# Two-sided productivity heterogeneity, firm boundaries and assortative matching

Kaniṣka Dam*  Konstantinos Serfes†

**Abstract**

We study two-sided markets where production requires the coordination of two complementary units. The market consists of a continuum of units on each side that are vertically ranked with respect to their productivities. Units match one-to-one to form an enterprise. An enterprise can choose a centralized or a decentralized form of organization and managerial actions are non-contractible. The choice of the organization mode gives rise to non-smooth Pareto frontiers. There is ample empirical evidence that this is the prevailing structure in many markets and matching is positive assortative. We provide a novel condition, in addition to the standard *type-type* and *type-payoff* complementarities, for positive assortative matching in this non-smooth imperfectly trasferable utility (ITU) environment, which we term *type-organization* complementarity. Higher enterprise revenue makes integration more likely, but how the surplus is shared within an enterprise also affects the incentives to integrate. There are equilibria where more productive enterprises integrate and equilibria where high productivity enterprises choose non-integration. These patterns have implications about the enterprise productivity. We investigate the effect of model primitives on the market productivity distribution.

*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

Modern enterprises usually consist of multiple complementary units where the form of organization varies across enterprises within the same market. On the one end, units adopt a centralized form of organization, while on the other, each unit can have some form of independence and formal contracts are used to coordinate actions across units. For example, in the airline industry, 'major' and 'regional' airlines must coordinate in order to connect smaller cities to the major hubs, while in the healthcare industry, hospitals and skilled nursing facilities (SNFs) must coordinate to guarantee the best health outcome for each patient. Some major airlines are vertically integrated, and some hospitals are integrated with some SNFs. Nevertheless, there are enterprises in these markets that are not integrated. In addition, there is ample empirical evidence of productivity heterogeneity and positive assortative sorting among the units (we review the empirical evidence later). High-productivity units are more likely to match with high-productivity complementary units to form enterprises, and are also more likely, compared with the low-productivity ones, to choose a more centralized form of organization.

Given that firm boundary decisions (i.e., integration versus non-integration) affect the productivity of the enterprises, and a possible assortative matching amplifies individual productivities, understanding how heterogeneous firms make decisions and interact with each other in these markets is crucial for assessing industry performance and welfare. Empirical research in this area has benefited tremendously from the recent availability of detailed firm-level data. However, there is no prior theoretical work examining the equilibrium matching pattern in a model that features two-sided productivity heterogeneity, endogenous matching and the choice of organizational structure. Our paper aims to contribute in this direction.

Guided by recent empirical evidence, and building on the theoretical work of Legros and Newman (2013), we posit a parsimonious model of enterprise formation in a two-sided market. Let us call the sides $A$ side and $B$ side, respectively. On each side of the market, there is a continuum of firms (e.g. input suppliers) whose assets are overseen by (cash constrained) managers. One $A$ firm is matched with one $B$ firm to form an enterprise where output is produced by combining two complementary assets of the supplier units. Given the two-sided heterogeneity of the (input) market, in equilibrium there is endogenous sorting of productivity types of $A$ and $B$ units. Within each enterprise the unit managers decide either to work as separate firms (decentralized organization) or to integrate (centralized organization). In a non-integrated enterprise, managers retain the decision rights, whereas in an integrated organization the decision rights are conferred to an outsider, called the *headquarter*. Managerial actions are not verifiable and entail private costs. Coordination of managerial actions enhances the likelihood of high revenue. No single organizational structure fully dominates the other. Non-integration places higher weights on private managerial costs, and hence, are conducive to poor coordination. Under integration, on the other hand, private costs are ignored as revenue maximization becomes the objective of the integrated entity. As a result, both organizational modes entail inefficiency.

The main objective of our paper is to provide conditions under which the matching is positive assortative (PAM), the most empirically plausible pattern. There are two particular characteristics while firms form enterprises through endogenous matching—(a) because managerial actions are non-contractible, utility can only be imperfectly transferred (ITU) between the production units, and (b) in each enterprise, the choice between two organizational modes induces the preferences of the units to be non-differentiable. An important feature of our model is how non-contractibility of managerial actions and firm boundary decisions interact with each other in the endogenous formation of enterprises. The set of sufficient conditions for PAM that we provide encompasses various degrees of *complementarity* between the units. First, the usual *type-type complementarity*, which is the sole determinant of PAM in a smooth transferable utility (TU) environment, asserts that the productivities of the $A$ and $B$ firms are complementary in producing enterprise surplus (e.g. Sattinger, 1993). Second, the *type-payoff complementarity* which captures the idea that the two units are complementary not only in creating surplus, but also in transferring enterprise surplus. This second sort of complementarity is a consequence of the ITU environment (e.g. Legros and Newman, 2007). Finally, a third type, which we term *type-organization complementarity*, is a novel complementarity condition that is induced naturally in our framework. This condition postulates that the marginal contributions of both units towards the change in enterprise surplus, arising from organizational restructuring, point in the same direction. In other words, type-organization complementarity emerges if both $A$ and $B$ firms gain at the margin by switching to the *same* organizational structure.

We begin by analyzing a tractable model of organizational choice where utility is perfectly transferable between the two units. In a TU environment, the issue of type-payoff complementarity does not play any role in determining the equilibrium sorting pattern. However, the choice between two production modes introduces non-differentiability in the preferences (represented by the indifference curves of the $A$ units). As is well-known, a supermodular production technology implies positive assortative matching (e.g. Sattinger, 1993; Topkis, 1998). However, supermodularity of production function, i.e., type-type complementarity under each organizational mode does not guarantee supermodularity under the combined organization. Thus, the type-organization complementarity is required to guarantee PAM. This difficulty exacerbates in our ITU environment. Legros and Newman (2007) assert that the Pareto frontier (of each enterprise) satisfying the *generalized increasing difference* (GID) is a necessary and sufficient condition for PAM in a smooth economy under ITU. GID is a 'single-crossing' condition which deals with both type-type and type-payoff complementarities—if an $A$ unit with a given productivity is indifferent between two productivity-utility combinations of the $B$ firms, then a higher-productivity $A$ firm is willing to transfer more surplus than her low-productivity counterpart to the higher-productivity $B$ firm. However, as in the case of TU, GID under each organizational structure does not necessarily implies GID under the combined structure, and hence, our type-organization complementarity becomes important in determining the equilibrium matching pattern.

We first offer a simple characterization of GID when each enterprise chooses between two organizational modes. Our principal contribution is to show that under type-organization complementarity, GID can be easily extended to a non-smooth environment. In other words, GID under each organizational mode implies that under the combined structure if the type-organization complementarity condition holds the equilibrium exhibits PAM. Our contribution is methodological, and thus, can be applied to situations other than choice of organizational modes. One such application (borrowed from Alonso-Paulí and Pérez-Castrillo, 2012) is analyzed in Section 6 where firms, comprising heterogeneous shareholders and managers, choose between two different contractual schemes for the managers.

Our paper is related to Legros and Newman (2013) and Dam and Serfes (2020) which contribute to the recent literature on Organizational Industrial Organization which is concerned with how market structure affects firm boundaries decisions. In other words, both the aforementioned papers analyze the structure of the product market (e.g. competition) as an external driver of enterprise heterogeneity in the sense that an exogenous change in market price affects firms' decision whether to integrate or stay as separate production units. In contrast, the present paper focuses on the internal drivers of productivity heterogeneity, in particular, how differences in managerial practices translate into differences in firm productivity. Our results conform to the recent empirical evidence on sorting of productivity in vertically or horizontally related markets. On the methodological side, ours is one of the first few applications of recent results in the theory of assortative matching under ITU to a non-smooth environment. On the technical side, the present paper is also close in spirit to Eekhout and Kircher (2018) who consider matching between firm size and both quality and quantity of workers. Their model analyzes equilibrium sorting under many-to-one matching. Complementarities between types, between quantities, and across qualities and quantities determine whether the equilibrium exhibits PAM.

## *1.1 Empirical relevance*

We briefly review the empirical literature that presents evidence on the main ingredients of our theoretical model—namely, productivity heterogeneity, firm boundary decisions based on productivity heterogeneity and diverse motives for integration such as enhanced coordination of management practices (intangible assets), and endogenous sorting.

*Productivity heterogeneity.* Productivity heterogeneity among firms is pervasive in almost every industry. Syverson (2011) identifies a number of factors within and outside the organizations that contribute to productivity differences across businesses. Differences in management practices (e.g. Bloom and Van Reenen, 2007, 2010), quality of labor inputs (e.g. Abowd, Haltiwanger, Jarmin, Lane, Lengermann, McCue, McKinney, and Sandusky, 2005), use of information technology as a special type of capital inputs (e.g. Jorgenson, Ho, and Stiroh, 2008) and firm boundary decisions (e.g. Forbes and Lederman, 2011), among others, are the main *internal drivers* of firm productivity differences. On the other hand, product market competition (e.g. Eaton and Kortum, 2002; Melitz, 2003; Legros and Newman, 2013; Dam and Serfes, 2020), productivity and knowledge spillovers (e.g. Moretti, 2004; Griffith, Harrison, and Van Reenen, 2006) and (de)regulation (e.g. Knittel, 2002) have been recognized as the principal *external drivers* of productivity heterogeneity. As we mentioned above, our focus in this paper is on the internal drivers of productivity heterogeneity.

*Firm boundary decision and coordination.* We first expand on the airline industry example we presented earlier, and we discuss some trade-offs between integration and non-integration. In the airline industry, regional airlines operate as 'subcontractors' for major U.S. network carriers in order to connect smaller cities to the major airline hubs. Forbes and Lederman (2011) show that airlines that own their regional affiliates (integration) experience shorter delays and fewer cancellations than those who outsource services to the regional airline firms (non-integration). As setting out the decision rights within the organization clearly improves coordination, a vertically integrated firm copes better with unforeseen adversities such as unexpected scheduling issues. However, this enhanced coordination comes at an expense—mainly the higher transaction costs of integration, e.g. higher wages for the employees in order to compensate for changes in the their daily routine integration introduces.[1] Grossman and Help-

[1]In the healthcare industry, as of 2015, approximately 20% of all Medicare fee-for-service hospital admissions ended in skilled nursing facility (SNF) stays (see 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 SNF's received 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.

man (2002), Hart and Hölmstrom (2010) and Legros and Newman (2013) posit tractable models of firm boundary decisions, on which we build, where such trade-offs naturally emerge.

Atalay, Hortaçsu, and Syverson (2014) argue that vertical integration facilitates efficient transfers of *intangible inputs* such as managerial oversight within firms rather than intra-establishment shipment of physical goods. As the authors assert " ... if this explanation is correct, there may not be anything particular about vertical structure within firms; intangible inputs can flow in any direction across a firm's production units. Vertical firm structures and expansions may not be fundamentally different from horizontal structures and expansions." Thus, our framework comprehends both vertical and horizontal production relationships.

*Positive assortative matching.* When heterogeneous firms match with each other in vertically and horizontally related market, evidence on positive sorting is also widespread. A number of papers find PAM of productivities in the context of intra-industry trade where buyers and suppliers maintain vertical relations through contracts (non-integration). Dragusanu (2014) analyzes matched importer-exporter data to show that more capable Indian manufacturing suppliers (exporters) match with more capable U.S. buyers (importers) following a positive assortative matching pattern with capability being proxied by firm size. Benguria (2015) finds similar evidence from U.S.-Colombia trade relations—more productive Colombian distributors import from more productive U.S. suppliers. Here, the firm productivity is measured by the residual revenue. Both Dragusanu (2014) and Benguria (2015) consider models of search and matching where firms invest in costly search for trading partners. In contrast, Sugita, Teshima, and Seira (2020) consider a classical matching model as ours where firms can seek alternative trading partners (partner switching). They also find evidence of PAM, i.e., more productive Mexican suppliers are matched with more productive U.S. distributors. Atalay, Hortaçsu, and Syverson (2014) analyze the role of productivity differences in the context of firm boundary decisions. Vertically integrated firms are not only larger and more productive on average, but also the authors present evidence of positive sorting of productivity and size of both upstream and downstream firms. In the context of horizontal merger, Braguinsky, Ohyama, Okazaki, and Syverson (2015) find evidence of positive assortative matching among heterogeneous firms with respect to productivity in the Japanese cotton spinning industry.

## 2 Choice of organization and PAM: An illustrative example

*Type-organization complementarity under transferable utility.* Consider two sides of a market—$A$ side and $B$ side, where on each side there is a continuum of agents, heterogeneous with respect to their productivities. The agents on side $A$ have productivities $a \in [\underline{a}, \bar{a}]$, and those on side $B$ have productivities $b \in [\underline{b}, \bar{b}]$. A partnership comprises one $A$ agent and one $B$ agent. Output or revenue (normalizing the market price to 1) of each partnership depends on the productivities of the agents, and thus denoted by $(a, b)$. Production of a single homogenous good in each partnership can be organized using one of the two technologies—organizational mode $\alpha$ entails a production $z^{\alpha}(a, b)$, whereas the production function associated with organization $\beta$ is given by $z^{\beta}(a, b)$. Both the production functions are assumed to be strictly monotone and supermodular, i.e., $z^{\alpha}_{ab}(a, b)$, $z^{\beta}_{ab}(a, b) \geq 0$.[2] Let $u$ and $v$ be the utilities of the

[2]Subscripts denote partial derivatives.

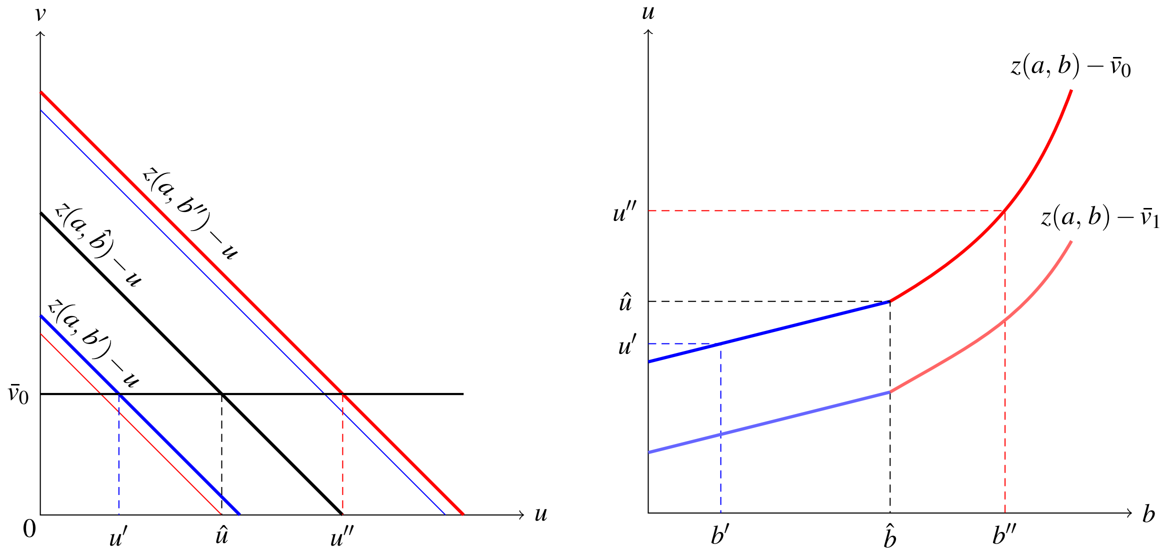


Figure 1: The left panel depicts the bargaining frontiers of a partnership $(a, b)$, obtained by varying $b$. The right panel depicts the indifference curves of a given $a$ that are derived from the bargaining frontiers.

$B$ and $A$ agents, respectively in a partnership $(a, b)$ such that

$$u + v = z(a, b) = \max\left\{z^{\alpha}(a, b), z^{\beta}(a, b)\right\}. \tag{1}$$

While in this example the two technologies are to some extent ad hoc, so that we can clearly illustrate our new condition in simple TU setting, in the main model that follows the two organizations arise naturally.

Let us assume that, for each $a$, there is a unique $\hat{b}$ such that $z^{\alpha}(a, b) < z^{\beta}(a, b)$ if and only if $b < \hat{b}$.[3] Any partnership $(a, \hat{b})$ is indifferent between the two organizational modes because $z^{\alpha}(a, \hat{b}) = z^{\beta}(a, \hat{b})$. Equation (1) represents the Pareto or bargaining frontier of an arbitrary match $(a, b)$, which are drawn in the left panel of Figure 1 with fixed $a$ and varying $b$. In each partnership utility is fully transferable (TU), and hence, the bargaining frontiers are linear [with slope equal to $-1$]. For low values of $b$, e.g. $b''$, organization $\alpha$ dominates organization $\beta$ (the red frontier lies above the blue one), and hence, $v = z(a, b') - u = z^{\alpha}(a, b') - u$. For high values of $b$, e.g. $b'$, on the other hand, $\beta$ dominates $\alpha$, i.e., the blue frontier lies above the red one. Therefore, $v = z(a, b'') - u = z^{\beta}(a, b'') - u$. Only at $b = \hat{b}$, the two frontiers associated with organizations $\alpha$ and $\beta$ coincide, i.e., $a$ is indifferent between the two modes if she is matched with $\hat{b}$.

Our object of interest is the indifference curve of each $a$ in her partner's productivity-utility space, which we derive taking the Pareto frontiers as the primitive, and are depicted in the right panel of Figure 1. Fix a utility level $\bar{v}_0$ of $a$. The intersection of the horizontal line at $\bar{v}_0$ of $a$ and the associated frontier determines the utility accruing to $b$. For example, $u'$, $\hat{u}$, and $u''$ correspond to $b'$, $\hat{b}$ and $b''$, respectively so that $a$ consumes a constant utility $\bar{v}_0$. In the right panel of Figure 1, the curve labeled $z(a, b) - \bar{v}_0$ is the indifference curve of $a$ on which all the productivity-utility combinations $(b', u')$, $(\hat{b}, \hat{u})$ and $(b'', u'')$ yields the same level $\bar{v}_0$. Each indifference curve has a kink at $\hat{b}$ because $(a, \hat{b})$ is indifferent between the two organizational modes. The slopes of the two portions of each indifference curve are in general

[3]Clearly $\hat{b}$ depends on $a$.

different because $z_b^\alpha(a, b) \neq z_b^\beta(a, b)$.[4] For a given $a$, higher indifference curves lie to the south-east (e.g. the one corresponds to a higher level of utility, $\bar{v}_1$). Note that, under TU, $\hat{b}$ remains the same irrespective of the constant utility level of $a$, i.e., the kinks of all the indifference curves lie on the same vertical line because the bargaining frontiers under the two modes coincide with each other at $b = \hat{b}$ for any utility allocation.

We are interested in a market with positive assortative matching (PAM). It is well-known that, under perfectly transferable utility, supermodularity of $z(a, b)$ is a sufficient condition for PAM. However, when $z(a, b) = \max\left\{z^\alpha(a, b), z^\beta(a, b)\right\}$, the supermodularity of each production technology does not guarantee the same property of the combined production function, and hence, PAM may fail to hold. The main reason of the failure of PAM is that $z(a, b)$ is not differentiable everywhere. To understand the intuition, assume that each production function is differentiable. Then the slope of the indifference curve under organizational mode $d = \alpha, \beta$ is given by:

$$\left.\frac{du}{db}\right|_{z^d} = z_b^d(a, b)\,.$$

Thus, supermodularity of $z^d(a, b)$ is equivalent to the fact that the indifference curves of $a''$ is everywhere steeper than those of $a'$ whenever $a'' > a'$ so that they cross only once (type-type complementarity). However, this not sufficient to guarantee the 'single-crossing' of the combined indifference curve. In order to illustrate further we consider the following two examples.

**Example 1** Let the production function be given by (as in Cole, Mailath, and Postlewaite, 2001):

$$z(a, b) = \max\left\{z^\alpha(a, b), z^\beta(a, b)\right\} \equiv \max\left\{ab, 2a^2b^2\right\}.$$

Each of the above production functions is supermodular. The indifference curves of two distinct $A$ units with $a' > a$ are depicted in the left panel of Figure 2. Note that the kink in a given $a$'s indifference curve is given by $\hat{b} = 1/2a$. Therefore, as $a$ increases the kink moves to the left meaning that the indifference curve of an $A$ agent with higher productivity is everywhere steeper than that of her lower-productivity counterpart. This ensures that $z(a, b)$ is supermodular and the economy exhibits PAM.

**Example 2** Let the production function be given by (as in Kremer and Maskin, 1996):

$$z(a, b) = \max\left\{z^\alpha(a, b), z^\beta(a, b)\right\} \equiv \max\left\{a^2b, ab^2\right\}.$$

Note first that each of $z^\alpha(a, b)$ and $z^\beta(a, b)$ is supermodular, i.e., holding the organizational structure fixed, both technologies satisfy type-type complementarity. We first show that, under the above production organization, the matching is not PAM. Let $a' = 1.5$, $a'' = 1.55$, $b' = 1.4$, $b'' = 1.6$, $u' \equiv u(b')$ and $u'' \equiv u(b'')$. Note first that $\hat{b}(a) = a$. Thus, both $a'$ and $a''$ would choose organization $\alpha$ if matched with $b'$, but organization $\beta$ if matched with $b''$. If there is PAM, i.e., the matches are $(1.5, 1.4)$ and $(1.55, 1.6)$, then it must be the case that

$$z^\alpha(1.5, 1.4) - u' \geq z^\beta(1.5, 1.6) - u'' \iff u'' - u' \geq 0.69,$$
$$z^\beta(1.55, 1.6) - u'' \geq z^\alpha(1.55, 1.4) - u' \iff u'' - u' \leq 0.6045.$$

[4] The approximate shapes of the indifference curves are consistent with the two examples analyzed in this section.

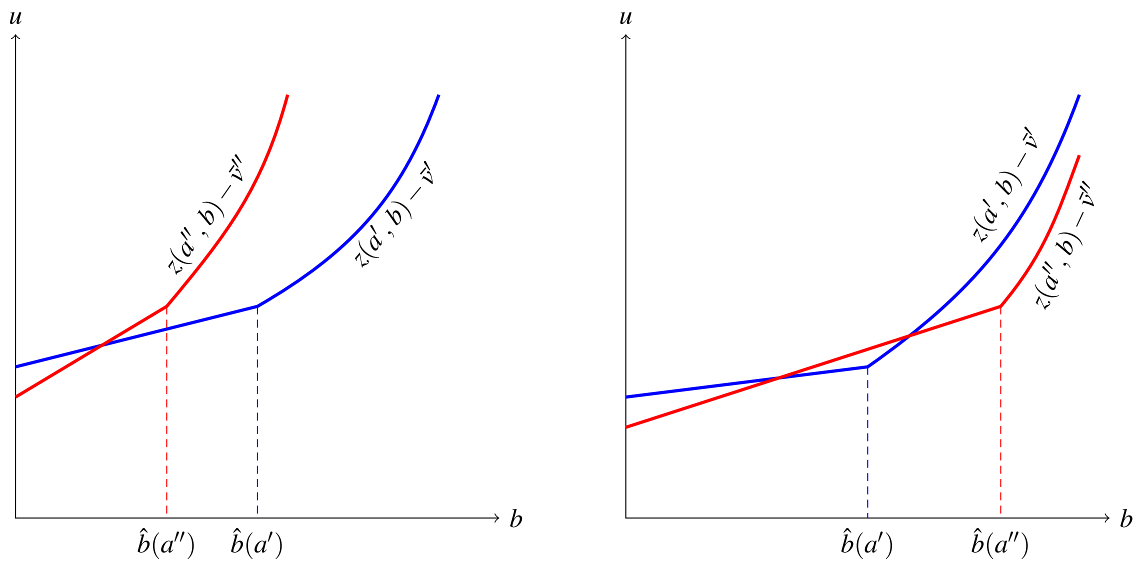


Figure 2: The left panel depicts the single-crossing of the indifference curves of two $A$ agents with productivities $a'$ and $a''$ with $a'' > a'$. The right panel shows the failure of single-crossing.

Clearly, there are no feasible utility allocations $(u', u'')$ for the $B$ units which satisfy the above two inequalities simultaneously, and hence, we cannot have PAM. The reason is that the indifference curves of $a'$ and $a''$ do not satisfy single-crossing property for $b \in [1.35, 1.65]$ because $\hat{b}$ moves to the right as $a$ increases, as shown in the right panel of Figure 2. Note also that

$$\begin{aligned} & z(a', b') - z(a', b) \leq z(a, b') - z(a, b) \\ \iff \quad & z^{\beta}(1.55, 1.6) - z^{\alpha}(1.55, 1.4) \leq z^{\beta}(1.5, 1.6) - z^{\alpha}(1.5, 1.4), \end{aligned}$$

i.e., $z(a, b)$ is not supermodular, i.e., type-type complementarity does not hold for the combines organizational structure.

The main difference between the two examples is that in the first one the kink of an indifference curve is moving to the left as we move to higher $a$ types, while in the second example it moves to the right. To understand this, define by $\Delta(a, b) \equiv z^{\alpha}(a, b) - z^{\beta}(a, b)$ the change in revenue due to a change in the organizational mode. Recall that $z^{\alpha}(a, \hat{b}) = z^{\beta}(a, \hat{b})$. Differentiating this condition with respect to $a$, we obtain

$$\frac{d\hat{b}}{da} = -\frac{\Delta_a(a, \hat{b})}{\Delta_b(a, \hat{b})}.$$

Clearly, $\hat{b}$ decreases with $a$ if and only if both the numerator and denominator of the above expression have the same sign, i.e., the marginal contributions of $a$ and $b$ towards the gain in revenue due to organizational restructuring point in the same direction, which we refer to as *type-organization complementarity*. In Example 1, $\Delta_a(a, \hat{b})$ and $\Delta_b(a, \hat{b})$ are of the same sign, i.e., type-organization complementarity holds, which reinforces type-type complementarity of each technology. As a result, we obtain PAM. By contrast, in Example 2, type-organization complementarity fails to hold as we have $\Delta_a(a, \hat{b}) = a^2 > 0$ and $\Delta_b(a, \hat{b}) = -a^2 < 0$. In other words, the marginal gain by switching from technology $\beta$ to technology $\alpha$

is increasing in $a$ but decreasing in $b$ around $(a, \hat{b})$. Thus, the matching may fail to be PAM even if each organizational mode exhibits type-type complementarity.

*Extension to ITU.* The main objective of the present paper is to extend the above argument to the case when utility is imperfectly transferable between the matched partner. Under ITU, a mere type-type complementarity such as the supermodularity of the production functions is often not sufficient to guarantee PAM. This is because the way the match surplus is divided between the two units also becomes a crucial determinant of the matching pattern. Thus, the notion of type-payoff complementarity emerges. In other words, one has to consider the non-linear bargaining frontier of each match $(a, b)$, denoted by $\phi(a, b, u)$, where $u$ is the utility of $b$, instead of the production function $z(a, b)$. Legros and Newman (2007) characterize PAM under ITU in a smooth environement—the bargaining frontier satisfying the following *generalized increasing difference* (GID) condition, i.e.,

$$\phi(a', b'', u'') = \phi(a', b', u') \quad \Longrightarrow \quad \phi(a'', b'', u'') \geq \phi(a'', b', u') \tag{2}$$

for any $a'' > a'$, $b'' > b'$ and $u'' > u'$ is a necessary and sufficient condition for PAM. In words, if $a'$ is indifferent between $b'$ and $b''$, when $b'$ receives utility $u'$ and $b''$ receives utility $u''$, then a higher type, $a''$, must receive a higher utility matching with the higher type $b''$ and offering utility $u''$ than matching with the lower type $b'$ and offering $u'$, for any utility levels $u$.

In what follows, we analyze economies with two-sided heterogeneity. The lack of contractibility of managerial actions gives rise to ITU. On the other hand, the choice between two distinct organizational structures induces non-differentiable indifference curves of the $A$ firms. We will show that GID of the bargaining frontier under each organizational mode is not sufficient to guarantee PAM, and type-organization complementarity (similar to the one developed above under TU) will imply GID of the combined bargaining frontier, and hence, PAM.

# 3 A Matching model of firm boundary decisions

## *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. 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 or revenue of an enterprise

$(a, b)$ is given by:

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

We assume that $z(a, b)$ is twice continuously differentiable, strictly increasing in $a$ and $b$, and supermodular in $(a, b)$, i.e., $z_{ab}(a, b) \geq 0$. 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 overseeing 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 $\frac{1}{2} e_A^2$ and that of a $B$ unit is $\frac{1}{2}(1 - e_B)^2$. Clearly, there is a disagreement about the direction of the decisions, what is easy for one is hard for the other. Also, managers, who have zero cash endowments, are protected by limited liability, i.e., their state-contingent incomes must always be non-negative. This assumption gives rise to two-sided moral hazard problem in each enterprise, and thus the division of surplus between the managers will affect the organizational choice.

### 3.2 *Organizational modes and contracts*

The organizational 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.[5] 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.[6] A contract $(s, d) \in [0, 1] \times \{N, I\}$ specifies a revenue share $s$ for the $B$ unit and an organizational mode $d$.

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

[5] 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.

[6] 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 continuum 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.

liability, the units get nothing in the case of failure. When the two units integrate, $HQ$ buys the assets of $A$ 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 $s_A + s_B + s_{HQ} = 1$. The $HQ$ market is assumed to be perfectly competitive with an opportunity cost normalized to zero.

### *3.3 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.

### *3.4 Equilibrium*

An equilibrium of the input or supplier market consists of a set of enterprises formed through feasible contracts, i.e., organizational structures and corresponding revenue shares, for each enterprise. Recall that there are two possible organizational modes for each enterprise—integration ($I$) and non-integration ($N$). In general, choice of organization depends on the revenue share that accrues to each member of an enterprise and the output of each enterprise. An allocation for the supplier 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 $\langle \alpha, v, u \rangle$ constitutes an equilibrium 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 for the enterprise;*

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

$$v = \max_b \phi(a, b, u). \tag{3}$$

*The function $\phi(a, b, u)$ 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.*

(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*

$$j_1 - j_0 = F(b_1) - F(b_0) = G(a_1) - G(a_0) = i_1 - i_0. \tag{4}$$

Definition 1-(b) asserts that each $A$ firm chooses her partner optimally which is equivalent to the notion of 'stability' in an assignment model. 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 The bargaining frontier of an arbitrary enterprise

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

*Non-integration.* Under this organizational mode the shares affect both the size and the distribution of surplus between the two units. In other words, utility is imperfectly transferable (ITU) between the unit managers. An optimal contract for a non-integrated enterprise $(a, b)$ solves the following maximization problem:

$$\max_{s} \quad V_A \equiv \pi(e_A, e_B)(1-s)z(a, b) - \frac{1}{2}e_A^2, \tag{5}$$

$$\text{subject to} \quad U_B \equiv \pi(e_A, e_B)sz(a, b) - \frac{1}{2}(1-e_B)^2 = u, \tag{6}$$

$$e_A = \underset{e}{\operatorname{argmax}} \left\{\pi(e, e_B)(1-s)z(a, b) - \frac{1}{2}e^2\right\} = (1-s)z(a, b), \tag{7}$$

$$e_B = \underset{e}{\operatorname{argmax}} \left\{\pi(e_A, e)sz(a, b) - \frac{1}{2}(1-e)^2\right\} = 1 - sz(a, b), \tag{8}$$

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 participate in any enterprise. Constraint (6) is the *participation constraint* of the $B$ unit, whereas constraints (7) and (8) are the *incentive compatibility constraints* of the $A$ firm and the $B$ firm, respectively. We analyze the optimal contracts under non-integration when $z(a, b) \leq 1$. Note that (7) and (8) together imply that $\pi(e_A, e_B) = z(a, b)$. When the firms in an arbitrary enterprise $(a, b)$ stay separate, the maximum payoff that accrues to the $A$ unit given that the $B$ unit consumes $u$ is given by:

$$v = \phi^N(a, b, u) = \frac{1}{2}\left[2z(a, b)\sqrt{z(a, b)^2 - 2u} - \left(z(a, b)^2 - 2u\right)\right] \text{ for } 0 \leq u \leq \frac{1}{2} \cdot z(a, b)^2. \tag{9}$$

The function $\phi^N(a, b, u)$ is the bargaining frontier of $(a, b)$ under non-integration. The participation constraint of the $B$ unit determines the optimal revenue share $s = s(a, b, u)$ that accrues to him. Note that $u$ must lie between 0, which corresponds to $s = 0$, and $z(a, b)^2/2$, the level corresponding to $s = 1$. The bargaining frontier is strictly concave and symmetric with respect to the $45^0$ line, on which $v = u$ 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 coordinate better their decisions, i.e., $e_A$ and $e_B$ move closer to each other. Finally, higher revenue $z$, holding the shares fixed, also induces better coordination.

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

$$v = \phi^I(a, b, u) = z(a, b) - \frac{1}{4} - u \quad \text{for } 0 \leq u \leq z(a, b) - \frac{1}{4}. \tag{10}$$

The above function is linear in $u$, i.e., utility is fully transferable (TU) between the two managers because neither the action taken by $HQ$ nor the costs borne by the managers depends on the revenue shares. The frontier is strictly increasing in $a$ and $b$, strictly decreasing in $u$ (with slope -1) and symmetric with respect to the $45^0$ line. Although utility 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 revenue, places too little weight on private managerial costs while maximizing expected revenue.

*Choice of organization.* We now analyze the optimal choice of organizational structure by a given enterprise $(a, b)$. It is worth noting that, relative to the first-best, both organizations are inefficient. Non-integration is conducive to poor coordination as this organization places too much weight on the private managerial costs. Integration, on the other hand, places too much importance on coordination, and tends to ignore private costs.[7] The optimal choice of organization thus depends on the revenue of the enterprise, $z(a, b)$ as well as the way the enterprise revenue, or, equivalently, the aggregate surplus is shared between the two units. At any level of utility $u$ accruing to the $B$ firm, an arbitrary enterprise $(a, b)$ would choose $N$ over $I$ if and only if $\phi^N(a, b, u) > \phi^I(a, b, u)$. The following lemma characterizes the optimal ownership structure.

**Lemma 1** *There are threshold values of enterprise revenue $z_L$ and $z_H$ with $0 < z_L < z_H < 1$ and threshold values of utility of the B firm $u_L$ and $u_H$ with $u_L < u_H$ such that an arbitrary enterprise*

(a) *chooses N over I if $z \in [0, z_L)$, or $z \in [z_L, z_H]$ and $u \in (u_L, u_H)$;*

(b) *chooses I over N if $z \in [z_L, z_H]$ and $u \notin [u_L, u_H]$, or $z \in (z_H, 1]$;*

(b) *is indifferent between N and I if $z \in [z_L, z_H]$ and $u \in \{u_L, u_H\}$.*

The bargaining frontier of an arbitrary enterprise $(a, b)$ is thus given by:

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

Low output, i.e., $z(a, b) < z_L$ implies benefits from coordination, and hence, the aggregate surplus from non-integration is strictly higher than that under integration, i.e., $\phi^N(a, b, u) > \phi^I(a, b, u)$ for all $u$. By contrast, for the high-output enterprises with $z(a, b) > z_H$, it is easier for the managers to ignore private costs, and hence, integration is the optimal choice as $\phi^I(a, b, u) > \phi^N(a, b, u)$ for all $u$. Interestingly, for medium-revenue enterprises, i.e., $z(a, b) \in [z_L, z_H]$, there is no clear dominance of one mode of organization over the other and the choice of organizational modes depends on how the surplus of the enterprise is distributed between the two units. For intermediate values of $u$, i.e., $u_L \leq u \leq u_H$, 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

[7] The first-best surplus, $z^2$, is strictly higher than $z - \frac{1}{4}$, the surplus accrued to an integrated firm as well as $\frac{3z^2}{4}$, the maximum surplus in a non-integrated firm, which corresponds to $s = \frac{1}{2}$.

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 strong. This case is depicted in the left panel of Figure 3 where the strictly concave frontier is the one associated with $N$ and the linear frontier is associated with $I$. Because both frontiers are symmetric with respect to the $45^0$ line, they intersect exactly twice at $u_L$ and $u_H$.[8,9]

## 5 The market equilibrium

We analyze the equilibrium of the input market with a continuum of firms and types. In particular, we extend the analysis presented in Section 2 to an environment in which utility is imperfectly transferable between the supplier units.

### *5.1 The indifference curves*

We first derive the indifference curve of each type $a$ $A$ firm from the bargaining frontiers, which are taken as the primitives. The analysis is very similar to that in Section 2. For each given $a$, an increase in $b$ implies an increase in the enterprise revenue $z(a, b)$, and hence, an expansion of the bargaining frontier (depicted by the 'higher' bargaining frontiers in the left panel of Figure 3) where we draw five frontiers, keeping $a$ fixed and varying $b$. To save on notation, each frontier is denoted by the productivity of the $b$ agent instead of $\phi(a, b, u)$. The concave frontier labeled $\underline{b}$ corresponds to the lowest-revenue enterprise $(a, \underline{b})$. We assume that $z(a, \underline{b}) < z_L$ so that for this enterprise $N$ is the dominant choice. On the other hand, the linear bargaining frontier labeled $\bar{b}$ is the one associated with $(a, \bar{b})$, the highest-revenue match. Because $z(a, \bar{b}) > z_H$, this enterprise prefers $I$ over $N$. The frontier labeled $b_L$ is such that $z(a, b_L) = z_L$. For any enterprise $(a, b)$ with $b_L < b < b_H$ where $z(a, b_H) = z_H$ there is no clear dominance of one organizational mode over the other. Matches $(a, b')$ and $(a, \hat{b})$ are two such enterprises, each one is indifferent between $N$ and $I$ at the utility allocations $u = u_L, u_H$.[10] The curve $EEE$, which we call the *type expansion path*, is drawn by joining the indifference points $u_L$ and $u_H$ by varying $b$. The starting point of the expansion path on each axis is $z_L^2/2$, the maximum utilities accruing to the $A$ or the $B$ unit in the enterprise $(a, b_L)$. Within the region enclosed by $EEE$, the corresponding enterprise chooses non-integration, whereas the units choose to integrate if the utility allocation lies outside this region.

Because $\phi(a, b, u)$ is strictly increasing in $b$ and strictly decreasing in $u$, the indifference curves $u = \psi(b, a, \bar{v})$ [drawn in the $b$-$u$ space] of $a$ are well-defined and satisfies $\phi(a, b, \psi(b, a, \bar{v})) = \bar{v}$ for some constant utility level $\bar{v}$ of $a$. In the left panel of Figure 3, consider the horizontal line at $\bar{v}_0$ which intersects all the bargaining frontiers at distinct utility allocations for the $B$ units (represented by the filled circles). Therefore, the five type-utility combinations $(\underline{b}, \underline{u})$, $(b_L, u_L)$, $(b', u')$, $(\hat{b}, \hat{u})$ and $(\bar{b}, \bar{u})$ are points

[8] Although the threshold outputs $z_L$ and $z_H$ do not depend on $a$ and $b$, the utility thresholds $u_L$ and $u_H$ do.

[9] When $z \geq 1$, the bargaining frontier under non-integration is different from the one in (9) because in this case the probability of success is given by $\pi(e_A, e_B) = 1$. When $z > 1$, the frontier under integration is everywhere above that under non-integration, and hence, integration is the preferred choice. When $z = 1$, 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 organizations, 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.

[10] Both $u_L$ and $u_H$ depend on $a$ and $b$; so, $u_L(a, b') \neq u_L(a, \hat{b})$ and $u_H(a, b') \neq u_H(a, \hat{b})$.

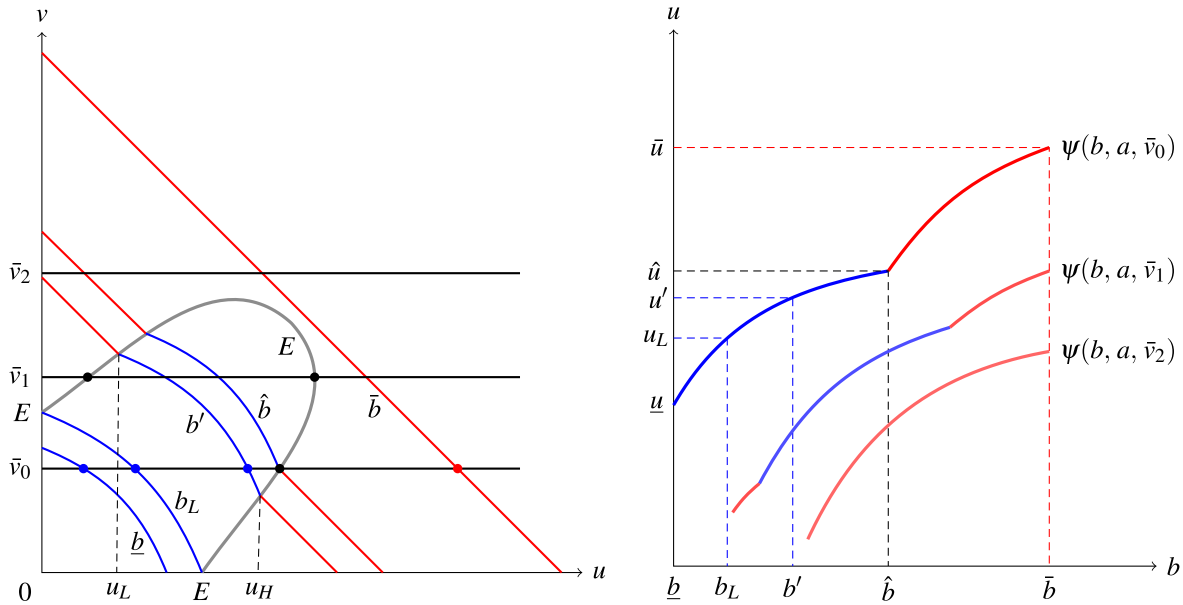


Figure 3: The left panel depicts the bargaining frontiers of a given type $a$ of $A$ firm by varying her partner's type $b$. The right panel depicts the corresponding indifference map of $a$ in her matched partner's type-utility space.

on $a$'s indifference curve at level $\bar{v}_0$ as depicted in the right panel of Figure 3. If the intersections of the horizontal line at $\bar{v}_0$ with a given bargaining frontier lies strictly inside the region enclosed by $EEE$, e.g. that with the frontiers labeled $\underline{b}$, $b_L$ and $b'$, then the enterprise chooses $N$. By contrast, if such intersection lies outside this region, then the enterprise prefers to integrate, e.g. enterprise $(a, \bar{b})$. Therefore,

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

where $\psi^d(b, a, \bar{v})$ denotes the indifference curve of $a$ at level $\bar{v}$ under organizational 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, then the horizontal line at $\bar{v}_0$ intersects the revenue expansion path at only one point, and the corresponding enterprise $(a, \hat{b})$ is indifferent between $N$ and $I$. Thus, the indifference curve at level $\bar{v}_0$ has a unique kink at $(\hat{b}, \hat{u})$ as shown in the left panel of Figure 3. In other words, $\hat{b}$ is the unique solution to $\psi^I(b, a, \bar{v}_0) = \psi^N(b, a, \bar{v}_0)$. Evidently, $\psi(b, a, \bar{v}_0) = \psi^N(b, a, \bar{v}_0)$ if and only if $b < \hat{b}$. By contrast, if the constant utility level of $a$ is high enough, i.e., $\bar{v}_1$, then the horizontal line at $\bar{v}_1$ intersects the type expansion path twice, and hence, the indifference curve at level $\bar{v}_1$ has two kinks. Note that $\bar{v}_1$ is not feasible for neither $\underline{b}$ nor $b_L$, and hence, this indifference curve starts at a $b > b_L$. Moreover, one kink lies to the left of $\hat{b}$ and the other kink to the right.[11] Finally, if the constant utility level of $a$ is even higher, e.g. $\bar{v}_2$, then it does not intersect the revenue expansion path at all, and hence, the corresponding indifference curve is smooth as $a$ chooses to integrate with all $B$ firms, i.e.,

[11]Under TU, each pair of bargaining frontiers of a given $a$ associated with two different organizations are parallel to each other, and hence, there is a unique $\hat{b}$ such that $(a, \hat{b})$ is indifferent between the two modes of organization for all levels $\bar{v}$, i.e., any indifference curve of each $a$ has a single kink which is invariant with respect to $\bar{v}$.

$\psi(b, a, \bar{v}_2) = \psi^I(b, a, \bar{v}_2)$. Note also that $\bar{v}_2$ is not feasible for a low-revenue enterprise, say $(a, b')$, and this indifference curve thus starts at a $b > b'$.[12]

### 5.2 A characterization of GID

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. Recall that GID [cf. (2)] is a necessary and sufficient condition for PAM, and hence, it suffices to prove that our economy satisfies GID. In order to show that the bargaining frontier of each enterprise satisfies GID, we first characterize GID in terms of *supermodularity* of the indifference curves of the $A$ units.

**Proposition 1** *The bargaining frontier $\phi(a, b, u)$ satisfies GID if and only if the indifference curve $\psi(b, a, \bar{v})$ is supermodular in $(b, a)$ for each $\bar{v}$, i.e., for any $a'' > a'$ and $b'' > b'$,*

$$\psi(b'', a'', \bar{v}'') - \psi(b', a'', \bar{v}'') \geq \psi(b'', a', \bar{v}') - \psi(b', a', \bar{v}'). \tag{11}$$

*for each constant utility levels $\bar{v}'$ of $a'$ and $\bar{v}''$ of $a''$.*

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, supermodularity of $\psi(b, a, \bar{v})$ 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. Therefore, GID is a *single-crossing condition*. The intuition is simple. In the left panel of Figure 4, $\psi(b, a'', \bar{v}'')$ is everywhere steeper than $\psi(b, a', \bar{v}')$, i.e., (11) 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 supermodularity 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.[13] It is worth noting that the above proposition extends the well-known result that "the 'Spence-Mirrlees condition', i.e., $-(\phi_2/\phi_3)$ is increasing in $a$ is a necessary and sufficient condition for GID" (see Chade, Eeckhout, and Smith, 2017) to a non-smooth environment.

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

$$\frac{3zz_b}{2z^2 - 2u - z\sqrt{z^2 - 2u}} \leq -\frac{z_{bb}}{z_b},$$

i.e., $z(a, b)$ is sufficiently concave in $b$.

[13] The right panel of Figure 4 depicts one of the many possible ways to violate supermodularity of $\psi(b, a, \bar{v})$. The only thing we require is that $\psi(b, a'', \bar{v}'')$ is flatter relative to $\psi(b, a', \bar{v}')$ at some $(b', u')$, i.e., (11) does not hold for the types $a'$ and $a''$ at this point. Then, we can always find a point on $\psi(b, a', \bar{v}')$, 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 $b'$, $\psi(b, a', \bar{v}')$ lies above $\psi(b, a'', \bar{v}'')$. Also, in Figure 4 the indifference curves are drawn with a single kink, but the result of Lemma 1 holds even if an indifference curve has more than one kinks.

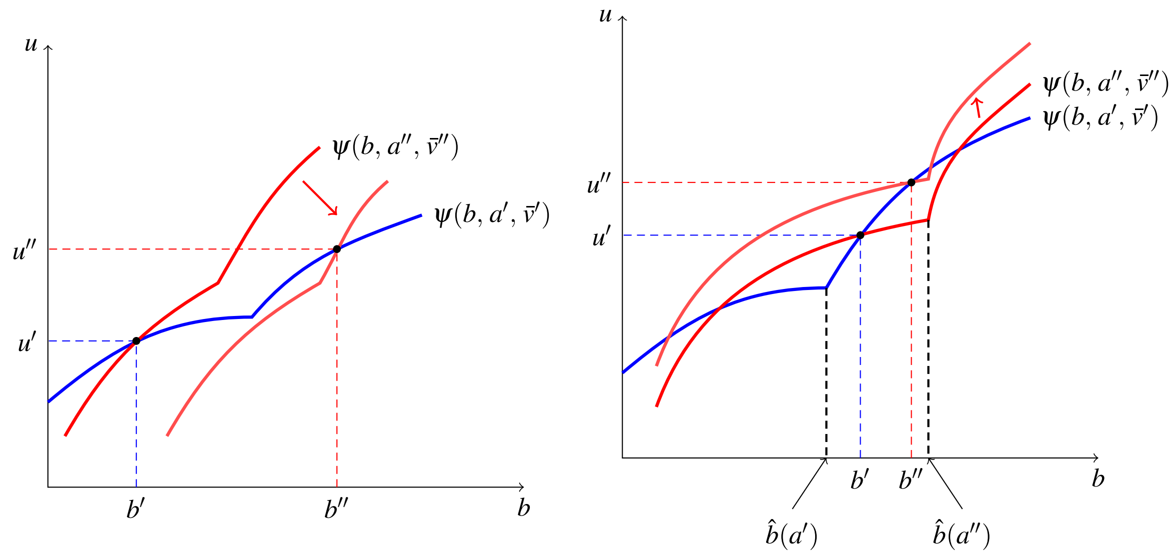


Figure 4: The left panel depicts that the indifference curves are supermodular in $(b, a)$, i.e., $\psi(b, a'')$ is everywhere steeper than $\psi(b, a')$ for $a'' > a'$, and hence, they cross only once. On the other hand, the right panel depicts a possible violation of supermodularity, which leads to that GID must also be violated.

### 5.3 *Sufficient conditions for GID*

Whether GID holds in any given model depends on the model specifications. If the bargaining frontier $\phi(a, b, u)$ satisfies GID for each partnership, then the equilibrium matching must be PAM. On the other hand, if the equilibrium matching is PAM, then it must be the case that GID is not violated for any matches. Having established Proposition 1, we now require to provide *sufficient conditions* under which an economy characterized by imperfectly transferable utility satisfies GID and hence, exhibits PAM. The difficulty is proving the supermodularity of the indifference curves $\psi(b, a, \bar{v})$ lies in the fact that, in our supplier economy, they are not differentiable everywhere. Thus, $\psi(b, a, \bar{v})$ may fail to satisfy supermodularity even if each of $\psi^N(b, a, \bar{v})$ and $\psi^I(b, a, \bar{v})$ is supermodular.

The right panel of Figure 4 depicts a possible violation of supermodularity of $\psi(b, a, \bar{v})$. Let the kink of a given indifference curve is defined by $\psi^N(\hat{b}, a, \bar{v}) = \psi^I(\hat{b}, a, \bar{v})$. Clearly, a sufficient condition for $\psi(b, a, \bar{v})$ to satisfy supermodularity is that the kink in the indifference curve of the $A$ agents moves to the left as $a$ increases, i.e., $d\hat{b}/da \leq 0$. If the indifference curves have multiple kinks at $\hat{b}_1, \ldots, \hat{b}_m$, then we require that all the kinks to move to the left as $a$ increases in order to obtain single-crossing of the indifference curves. Note that

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

Thus, if each segment of the indifference curve is steeper for higher $a$, and the each kink moves to the left as $a$ increases (as in the left panel of Figure 4), then there is no way that the indifference curves of two distinct types $a'$ and $a''$ can cross each other more than once.

**Theorem 1** *Let $\psi^d(b, a, \bar{v})$ is supermodular in $(a, b)$ for each $d = N, I$ and for each $\bar{v}$. Given the com-*

*bined indifference curve $\psi(b, a, \bar{v})$, if for all a we have*

$$sign[\psi_a^I(\hat{b}(a), a, \bar{v}) - \psi_a^N(\hat{b}(a), a, \bar{v})] = sign[\psi_b^I(\hat{b}(a), a, \bar{v}) - \psi_b^N(\hat{b}(a), a, \bar{v})], \tag{12}$$

*then $\psi(b, a, \bar{v})$ is supermodular in $(a, b)$ for each $\bar{v}$.*

The intuition for the result is simple. The term $\psi_a^I(\hat{b}, a, \bar{v}) - \psi_a^N(\hat{b}, a, \bar{v})$ is the contribution of $a$ to a marginal change in the match surplus [in a neighborhood of $\hat{b}$] when the match $(a, b)$ switches from $N$ to $I$, whereas the term $\psi_b^I(\hat{b}, a, \bar{v}) - \psi_b^N(\hat{b}, a, \bar{v})$ 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 $B$ units gain at the margin by switching to the same organizational mode as their productivity increases. When these marginals point in opposite directions, i.e., $\hat{b}$ is increasing in $a$, GID may fail to hold as depicted in the right panel of Figure 4.

Legros and Newman (2007) show that $\phi_{12}(a, b, u) \geq 0$ and $\phi_{13}(a, b, u) \geq 0$ are sufficient conditions under which $\phi(a, b, u)$ satisfies GID in a smooth economy. The first inequality refers to the role of complementarity between partners' types, whereas the second one indicates the complementarity between an agent's type and her matched partner's payoff. Under ITU, both sorts of complementarity play a crucial role in determining GID, and hence, PAM. However, verifying the type-type and the type-payoff complementarities is not an easy task when bargaining frontiers are not differentiable everywhere (as is the case for intermediate values of $z(a, b)$, i.e., on $[z_L, z_H]$). Theorem 1 thus offers a tractable and intuitive condition [cf. (12)] to verify GID in non-smooth economies.[14]

It is also worth mentioning that the condition that $\psi^d(b, a, \bar{v})$ is supermodular in $(a, b)$ or equivalently, $\phi^d(a, b, u)$ satisfies GID for each $d = N, I$ is crucial for the GID of the combined bargaining frontier. If this property does not hold under one of the two organizational modes, then $\phi(a, b, u)$ may fail to have GID. In other words, condition (12) alone is not sufficient to guarantee GID. Consider the following example:

**Example 3** Let the production function be given by:

$$z(a, b) = \max\left\{a + b - ab,\, a^2b^2\right\}.$$

Further let $a' = 0.65$, $a'' = 0.75$, $\bar{v}' = \bar{v}'' = 0.1$ and $b \in [0.5, 2]$. It is immediate to see that $\hat{b}(0.65) = 1.72$ and $\hat{b}(0.75) = 1.4$, i.e., condition (12) holds. However, the indifference curve $\psi(b, 0.75, 0.1)$ crosses the indifference curve of the lower type, $\psi(b, 0.65, 0.1)$ at $b = 0.81$ and $b = 1.43$, and hence, GID does not hold. The reason is that the production function $a + b - ab$ is submodular in $(a, b)$. As a consequence, the economy fails to have PAM.

In the above example, for tractability, we have considered an economy with perfectly transferable utility. One can construct examples under ITU that if the bargaining frontier associated with one of the two organizational modes does not satisfy GID, then the combined frontier may not satisfy GID even if $\hat{b}$ is decreasing in $a$.[15]

[14] In a smooth economy, either $\phi_{12}$ or $\phi_{13}$ may be negative. However, this does not preclude the Spence-Mirrlees condition to hold. In fact, in our input market, $\phi^N$ satisfies GID in spite of the fact that $\phi_{12}^N$ is not positive for the entire parameter space. Thus, if for a subset of parameter values either the type-type complementarity or type-payoff complementarity is violated, then verifying GID would be even harder for a non-smooth economy.

[15] An example of such situation would be when the probability of success function in our supplier market is given by $\pi(e_A, e_B) = 1 - (e_B - e_A)^2$ (as in Legros and Newman, 2013; Dam and Serfes, 2020). In this case $\phi^N(a, b, u)$ does not always satisfy GID, and hence, the equilibrium matching may not be PAM.

### 5.4 *Positive sorting and firm boundary decisions*

We now establish that the equilibrium of the input market with two-sided heterogeneity exhibits PAM, i.e., more productive $A$ units are matched with more productive $B$ units to form enterprises. Consequently, in our economy, enterprise heterogeneity becomes endogenous.

**Proposition 2** *The bargaining frontier $\phi(a,b,u)$ satisfies GID, and hence, in the equilibrium of the input market more productive A firms match with more productive B firms to form enterprises following a positive assortative matching pattern.*

In our economy, the bargaining frontier under both ownership structure satisfies GID and condition (12) holds, and hence, it follows from Theorem 1 that the equilibrium exhibits PAM. The derivative of $\hat{b}$ with respect to $a$ takes a very simple form:

$$\frac{d\hat{b}}{da} = -\frac{z_a(a,\hat{b})}{z_b(a,\hat{b})}.$$

Because the units produce the same output $z(a,b)$ under both organizational modes, the marginal contribution of each unit toward the enterprise revenue does not alter from one organization to the other. As a result, condition (12) holds. This intuition does not go through in Example 2 where in one organizational mode, i.e., $a^2b$ the $A$ firm contributes more than the $B$ firm at the margin, whereas under the production technology, $ab^2$ the ranking of $A$ and $B$ in terms of marginal contribution reverses. Consequently, PAM fails to hold. In contrast, in Example 1, although the production function under the two organizations are different, this ranking remains unaltered. In fact, both organizations treat the two units symmetrically. This motivates the following result.

**Proposition 3** *Let $z^I(a,b) = z(a,b)$ and $z^N(a,b) = h(z(a,b))$ where $h : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function. Then, the bargaining frontier $\phi(a,b,u)$ satisfies GID, and hence, the equilibrium of the economy exhibits PAM.*

The above result asserts that Proposition 2 continues to hold if the production function under one organizational structure is obtained by a monotonic transformation of that in the other structure. Clearly, a monotonic transformation preserves the ranking of the two units in terms of their marginal contributions toward the enterprise revenue (as is the case with Example 1). Such assumption can be rationalized as vertical or lateral integration may often imply a loss of the enterprise output due to the presence of diverse transaction costs (e.g. Grossman and Helpman, 2002) or the unit managers may have to give up a fraction of output produced to $HQ$ in the pursuit of a more favorable action.

Our principal focus in the present paper has been to show that the equilibrium of the input market exhibits positive assortative matching, and the enterprise heterogeneity comes to be endogenous—the highest-productivity matched units form the highest-revenue enterprise, the second highest-productivity units constitute the second highest-revenue enterprise, and so on. It is easy to show that, with respect to the firm boundary decisions, the input market is segregated in the sense that high-revenue enterprises choose to integrate, while the low-revenue ones stay separate with enterprises with a unique productivity combination $(a^*, b^*)$ being indifferent between the two organizational modes. This result can be easily derived from Dam and Serfes (2020, Proposition 1) because the equilibrium distribution of integrated enterprises depends only on the enterprise output.[16]

[16]In Dam and Serfes (2020), we also show that the equilibrium choice of organization may be non-monotonic in that the

# 6 Sorting in the shareholder-manager market

In this section, we analyze another application of Theorem 1. We consider the matching market described in Alonso-Paulí and Pérez-Castrillo (2012) who analyze partnerships between shareholders (say, $A$ agents) and managers (say, $B$ agents). The stochastic production function of an arbitrary partnership $(a, b)$ is given by:

$$z = ahe,$$

where $e$ is the managerial effort and $h$ is a random variable with mean $\mu$ and variance $\sigma^2$. The effort cost function of a type $b$ manager is given by:

$$C(e, b) = \frac{e^2}{2b}, \quad b > 0.$$

Thus, higher values of $a$ represent more 'productive' shareholders, and higher values of $b$ imply more 'talented' managers. Each matched pair chooses between two contractual structures—namely, an *incentive scheme*, denoted by *IS*, and a *code of best practice* contract, denoted by *CBP*. The Pareto frontier of each arbitrary partnership $(a, b)$ is given by $\phi(a, b, u) = \max\{\phi^{IS}(a, b, u), \phi^{CBP}(a, b, u)\}$, where

$$\phi^{CBP}(a, b, u) = \frac{1}{2}\mu^2 a^2 b - u,$$
$$\text{and} \quad \phi^{IS}(a, b, u) = \min\left\{\frac{1}{2}(\mu^2+\sigma^2)a^2 b,\, a\sqrt{2(\mu^2+\sigma^2)bu} - 2u\right\}.$$

We first derive the indifference map of a given type $a$ of shareholder (see Figure 5). We assume that

$$\frac{\mu}{\sigma} \geq \sqrt{3}.$$

Notice that the first segment of the frontier $\phi^{IS}(a, b, u)$ is horizontal, whereas the second segment is strictly decreasing in $u$. The above assumption guarantees that $\phi^{CBP}(a, b, u)$ intersects $\phi^{IS}(a, b, u)$ in its downward-sloping portion.[17]

In the left panel of Figure 5 we depict the combined Pareto frontiers of a given type $a$ shareholder. The three frontiers correspond to the shareholder-manager pairs $(a, b')$, $(a, \hat{b})$ and $(a, b'')$ with $b' < \hat{b} < b''$. The higher the managerial type, the higher is the frontier. Consider an arbitrary match $(a, b)$. The utility allocation $(u^*, v^*)$ is the one at which this pair is indifferent between signing a *CBP* contract and an *IS* contract. This indifference allocation is given by:

$$\phi^{CBP}(a, b, u^*) = \phi^{IS}(a, b, u^*) \quad \Longleftrightarrow \quad u^* = \frac{1}{2}\left(\sqrt{\mu^2+\sigma^2} - \sqrt{\sigma^2}\right)^2 a^2 b,$$
$$v^* = \frac{1}{2}\left[\mu^2 - \left(\sqrt{\mu^2+\sigma^2} - \sqrt{\sigma^2}\right)^2\right] a^2 b.$$

The line labeled *EE* is the locus of all such utility allocations $(u^*, v^*)$ by varying the managerial type and keeping the shareholder type fixed, which is the *type expansion path*. It is a straight line because

low- and high-revenue enterprises choose to integrate, whereas medium-revenue enterprises choose to stay separate under an alternative set of assumptions.

[17] When the Pareto frontier under the *IS* contract is horizontal, the indifference curves are not well-defined because $\phi^{IS}(\cdot, \cdot, u)$ is not invertible. Intuitively, given a horizontal frontier, any $a$ strictly prefers $b'$ to $b$ if $b' > b$ irrespective of the values of $u$ and $u'$. In other words, no $a$ is indifferent between any two distinct $b$ and $b'$.

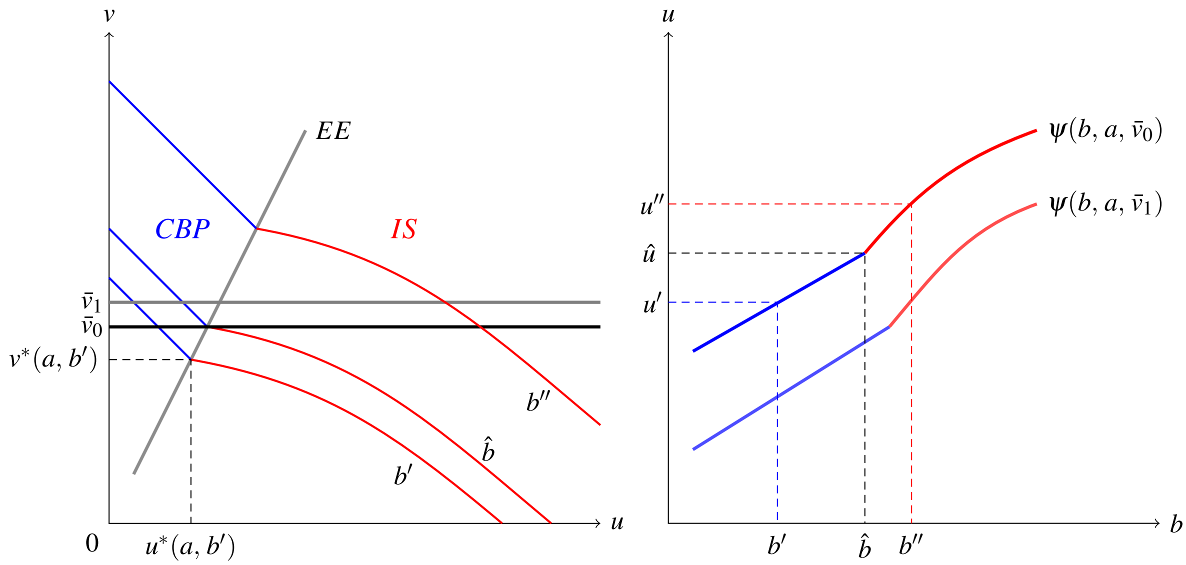


Figure 5: The left panel depicts the bargaining frontiers of a given type $a$ of shareholders by varying her partner's type $b$ when $\mu/\sigma \geq \sqrt{3}$. The right panel depicts the corresponding indifference curves in the type-utility space.

$v^*/u^*$ is independent of $a$ and $b$, and it divides the utility space into two regions—(i) above $EE$ (the region labeled *CBP*), each type $a$ shareholder prefers to offer a *CBP* contract to the managers, and (ii) below $EE$ (the region labeled *IS*), an *IS* contract is the preferred choice.[18] Now, fix a utility [of type $a$] level $\bar{v}_0$. This shareholder prefers to sign a *CBP* contract with a manager of type $b'$, but prefers to offer an *IS* contract to $b''$. However, she is indifferent between the two contractual modes if she is matched with $\hat{b}$. Because all these three matching-contract combinations yield the same utility level $\bar{v}_0$ for $a$, the three managerial type-utility combinations, $(b', u')$, $(\hat{b}, \hat{u})$ and $(b'', u'')$ lie on the same indifference curve $\psi(b, a, \bar{v}_0)$ of $a$, which is drawn in the right panel. Each indifference curve has a unique kink at $\hat{b}$ representing indifference between the two contractual structures.[19] Higher indifference curves, e.g. $\psi(b, a, \bar{v}_1)$, lies to the south-east. Because the $EE$ is strictly increasing, the kink $\hat{b}$ moves to right as one moves to a higher indifference curve of the same type $a$ shareholder.

In the following proposition, we show that the shareholder-manager economy satisfies GID, and hence, the equilibrium matching is PAM.

**Proposition 4** *Assume that $\mu/\sigma \geq \sqrt{3}$. The bargaining frontier $\phi(a, b, u)$ of the shareholder-manager matching market satisfies GID, and hence, in equilibrium more talented managers are lured into more productive fimrs following a positive assortative matching pattern.*

[18]A *CBP* contract is characterized by the shareholders owning a monitoring technology that obliges a manager to choose a stipulated effort level, whereas monitoring is not feasible under the *IS* contract. Thus, each firm prefers to offer a *CBP* contract to managers who entail greater marginal effort cost, i.e., for whom the moral hazard problem is more severe.

[19]The segment of the indifference curve corresponding to *CBP* is linear because

$$\psi_b(b, a, \bar{v}) = -\frac{\phi_1^{CBP}(a, b, u)}{\phi_3^{CBP}(a, b, u)} = \frac{\mu^2 a^2}{2}$$

is independent of $b$. On the other hand, the segment associated with *IS* is non-linear.

The shareholder-manager economy satisfies the sufficient conditions for GID in Theorem 1, and hence, the above proposition follows. The assumption that $\mu/\sigma \geq \sqrt{3}$ is crucial for the above result to hold. Alonso-Paulí and Pérez-Castrillo (2012) show that there may be equilibria without PAM when $\phi^{CBP}$ intersects the horizontal segment of $\phi^{IS}$ for $\mu/\sigma \in (1, \sqrt{3})$. A possible explanation of this is that GID is not well-defined along the horizontal part of the frontier, for each $a$ it is always the case that $\phi(a, b', u') > \phi(a, b, u)$ for any $b' > b$, and any $u$ and $u'$.

## 7 Concluding remarks

The present paper contributes to the literature on heterogeneous matching markets wherein positive assortative matching is established under imperfectly transferable utility when 'easy' differential methods are not available. In particular, we provide an alternative characterization of the generalized increasing difference condition analyzed by Legros and Newman (2007), which is a single-crossing condition, in terms of the *indifference curves* of each the market participants on one side. Furthermore, we provide tractable sufficient conditions under which GID holds for non-smooth economies, and analyze two distinct matching markets (with non-differentiability) where our characterization result can be applied.

Our results are also useful in characterizing equilibrium outcomes in economies where bargaining frontiers are not concave. Certainly, there are markets where non-convex utility possibility sets can be convexified using lotteries. However, there are situations where such randomization does not make much sense. We have analyzed two such economies where discrete decisions give rise to non-convexity.

Our characterization result, Proposition 1 has been established by taking the bargaining frontiers as the primitives (from which the indifference curves are derived). In many other contexts, in which monotone optimal solutions are sought (see Milgrom and Shannon, 1994), the Pareto frontiers are not taken as a fundamental; however, indifference curves are well-defined, which are derived from type-dependent utility functions. Our result can be applied to such contexts where easy differential methods are not available. One such context is the *monopolistic screening problems* (e.g. Maskin and Riley, 1984). Let $\phi(\theta, q, t)$ be the utility of the buyer of a monopolist's product which depends on his type $\theta$, the quantity consumed $q$ and the transfer to the seller $t$. As long as $\phi$ is increasing in the first two arguments, and strictly decreasing in the third argument, Theorem 1 provides a set of easily verifiable conditions under which optimal screening involves quantity $q$ being increasing in $\theta$ if $\phi(\theta, q, t)$ gives rise to non-smooth indifference curves of each type $\theta$ in the $q$-$t$ space. Another context is *sender-receiver games* with a continuum of types (e.g. Mailath, 1987; Sobel, 2009). Let $\phi(\theta, a, s)$ be the sender's utility which is increasing in his type $\theta$ and receiver's action $a$, and strictly decreasing in the costly signal $s$. Existence of separating equilibria typically requires that higher sender types send costlier signals. If the sufficient conditions in Theorem 1 hold, then a separating equilibrium exists, which is guaranteed by the single-crossing of the indifference curves even if they are not differentiable in the action-signal space.

## Appendix

*Proof of Lemma 1.* Substituting for $e_A$ and $e_B$ from the incentive compatibility constraints (7) and (8), the optimal contracting problem of an arbitrary enterprise $(a, b)$ reduces to:

$$\max_{s\in[0,1]} \; V_A(s;z) \equiv \frac{z^2}{2}(1-s^2), \tag{A1}$$

$$\text{subject to } \; U_B(s;z) \equiv \frac{z^2}{2}s(2-s) = u. \tag{A2}$$

From (A2) it follows that

$$s(a,b,u) = 1 - \frac{\sqrt{z(a,b)^2-2u}}{z(a,b)}.$$

We ignore the other root as it is strictly larger than 1. Using the expression in (A1), the bargaining frontier under $N$ is given by:

$$\phi^N(a,b,u) = \frac{1}{2}\left[2z(a,b)\sqrt{z(a,b)^2-2u} - \left(z(a,b)^2-2u\right)\right]. \tag{A3}$$

Now 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.,

$$\frac{z^2}{2} > z - \frac{1}{4} \quad \Longleftrightarrow \quad \left[z - \frac{1}{2}\left(2-\sqrt{2}\right)\right]\left[z - \frac{1}{2}\left(2+\sqrt{2}\right)\right] > 0$$

Because $z < 1$, the above holds for $z < \frac{1}{2}\left(2-\sqrt{2}\right) \approx 0.29 \equiv z_L$. 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.,

$$z - \frac{1}{4} > \frac{3z^2}{4} \quad \Longleftrightarrow \quad \left(z-\frac{1}{3}\right)(z-1) < 0.$$

The above holds for $z > \frac{1}{3} \approx 0.33 \equiv z_H$. Now consider the case when $z_L \leq z \leq z_H$. Note first that $\phi^N(a,b,u)$ intersects the linear function $\phi^I(a,b,u)$ 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:

$$u_L(a,b) = \frac{1}{8}\left(4z(a,b) - 1 - 2z(a,b)\sqrt{(1-z(a,b))(1-3z(a,b))}\right),$$
$$u_H(a,b) = \frac{1}{8}\left(4z(a,b) - 1 + 2z(a,b)\sqrt{(1-z(a,b))(1-3z(a,b))}\right).$$

Note that $u_H(a,b) = z(a,b) - \frac{1}{4} - u_L(a,b)$. Thus, for the existence of the two intersection points it suffices to show that $u_L(a,b) \geq 0$, which holds for $z_L \leq z(a,b) \leq z_H$. Clearly, $\phi^N(a,b,u) > \phi^I(a,b,u)$ is and only is $u \in (u_L(a,b), u_H(a,b))$.

*Proof of Proposition 1.* We first prove the sufficiency of (11). Take any $a'' > a'$, $b'' > b'$ and $u'' > u'$, and let $\phi(a',b',u') = \phi(a',b'',u'') = \bar{v}'$. Thus, $u' = \psi(b',a',\bar{v}')$ and $u'' = \psi(b'',a',\bar{v}')$, i.e., $(b',u')$ and $(b'',u'')$ are on the same indifference curve of $a'$ that yields $\bar{v}'$. 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., $u' = \psi(b', a'', \bar{v}'')$ and $\hat{u} = \psi(b'', a'', \bar{v}'')$. Then, it follows from (11) 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 utilities of type $b$ are arbitrarily chosen, the last inequality implies that $\phi(a, b, u)$ satisfies GID.

Next, we show the necessity of (11). Given two arbitrary types $a'$ and $a''$ with $a' < a''$, suppose that (11) is violated for these two types at $b = b'$, i.e., $\psi(b, a'', \bar{v}'')$ crosses $\psi(b, a', \bar{v}')$ from above at $b = b'$ for constant levels of utility $\bar{v}'$ of $a'$ and $\bar{v}''$ of $a''$. Let $u' = \psi(b', a', \bar{v}') = \psi(b', a'', \bar{v}'')$, i.e., the indifference curve of $a'$ at level $\bar{v}'$ and that of $a''$ at level $\bar{v}''$ cross each other at $(b', u')$. By continuity of $\psi(b, a, \bar{v})$, there is a $\bar{b} > b'$ such that $\psi(b, a', \bar{v}') > \psi(b, a'', \bar{v}'')$ 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 $\psi(b, a', \bar{v}')$ is strictly increasing in $b$, we have $u'' > u'$. On the other hand, $\hat{u} = \psi(b'', a'', \bar{v}'') > \psi(b'', a', \bar{v}_1'') = u''$ for $\bar{v}_1'' > \bar{v}''$ as $b'' \in (b', \bar{b}]$. Thus, $\bar{v}_1'' = \phi(a'', b'', u'') \leq \phi(a'', b'', \hat{u}) = phi(a'', b', u') = \bar{v}''$ because $\phi(a, b, u)$ is strictly decreasing in $u$. Hence, GID is violated for these two given types $a'$ and $a''$ at $(b'', u'')$.

*Proof of Proposition 2.* We first show that both $\psi^N(b, a, \bar{v})$ and $\psi^I(b, a, \bar{v})$ are supermodular in $(b, a)$. Because both functions are differentiable, for each $d = I, N$, supermodularity of $\psi^d(b, a, \bar{v})$ is equivalent to $\psi^d_{ab}(b, a, \bar{v}) \geq 0$. Recall that $\phi^d(a, b, \psi^d(b, a, \bar{v})) = \bar{v}$. Implicitly differentiating the last equality we obtain

$$\psi^d_b(b, a, \bar{v}) = -\frac{\phi^d_2(a, b, u)}{\phi^d_3(a, b, u)}.$$

First, consider $d = N$. Differentiating (9) with respect to $a$, $b$ and $u$ we get

$$\phi^N_1(a, b, u) = z_a(a, b)\left(\frac{z(a, b)^2}{\sqrt{z(a, b)^2 - 2u}} - z(a, b) + \sqrt{z(a, b)^2 - 2u}\right),$$

$$\phi^N_2(a, b, u) = z_b(a, b)\left(\frac{z(a, b)^2}{\sqrt{z(a, b)^2 - 2u}} - z(a, b) + \sqrt{z(a, b)^2 - 2u}\right),$$

$$\phi^N_3(a, b, u) = -\left(\frac{z(a, b)}{\sqrt{z(a, b)^2 - 2u}} - 1\right).$$

Thus,

$$\psi^N_b(b, a, \bar{v}) = -\frac{\phi^N_2(a, b, u)}{\phi^N_3(a, b, u)} = z_b(a, b) \cdot \underbrace{\left(z(a, b) + \frac{z(a, b)^2 - 2u}{z(a, b) - \sqrt{z(a, b)^2 - 2u}}\right)}_{H(z(a,b),u)}.$$

Note that

$$\frac{\partial H}{\partial z}(z, u) = \frac{3z}{z - \sqrt{z^2 - 2u}} > 0$$

because $u \leq z^2/2$. Therefore,

$$\psi^N_{ab}(b, a, \bar{v}) = \frac{\partial H}{\partial z}(z(a, b), u) z_a(a, b) z_b(a, b) + H(z(a, b), u) z_{ab}(a, b) > 0$$

because $z_a, z_b > 0$ and $z_{ab} \geq 0$. Thus, $\psi^N(b, a, \bar{v})$ is supermodular in $(b, a)$. Next, consider $d = I$. Differentiating (10) with respect to $a$, $b$ and $u$ we get

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

Therefore, $\psi^I_b(b, a, \bar{v}) = z_b(a, b)$, and $\psi^I_{ab}(b, a, \bar{v}) = z_{ab}(a, b) \geq 0$. Therefore, $\psi^I(b, a, \bar{v})$ is supermodular in $(b, a)$. Thus, the combined indifference curves satisfy the first set of sufficient conditions in Theorem 1.

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

$$\psi^N_a(b, a, \bar{v}) = -\frac{\phi^N_1}{\phi^N_3} = H(z(a, b), u)z_a(a, b),$$
$$\psi^I_a(b, a, \bar{v}) = -\frac{\phi^I_1}{\phi^I_3} = z_a(a, b).$$

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

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

Thus, condition (12) holds, and hence, the indifference curve $\psi(b, a, \bar{v})$ is supermodular in $(b, a)$, 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 Proposition 3.* Note first that the expressions for $\psi^I_b(b, a, \bar{v})$ and $\psi^I_{ab}(b, a, \bar{v})$ remain the same in the proof of Proposition 2. It follows from the expression of $\psi^N_{ab}$ in the proof of Proposition 2 that $\psi^N_{ab}(b, a, \bar{v}) \geq 0$ if $h'(\cdot) > 0$. Next, it is easy to see that

$$\psi^N_x(a, b, \bar{v}) = h' \cdot z_x(a, b)H(h(z(a, b), u) \quad \text{for } x = a, b.$$

Thus,

$$\frac{d\hat{b}}{da} = -\frac{\psi^I_a(\hat{b}, a, \bar{v}) - \psi^N_a(\hat{b}, a, \bar{v})}{\psi^I_b(\hat{b}, a, \bar{v}) - \psi^N_b(\hat{b}, a, \bar{v})} = -\frac{z_a(a, \hat{b})[1 - h'H(z(a, \hat{b}), u)]}{z_b(a, \hat{b})[1 - h'H(z(a, \hat{b}), u)]} = -\frac{z_a(a, \hat{b})}{z_b(a, \hat{b})} < 0.$$

The result thus follows from Theorem 1.

*Proof of Proposition 4.* The proof is very similar to that pf Proposition 2, and hence, we omit many details. It is easy to derive that

$$\psi^{CBP}_{ab}(b, a, \bar{v}) = \frac{\partial}{\partial a}\left(-\frac{\phi^{CBP}_2(a, b, u)}{\phi^{CBP}_3(a, b, u)}\right) = \mu^2 a \geq 0,$$
$$\psi^{IS}_{ab}(b, a, \bar{v}) = \frac{\partial}{\partial a}\left(-\frac{\phi^{IS}_2(a, b, u)}{\phi^{IS}_3(a, b, u)}\right) = \frac{4(\mu^2 + \sigma^2)u\sqrt{2(\mu^2 + \sigma^2)bu}}{\left(\sqrt{2}(\mu^2 + \sigma^2)ab - 4\sqrt{(\mu^2 + \sigma^2)bu}\right)^2} \geq 0.$$

Thus, each contractual structure has GID. Next, condition (12) holds because

$$\frac{d\hat{b}}{da} = -\frac{2\hat{b}}{a} < 0.$$

Thus, the result follows from Theorem 1.

## References


Abowd, John, John Haltiwanger, Ron Jarmin, Julia Lane, Paul Lengermann, Kristin McCue, Kevin McKinney, and Kristin Sandusky (2005), "The Relation among Human Capital, Productivity, and Market Value: Building Up from Micro Evidence." In *Measuring Capital in the New Economy* (Carol Corrado, John Haltiwanger, and Daniel Sichel, eds.), 153–198, University of Chicago Press.

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.

Benguria, Felipe (2015), "The Matching and Sorting of Exporting and Importing Firms: Theory and Evidence." Mimeo, University of Kentucky.

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

Bloom, Nicholas and John Van Reenen (2010), "Why Do Management Practices Differ across Firms and Countries?" *Journal of Economic Perspectives*, 24, 203–224.

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.

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

Dam, Kaniṣka and Konstantinos Serfes (2020), "A Price Theory of Vertical and Lateral Integration under Productivity Heterogeneity." Forthcoming, *The Journal of Industrial Economics*.

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

Eaton, Jonathan and Samuel Kortum (2002), "Technology, Geography, and Trade." *Econometrica*, 70, 1741–1779.

Eekhout, Jan and Philipp Kircher (2018), "Assortative matching with large firms." *Econometrica*, 86, 85–132.

Forbes, Silke and Mara Lederman (2011), "Does Vertical Integration Affect Firm Performance? Evidence from the Airline Industry." *RAND Journal of Economics*, 41, 765–790.

Griffith, Rachel, Rupert Harrison, and John Van Reenen (2006), "How Special is the Special Relationship? Using the Impact of U.S. R&D Spillovers on U.K. Firms as a Test of Technology Sourcing." *The American Economic Review*, 96, 1859–1875.

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

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

Jorgenson, Dale, Mun Ho, and Kevin Stiroh (2008), “A Retrospective Look at the U.S. Productivity Growth Resurgence.” *Journal of Economic Perspectives*, 22, 3–24.

Knittel, Christopher (2002), “Alternative Regulatory Methods and Firm Efficiency: Stochastic Frontier Evidence from the U.S. Electricity Industry.” *The Review of Economics and Statistics*, 84, 530–540.

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.

Mailath, George (1987), “Incentive Compatibility in Signaling Games with a Continuum of Types.” *Econometrica*, 55, 1349–1365.

Maskin, Eric and John Riley (1984), “Monopoly with Incomplete Information.” *The RAND Journal of Economics*, 15, 171–196.

Melitz, Marc (2003), “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, 71, 1695–1725.

Milgrom, Paul and Chris Shannon (1994), “Monotone Comparative Statics.” *Econometrica*, 62, 157–180.

Moretti, Enrico (2004), “Workers’ Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions.” *The American Economic Review*, 94, 656–690.

Sattinger, Michael (1993), “Assignment Models and the Distribution of Earnings.” *Journal of Economic Literature*, 31, 831–880.

Sobel, Joel (2009), “Signaling Games.” In *Encyclopedia of Complexity and Systems Science* (Robert Meyers, ed.), 8125–8139, Springer New York.

Sugita, Yoichi, Kensuke Teshima, and Enrique Seira (2020), “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.

Topkis, Donald (1998), *Supermodularity and Complementarity*. Princeton University Press.

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.