lasts for two dates, $t = 1, 2$. At date 1, one A firm and one B firm match one to one to form an enterprise $(a,\, b)$ and each A unit makes a take-it-or-leave-it contract offer $(s,\, d)$ to each B unit. At date 2, the manager of each unit chooses e_A and e_B. We solve the model by backward induction.

3.5 Equilibrium

An equilibrium of the input market consists of a set of enterprises formed through feasible contracts, i.e., ownership structures and corresponding revenue shares, for each enterprise and a market-clearing price. Recall that there are two possible ownership structures for each enterprise – integration (I) and non-integration (N). In general, choice of ownership structures depends on the revenue share that accrues to each member of an enterprise, the output of each enterprise and the market price. An allocation for the market $\langle \alpha,\, v,\, u \rangle$ specifies a one-to-one matching rule $\alpha : B \to A$, and payoff functions $v : A \to \mathbb{R}_+$ and $u : B \to \mathbb{R}_+$ for the A and B firms, respectively.

Definition 1 (Equilibrium) *Given the type distributions $G(a)$ and $F(b)$, an allocation $\langle \alpha,\, v,\, u \rangle$ and a product-market price P constitute an equilibrium allocation of the economy if they satisfy the following conditions:*

(a) **Feasibility:** *The revenue shares and the corresponding payoffs to the agents in each equilibrium enterprise are feasible given the output of the enterprise and the equilibrium price P;*

(b) **Optimization:** *Each A firm of a given type chooses optimally a B firm to form an enterprise, i.e., given $u(b)$ for $b \in B$, each $a \in A$ solves*

$$v(a) = \max_{b} \, \phi(a,\, b,\, u(b),\, P). \qquad (\mathcal{M}_a)$$

The function $\phi(a,\, b,\, u(b),\, P)$ is the bargaining frontier *or* Pareto frontier *of the enterprise $(a,\, b)$, which is the maximum payoff that can be achieved by a type a A unit given that the B unit of type b consumes $u(b)$ at a given market price P.*

(c) **Input market clearing:** *The equilibrium matching function satisfies the following 'measure consistency' condition. For any subinterval $[i_0,\, i_1] \subseteq J_A$, let $i_k = G(a_k)$ for $k = 0,\, 1$, i.e., a_k is the productivity of the A firm at the i_k-th quantile. Similarly, for any subinterval $[j_0,\, j_1] \subseteq J_B$, let $j_h = F(b_h)$ for $h = 0,\, 1$. If $[a_0,\, a_1] = \alpha([b_0,\, b_1])$, then it must be the case that*

$$j_1 - j_0 = F(b_1) - F(b_0) = G(a_1) - G(a_0) = i_1 - i_0. \qquad \text{(MC)}$$

(d) **Product market clearing:** *The aggregate (expected) supply in the industry $Q(P)$ is equal to the demand $D(P)$.*

Definition 1-(b) asserts that each A firm chooses her partner optimally. Part (c) of the above definition simply says that one cannot match say two-third of the A units to one-third of the B units because the matching is constrained to be one-to-one.

4 Optimal ownership structure for an arbitrary enterprise

We analyze the optimal contract for an arbitrary enterprise $(a,\, b)$. We first analyze each ownership structure separately.

4.1 Non-integration

Under this organizational mode the shares affect both the size and the distribution of surplus between the two units (imperfectly transferable utility). An optimal contract for a non-integrated enterprise $(a,\, b)$ solves the following maximization problem:

$$\max_{s} \quad V_A \equiv \pi(e_A,\, e_B)(1-s)R(a,\, b) - ce_A^2, \tag{$\mathcal{P}_N$}$$

$$\text{subject to} \quad U_B \equiv \pi(e_A,\, e_B)sR(a,\, b) - c\,(1-e_B)^2 = u, \tag{PC$_B$}$$

$$e_A = \underset{e}{\text{argmax}} \,\left\{\pi(e,\, e_B)(1-s)R(a,\, b) - ce^2\right\} = \frac{(1-s)R(a,\, b)}{2c}, \tag{IC$_A$}$$

$$e_B = \underset{e}{\text{argmax}} \,\left\{\pi(e_A,\, e)sR(a,\, b) - c(1-e)^2\right\} = 1 - \frac{sR(a,\, b)}{2c}, \tag{IC$_B$}$$

where u is the outside option of the B unit. We assume that $u \geq u_0$, where $u_0 \geq 0$ is the reservation utility of all B firms, i.e., the utility any B firm obtains if he does not form a parnership with any A firm. Constraint (PC$_B$) is the *participation constraint* of the B unit, whereas constraints (IC$_A$) and (IC$_B$) are the *incentive compatibility constraints* of the A firm and the B firm, respectively. We analyze the optimal contracts under non-integration when $R(a,\, b) < 2c$. Note that (IC$_A$) and (IC$_B$) together imply that $\pi(e_A,\, e_B) = R(a,\, b)/2c$. When the firms in an arbitrary enterprise $(a,\, b)$ stay separate, at a given product market price P, the maximum payoff that accrues to the A unit given that the B unit consumes u is given by:

$$\phi^N(a,\, b,\, u,\, P) = \frac{1}{4c}\left[2R(a,\, b)\sqrt{R(a,\, b)^2 - 4cu} - \left(R(a,\, b)^2 - 4cu\right)\right] \text{ for } 0 \leq u \leq \frac{R(a,\, b)^2}{4c}. \tag{1}$$

The details about the derivation of (1) are relegated to the Appendix. The function $\phi^N(a,\, b,\, u;\, P)$ is the *bargaining or Pareto frontier* under non-integration. The participation constraint of the B unit determines the optimal revenue share $s = s(a,\, b,\, u,\, P)$ of each type b. Note that u must lie between 0, which corresponds to $s = 0$, and $R(a,\, b)^2/4c$, the level corresponding to $s = 1$. The bargaining frontier is strictly concave and symmetric with respect to the 45^0 line, on which $\phi^N(a,\, b,\, u,\, P) = u$ and $s = 1/2$. This implies that total surplus is maximized when the shares across the two non-integrated units are equal. Equal, or more broadly 'balanced', shares yields strong incentives for the managers to better coordinate their decisions, i.e., e_A and e_B move closer to each other. Finally, higher revenue R, holding the shares fixed, also induces better coordination. At a given product market price P, any non-integrated enterprise $(a,\, b)$ produces an expected output which is given by:

$$q^N(P,\, z(a,\, b)) = \underbrace{\pi(e_A(a,\, b),\, e_B(a,\, b))}_{R(a,\, b)/2c} \cdot z(a,\, b) = \frac{P[z(a,\, b)]^2}{2c}.$$

The above quantity $q^N(P,\, z(a,\, b))$ is lower than $z(a,\, b)$, i.e., a non-integrated enterprise does not reach its full potential.

4.2 Integration

When the units integrate, the enterprise is acquired by HQ who is conferred with the decision making right. Motivated entirely by incomes, HQ will choose e_A and e_B to maximize the expected revenue $\pi(e_A, e_B)R(a, b)$ as long as $s_{HQ} > 0$. This induces $e_A = e_B$, and hence, $\pi(e_A, e_B) = 1$ for all integrated enterprises. The HQ breaks even as the market for headquarters is perfectly competitive. The private costs of managerial actions are still borne by the individual units. The aggregate managerial cost, $ce_A^2 + c(1-e_B)^2$ is minimized when $e_A = e_B = 1/2$. Thus, the bargaining frontier under integration is given by:

$$\phi^I(a, b, u, P) = R(a, b) - \frac{c}{2} - u \quad \text{for } 0 \leq u \leq R(a, b) - \frac{c}{2}. \tag{2}$$

The above function is linear in u, i.e., surplus is fully transferable between the two managers because neither the action taken by HQ nor the costs borne by the managers depends on the revenue shares. The function $\phi^I(a, b, u; P)$ is strictly increasing in a and b, strictly decreasing in u (with slope -1) and symmetric with respect to the 45^0 line. The expected output produced by an arbitrary integrated enterprise (a, b) is given by:

$$q^I(P, z(a, b)) = \underbrace{\pi(e_A(a, b), e_B(a, b))}_{1} \cdot z(a, b) = z(a, b).$$

Although surplus is fully transferable between the A and B firms, this form of organization is in general not efficient as HQ, having a stake in the enterprise's revenue, places too little weight on private managerial costs while maximizing expected revenue.[13]

4.3 Choice of ownership structures

We now analyze the optimal choice of ownership structure by a given enterprise. At any given product market price P and utility u accruing to the B firm, an arbitrary enterprise (a, b) would choose N over I if and only if $\phi^N(a, b, u, P) > \phi^I(a, b, u, P)$. In an arbitrary enterprise (a, b), the optimal choice of ownership structure depends on the revenue of the enterprise, $R(a, b)$ as well as the way the enterprise revenue, or, equivalently, the aggregate surplus is shared between the two units. Low revenue, i.e., $R(a, b) < R^- \equiv \left(2 - \sqrt{2}\right) c$ implies that an enterprise places more emphasis on private costs relative to the benefits accruing from coordination, and hence, the aggregate surplus from non-integration is strictly higher than that under integration, i.e., $\phi^N(a, b, u, P) > \phi^I(a, b, u, P)$ for all levels of u. Thus, the enterprise chooses non-integration over integration. By contrast, for the high-revenue (or high-productivity) enterprises with $R(a, b) > R^+ \equiv 2c/3$, more weight is placed on coordination and revenue maximization, and hence, integration is the optimal choice as $\phi^I(a, b, u, P) > \phi^N(a, b, u, P)$ for all u. Both of the aforementioned choices depend only on the level of revenues (low or high), and are independent of how the revenue is shared. Interestingly, for intermediate-productivity enterprises, i.e., $R(a, b) \in [R^-, R^+]$, there is no clear dominance of one mode of organization over the other, and hence, the choice of organizational modes depends on how the surplus of the enterprise is distributed between the two units, i.e., on the levels of u. This case is depicted in Figure 1 where the strictly concave frontier is the one associated with N, i.e., $\phi^N(a, b, u, P)$, and the linear frontier is $\phi^I(a, b, u, P)$, the frontier

[13] The first-best surplus, $\frac{R^2}{2c}$, is strictly higher than $R - \frac{c}{2}$, the surplus accrued to an integrated firm as well as $\frac{3R^2}{8c}$, the maximum surplus in a non-integrated firm, which corresponds to $s = \frac{1}{2}$. The diminished output under non-integration reflects the distortionary effect of incentive contracting.

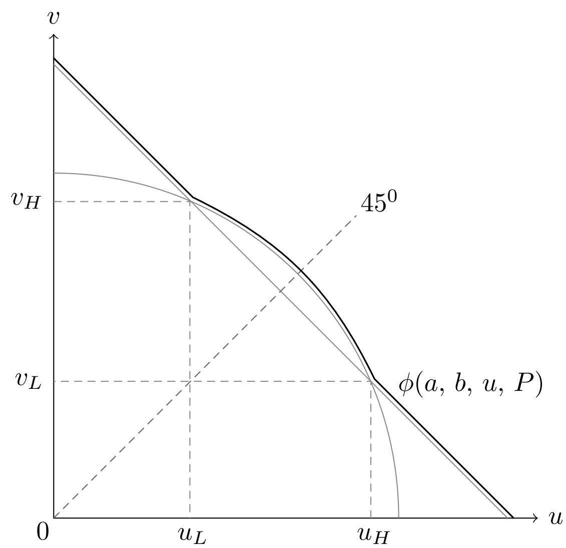


Figure 1: For $R^- \leq R(a,\, b) \leq R^+$, the bargaining frontier $\phi(a,\, b,\, u.\, P)$ of a given enterprise $(a,\, b)$ is the upper envelope of $\phi^N(a,\, b,\, u,\, P)$, the concave frontier and $\phi^I(a,\, b,\, u,\, P)$, the linear frontier. Non-integration is preferred over integration for intermediate values of u, i.e., $u_L \leq u \leq u_H$. By contrast, integration is the preferred choice for low or high values of u, i.e., $u < u_L$ or $u > u_H$.

associated with I. Because both frontiers are symmetric with respect to the 45^0 line, they intersect exactly twice at $u_L(z(a,\, b),\, P)$ and $u_H(z(a,\, b),\, P)$ (see the Appendix where these two thresholds are explicitly derived). The (combined) bargaining frontier of an enterprise $(a,\, b)$ is given by:[14]

$$\phi(a,\, b,\, u,\, P) = \max\left\{\phi^I(a,\, b,\, u,\, P),\, \phi^N(a,\, b,\, u,\, P)\right\}. \tag{BF}$$

It is easy to show that $\phi(a,\, b,\, u,\, P)$ is strictly increasing in a and b, and strictly decreasing in u. When $R^- \leq R(a,\, b) \leq R^+$, for intermediate values of u, i.e., $u_L(z(a,\, b),\, P) \leq u \leq u_H(z(a,\, b),\, P)$, an enterprise prefers to stay separate because the corresponding revenue shares s and $1-s$ are more balanced and so coordination among the two units can be achieved without being integrated. On the other hand, for the extreme values of u, either high or low, integration is preferred because the shares are tilted in favor of one of the two units and the incentives for revenue maximization are weak.[15]

[14]When $R \geq 2c$, the bargaining frontier under non-integration is different from the one in (1) because in this case the probability of success is given by $\pi(e_A,\, e_B) = 1$. When $R > 2c$, the frontier under integration is everywhere above that under non-integration, and hence, integration is the preferred choice. When $R = 2c$, the frontier under non-integration lies below that under integration, except the linear frontier is tangent to the non-linear one on the 45^0-line. In this case, the total expected output is the same under both ownership structure, but under unequal surplus sharing the incentive costs are shifted from one side to the other implying a loss of efficiency relative to equal surplus sharing. The detailed analysis of this case is available upon request to the authors.

[15]Notice that, for $R(a,\, b) < R^-$, $u_L(z(a,\, b),\, P) < 0$ and $u_H(z(a,\, b),\, P) > R(a,\, b)^2/4c$. On the other hand, for $R(a,\, b) > R^+$, $\phi^N(a,\, b,\, u,\, P)$ and $\phi^I(a,\, b,\, u,\, P)$ do not intersect each other at any u, and hence, $u_L(z(a,\, b),\, P)$ and $u_H(z(a,\, b),\, P)$ are complex numbers for all $(a,\, b)$.

5 The market equilibrium

We analyze the equilibrium of the input market with a continuum of units and types. We first show that the equilibrium matching is positive assortative (PAM), i.e., high- (low-) productivity A firms match with high- (low-) productivity B firms. Next we analyze the equilibrium ownership structures, i.e., whether high-revenue enterprises choose to stay separate or integrate.

5.1 Equilibrium matching

The indifference curve. For the analysis of this section, our object of interest is the *indifference curve* of each type a A firm when $R^- \leq R(a,\, b) \leq R^+$ for any b, which is derived as follows. The revenue of a given enterprise, $Pz(a,\, b)$ can be increased in three equivalent ways — an exogenous increase in the product market price, P, an increase in the productivity of the A unit, a, or an increase in the productivity of the B unit, b. Each of such changes would lead to an expansion of the bargaining frontier of the given enterprise (depicted by the 'higher' bargaining frontiers in the left panel of Figure 2). Clearly, each expanded frontier will continue to have two kinks, at $u = u_L,\, u_H$ at which an enterprise would be indifferent between N and I. The curve RRR, which we call the *revenue expansion path*, is drawn by joining the indifference points by varying the levels of enterprise revenue. The starting point of the revenue-expansion path on each axis is $R^- - c/2$. So, in the region enclosed by RRR, the corresponding enterprise chooses non-integration, and the units choose to integrate if the utility allocation lies outside this region.[16]

We suppress the argument P from all functions until Section 5.2. Because $\phi(a,\, b,\, u)$ is strictly increasing in b, strictly decreasing in u, and differentiable almost everywhere with respect to u (except at the two indifference points u_L and u_H), the indifference curves $u = \psi(b;\, a)$ are well-defined, and must satisfy $\phi(a,\, b,\, \psi(b;\, a)) = \bar{v}$ for some constant utility level $\bar{v}$ of a. In the left panel of Figure 2, consider the horizontal line at $\bar{v}$ which intersects the three bargaining frontiers associated with enterprises $(a,\, b')$, $(a,\, \hat{b})$ and $(a,\, b'')$ with $b' < \hat{b} < b''$ at three distinct utility allocations, say $(u',\, \bar{v})$, $(\hat{u},\, \bar{v})$ and $(u'',\, \bar{v})$. Therefore, the three type-utility combinations $(b',\, u')$, $(\hat{b},\, \hat{u})$ and $(b'',\, u'')$ are three points on a's indifference curve $\phi(a,\, b,\, u) = \bar{v}$ as depicted in the right panel of Figure 2. If the intersection of $\bar{v}$ with a given bargaining frontier lies strictly inside the region enclosed by RRR, e.g. that with the frontier labeled b', then the enterprise prefers to stay separate. By contrast, if such intersection lies outside this region, then the enterprise prefers to integrate, e.g. enterprise $(a,\, b'')$. Therefore,

$$\psi(b;\, a) = \max\left\{\psi^I(b;\, a),\, \psi^N(b;\, a)\right\},$$

where $\psi^d(b;\, a)$ denotes the indifference curve of a under ownership structure $d = N,\, I$. Because at a given b, higher u implies lower utility for the A firm, "higher" indifference curves lie to the southeast.

If the constant utility level of a is low (below point R on the vertical axis, i.e., $\bar{v} \leq R^- - c/2$), then the horizontal line at $\bar{v}$ intersects the revenue expansion path at only one point, the corresponding enterprise being $(a,\, \hat{b})$ that is indifferent between N and I. Thus, the indifference curve $\phi(a,\, b,\, u) = \bar{v}$

[16]The revenue expansion path for any given enterprise is qualitatively the same as the one derived by Legros and Newman (2013, Figure I), who assume a quadratic success probability. The only difference is that, under a linear probability of success function, the two paths meet on the 45^0-line when the enterprise revenue is equal to R^+ so that $u_L(z(a,\, b),\, P) = u_H(z(a,\, b),\, P)$.

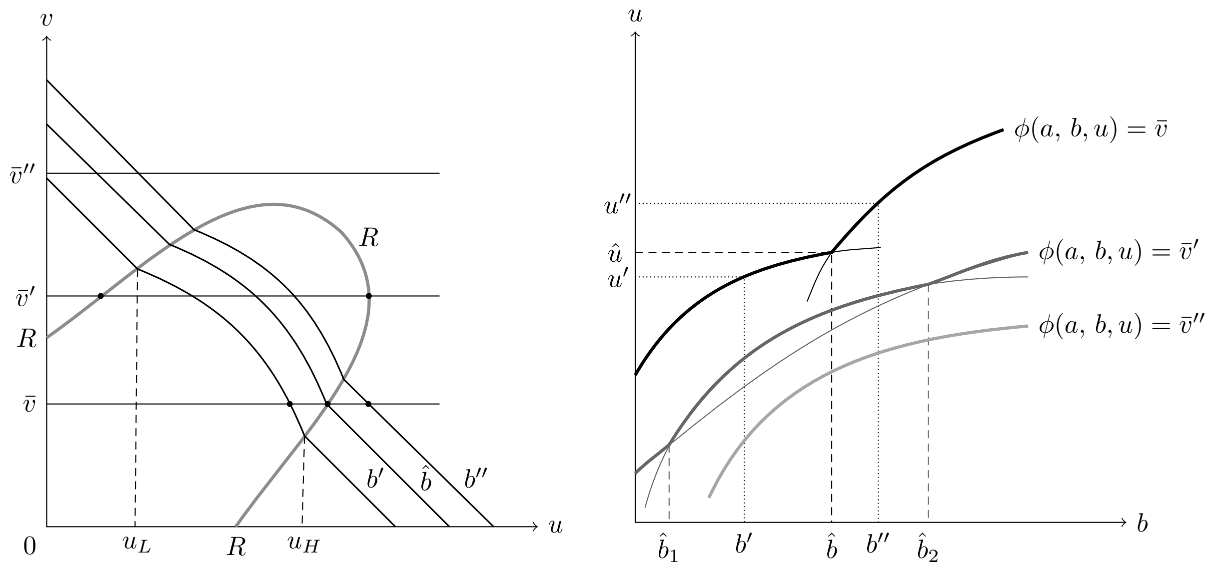


Figure 2: The left panel depicts the bargaining frontiers of a given type a of A firm by varying her partner's type b when $R(a,\, b) \in [R^-, R^+]$. The right panel depicts the corresponding indifference map in the type-utility space.

has a unique kink at $(\hat{b},\, \hat{u})$ as shown in the left panel of Figure 2. In other words, $\hat{b}$ is the unique solution to $\psi^I(b;\, a) = \psi^N(b;\, a)$. Evidently, $\psi(b;\, a) = \psi^N(b;\, a)$ if and only if $b < \hat{b}$. By contrast, if the constant utility level of a is high enough, i.e., $\bar{v}' > R^- - c/2$, then the horizontal line at $\bar{v}'$ intersects the revenue expansion path twice, and hence, the indifference curve $\phi(a,\, b,\, u) = \bar{v}'$ has two kinks at $b = \hat{b}_1,\, \hat{b}_2$. In this case, $\psi(b;\, a) = \psi^N(b;\, a)$ if and only if $\hat{b}_1 \le b \le \hat{b}_2$ as shown in the right panel of Figure 2.[17] Finally, if the constant utility level of a is even higher, e.g. at a level $\bar{v}''$, then it does not intersect the revenue expansion path, and hence, the corresponding indifference curve is smooth as a chooses to integrate with all b, i.e., $\psi(b;\, a) = \psi^I(b;\, a)$.

GID and PAM. We now analyze the equilibrium matching function $a = \alpha(b)$ and show that the matching is positive assortative, i.e., $\alpha(b)$ is an increasing function. PAM in an equilibrium allocation relies on the properties of the bargaining frontier $\phi(a,\, b,\, u)$. Recall that in the enterprise formation stage at date 1, each type a A firm chooses a b to maximize

$$\phi(a,\, b,\, u(b)) = \max\left\{\phi^N(a,\, b,\, u(b)),\, \phi^I(a,\, b,\, u(b))\right\}.$$

Legros and Newman (2007) show that the bargaining frontier satisfying the *generalized increasing difference* (GID) condition, i.e., for any $a'' > a'$, $b'' > b'$ and $u'' > u'$,

$$\phi(a',\, b'',\, u'') = \phi(a',\, b',\, u') \quad \Longrightarrow \quad \phi(a'',\, b'',\, u'') \ge \phi(a'',\, b',\, u') \tag{GID}$$

[17] We have drawn the indifference curves concave under both ownership structures, which is innocuous for any of our subsequent results. The function $\psi^I(b;\, a)$ is concave if $z(a,\, b)$ is concave in b. On the other hand, the indifference curve under non-integration, $\psi^N(b;\, a)$ is concave if

$$\frac{3RR_b}{2R^2 - 4cu - R\sqrt{R^2 - 4cu}} \le -\frac{R_{bb}}{R_b},$$

i.e., $R(a,\, b)$, or equivalently, $z(a,\, b)$ is sufficiently concave in b.

is a necessary and sufficient condition for PAM. In order to show that our bargaining frontier, $\phi(a,\, b,\, u)$ satisfies GID, we first characterize GID in terms of the property of *increasing differences* of the indifference curves of the A units.

Lemma 1 *The bargaining frontier $\phi(a,\, b,\, u)$ satisfies GID if and only if the indifference curve $\psi(b;\, a)$ has increasing differences in $(b;\, a)$, i.e., for any $a'' > a'$ and $b'' > b'$,*

$$\psi(b'';\, a'') - \psi(b';\, a'') \geq \psi(b'';\, a') - \psi(b';\, a'). \tag{ID}$$

Intuitively, GID asserts that if for a lower type a', two type-payoff combinations $(b',\, u')$ and $(b'',\, u'')$ with $(b'',\, u'') > (b',\, u')$ are on the same indifference curve, then $(b'',\, u'')$ must lie on a higher indifference curve of any higher type a''. On the other hand, increasing differences of $\psi(b;\, a)$ means that the indifference curve of a higher type A unit is everywhere steeper than that of her lower-type counterpart, and hence, they cross each other only once. The intuition is simple. In panel A of Figure 3, $\psi(b;\, a'')$ is everywhere steeper than $\psi(b;\, a')$, i.e., (ID) holds for these two arbitrarily chosen types a' and a'' with $a'' > a'$. The points $(b',\, u')$ and $(b'',\, u'')$ with $(b'',\, u'') > (b',\, u')$ lie on the same indifference curve of a', but $(b'',\, u'')$ lies on the indifference curve of a'', which is higher than the one passing through $(b',\, u')$ implying that (GID) holds. On the other hand, if (ID) is violated, i.e., the indifference curve of a'' is flatter than that of the lower type a' at some $(b',\, u')$, then one can find points like $(b'',\, u'') > (b',\, u')$ such that type a' is indifferent between $(b',\, u')$ and $(b'',\, u'')$, but $(b'',\, u'')$ yields lower utility to a'' than $(b',\, u')$ does [because $(b'',\, u'')$ lies on a lower indifference curve], which contradicts GID.[18]

Having established Lemma 1, we now require to prove that $\psi(b;\, a)$ has increasing differences in $(b;\, a)$ in order show that the equilibrium matching is PAM. We proceed in the following steps:

1. Using the expressions of the bargaining frontiers (1) and (2), we show that, for each $d = N,\, I$, $\psi^d(b;\, a)$ has increasing differences in $(b;\, a)$. This property is easily established because $z(a,\, b)$ is supermodular in $(a,\, b)$, and each of $\psi^N(b;\, a)$ and $\psi^I(b;\, a)$ is a differentiable function. Because $\psi(b;\, a) \equiv \psi^N(b;\, a)$ for all $R(a,\, b) < R^-$, and $\psi(b;\, a) \equiv \psi^I(b;\, a)$ for all $R(a,\, b) > R^+$, this also establishes PAM for all $R(a,\, b) < R^-$ and $R(a,\, b) > R^+$.

2. However, when $R(a,\, b) \in [R^-,\, R^+]$, the combined indifference curve $\psi(b;\, a)$ does not necessarily have increasing differences even if each of $\psi^N(b;\, a)$ and $\psi^I(b;\, a)$ has increasing differences in $(b;\, a)$. Panel B of Figure 3 depicts a possible violation of increasing differences property of $\psi(b;\, a)$. Clearly, a sufficient condition for $\psi(b;\, a)$ to satisfy (ID) is that the kink in the indifference curve of the A unit moves to the left as a increases, i.e., $d\hat{b}/da \leq 0$. Recall that $\hat{b}$ is not always unique as an indifference curve has two kinks whenever $\bar{v}$ is high enough. In this case, the condition $d\hat{b}/da \leq 0$ requires that both kinks move to the left as a increases. Note that

$$\frac{d\hat{b}}{da} = -\frac{\psi^I_a(\hat{b};\, a) - \psi^N_a(\hat{b};\, a)}{\psi^I_b(\hat{b};\, a) - \psi^N_b(\hat{b};\, a)}.$$

[18] Panel B of Figure 3 depicts one of the many possible ways to violate (ID). The only thing we require is that $\psi(b;\, a'')$ is flatter relative to $\psi(b;\, a')$ at some $(b',\, u')$, i.e., (ID) does not hold for the types a' and a'' at this point. Then, we can always find a point on $\psi(b;\, a')$, which lies to the northeast of $(b',\, u')$ at which GID is violated because for a range of values of b to the right of b', $\psi(b;\, a')$ lies above $\psi(b;\, a'')$. Also, in Figure 3 the indifference curves are drawn with a single kink, but the result of Lemma 1 holds even if an indifference curve has two kinks.

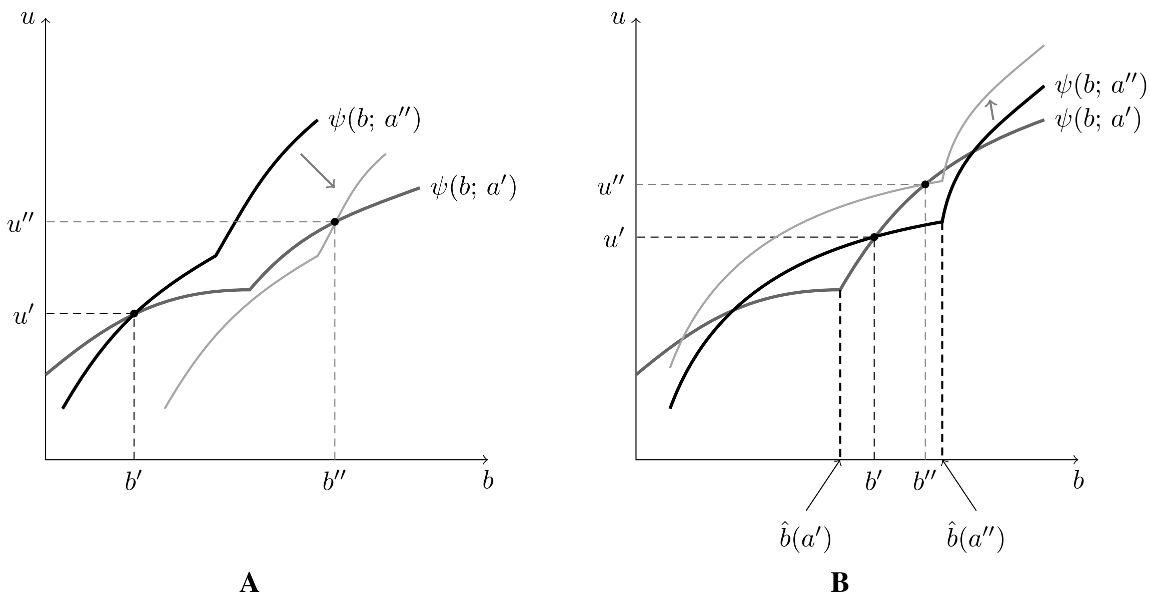


Figure 3: Panel A depicts that the indifference curves have increasing differences in $(b;\, a)$, i.e., $\psi(b;\, a'')$ is everywhere steeper than $\psi(b;\, a')$, and hence, they cross only once. As a result, the bargaining frontier satisfies GID. Panel B depicts a possible violation of (ID), which leads to that GID must also be violated.

> It turns out that $\text{sign}[\psi_a^I(b;\, a) - \psi_a^N(b;\, a)] = \text{sign}[\psi_b^I(b;\, a) - \psi_b^N(b;\, a)]$ for all $(a,\, b)$, and hence, $d\hat{b}/da \le 0$. The intuition is simple. The term $\psi_a^I(\hat{b};\, a) - \psi_a^N(\hat{b};\, a)$ is the contribution of a to a marginal change in the match surplus [in a neighborhood of $\hat{b}$] when the enterprise $(a,\, b)$ switches from N to I, whereas the term $\psi_b^I(\hat{b};\, a) - \psi_b^N(\hat{b};\, a)$ is the contribution of b to such a marginal change. Because the marginal contributions on behalf of a and b point in the same direction, both a and b gain at the margin by switching to the same ownership structure. When these marginals point in opposite directions, PAM may fail to hold around the kink (as in Panel B of Figure 3).

Two popularly used functions are $\psi(b;\, a) = \max\{ba,\, 2b^2a^2\}$ (in Cole, Mailath, and Postlewaite, 2001), and $\psi(b;\, a) = \max\{ba^2,\, b^2a\}$ (in Kremer and Maskin, 1996). In the first case, (GID) holds because $\hat{b}(a) = 1/2a$, and hence, $d\hat{b}/da < 0$; whereas in the second case, (GID) may fail to hold for all $(a,\, b)$ because $\hat{b}(a) = a$ implying that the kink moves to the right as a increases. Intuitively, a and b are treated symmetrically under two distinct technologies in Cole et al. (2001), whereas in the other case, they are treated asymmetrically. It follows from the above discussion that

Proposition 1 *At any given product market price the bargaining frontier $\phi(a,\, b,\, u)$ satisfies GID, and hence, the equilibrium matching is positive assortative.*

We denote by $(\alpha(b),\, b)$ a typical enterprise formed in the input market equilibrium. The above proposition together with the measure consistency condition (MC) imply that if $a = \alpha(b)$, then $G(a) = F(b)$, and hence, $\alpha'(b) = f(b)/g(\alpha(b)) > 0$ for all $b \in B$.

5.2 Equilibrium ownership structures

Having shown that there is PAM with respect to firm productivity, we now analyze the choice of ownership structure of each equilibrium enterprise $(\alpha(b),\, b)$. In order to determine the equilibrium organizational structure of the input market, we will require two equilibrium objects – namely, the *utility function* [of the B units] and the *indifference locus*. Because the equilibrium matching is PAM irrespective of the choice of ownership structures, each equilibrium enterprise $(\alpha(b),\, b)$ can be uniquely identified by b, the productivity of the B unit.

The utility function. We first derive the equilibrium utility function of the B firms, which is denoted by $u(b,\, P)$. The first-order condition of the optimization problem $(\mathcal{M}_a)$ is given by the following *ordinary differential equation* (ODE):

$$u_b(b,\, P) = \begin{cases} -\frac{\phi_2^N(\alpha(b), b, u(b), P)}{\phi_3^N(\alpha(b), b, u(b), P)} & \text{if } u(b) \in [u_L(z(\alpha(b),\, b),\, P),\, u_H(z(\alpha(b),\, b),\, P)]; \\ \\ -\frac{\phi_2^I(\alpha(b), b, u(b), P)}{\phi_3^I(\alpha(b), b, u(b), P)} & \text{otherwise,} \end{cases} \tag{3}$$

where u_b denote the derivative of u with respect to b. Given that $-\phi_2^d/\phi_3^d > 0$ for each $d = N,\, I$, the equilibrium utility of each B firm, which is given by:

$$u(b,\, P) = u(\underline{b},\, P) + \int_{\underline{b}}^{b} u_b(x,\, P)dx = u_0 + \int_{\underline{b}}^{b} u_b(x,\, P)dx, \tag{4}$$

is strictly increasing in b. Because the reservation utility u_0 of all the B firms is the outside option of a B firm with type $\underline{b}$, in equilibrium we have $u(\underline{b},\, P) = u_0$ at any P. According to the Picard-Lindelöf Theorem (see Birkhoff and Rota, 1989), a unique solution to the ODE (3) exists (at least in the neighborhood of the initial condition) provided that $u_b(b,\, P)$ is bounded, Lipschitz continuous in u and continuous in b.[19]

Under integration (4) takes the form

$$u^I(b,\, P) \equiv u_0 + \int_{\underline{b}}^{b} P z_b(\alpha(x),\, x)dx, \tag{5}$$

while under non-integration it becomes

$$u^N(b,\, P) \equiv u_0 + \int_{\underline{b}}^{b} P z_b(\alpha(x),\, x) H(x,\, u,\, P)dx, \tag{6}$$

where[20]

$$H(b,\, u,\, P) \equiv \frac{1}{2c}\left(Pz(\alpha(b),\, b) + \frac{(Pz(\alpha(b),\, b))^2 - 4cu}{Pz(\alpha(b),\, b) - \sqrt{(Pz(\alpha(b),\, b))^2 - 4cu}}\right).$$

[19]Whenever $u(b,\, P) < u_L(z(\alpha(b),\, b),\, P)$ and $u(b,\, P) > u_H(z(\alpha(b),\, b),\, P)$, the ODE assumes a simple form; all the aforementioned properties are satisfied and an analytical solution can be easily obtained. However, for values of u in the interval $[u_L(z(\alpha(b),\, b),\, P),\, u_H(z(\alpha(b),\, b),\, P)]$, the ODE is much more complicated and an analytical solution to it does not exist. In this region, we require to establish the existence and uniqueness of a solution. Our assumptions ensure that it is continuous in b, because b enters u_b through $z(\alpha(b),\, b)$ which is a continuous function. If a differentiable function has a derivative that is bounded everywhere by a real number, then the function is Lipschitz continuous. If $u_0 = 0$, then u_b becomes unbounded. Hence, we require u_0 to be strictly positive.

[20]More details about the derivations are provided in the Appendix in the Proof of Proposition 1.

Under I, (5) reveals that the equilibrium utility of firm b is the outside option u_0 plus the area under the marginal productivities z_b, on the equilibrium path, of all the inframarginal B firms, times the market price. Under N, however, the area under the marginal productivity function of the inframarginal B firms is adjusted by the $H(\cdot)$ function, see (6). This function captures the issue of imperfect transferability of surplus under non-integration. Under I there is a one-to-one transferability of utility from a to b and so H is always 1. But under N, H is greater than 1 for low u, where the Pareto frontier is relatively flat and hence a can easily transfer surplus to b and it becomes less than 1 for high u's where the frontier is steep and thus it is costly for a to transfer utility to b.

An important relationship, and a one that we will refer to later, is the association between the market price and the utility. A higher price, as we will show later, increases $u(b)$ and $v(a)$, but the relative utilities, v/u, matter for the incentive of enterprise $(\alpha(b), b)$ to integrate. The utility of b under I, (5), increases linearly with P, but under N, (6), the relationship is non-linear, as higher P also affects the easiness with which the extra surplus can be transferred from a to b.

The indifference locus. Given any product market price P, for each equilibrium enterprise $(\alpha(b), b)$ there are exactly two utility points $u_L(z(\alpha(b), b), P)$ and $u_H(z(\alpha(b), b), P)$ of b at which the enterprise is indifferent between staying separate and integrate. Therefore, when the types vary continuously on $[\underline{b}, \bar{b}]$, both $u_L(z(\alpha(b), b), P)$ and $u_H(z(\alpha(b), b), P)$ are continuous functions of b at any given price P as drawn in Figure 4, which we will call the ***equilibrium indifference locus***. To simplify notation, in Figure 4, the curve $u_L(z(\alpha(b), b), P)$ is labeled as u_L, and $u_H(z(\alpha(b), b), P)$ is labeled as u_H.

The characteristics of the equilibrium indifference locus depend on the output of each enterprise, $z(\alpha(b), b)$ and the product market price, P. Note that $z(\alpha(b), b)$ is an equilibrium object because it involves the matching function $\alpha(b)$. However, it can be easily determined once the type distributions, $G(a)$ and $F(b)$ are known. Moreover, $z(\alpha(b), b)$ is strictly increasing in b because z_a, z_b, $\alpha'(b) > 0$ for all b. Let $\underline{z} \equiv z(\underline{a}, \underline{b}) = z(\alpha(\underline{b}), \underline{b})$ and $\bar{z} \equiv z(\bar{a}, \bar{b}) = z(\alpha(\bar{b}), \bar{b})$. Further, define $b^-(P)$ and $b^+(P)$ which respectively solve

$$z(\alpha(b), b) = \frac{R^-}{P}, \quad \text{and} \quad z(\alpha(b), b) = \frac{R^+}{P}.$$

Clearly, $b^-(P) < b^+(P)$. At $b = b^-(P)$ we have $u_L(z(\alpha(b), b), P) = 0$, and $u_L(z(\alpha(b), b), P) = u_H(z(\alpha(b), b), P)$ at $b = b^+(P)$.

In principle the product market price P can vary on the entire interval $[0, \infty)$. Nevertheless, it would be convenient to divide the whole range of variation of price into a number of disjoint subintervals. The equilibrium choice of ownership structures by each equilibrium enterprise would depend on which subinterval P belongs to. Consider the following cases:

1. If $0 \leq P < P^- \equiv R^-/\bar{z}$, i.e., $P\bar{z} < R^-$, then all enterprises strictly prefers non-integration over integration as none of them is able to generate sufficient revenue. In this case, we have $u_L(z(\alpha(b), b), P) < 0$ for any b.

2. Define $\underline{P} = \min\{R^-/\underline{z}, R^+/\bar{z}\}$. If $P^- \leq P < \underline{P}$, then R^-/P intersects $z(\alpha(b), b)$ at a unique point $b^-(P) \in (\underline{b}, \bar{b})$, but R^+/P lies strictly above $z(\alpha(b), b)$, and hence, $b^+(P)$ does not exist. In this case, $u_L(z(\alpha(b), b), P)$ and $u_H(z(\alpha(b), b), P)$ are as drawn in Panel A of Figure 4. Note that u_L and u_H do not meet at any $b \leq \bar{b}$.

3. Define $\bar{P} = \max\{R^-/\underline{z},\ R^+/\bar{z}\}$. If $\underline{P} \le P \le \bar{P}$, then there are two sub-cases. First, if $R^-/\underline{z} < R^+/\bar{z}$, then R^-/P lies strictly below $z(\alpha(b),\ b)$ and R^+/P lies strictly above $z(\alpha(b),\ b)$, and hence, neither $b^-(P)$ nor $b^+(P)$ exists. In this case, $u_L(z(\alpha(b),\ b),\ P)$ and $u_H(z(\alpha(b),\ b),\ P)$ are as drawn in Panel B of Figure 4. Note that $u_L(\underline{z},\ P) > 0$. Second, if $R^-/\underline{z} < R^+/\bar{z}$, both R^-/P and R^+/P intersect $z(\alpha(b),\ b)$, and hence, we have $\underline{b} < b^- < b^+ < \bar{b}$. In this case, $u_L(\alpha(b),\ b,\ P)$ and $u_H(\alpha(b),\ b,\ P)$ are as drawn in Panel C of Figure 4. Note that u_L and u_H meet each other at $b = b^+$.

4. If $\bar{P} < P \le P^+ \equiv R^+/\underline{z}$, then $z(\alpha(b),\ b)$ intersects R^+/P at a unique point $b^+(P) \in (\underline{b},\ \bar{b})$, but R^-/P lies strictly below $z(\alpha(b),\ b)$, and hence, $b^-(P)$ does not exist. In this scase, $u_L(z(\alpha(b),\ b),\ P)$ and $u_H(z(\alpha(b),\ b),\ P)$ are as drawn in Panel D of Figure 4. For these prices, $u_L(\underline{z},\ P) > 0$, and u_L and u_H meet each other at $b = b^+$.

5. Finally, if $P > P^+$, i.e., $P\underline{z} > R^+$, then all enterprises strictly prefers integration over non-integration as even the lowest-productivity enterprise generates sufficiently high revenue (more than R^+). In this case both $u_L(z(\alpha(b),\ b),\ P)$ and $u_H(z(\alpha(b),\ b),\ P)$ are complex numbers.

Equilibrium choice of organization. Depending on which subinterval the product market price P lies in, there are many possible equilibria with respect to the choice of ownership structures. We will focus on equilibria in which the lowest productivity enterprise $(\underline{a},\ \underline{b})$ chooses to stay separate in all equilibria at any given price P. This rules out the possibility of having a product market price in the subinterval $(P^+,\ \infty)$ where all enterprises choose to integrate. Because $\underline{a}$ must consume non-negative utility, we must have

$$\phi^N(\underline{z},\ P) \ge 0 \quad \Longleftrightarrow \quad u_0 \le \frac{P^2\underline{z}^2}{4c} \quad \Longleftrightarrow \quad P \ge \frac{\sqrt{4cu_0}}{\underline{z}} \equiv P_0.$$

Note that $P_0 > 0$ if $u_0 > 0$.[21] Moreover, the type of equilibria of our interest is monotone equilibria, i.e., there is a unique $b^*(P) \in (\underline{b},\ \bar{b})$, whenever it exists, such that any equilibrium enterprise with $b < b^*(P)$ chooses to stay separate, while any equilibrium enterprise with $b > b^*(P)$ integrates. Note that $b^*(P)$ is defined by the intersection between the equilibrium utility function $u(b,\ P)$ and the indifference locus, i.e., enterprise $(\alpha(b^*(P)),\ b^*(P))$ is indifferent between N and I at equilibrium under market price P. Let

$$r(b) \equiv \frac{z_a(\alpha(b),\ b)/g(\alpha(b))}{z_b(\alpha(b),\ b)/f(b)},$$

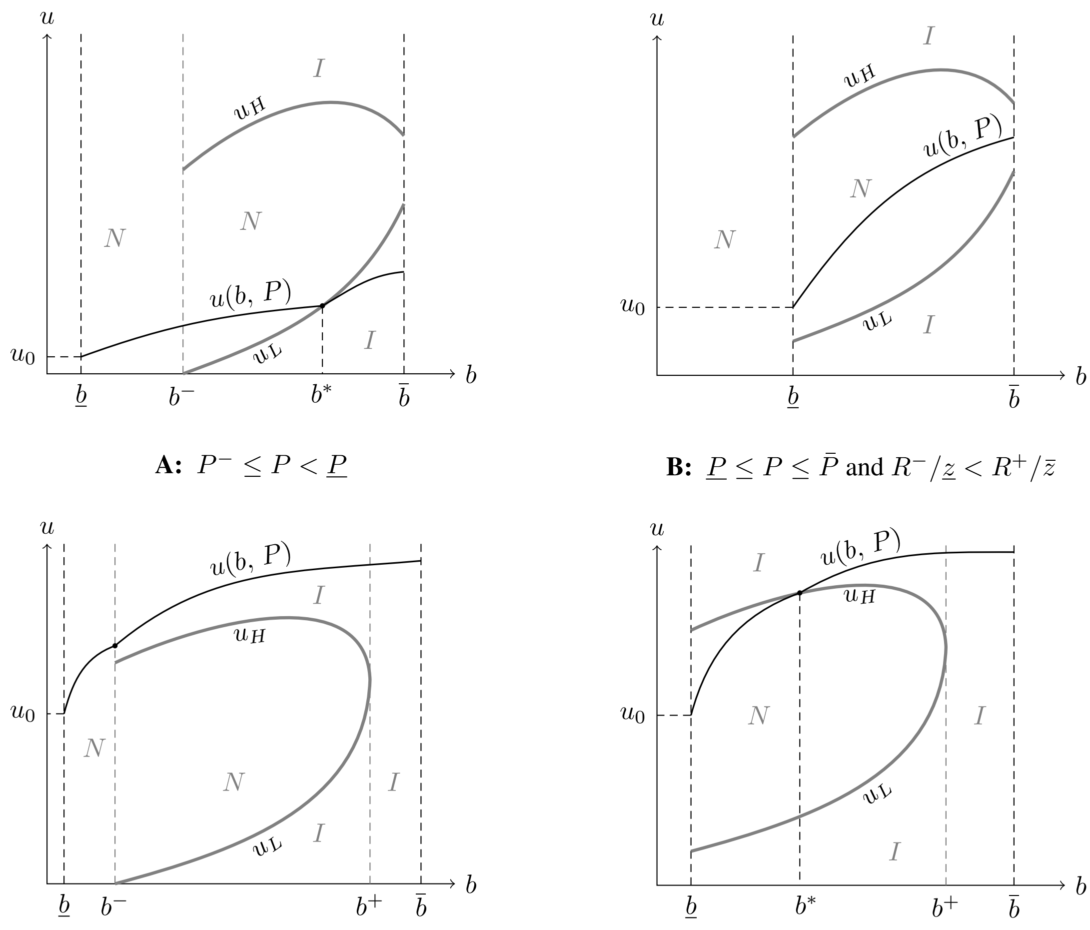
which is the ratio of the average marginal productivities of the A and B firms in a typical equilibrium enterprise $(\alpha(b),\ b)$. The following lemma provides sufficient conditions under which $b^*(P)$ is unique.

Lemma 2 *If $\sqrt{2} - 1 < r(b) < \sqrt{2} + 1$ for all $b \in [\underline{b},\ \bar{b}]$, then the threshold productivity $b^*(P) \in (\underline{b},\ \bar{b})$ of the B firms is unique at a given product market price P.*

In Figure 4, the regions labeled N are the regions in which an equilibrium enterprise $(\alpha(b),\ b)$ chooses to stay separate at a price P because $u_L(z(\alpha(b),\ b),\ P) \le u(b,\ P) \le u_H(z(\alpha(b),\ b),\ P)$, whereas at a region labeled I, an equilibrium enterprise chooses to integrate. Once $u(b,\ P)$ intersects the indifference

[21] In order to have

$$P_0 \le P^- \quad \Longleftrightarrow \quad u_0 \le \frac{(2-\sqrt{2})^2 c}{4}\left(\frac{\underline{z}}{\bar{z}}\right)^2.$$



A: $P^- \leq P < \underline{P}$

B: $\underline{P} \leq P \leq \bar{P}$ and $R^-/\underline{z} < R^+/\bar{z}$

C: $\underline{P} \leq P \leq \bar{P}$ and $R^-/\underline{z} > R^+/\bar{z}$

D: $\bar{P} < P \leq P^+$

Figure 4: Equilibrium choice of ownership structures.

locus (either at u_L, or at u_H), it enters the region I. For $b^*(P)$ to be unique, we require that at any $b > b^*(P)$, $u(b, P)$ does not cross the indifference locus again. The first inequality in the above lemma guarantees that $u(b, P)$ is flatter everywhere than u_L at any $b \geq b^*(P)$ so that if there is an intersection on u_L, then it is unique (as drawn in Panel A of Figure 4). The second inequality, on the other hand, implies that $u(b, P)$ is steeper than u_H at any $b > b^*(P)$, and hence, the intersection of $u(b, P)$ with u_H is unique (Panel D of Figure 4). We may have a situation as depicted in Panel C of Figure 4 where $u(b, P)$ does not intersect the indifference locus as $u(b^-, P) > u_H(z(\alpha(b^-), b^-), P)$. In this case, all enterprises $(\alpha(b), b)$ with $b < b^-$ choose to stay separate, and any equilibrium enterprise with $b \geq b^-$ integrates, i.e., there is a discrete jump between N and I in the sense that there is no enterprise indifferent between N and I. In this case, we would assign $b^*(P) = b^-(P)$, and the ownership structure is still monotonic. The only exception is that when $u(b, P)$ lies strictly between u_L and u_H when u_L and u_H do not meet at any $b \leq \bar{b}$ (as drawn in Panels A and B of Figure 4) in which case all equilibrium enterprises choose to stay separate, and hence, $b^*(P)$ does not exist. It is difficult a priori to rule out such equilibrium unless $u(b, P)$ is either very steep or very flat.

Intuitively, $r(b)$ can be very large if in an equilibrium enterprise $(\alpha(b),\, b)$, either the marginal productivity of the A units is very high relative to that of the B units, or the B firms of this particular type are abundant relative to the A firms of the matched type, i.e., $g(\alpha(b))/f(b)$ low enough (a relatively flat matching function). When $r(b)$ is low (high), it is more likely that the equilibrium utility function intersects the u_H (u_L) portion of the indifference locus. Note that even though the above sufficient condition is written in terms of the equilibrium variable $\alpha(b)$, it is easy to pin down the types a and b along the equilibrium path because the equilibrium exhibits PAM and the type distributions $G(a)$ and $F(b)$ are known.

Proposition 2 (Equilibrium ownership structures) *Let $\sqrt{2}-1 < r(b) < \sqrt{2}+1$ so that $b^*(P)$ is unique whenever it exists, and $u_L(\underline{z},\, P) \leq u_0 \leq u_H(\underline{z},\, P)$ at any $P \in [P_0,\, P^+]$. Then there are price thresholds P^-, $\underline{P}$ and $\bar{P}$ with $P^- < \underline{P} < \bar{P}$ such that in equilibrium,*

(a) *all enterprises choose to stay separate at any price $P \in [P_0,\, P^-)$ irrespective of how the enterprise revenue is shared;*

(b) *If $P \in [P^-,\, \underline{P})$, then there are two possible equilibria: (i) if $u_L(\bar{z},\, P) \leq u(\bar{b},\, P) \leq u_H(\bar{z},\, P)$, then all enterprises choose to stay separate; (ii) otherwise, the equilibrium ownership choice is monotone with enterprises with productivity b chose to integrate if and only if $b > b^*(P)$;*

(c) *If $P \in [\underline{P},\, \bar{P}]$, then the equilibrium choice of ownership structure depends on the values of $R^-/\underline{z}$ and $R^+/\bar{z}$: (i) if $R^-/\underline{z} < R^+/\bar{z}$ and $u_L(\bar{z},\, P) \leq u(\bar{b},\, P) \leq u_H(\bar{z},\, P)$, then all enterprises chooses to stay separate; (ii) otherwise, the equilibrium ownership choice is monotone with enterprises with productivity b chose to integrate if and only if $b > b^*(P)$;*

(d) *Finally, the equilibrium ownership choice is monotone, i.e., enterprises with productivity b choose to integrate if and only if $b > b^*(P)$ at any price $P \in (\bar{P},\, P^+]$ irrespective of how the enterprise revenue is shared.*

A low product market price, i.e., $P_0 \leq P < P^-$ is not able to generate sufficient revenue (more than R^-) at any enterprise, and hence, all of them choose to stay separate. On the other hand, a very high price, i.e., $P > P^+$ yields very high revenue (higher than R^+) for every enterprise so that all of them choose to integrate. We would focus on monotonic ownership structures of the type $\langle N,\, I\rangle$, i.e., there is a unique $b^*(P)$ at a given price P such that any enterprise chooses to integrate if and only if the productivity $(\alpha(b),\, b)$ is such that $b > b^*(P)$, because they are more plausible given the existing empirical evidence (we offer a discussion about the empirical evidence in Section 8).[22]

[22] Other equilibrium patterns of ownership structures are also possible. It depends on how the enterprise revenues are shared between the A and B units, i.e., whether $u(b,\, P)$ lies in or outside the interval $[u_L(z(\alpha(b),\, b),\, P),\, u_H(z(\alpha(b),\, b),\, P)]$. It is possible to have monotone equilibria in which low-productivity enterprises integrate and high-productivity enterprises stay separate. For example, in Panel B of Figure 4, if $u_0 < u_L(\underline{z},\, P)$, then either all enterprise integrate (if $u(b,\, P)$ always lies below u_L), or low-productivity enterprises integrate and high-productivity ones stay separate (if $u(b,\, P)$ intersects u_L only once). Finally, it is possible to have non-monotone equilibria if the equilibrium utility function crosses the indifference locus more than once. However, such non-monotonicity does not emerge if $\sqrt{2}-1 < r(b) < \sqrt{2}+1$.

6 Effect of price changes on the equilibrium

In this section, we intend to analyze the effects of an exogenous increase in the product market price on (a) the decision to integrate, and (b) the aggregate industry output.

6.1 Incidence of integration

We examine how a change in the product market price P affects the fraction of integrated firms in the market, i.e., how $b^*(P)$ changes as P goes up. It is clear from the analysis presented in Section 5.2 that there is no $b^*(P)$ for $P < P^-$ (at these prices all enterprises choose to stay separate) and $P > P^+$ (when all enterprises integrate), and hence, we only consider the variation of price in the interval $[P^-, P^+]$.[23] Further, we assume that $b^*(P)$ is unique for each P [cf. Lemma 2]. Let $U^L(b, P) \equiv u_L(z(\alpha(b), b), P)$ and $U^H(b, P) \equiv u_H(z(\alpha(b), b), P)$. Recall that $b^*(P)$ solves

$$\text{either} \quad u(b^*(P), P) = U^L(b^*(P), P), \tag{E$_L$}$$

$$\text{or} \quad u(b^*(P), P) = U^H(b^*(P), P). \tag{E$_H$}$$

As it can be seen from Figure 4, there are two different cases. The first is when the intersection is between $u(b, P)$ and the u_L portion of the indifference locus, and the second case is when the intersection is between $u(b, P)$ and the u_H portion of the indifference locus. When market price increases, the indifference locus shifts. The u_L portion of the indifference locus moves up as price increases. However, u_H shifts up for low values of b, but shifts down for high values of b. On the other hand, the equilibrium utility is higher for all b at the higher price except at $b = \underline{b}$ as $u(\underline{b}, P) = u(\underline{b}, P') = u_0$ for any $P \neq P'$.

The fraction of integrated enterprises increases, following an increase in the market price, if and only if the indifferent enterprise $(\alpha(b^*(P)), b^*(P))$ at the initial price P chooses to integrate at the increased price P'. As an increase in price shifts both the indifference locus and the utility function $u(b, P)$, the effect of a price change on the decision made by $(\alpha(b^*(P)), b^*(P))$ is in general ambiguous. Formally, differentiating respectively (E$_L$) and (E$_H$) with respect to price we obtain

$$\text{either} \quad \frac{db^*(P)}{dP} = \frac{u_P(b^*(P), P) - U_P^L(b^*(P), P)}{U_b^L(b^*(P), P) - u_b(b^*(P), P)},$$

$$\text{or} \quad \frac{db^*(P)}{dP} = \frac{U_P^H(b^*(P), P) - u_P(b^*(P), P)}{u_b(b^*(P), P) - U_b^H(b^*(P), P)}.$$

Note that the denominators of both the above expressions are positive because $u(b, P)$ is flatter than $U^L(b, P)$ at E_L as $u(b, P)$ intersects $U^L(b, P)$ from above, and $u(b, P)$ is steeper than $U^H(b, P)$ at E_H. Therefore, the sign of db^*/dP is completely determined by the sign of the numerator of each of the above expressions, which precisely describes the relative shifts of the equilibrium utility function and the indifference locus.

An increase in the product market price has two possible countervailing effects on the decision to integrate of any equilibrium enterprise $(\alpha(b), b)$, in particular that of $(\alpha(b^*(P)), b^*(P))$. First, a higher price implies a higher enterprise revenue, and hence, it is easier for the unit managers to ignore private costs. As a result, the higher the price, the smaller is the dominance of non-integration over integration.

[23] Even in this range of price variation, $b^*(P)$ may not exist as shown in Panels A and B of Figure 4.

For enterprise $(\alpha(b^*(P)),\, b^*(P))$, this effect of price change is nothing but a movement along its revenue expansion path, RRR drawn in Figure 5. The shape of RRR indicates that $u_H(z(\cdot),\, P) - u_L(z(\cdot),\, P)$ decreases with the price level.[24]

There is a second effect of price increase that has to do with how the extra surplus is allocated at enterprise $(\alpha(b^*(P)),\, b^*(P))$. If a higher price leads to a more balanced allocation of utilities, then non-integration is more likely since coordination is easier using incentive contracts. By contrast, if the allocation of utilities becomes more unbalanced, integration is more likely. The effect of P on the equilibrium utility of the B unit in enterprise $(\alpha(b^*(P)),\, b^*(P))$ is given by:

$$u_P^N(b^*(P),\, P) = \int_{\underline{b}}^{b^*(P)} z_b(\alpha(x),\, x)h(x,\, u,\, P)dx, \tag{7}$$

where

$$h(b,\, u,\, P) \equiv \frac{1}{2c}\left(\frac{5(Pz(\alpha(b),\, b))^2 - 4cu - Pz(\alpha(b),\, b)\sqrt{(Pz(\alpha(b),\, b))^2 - 4cu}}{Pz(\alpha(b),\, b) - \sqrt{(Pz(\alpha(b),\, b))^2 - 4cu}}\right) > 0$$

is the derivative of $PH(b,\, u,\, P)$ with respect to P.[25] It is sufficient to see how a change in price affects the utility of $b^*(P)$ under non-integration because at any price P we have $u^N(b^*(P),\, P) = u^I(b^*(P),\, P)$. Recall that $H(b,\, u,\, P)$ [cf. equation (6)] captures how easy it is for a to transfer surplus to b at any price P. For a given u, a higher price makes transferability easier under non-integration, and hence, $h(\cdot)$ is positive. This is because it is easier for a to incentivize b to attach less weight to private costs when the value from coordination is higher due to a higher market price.[26] By how much the utility of $b^*(P)$ changes when P increases depends on how the extra surplus is shared at all inframarginal enterprises who are non-integrated.

The two aforementioned effects of price increase on the decision to integrate of the initially indifferent enterprise $(\alpha(b^*(P)),\, b^*(P))$ is best understood in terms of Figure 5 in which the bargaining frontier labeled P denotes the one at the initial market price P, whereas the higher one labeled P' is the frontier associated with $(\alpha(b^*(P)),\, b^*(P))$ at the new price P'. Depending on the indifferent enterprise before the price increase being at E_L or E_H, we have $u(b^*(P),\, P) = u_L^*$ or u_H^*.

To fix ideas, suppose that the B side is homogeneous, but the A side exhibits productivity heterogeneity, and both sides being of equal measure. In this limiting case, all B's receive the same utility in equilibrium which is equal to their outside option u_0. Suppose that $u_0 = u_H^*$, i.e., the utility allocation between A and B units in enterprise $(\alpha(b^*(P)),\, b^*(P))$ is given by the point E_H. At E_H, the B firm clearly receives higher utility than the A firm (because E_H lies below the 45^0-line). Following an increase in the market price, all the additional surplus would accrue to the A firm which would make the utility allocation more balanced (the utility of the A firm, $v(\alpha(b^*(P)))$ increases to V'). Thus, the initially indifferent enterprise would prefer non-integration at the higher price. By contrast, if $u_0 = u_L^*$, the enterprise $(\alpha(b^*(P)),\, b^*(P))$ is at E_L. All the extra surplus due to price increase would again be

[24] Note that
$$\frac{\partial[u_H(\cdot,\, P) - u_L(\cdot,\, P)]}{\partial P} = \frac{z\left(2c^2 - 6cR + 3R^2\right)}{2c\sqrt{(2c-R)(2c-3R)}}\,.$$
The above expression is strictly negative for $R^- < R < R^+$.

[25] The details of the derivations are provided in the Appendix in the Proof of Proposition 2.

[26] Under integration the utility of b increases proportionately with the price, as it can be easily observed from (5).

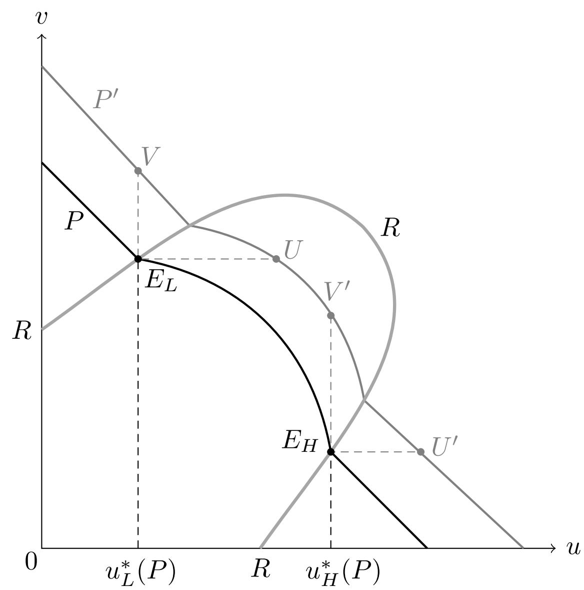


Figure 5: The bargaining frontiers of the marginal enterprise before and after the price increase. If the initial utility allocation is at E_L, and u is fixed at u_L^*, then this enterprise prefers to integrate at the new utility allocation at V. If, on the other hand, the initial utility allocation is at E_H, and u is fixed at u_H^*, then this enterprise prefers to stay separate at the new utility allocation at V'.

consumed by the A unit. However, now the utility allocation would become more unbalanced in favor of A as $v(\alpha(b^*(P))$ settles at point V. As a result, the initially indifferent enterprise prefers integration at the higher price P'.

The above intuition does not change qualitatively if we add a little heterogeneity on the B side. Suppose the B firms are heterogeneous with respect to their productivities and the initially indifferent enterprise is at E_H. Starting at a u_0 close to u_H^* a higher price would induce this enterprise to choose non-integration provided that the B firm receives very little of the extra surplus relative to the A side. The derivative that measures this response is given by (7). Following this expression, the increase of $u^N(b,\, P)$ is small when P increases if the area under the 'adjusted marginal productivity' $h(b,\, u,\, P)z_b(\alpha(b),\, b)$ up to the threshold productivity $b^*(P)$, is small. This depends on the marginal productivities of all the inframarginal B firms, the relative scarcity of types (which affects the steepness of the matching function $\alpha(b)$, given by the ratio of the densities, $f(b)/g(\alpha(b))$), and the level of $b^*(P)$.[27] But if the area under $z_b(\alpha(b),\, b)$ is high, then the B units get most of the additional surplus and a higher price may imply more integration.

In the limiting case where the A firms are homogeneous, the entire additional surplus would accrue to the B unit at the initially indifferent enterprise with the utility allocation at point U' in Figure 5 where integration is the preferred choice of $(\alpha(b^*(P)),\, b^*(P))$ at the increased price P'. A similar intuition can be told if the utility allocation at the initially indifferent enterprise is at E_L. Whether such an enterprise is at E_L or at E_H depends on u_0 and the area under the adjusted value of marginal product

[27]For example, $z_b(\alpha(b),\, b)$ may be decreasing in b for a fixed a, but if the matching function is steep it can turn to be an increasing function in equilibrium.

curve, $Pz_b(\alpha(b),\, b)H(\cdot)$ [cf. equation (6)]. A high u_0, very heterogeneous B units, or high marginal productivity of the B firms makes it more likely that the initially indifferent enterprise will have a utility allocation be at E_H because the utility of the B firms would be relatively high. We summarize the above discussion in the Proposition below,

Proposition 3 *The effect of an exogenous increase in the product market price on the fraction of integrated enterprises can be positive or negative.*

The main difficulty in coming up with a tractable sufficient condition for the behavior of $b^*(P)$ is that, given the nature of the ODE under non-integration, (6), it is not possible to solve for $u(b,\, P)$ analytically. In the following numerical example we show that higher product market price can lead to less integration under two-sided heterogeneity.

Example 1 (higher prices leading to less integration) Let $z(a,\, b) = a^{0.21}b^{0.28}$. We set $c = 0.25$ and $u_0 = 0.0165$. Furthermore, we assume that a follows Beta distribution on the support $[0.24,\, 0.3]$ with shape parameters $(10,\, 1)$, and b follows Beta distribution on $[0.3,\, 0.4]$ with shape parameters $(2,\, 2)$. Using the above data we solve the ODE in (3) numerically to find $u(b,\, P)$.[28] The product market price has been varied between $P^- \approx 0.24$ and $P^+ \approx 0.32$ as $b^*(P)$ does not exist outside this range.[29]

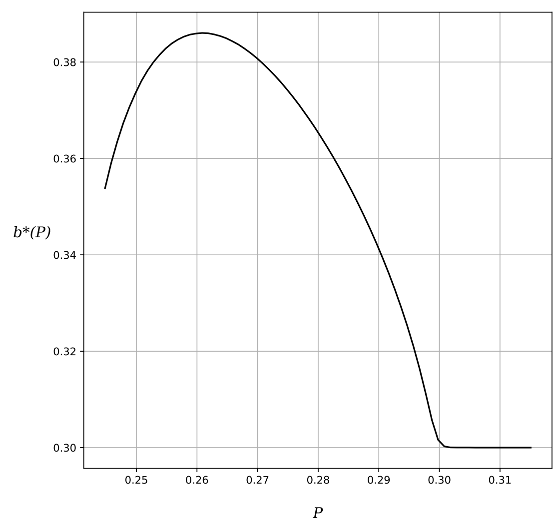


Figure 6: The threshold productivity increases with the market price if price is low, whereas it is decreasing in P for higher values.

In the above example, the initial equilibrium is at E_H. Figure 6 shows that more enterprises shift to non-integration following a price increase if the price level is low, whereas the incidence of integration increases with price for high values. For very high values of P, $b^*(P)$ almost coincides with $\underline{b}$ as nearly all enterprises choose to integrate at these prices. As $u_H(\cdot,\, P) - u_L(\cdot,\, P)$ at any enterprise decreases

[28]We use *Julia* for the numerical simulation. The codes are available upon request.

[29]Note that $b^*(P)$ may not exist even for prices in $[P^-,\, P^+]$ as shown in panels A and B of Figure 4.

with the price level, for high prices (close to P^+) the effect of price increase in favor of integration is more likely to dominate, whereas for prices close to P^-, this positive effect is weak. Consequently, in the above example, we observe an inverted-U shape of the $b^*(P)$ function.

6.2 Organizationally augmented industry supply

We now derive the industry supply curve (OAS) as a function of the product market price. Recall that $q^N(P,\, z(a,\, b)) = (P/2c)[z(a,\, b)]^2$ and $q^I(P,\, z(a,\, b)) = z(a,\, b)$. The organizationally augmented industry supply is the expected output aggregated across all the enterprises in the market equilibrium, which is given by:

$$Q(P) = \begin{cases} \dfrac{P}{2c}\displaystyle\int_{\underline{b}}^{\bar{b}} [z(\alpha(b),\, b)]^2 dF(b) & \text{if } P < P^-, \\ \dfrac{P}{2c}\displaystyle\int_{\underline{b}}^{b^*(P)} [z(\alpha(b),\, b)]^2 dF(b) + \displaystyle\int_{b^*(P)}^{\bar{b}} z(\alpha(b),\, b) dF(b) & \text{if } P^- \leq P \leq P^+. \end{cases}$$

Note that when $P < P^-$, all enterprises choose to stay separate, and hence, $Q(P)$ is a linear upward-sloping function. Interestingly, for prices in the interval $[P^-,\, P^+]$, $Q(P)$ is in general non-linear, and possibly non-monotonic.[30] To see this, consider any P in $[P^-,\, P^+]$, and differentiate $Q(P)$ with respect to P to get

$$Q'(P) = -\underbrace{f(b^*(P))[q^I(P,\, z(\alpha(b^*(P)),\, b^*(P))) - q^N(P,\, z(\alpha(b^*(P)),\, b^*(P)))]}_{(+)} \cdot \underbrace{\frac{db^*(P)}{dP}}_{(-,+)} + \underbrace{\int_{\underline{b}}^{b^*(P)} \frac{[z(\alpha(b),\, b)]^2}{2c} f(b) db}_{(+)}. \tag{8}$$

For this range of price variation, a change in the product market price P affects $Q(P)$ via two channels – a rise in P (i) changes the fraction of integrated enterprises by changing the threshold productivity $b^*(P)$ of the indifferent enterprise which is captured by the first term of the above expression, and (ii) augments the output $q^N(P,\, z(\alpha(b),\, b))$ of each non-integrated enterprise, but leaves the integrated output $q^I(P,\, z(\alpha(b),\, b))$ unaltered – captured by the second term. Moreover, these two effects may point in opposite directions because the first term of the above expression may be positive or negative depending on the sign of $db^*(P)/dP$. If there is a positive association between the product market price and integration, i.e., $db^*(P)/dP < 0$ for all price levels, then $Q'(P) > 0$, and hence, the industry supply curve is upward-sloping for all P. On the other hand, a negative association between the product market price and integration, i.e., $db^*(P)/dP > 0$ is only a necessary but not sufficient condition for $Q'(P) < 0$ because the second term of the above expression is always positive.[31]

Proposition 4 *A positive association between the product market price and integration is a sufficient condition for the organizationally augmented industry supply $Q(P)$ being upward-sloping. However,*

[30] If we assume away that the lowest-productivity enterprise only chooses non-integration, then for $P > P^+$ all enterprises integrate. Because $q^I(P,\, z(a,\, b))$ does not depend on P, the OAS is vertical for this price range.

[31] The integrand in the second term is the partial derivative of $f(b)q^N(P,\, z(\alpha(b),\, b))$ with respect to P.

$Q(P)$ may have a downward-sloping segment. A necessary condition for the downward-sloping supply curve is a negative association between the product market price and integration, i.e., $db^*(P)/dP > 0$.

It is hard to write a simple sufficient condition under which the negative term in the expression of $Q'(P)$ would dominate the positive one whenever $db^*/dP > 0$. However, there are two ways in which the negative term can be large in magnitude. First, if $f(b^*(P))$ is very high, i.e., the equilibrium matching function $\alpha(b)$ is very steep at $b = b^*(P)$, then whenever $db^*/dP > 0$, a very large fraction of enterprises [around $b^*(P)$, initially indifferent ones] would switch to non-integration under an increased market price, thereby inducing a concomitant drop in the aggregate industry output. Second, for prices close to P^-, db^*/dP is more likely to be positive. Moreover, the impact of the negative term is significant if the difference between the integration and non-integration output is large enough around $b^*(P)$. Note that

$$q^I(P,\, z(a,\, b)) - q^N(P,\, z(a,\, b)) = z(a,\, b) - \frac{P}{2c}[z(a,\, b)]^2.$$

So, the above difference is more likely to be large when the product market price is close to P^-, and hence, one would expect to have a downward-sloping OAS for this price range, and an upward-sloping supply curve for prices close to P^+. Clearly, at $P \approx P^+$, the OAS would be almost vertical as most of the equilibrium enterprises except the very low-productivity ones would integrate. In the following example we show that the industry supply curve can be non-monotonic on $[P^-,\, P^+]$.

To sum up, when prices are either very low or very high all enterprises either remain separate or integrate. In both cases, the OAS curve is upward sloping. A mix of organizational structures in the market is predicted for prices that are neither too low nor too high. Within this price range, a backward-bending supply curve is likely to be observed for lower prices, as low prices make the negative association between price and industry output stronger.

Example 2 (Non-monotonic OAS) We maintain the same parameter specifications as in Example 1. Figure 7 depicts a non-monotonic organizationally augmented supply curve. We have taken $[P^-,\, P^+] = [0.243,\, 0.32]$ as the range of price variation. As expected, for prices close to P^-, the OAS is downward-sloping, whereas for prices close to P^+, it is upward-sloping. For very high prices, the aggregate industry output is not very sensitive to price changes as almost every enterprise integrate. It is worth noting that about 21.3% of the entire range of price variation attributes to the downward-sloping portion of the OAS.

Some aspects of Example 2 are worth mentioning. First, while the density of b is symmetric, that of a is skewed to the right, which makes the matching function very steep around $b^*(P)$ for prices close to P^-. Second, the range of price variation is $[P^-,\, P^+]$ is only a subset of the entire range $[P^0,\, \infty)$ where $P_0 = 0.241$ under the parameter specifications. Clearly, $Q(P)$ is linearly increasing on $[P_0,\, P^-]$ as for these prices all enterprises choose to stay separate. Moreover, around P^- there is a discontinuous jump in the aggregate industry output as $b^*(P)$ exists for prices strictly higher that P^-. On the other hand, for any price $P > P^+$ all enterprises integrate, and the OAS is vertical. We have not considered such prices because we have assumed that at least the lowest-productivity enterprise, i.e., $(\underline{a},\, \underline{b})$ chooses to stay separate at all prices. Thus on $[P^0,\, \infty)$, the OAS is actually *backward-bending*. To close our model, the equilibrium product market price P^* is determined by the intersection of the OAS, $Q(P)$ and the demand curve, $D(P)$.

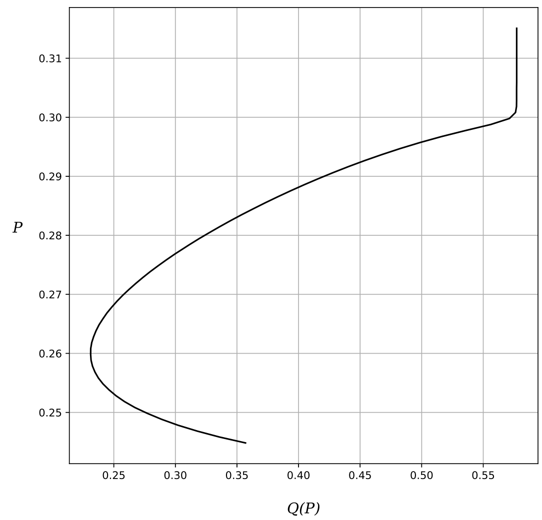


Figure 7: A non-monotonic industry supply. The OAS is downward-sloping for low levels of market price, whereas it is upward-sloping for high price levels.

7 Free entry: effect on integration and the long-run OAS

Our model so far has a fixed mass of firms. We now extend the analysis to allow for entry of firms in the long-run. There is a fixed cost of entry $\tau > 0$ for all A firms, whereas the entry cost of the B firms is normalized to zero. Suppose at an initial price P, A firms break-even, and B firms earn u_0 in the lowest-productivity enterprise $(\underline{a},\, \underline{b})$. When the product market price goes up, strictly positive profits at this enterprise will attract entry of new A firms with productivities lower than $\underline{a}$. To balance the input market, a positive mass of B firms will enter with productivities less than $\underline{b}$.

Entry at the bottom will continue at a given price P until the profits of the *new* lowest-productivity A firms dissipate. Hence, in the free-entry equilibrium, the productivity levels of the least-productive units are functions of the market price, i.e., $\underline{a}(P)$ and $\underline{b}(P)$. There will not be any re-match among the incumbents, and the long-run matching function will just be an extension towards the bottom of the short-run one. We continue assuming that the lowest-productivity enterprise following entry will stay separate. Thus, $\underline{a}(P)$ and $\underline{b}(P)$ are determined by solving the following two (long-run) equilibrium conditions:

$$\phi^N(\underline{a}(P),\, \underline{b}(P),\, u_0,\, P) = \tau, \tag{ZP}$$

$$\underline{a}(P) = \alpha(\underline{b}(P)). \tag{M}$$

The first condition is the zero-profit condition for $\underline{a}(P)$, and the second one is the matching function $\alpha(\cdot)$ obtained from the type distributions $G(a)$ and $F(b)$ defined on $[\underline{a}(P),\, \bar{a}]$ and $[\underline{b}(P),\, \bar{b}]$, respectively. Clearly, $\underline{a}'(P),\, \underline{b}'(P) < 0$.

Effect of entry on integration. Because the lowest-productivity B unit following entry consumes u_0, the least productive B firm before entry would have a utility strictly higher than u_0. As a result the equilibrium utility function would shift up relative to the one in the short-run. However, the equilibrium indifference locus would remain the same for all incumbent types b. Therefore, the indifferent enterprise in the short-run equilibrium would integrate in the long-run if the initial equilibrium is at E_H, whereas this enterprise would move to non-integration if the initial equilibrium is at E_L. Therefore, entry at the bottom generates a force towards non-integration if the equilibrium is at E_L. Nonetheless, it is difficult to predict whether the fraction of integrated firms in the long-run would go up because the total number of firms is higher in the long-run equilibrium and the entrants choose non-integration.

The long run OAS. The aggregate industry output surely increases in the long run due to more firms in the input market following entry. The long run OAS is similar to the short run where $\underline{b}$ is replaced by $\underline{b}(P)$. Recall that, for low prices, i.e., $P < P^-$, the short-run OAS is an increasing linear function of P. In the long run, in the expression of $Q'(P)$, there is an additional positive term induced by an increase in the non-integration output, and hence, for this range of prices, the long-run OAS would more elastic, and possibly non-linear. Next, consider the range of price $[P^-, P^+]$, and let the indifferent enterprise be denoted by $(\alpha(b^{**}(P)), b^{**}(P))$ which we assume to be unique. Differentiating the aggregate industry output with respect to price we obtain

$$\begin{aligned} Q'(P) = & \underbrace{-f(\underline{b}(P))q^N(P, z(\alpha(\underline{b}(P)), \underline{b}(P))\, \underline{b}'(P)}_{+} \\ & - f(b^{**}(P))[q^I(P, z(\alpha(b^{**}(P)), b^{**}(P))) - q^N(P, z(\alpha(b^{**}(P)), b^{**}(P)))] \cdot \frac{db^{**}(P)}{dP} \\ & + \int_{\underline{b}(P)}^{b^{**}(P)} \frac{[z(\alpha(b), b)]^2}{2c} f(b)db. \end{aligned}$$

Relative to the expression for the slope of the OAS in the short-run [cf. expession (8)], there is an additional term (the first term of the above expression) which is positive, and captures the effect of entry of lower productivity firms. Clearly, this reduces the range of price variation over which the OAS may be downward-sloping. Moreover, whenever the industry supply in the long run is upward-sloping, it is more elastic relative to the short run. By contrast, whenever the industry supply in the long run is downward-sloping, it is steeper than that in the short run because of the positive effect entry has on the output of the non-integrated firms.[32] To summarize,

Proposition 5 *Relative to the short run, in the long run following free entry at the bottom, (a) the aggregate industry output is higher; (b) the backward-bending segment of the OAS continues being a possibility although this segment shrinks; and (c) the upward-sloping part becomes more elastic, whereas the downward-sloping segment (if it exists) becomes less elastic.*

[32] For the other extreme, i.e., for prices $P > P^+$, the long run OAS would be upward-sloping, whereas the short-run OAS is vertical. To see this, note that for $P > P^+$ the OAS is given by:

$$Q(P) = \int_{\underline{b}(P)}^{\bar{b}} z(\alpha(b), b) f(b)db.$$

Therefore,

$$Q'(P) = -f(\underline{b}(P))z(\alpha(\underline{b}(P)), \underline{b}(P))\, \underline{b}'(P) + \underbrace{\frac{\partial}{\partial P}\int_{\underline{b}(P)}^{\bar{b}} z(\alpha(b), b) f(b)db}_{0}.$$

The above expression is strictly positive because $\underline{b}'(P) < 0$.

8 Empirical implications

The model makes reasonable assumptions under which: (i) there is positive assortative matching of firms with heterogeneous productivities who are supplying inputs that are complementary for the production of output, and (ii) high productivity enterprises, comprised of high productivity firms, integrate while low productivity ones remain separate (however, as we have discussed earlier, other equilibrium patterns of organizational modes are theoretically possible).

The prediction about positive assortative matching in general finds support in recent empirical literature, (e.g. Hortaçsu and Syverson, 2007; Atalay et al., 2014). Braguinsky et al. (2015) show, in the context of lateral mergers, that there may be negative sorting in the sense that high-profitability firms acquire low-profitability ones, but matching with respect to productivity is still positive assortative. The novelty of their work is a very detailed data set (from the Japanese cotton spinning industry) that allows them to accurately measure productivity, by distinguishing between quantity and revenue productivity.[33]

The empirical literature also offers supportive evidence about the likelihood of integration as a function of productivity. The results in Hortaçsu and Syverson (2007) and Atalay et al. (2014) suggest that high productivity firms are more likely to integrate.

Then, we predict that the association between the product market price and the incidence of integration may be positive or negative. The empirical evidence on such association is mixed — Hastings (2004), Alfaro et al. (2016) and McGowan (2017) find a positive relation, whereas the findings of Hortaçsu and Syverson (2007) suggest a negative association between price and integration. Our analysis brings together two forces: a higher price i) increases the integration benefits and ii) affects the relative allocation of the equilibrium utilities. If utilities become more balanced, incentives under non-integration are harmonized, and hence, non-integration is more likely. For instance, if units on the one side of the market either have low marginal productivity or there is not high productivity heterogeneity among them, then they will receive very little of the extra surplus due a higher market price. If this side has already a high utility, because their assets can also be employed productively in other markets, then utilities become more balanced when price increases and as a result integration in the market decreases.

Finally, we predict that the organizationally augmented supply curve (OAS) can have a downward-sloping segment. A negative relationship between market price and the fraction of integrated firms is a necessary condition for a higher price to lower aggregate output. This is the most surprising prediction of our analysis and, as far as we know, there is no empirical evidence on this.

Proposition 4 implies three possible patterns, due to price changes, with respect to integration and the aggregate industry output. As the product market price increases, we may have (a) a backward-bending OAS accompanied by less integration, (b) an upward-sloping OAS and less integration and (c) an upward-sloping OAS along with more integration. With this taxonomy in mind our model is amenable to analyze implications of changes in the product market price resulting from an exogenous (outward) shift of the market demand. If we are in case (a), then we would observe (i) a lower product price, (ii) a higher aggregate output and (iii) a greater number of integrated firms. This prediction is consistent with the findings of Hortaçsu and Syverson (2007). Legros and Newman (2013) consider two distinct levels (high and low) of exogenously given enterprise productivities and introduce a technology shock

[33] Of course, there are mergers and acquisitions where big firms acquire small firms, but first, size or profitability is not perfectly correlated with productivity, and second, the motives for these mergers are usually different from the ones our theoretical model is trying to capture.

that increases the measure of high productivity firms. This shifts the (upward-sloping) OAS out which, given a fixed downward-sloping demand, results in a lower price, higher output and more integration (because high productivity firms integrate). Thus, our model offers a complementary explanation of the same phenomenon in the sense that we rely on a demand rather than a technology shock.

9 Conclusion

We analyze the determinants of firm boundaries when firms, heterogeneous with respect to productivity, interact in a vertically differentiated input and a perfectly competitive product market. The choice of ownership structure in a given enterprise depends on the trade-off between the benefits from coordination and private costs — non-integration, a mode based on contingent revenue shares, puts too much weight on the private costs of managerial actions and hinders coordination; integration, which is based on delegation of decision rights to an outsider, facilitates coordination but ignores private costs. Neither mode of organization thus achieves full efficiency. Unbalanced utilities (due to unbalanced revenue shares) between the two units induce the managers to opt for integration because coordination is poor if they remain separate. Balanced utilities, on the other hand, harmonize incentives, and make non-integration more likely to dominate.

When supplier units are vertically ranked with respect to productivity, competition for high-quality units arises naturally in the input market. We model such competition as a two-sided matching game, which endogenizes revenue share or utility allocation in each enterprise. In an equilibrium allocation, the matching is positive assortative, i.e., more productive firms match together to form enterprises. Thus, ex-ante differences in input productivity imply ex-post differences in firm revenue. Moreover, high-revenue firms opt to integrate because the incentive to coordinate is high in such firms, whereas low-revenue enterprises stay separate.

Our paper contributes to the extant literature pertaining to OIO by introducing two-sided heterogeneity and sorting. Within each enterprise, a price change alters the endogenously determined distribution of surplus between each pair of firms. It can induce more or less coordination under non-integration, and consequently, the product market price may be positively or negatively associated with the decision to integrate. Moreover, the possible negative association between price and integration may give rise to a backward-bending industry supply curve. These findings generate interesting testable implications.

The present model also yields interesting implications with respect to *managerial firms*. When managers are partial revenue claimants, they tend to underweight enterprise revenues in favor of private costs because the perceived price is lower than the actual market price. Essentially, the impact of partial revenue claims by the managers on ownership structures is similar to the impact market price has when firms are non-managerial. The presence of managerial firms thus yields an equilibrium that is organizationally inefficient (e.g. Leibenstein, 1966) because the true market price (that reflects the social value of output) is unchanged, the decision to integrate depends on the fraction of the price that accrues to managers, and the equilibrium analyzed in the present context is efficient. Thus, the presence of managerial firms may imply either 'too little' or 'too much' integration in equilibrium relative to the social optimum under non-managerial firms. This can have interesting policy implications related to corporate governance. Stronger (weaker) CEO incentives, if it applies uniformly across all firms, can be viewed as equivalent to a higher (lower) price in our base model with no managerial firms, which may result in more (less) integration. If there is already too little (much) integration from the social perspective, then stronger (weaker) incentives can be welfare-enhancing. In the latter case, a cap on CEO pay may be a policy tool. Furthermore, taxes can also have implications on the efficiency of organizational choice that are similar to a price change.

Appendix: Proofs

Derivation of the bargaining frontier under non-integration. Substituting for e_A and e_B from the incentive compatibility constraints (IC_A) and (IC_B), the optimal contracting problem of an arbitrary enterprise $(a,\, b)$ reduces to:

$$\max_{s\in[0,\,1]} \;\; V_A(s;\, R) \equiv \frac{R^2}{4c}(1-s^2), \tag{$\mathcal{P}'_N$}$$

$$\text{subject to} \;\; U_B(s;\, R) \equiv \frac{R^2}{4c} s(2-s) = u. \tag{PC'_B}$$

From (PC'_B) it follows that

$$s(a,\, b,\, u;\, P) = 1 - \frac{\sqrt{R(a,\, b)^2 - 4cu}}{R(a,\, b)}.$$

We ignore the other root as it is strictly larger than 1. Using the expression in ($\mathcal{P}'_N$), the bargaining frontier under N is given by:

$$\phi^N(a,\, b,\, u;\, P) = \frac{1}{4c}\left[2R(a,\, b)\sqrt{R(a,\, b)^2 - 4cu} - \left(R(a,\, b)^2 - 4cu\right)\right]. \tag{A1}$$

Derivation of $u_L(z(a,\, b),\, P)$ and $u_H(z(a,\, b),\, P)$. First, consider the case when non-integration completely dominates integration in a given enterprise $(a,\, b)$. In this case the minimum non-integration surplus must be strictly higher than the maximum surplus under integration, i.e.,

$$\frac{R^2}{4c} > R - \frac{c}{2} \iff \left[R - c\left(2-\sqrt{2}\right)\right]\left[R - c\left(2+\sqrt{2}\right)\right] > 0$$

Because $R < 2c$, the above holds for $R < \left(2-\sqrt{2}\right)c \equiv R^-$. Next, consider the case when integration completely dominates non-integration in a given enterprise $(a,\, b)$. In this case the minimum non-integration surplus must be strictly higher than the maximum surplus under non-integration, i.e.,

$$R - \frac{c}{2} > \frac{3R^2}{8c} \iff \left(R - \frac{2c}{3}\right)(R-2c) < 0.$$

The above holds for $R > \frac{2c}{3} \equiv R^+$. Now consider the case when $R^- \leq R \leq R^+$. Note first that $\phi^N(a,\, b,\, u;\, P)$ intersects the linear function $\phi^I(a,\, b,\, u;\, P)$ exactly twice because both are symmetric with respect to the 45^0-line and the non-linear frontier is strictly concave. The two intersection points are given by:

$$u_L(z(a,\, b),\, P) = \frac{1}{8}\left[4R(a,\, b) - 2c - \frac{R(a,\, b)}{c}\cdot\sqrt{(2c - R(a,\, b))(2c - 3R(a,\, b))}\right], \tag{A2}$$

$$u_H(z(a,\, b),\, P) = \frac{1}{8}\left[4R(a,\, b) - 2c + \frac{R(a,\, b)}{c}\cdot\sqrt{(2c - R(a,\, b))(2c - 3R(a,\, b))}\right]. \tag{A3}$$

Note that $u_H(z(a,\, b),\, P) = R(a,\, b) - \frac{c}{2} - u_L(z(a,\, b)),\, P)$. Thus, for the existence of the two intersection points it suffices to show that $u_L(a,\, b) \geq 0$, which holds for $R^- \leq R(a,\, b) \leq R^+$.

Proof of Lemma 1. We first prove the sufficiency of (ID). Take any $a'' > a'$, $b'' > b'$ and $u'' > u'$, and let $\phi(a',\, b',\, u') = \phi(a',\, b'',\, u'')$. Thus, $u' = \psi(b';\, a')$ and $u'' = \psi(b'';\, a')$, i.e., $(b',\, u')$ and $(b'',\, u'')$ are on the same indifference curve of a'. Now, consider the indifference curve of a'' passing through $(b',\, u')$, and take the point $(b'',\, \hat{u})$ with $\hat{u} > u'$ on this indifference curve, i.e., $u' = \psi(b';\, a'')$ and $\hat{u} = \psi(b'';\, a'')$. Then, it follows from (ID) that $\hat{u} \geq u''$. Because $\phi(a,\, b,\, u)$ is strictly decreasing in u, we have $\phi(a'',\, b'',\, u'') \geq \phi(a'',\, b'',\, \hat{u}) = \phi(a'',\, b',\, u')$. As the types and utility of type b are arbitrarily chosen, the last inequality implies that $\phi(a,\, b,\, u)$ satisfies GID.

Next, we show the necessity of (ID). Given two arbitrary types a' and a'' with $a' < a''$, suppose that (ID) is violated for these two types at $b = b'$, i.e., $\psi(b;\, a'')$ crosses $\psi(b;\, a')$ from above at $b = b'$. Let $u' = \psi(b';\, a') = \psi(b';\, a'')$. By continuity of $\psi(b;\, a)$, there is a $\bar{b} > b'$ such that $\psi(b;\, a') > \psi(b;\, a'')$ for all $b \in (b',\, \bar{b}]$. Take $b'' \in (b',\, \bar{b}]$ such that $\phi(a',\, b',\, u') = \phi(a',\, b'',\, u'')$ and $\phi(a'',\, b',\, u') = \phi(a',\, b'',\, \hat{u})$. Because $\psi(b;\, a')$ is strictly increasing in b, we have $u'' > u'$. On the other hand, $\hat{u} = \psi(b'';\, a'') < \psi(b'';\, a') = u''$ as $b'' \in (b',\, \bar{b}]$. Thus, $\phi(a'',\, b',\, u') = \phi(a'',\, b'',\, \hat{u}) \geq \phi(a'',\, b'',\, u'')$ because $\phi(\cdot,\, \cdot,\, u)$ is strictly decreasing in u. Hence, GID is violated for these two given types a' and a'' at $(b'',\, u'')$.

Proof of Proposition 1. We prove the result in the following steps:

STEP 1: We first show that both $\psi^I(b;\, a)$ and $\psi^N(b;\, a)$ have increasing differences in $(b;\, a)$. Because both functions are differentiable, for each $d = I,\, N$, increasing differences of $\psi^d(b;\, a)$ is equivalent to $\psi^d_{ba}(b;\, a) \geq 0$. Note that for each $d = I,\, N$ as we have $\phi^d(a,\, b,\, \psi^d(b;\, a)) = \bar{v}$. Implicitly differentiating the last equality we obtain

$$\psi^d_b(b;\, a) = -\frac{\phi^d_2(a,\, b,\, u)}{\phi^d_3(a,\, b,\, u)} \quad \text{where} \;\; u = \psi^d_b(b;\, a).$$

First, consider $d = I$. Differentiating (2) with respect to a, b and u we get

$$\phi^I_1(a,\, b,\, u) = R_a(a,\, b), \;\; \phi^I_2(a,\, b,\, u) = R_b(a,\, b), \;\; \text{and} \;\; \phi^I_3(a,\, b,\, u) = -1. \tag{A4}$$

Therefore, $\psi^I_b(b;\, a) = R_b(a,\, b)$, and $\psi^I_{ab}(b;\, a) = R_{ab}(a,\, b) \geq 0$ as $R_{ab}(a,\, b) = Pz_{ab}(a,\, b) \geq 0$. Therefore, $\psi^I(b;\, a)$ has increasing differences in $(b;\, a)$. Next, consider $d = N$. Differentiating (1) with respect to a, b and u we get

$$\phi^N_1(a,\, b,\, u) = \frac{R_a(a,\, b)}{2c}\left(\frac{R(a,\, b)^2}{\sqrt{R(a,\, b)^2 - 4cu}} - R(a,\, b) + \sqrt{R(a,\, b)^2 - 4cu}\right), \tag{A5}$$

$$\phi^N_2(a,\, b,\, u) = \frac{R_b(a,\, b)}{2c}\left(\frac{R(a,\, b)^2}{\sqrt{R(a,\, b)^2 - 4cu}} - R(a,\, b) + \sqrt{R(a,\, b)^2 - 4cu}\right), \tag{A6}$$

$$\phi^N_3(a,\, b,\, u) = -\left(\frac{R(a,\, b)}{\sqrt{R(a,\, b)^2 - 4cu}} - 1\right). \tag{A7}$$

Thus,

$$\psi^N_b(b;\, a) = -\frac{\phi^N_2(a,\, b,\, u)}{\phi^N_3(a,\, b,\, u)} = R_b(a,\, b) \cdot \underbrace{\frac{1}{2c}\left(R(a,\, b) + \frac{R(a,\, b)^2 - 4cu}{R(a,\, b) - \sqrt{R(a,\, b)^2 - 4cu}}\right)}_{H(R(a,b),u)}.$$

Note that

$$\frac{\partial H}{\partial R}(R,\, u) = \frac{3R}{2c\left(R - \sqrt{R^2 - 4cu}\right)} > 0$$

because $u \leq R^2/4c$. Therefore,

$$\psi^N_{ab}(b;\, a) = \frac{\partial H}{\partial R}(R(a,\, b),\, u)R_a(a,\, b)R_b(a,\, b) + H(R(a,\, b),\, u)R_{ab}(a,\, b) > 0$$

because $R_a = Pz_a > 0$, $R_b = Pz_b > 0$ and $R_{ab} = Pz_{ab} \geq 0$. Thus, $\psi^N(b;\, a)$ has increasing differences in $(b;\, a)$.

STEP 2: Next, we show that $\hat{b}(a)$ decreases with a. First, by differentiating $\phi^d(a,\, b,\, \psi^N(b;\, a)) = \bar{v}$ for each $d = I,\, N$ with respect to a, we obtain

$$\psi^I_a(b;\, a) = -\frac{\phi^I_1}{\phi^I_3} = R_a(a,\, b),$$
$$\psi^N_a(b;\, a) = -\frac{\phi^N_1}{\phi^N_3} = H(R(a,\, b),\, u)R_a(a,\, b).$$

Recall that $\hat{b}$ solves $\psi^I(\hat{b};\, a) = \psi^N(\hat{b};\, a)$. Therefore, differentiating the last equation with respect to a we obtain

$$\frac{d\hat{b}}{da} = -\frac{\psi^I_a(\hat{b};\, a) - \psi^N_a(\hat{b};\, a)}{\psi^I_b(\hat{b};\, a) - \psi^N_b(\hat{b};\, a)} = -\frac{R_a(a,\, \hat{b})[1 - H(R(a,\, \hat{b}),\, u)]}{R_b(a,\, \hat{b})[1 - H(R(a,\, \hat{b}),\, u)]} = -\frac{z_a(a,\, \hat{b})}{z_b(a,\, \hat{b})} < 0.$$

The above inequality proves that the combined indifference curve $\psi(b;\, a)$ satisfies (ID), which is equivalent to $\phi(a,\, b,\, u)$ satisfying GID by Lemma 1. Therefore, it follows from Legros and Newman (2007) that the equilibrium matching is PAM.

Proof of Lemma 2. To guarantee uniqueness of the indifferent enterprise in equilibrium, suppose that $u(b,\, P)$ intersects the u_L portion of the indifference locus. By construction of the indifference point, $u_b(b,\, P) = R_b(\alpha(b),\, b)$ at any $b = b^*(P) + \varepsilon$ with $\varepsilon > 0$ but very small. If $u(b,\, P)$ intersects the indifference locus again at some $b' \in (b^*(P),\, b^+(P))$, then second intersection has to be with u_L. To avoid such a second intersection, a sufficient condition is that $u(b,\, P)$ is flatter than u everywhere. Note that

$$\frac{du_L}{db} = \frac{du_L}{dR} \cdot \left(R_a(\alpha(b),\, b)\alpha'(b) + R_b(\alpha(b),\, b)\right),$$

where

$$\frac{du_L}{dR} = \frac{1}{2} - \frac{3R^2 - 6cR + 2c^2}{4c\sqrt{(R - 2c)(3R - 2c)}}.$$

It is easy to show that u_L is strictly convex in R so that du_L/dR achieves its minimum value at $R^- = (2 - \sqrt{2})c$, which is given by:

$$\frac{du_L}{dR}(R^-) = \frac{1}{\sqrt{2}}.$$

Recall that $\alpha'(b) = f(b)/g(\alpha(b))$. Therefore, the sufficient condition for the uniqueness of $b^*(P)$ in this case is given by:

$$
\begin{aligned}
& \frac{1}{\sqrt{2}}\left(R_a(\alpha(b),\, b)\alpha'(b) + R_b(\alpha(b),\, b)\right) > R_b(\alpha(b),\, b) \\
\Leftrightarrow\ & \frac{1}{\sqrt{2}}\left(Pz_a(\alpha(b),\, b)\cdot\frac{f(b)}{g(\alpha(b))} + Pz_b(\alpha(b),\, b)\right) > Pz_b(\alpha(b),\, b) \\
\Leftrightarrow\ & r(b) > \sqrt{2} - 1. \qquad \text{(A8)}
\end{aligned}
$$

Now suppose that $u(b,\, P)$ intersects the u_H. If there is a second intersection it must be on u_H. To rule out such intersection, a sufficient condition would be that $du_H/db < u_b(b,\, P)$ for all b. Because for any $b = b^*(P) + \varepsilon$ with $\varepsilon > 0$ but very small, enterprise $(\alpha(b),\, b)$ integrate in equilibrium, we have $u_b(b,\, P) = R_b(\alpha(b),\, b)$ for $b = b^*(P) + \varepsilon$. On the other hand,

$$\frac{du_H}{db} = \frac{du_H}{dR}\cdot\left(R_a(\alpha(b),\, b)\alpha'(b) + R_b(\alpha(b),\, b)\right),$$

where

$$\frac{du_H}{dR} = \frac{1}{2} + \frac{3R^2 - 6cR + 2c^2}{4c\sqrt{(R-2c)(3R-2c)}}.$$

It is easy to show that u_H is strictly concave in R so that $du_H(R)/dR$ achieves its maximum value at $R^- = (2-\sqrt{2})c$, which is given by:

$$\frac{du_H}{dR}(R^-) = \frac{1}{2+\sqrt{2}}.$$

Therefore, the sufficient condition for the uniqueness of $b^*(P)$ in this case is given by:

$$
\begin{aligned}
& \frac{1}{2+\sqrt{2}}\left(R_a(\alpha(b),\, b)\alpha'(b) + R_b(\alpha(b),\, b)\right) < R_b(\alpha(b),\, b) \\
\Leftrightarrow\ & \frac{1}{2+\sqrt{2}}\left(Pz_a(\alpha(b),\, b)\cdot\frac{f(b)}{g(\alpha(b))} + Pz_b(\alpha(b),\, b)\right) < Pz_b(\alpha(b),\, b) \\
\Leftrightarrow\ & r(b) < \sqrt{2} + 1. \qquad \text{(A9)}
\end{aligned}
$$

Thus, conditions (A8) and (A9) together establish the proposition.

Proof of Proposition 2. The proof of part (a) is trivial because for prices $P < P^+$ all enterprises choose to stay separate because for each one of them N dominates I. In parts (b) and (c), all enterprises choose to stay separate because $u(b,\, P)$ never meets the indifference locus. Such possibility is depicted in Panel B of Figure 4. Finally, the result in part (d) emerges when the u_L and u_H portions of the indifference locus meet each other as in Panels C and D of Figure 4. In this case, $b^*(P)$ exists, and given the assumptions of the proposition, it is unique.

Sufficient condition for $db^*(P)/dP > 0$. Note that $u(b,\, P)$ consists of two parts — one under non-integration and the other under integration. Call them $u^N(b,\, P)$ and $u^I(b,\, P)$, respectively. It is easy to show that

$$
\begin{aligned}
u_P^N(b,\, P) &= \int_{\underline{b}}^{b} h(Pz(\alpha(x),\, x),\, u(x,\, P))z_b(\alpha(x),\, x)dx, \\
u_P^I(b,\, P) &= \int_{\underline{b}}^{b} z_b(\alpha(x),\, x)dx,
\end{aligned}
$$

where $h(R,\, u)$ is given by:

$$\frac{\partial}{\partial P}\left(-\frac{\phi_2^N}{\phi_3^N}\right) = \frac{\partial}{\partial P}[R_b H(R,\, u)] = \underbrace{\frac{1}{2c}\left(\frac{5R^2 - 4cu - R\sqrt{R^2 - 4cu}}{R - \sqrt{R^2 - 4cu}}\right)}_{h(R,u)} \cdot z_b.$$

The term $h(R,\, u)z_b$ is the marginal contribution of type b towards the additional surplus caused by a price increase, but adjusted for the problem of imperfect transferability. It is easy to show that $h(\cdot,\, u)$ is strictly decreasing and convex in u on $[0,\, R^2/4c]$ with $\lim_{u\to 0} h(R,\, u) = \infty$ and $h(R,\, R^2/4c) = 2R/c > 1$ for all $R \in [R^-,\, R^+]$. Therefore, $u_P^N(b,\, P) > u_P^I(b,\, P)$.

If the initial equilibrium is at E_L, for $b^*(P)$ to go up with the price level, we require that $u(b,\, P)$ shifts more than $U^L(b,\, P)$ following a price increase. Because $u_P^N(b,\, P) > u_P^I(b,\, P)$, a sufficient condition for this to hold is

$$\int_{\underline{b}}^{b} z_b(\alpha(x),\, x)dx \geq U_P^L(b,\, P).$$

As the equilibrium at E_L is reached for low-productivity B units, the above sufficient condition is hard to satisfy, and hence, it is more likely to have a positive association between price and integration. On the other hand, if the initial equilibrium is at E_H, for $b^*(P)$ to go up with the price level, we require that $u(b,\, P)$ shifts less than $U^H(b,\, P)$ following a price increase. Because $u_P^N(b,\, P) > u_P^I(b,\, P)$, and $h(R,\, u)$ is strictly decreasing in u, a sufficient condition for this to hold is

$$\int_{\underline{b}}^{b} h(Pz(\alpha(x),\, x),\, u_0)z_b(\alpha(x),\, x)dx \leq U_P^H(b,\, P).$$

As the equilibrium at E_H is reached for high-productivity B units, the above sufficient condition is more likely to hold when u_0 is high enough.

References

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

Aghion, Philippe, Rachel Griffith, and Peter Howitt (2006), "Vertical Integration and Competition." *The American Economic Review*, 96, 97–102.

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

Alonso-Paulí, Eduard and David Pérez-Castrillo (2012), "Codes of Best Practice in Competitive Markets for Managers." *Economic Theory*, 49, 113–141.

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

Bernard, Andrew, Andreas Moxnes, and Karen Helene Ulltveit-Moe (2014), "Two-sided Heterogeneity and Trade." NBER Working Paper No. 20136.

Birkhoff, Garret and Gian-Carlo Rota (1989), *Ordinary Difierential Equations*. John Wiley.

Bloom, Nicholas and John van Reenen (2007), "Measuring and Explaining Management Practices Across Firms and Countries." *The Quarterly Journal of Economics*, 122, 1351–1408.

Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson (2015), "Acquisitions, productivity, and profitability: Evidence from the Japanese cotton spinning industry." *American Economic Review*, 105, 2086–2119.

Chade, Hector, Jan Eeckhout, and Lones Smith (2017), "Sorting through Search and Matching Models in Economics." *Journal of Economic Literature*, 55, 493–544.

Chakraborty, Archishman and Alessandro Citanna (2005), "Occupational Choice, Incentives and Wealth Distribution." *Journal of Economic Theory*, 206–224.

Cole, Harold, George Mailath, and Andrew Postlewaite (2001), "Efficient Non-Contractible Investments in Large Economies." *Journal of Economic Theory*, 101, 333–373.

Dragusanu, Raluca (2014), "Firm-To-Firm Matching Along the Global Supply Chain." Mimeo, Harvard University.

Fallick, Bruce, Charles Fleischman, and James Rebitzer (2006), "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster." *The Review of Economics and Statistics*, 88, 472–481.

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

Forbes, Silke and Mara Lederman (2010), "Does Vertical Integration Affect Firm Performance? Evidence from the Airline Industry." *RAND Journal of Economics*, 41, 765–790.

Gibbons, Robert (2010), "Inside Organizations: Pricing, Politics and Path Dependence." *Annual Review of Economics*, 2, 337–365.

Gibbons, Robert, Richard Holden, and Michael Powell (2012), "Organization and Information: Firms' Governance Choices in Rational-Expectations Equilibrium." *The Quarterly Journal of Economics*, 127, 1813–1841.

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

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

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

Hastings, Justine (2004), "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California." *The American Economic Review*, 94, 317–328.

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

Kremer, Michael and Eric Maskin (1996), "Wage Inequality and Segregation by Skill." NBER Working Paper 5718.

Legros, Patrick and Andrew Newman (2007), "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferabilities." *Econometrica*, 75, 1073–1102.

Legros, Patrick and Andrew Newman (2013), "A Price Theory of Vertical and Lateral Integration." *The Quarterly Journal of Economics*, 128, 727–770.

Legros, Patrick and Andrew Newman (2014), "Contracts, Ownership, and Industrial Organization: Past and Future." *Journal of Law, Economics, and Organization*, 30, i82–i117.

Leibenstein, Harvey (1966), "Allocative Efficiency vs." X-Efficiency"." *The American Economic Review*, 56, 392–415.

Macho-Stadler, Inés, David Pérez-Castrillo, and Nicolás Porteiro (2014), "Coexistence of Long-term and Short-term Contracts." *Games and Economic Behavior*, 86, 145–164.

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

Sugita, Yoichi, Kensuke Teshima, and Enrique Seira (2015), "Assortative Matching of Exporters and Importers." Mimeo, Instituto Tecnológico Autónomo de México.

Syverson, Chad (2011), "What Determines Productivity?" *Journal of Economic literature*, 49, 326–365.

Zhu, Jane M, Viren Patel, Judy A Shea, Mark D Neuman, and Rachel M Werner (2018), "Hospitals using bundled payment report reducing skilled nursing facility use and improving care integration." *Health Affairs*, 37, 1282–1289.