A Two-sided Matching Approach for Partner Selection and Assessing Complementarities in Inter-firm Alliances

Denisa Mindruta
HEC, Paris
mindruta@hec.fr

Mahka Moeen
University of South Carolina
mahka.moeen@moore.sc.edu

Rajshree Agarwal*
University of Maryland
rajshree@umd.edu

Submitted to SMJ Special Issue: Question-Focused Innovations in Research Methods

This version: September 7, 2013

ABSTRACT

Partner selection in strategic alliances is an enduring and core question within strategic management. Binary choice models are the dominant empirical methodology used to test rich theoretical models of partner choice. We discuss how recent econometric advances in two-sided matching models may overcome inference limitations of the binary choice models, particularly in the dyadic context of assessing complementarities among partner attributes. We use simulations to assess the relative performance of both methodologies when estimating parameters generated by a known functional relationship, and demonstrate the method's applicability within the context of research alliances in the bio-pharmaceutical industry.

* Corresponding author. All authors contributed equally and the names are listed in random order. We would like to thank Martin Goossen for data sharing, and Heejung Byun and Justin Frake for research assistance.

Strategic alliances are key to creating value by combining resources and capabilities that reside across firm boundaries. For firms seeking to engage in strategic alliances, partner selection is a critical consideration, since value creation is contingent on whether prospective partners represent synergies in the relevant attributes (Alcacer, Cantwell & Gittelman, 2009; Ahuja, Polidoro & Mitchell, 2009; Mitsuhashi & Greve, 2009). Accordingly, a rich literature stream in strategic management has focused on factors—firm or dyad specific—that influence the partner choice decision, using both the resource based view and organizational economics lens (Dushnitsky & Shaver, 2009; Katila, Rosenberger & Eisenhardt, 2008; Rothaermel & Boeker, 2008). Since the twin considerations of maximizing value creation and minimizing contractual hazards when engaging in strategic alliances are interrelated in practice, scholars have also examined how signals based on prior relationships and network positions may inform partner choice decisions (Gimeno, 2004; Gulati, 1995; Gulati & Garguilo, 1999; Stuart, 1998). While our knowledge of the theoretical drivers of partner choice and complementarities across attributes has blossomed, the dominant empirical methodology—binary choice models—used to test the theoretical relationships have some key limitations, especially when addressing questions related to complementarities of partner attributes in the dyadic perspective of "who collaborates with whom."

Two-sided matching models, which have a long theoretical history in economics (Becker, 1973; Roth & Sotomayor, 1990), are particularly relevant to examining complementarities in partner characteristics, as critical to partner selection (Greve, Mitsuhashi & Baum, 2013; Mortensen, 1988). Importantly, recent empirical advances in matching methodologies (Fox, 2010a) enable scholars to directly assess complementarities and substitutions among partner characteristics, rather than relying on complicated and cumbersome, if not erroneous, interpretations of interaction effects in binary choice models (Hoetker, 2007; Zelner, 2009). While the methodology is taking off in related disciplines of finance and economics (Akkus, Cookson, & Hortacsu, 2012; Baccara et al., 2011; Levine, 2009; Pan, 2010; Yang, Shi & Goldfarb, 2009), its usage in strategic management is still in inception stages (Chatain, 2013; Mindruta, 2013).

We provide the theoretical rationale for why two-sided matching models represent a significant advancement in the study of partner selection and complementarities across partner attributes; briefly these relate to accommodating (a) both partners' preferences, and (b) competition in each side for the "best partner." We then discuss the use of maximum score estimator (Fox, 2010a) for two-sided matching models to empirically test theoretical propositions. Next, we articulate the differences in matching methodology and the traditional logit regression, and use simulations to assess their relative performance when estimating parameters in a known functional relationship. We find that while logit models have high predictive power in whether an alliance is formed, the signs and magnitude of the coefficients are very different from the true parameter values. The maximum score matching model estimator, in contrast shows a high degree of accuracy in estimating the extent of complementarities across partner attributes. Thus, while binary choice specifications are appropriate for predicting outcomes (i.e. whether two firms will select each other as partners), matching models are appropriate for uncovering the underlying relationships among partner attributes (i.e. whether complementarities among partners exist). We also demonstrate the method's applicability within the context of the bio-pharmaceutical industry, where we find evidence for complementarities in technological capabilities, and substitution effects in firm age.

In describing how the theoretical lens and the empirical estimation approach of matching models can be used to examine critical questions related to complementarities and substitutions in partner attributes, we contribute to the strategic management literature by creating greater awareness of recent advances in the related discipline of economics. While we have focused our attention to the question of partner selection and assessment of complementarities across alliance partners, two-sided matching models and their estimation are relevant to a wider range of questions that are of interest to strategic management scholars. Indeed, to the extent that strategy scholars focus on markets—whether factor, intermediate, or final product—that are *not* characterized by assumptions that underpin perfectly competitive market conditions, matching models may be applied to a whole host of questions, including but not limited to selection in mergers and acquisitions, buyer-supplier relationships, investor-investee relationships, matching of human capital and firms, and TMT and board composition of firms.

Importantly, we hope that our work helps strategy scholars improve answers, based on reliable research design that mirrors the assumptions that underpin questions that they seek to address. Recent efforts at identification of causal relationships have stressed the importance of addressing selection and sorting effects when inferring about performance. By using matching models, scholars can distinguish the role of the distribution of exogenous partner characteristics from the role of the alliance as a "match production function" in the sorting patterns observed in the data (Fox, 2010a). Thus, by assessing complementarities and substitutions of partner attributes, matching models permit better identification of the value creation through strategic alliances.

LITERATURE REVIEW: UNDERLYING THEORETICAL ASSUMPTIONS AND RESULTANT EMPIRICAL APPROACHES

Strategic alliances are voluntary cooperative relationships formed between two or more independent organizations for value creation through access to capabilities and inter-firm knowledge transfer (Ahuja, 2000; Hamel, 1991; Mowery, Oxley & Silverman, 1996; Rosenkopf & Almeida, 2003). Firms engage in alliances when confronted with a lack of fit between existing resources and capabilities and those required by the changing competitive landscape or their strategic goals. As drivers of alliance success, existing research has focused on two broad areas, ex ante partner selection and alliance formation, and ex post management of the alliance process (Alcacer et al., 2009; Ahuja et al., 2009; Khanna, Gulati & Nohria, 1998; Mitsuhashi & Greve, 2009; Oxley, 1997).

Overview of Work in Partner Selection in Strategic Alliances

Our focus is on the ex-ante partner selection issue, as it relates to capability matching and value creation. The question of partner choice relates to factors that impact how business actors pick each other from a pool of potential partners to form mutually acceptable, high-value relationships. Since strategic alliances represent an inherent tension between competition and cooperation (Hamel, 1991; Khanna et al., 1998; Agarwal, Croson & Mahoney, 2010), partner choice requires attention to both access to requisite capabilities, and attention to potential opportunism. Rich literature streams have extensively examined the twin considerations of maximizing value creation and minimizing contractual hazards (or addressing risks of value appropriation). Although interrelated in practice, these two aspects of alliances are studied by employing different theoretical lenses—a resource based view or organizational learning perspective (e.g. Rothaermel & Boeker, 2008; Sampson, 2007), and a game theoretic or transactions cost perspective (Khanna et al., 1998; Katila et al., 2008; Oxley, 1997). We review the existing literature of partner selection in strategic alliances (Table 1)[1], grouping them within the two dominant frames: a focal firm perspective, and a dyadic perspective.

[Table 1 about here]

Single Firm Perspective: Papers that adopt a single firm perspective seek to answer the critical question: "who does the focal firm collaborate with?" Thus, the question of partner selection is driven by the perspective of a focal firm that is seeking out alliance partners to overcome the lack of necessary resources and capabilities. Within the single firm perspective, studies represented in Panel A of Table 1 abstract away from considerations of fit between partners. While they may characterize potential partners as "friends," "acquaintances" and "strangers" based on whether they have prior alliance experience (Li et al., 2008), the focus of such studies is on the "main effects" of either focal or partner firm characteristics on partner selection. For example, studies have focused on how firms choose partners based on market- or firm-specific uncertainty (Beckman, Haunschild & Phillips, 2004), by balancing the need for resource access and enhanced appropriation concerns (Katila et al., 2008), or by assessing the potential partner's status and reputation (Stern, Dukerich & Zajac, 2013). Alternatively, scholars have examined inter-country differences in executives' focus on various types of partner capabilities (Hitt et al., 2000; Hitt et al., 2004).

[1] The literature review is by no means exhaustive, given hundreds of management articles on the subject. We chose representative articles that are discussed in this section based in the following manner: We first identified the five articles that received the highest Google Scholar citations for alliance formation. We then focused on those articles that examined factors explaining partner selection, rather than general propensity to form alliances. Thus, while Eisenhardt & Schoonhoven (1996) is widely cited, it is not included in Table 1 because its focus is on whether or not to form an alliance, rather than whom to partner with. We then expanded the list based on forward and backward citations of these articles, and included articles that represented not only high citations counts, but also diversity in terms of research questions (e.g. focal firm vs. dyadic perspective), empirical contexts, and econometric specifications. Finally, to ensure that our search included recent articles in addition to highly cited ones, we conducted a search in the *Strategic Management Journal* for all articles examining partner selection in the 2000-2013 period.

Studies represented in Panel B, while taking a focal firm perspective, additionally incorporate fit considerations. Studies examining the effects on partner selection of dyadic variables include a focus on geographic co-location (Alcacer et al., 2009; Gittelman, 2013), technological and market overlap (Diestre & Rajagopalan, 2012; Dushnitsky & Shaver, 2009; Hallen, 2008; Sorenson & Stuart, 2001), prior affiliation (Podolny, 1994) and similarity in status and reputation (Hallen, 2008; Podolny, 1994) Importantly, a focus on fit has enabled a few studies to include inferences about complementarities of partner attributes. Accordingly, scholars have hypothesized about complementarities in partners regarding the markets they stem from (Dushnitsky & Shaver, 2009; Li & Rowley, 2002) and compatibility or similarity among partners in status and network positions (Hallen, 2008; Podolny, 1994; Sorenson & Stuart, 2001).

Dyadic Perspective: Papers that employ a dyadic perspective (Panel C of Table 1) answer the critical question of "who collaborates with whom?" These studies note that partner selection is the function of mutual preferences of a dyad, and look at alliance formation as a joint decision of two or more partners. Given that dyadic perspectives emphasize mutual attractiveness and synergy between resources of the partners as an important element of alliance formation, they explicitly account not only for individual characteristics of each partner, but also their fit across multiple dimensions. The fit captures perceived interdependence between firms. Scholars have studied various aspects of the interdependence, including resource complementarity and co-specialization (Chung, Singh & Lee, 2000; Gimeno, 2004; Mitsuhashi & Greve, 2009; Rothearmel & Boeker, 2008); participation in technical communities (Rosenkopf, Metiu & George, 2001); trust and commitment due to prior relationships (Gulati, 1995; Gulati & Garguilo, 1999), and structural positions within networks (Ahuja et al., 2009; Chung et al., 2000; Gulati & Garguilo, 1999).

Even more than studies adopting a focal firm perspective, this set of papers explicitly assesses whether configurations of partners' attributes represent complementarities[2]: Mutual

[2] "Complementarity" here refers to the concept of super-modularity or Edgeworth complements (Milgrom & Roberts, 1995): having more of any one attribute increases the returns of having more of the other attribute.

attractiveness may stem from fit arising due to exchange partners compensating for the other's lack of attributes, and value creation occurs due to a match between partners where each partner is endowed with capabilities that the other finds valuable and difficult to develop internally in a timely manner, or difficult to acquire profitably through factor markets (Ahuja, 2000; Mitchell & Singh, 1996). Thus, young biotechnology firms may seek marketing and manufacturing capabilities of the more established pharmaceutical partner firms, who in turn benefit by diversifying into the technological domains and capabilities represented by the biotechnology firm (Mason & Drakeman, 2013; Rothearmel & Boeker, 2008). Or, mutual attractiveness may be due to attributes such as technological overlap, geographic co-location, social similarity, structural homophily, or shared history (Gulati & Garguilo, 1999; Rosenkopf & Padula, 2008; Rothearmel & Boeker, 2008), factors that add value due to reduction in communication costs, or increase in partner reliability and trust.

Theoretical Assumptions Underlying Extant Research in Partner Selection in Alliances

Our interest in articulating the theoretical assumptions underlying work on alliance formation stems from the fact that these have important implications for the appropriateness of the empirical methodology. The papers reviewed above make at least one of the following three assumptions, implicitly or explicitly: (1) Partner preferences do not impact the likelihood of partner selection, (2) Mutual fit of partner attributes is not a critical criteria in the partner selection decision, and (3) There is no competition among firms for the potential "best" partners, so firms are free to choose from the universe of available firms with requisite capabilities.

Studies that focus on a single firm's perspective make the first and third assumption, but vary regarding whether they make the additional assumption regarding importance of fit between potential partners. Studies adopting the dyadic perspective make the third assumption, since they too assume that the supply of linkage partners is "infinitely elastic" (Ahuja, 2000), and accordingly, that there is no competition among firms for the more desired partners (Greve et al., 2013). We note that some studies do acknowledge the constraints that partner availability pose in partner selection; for example, studies that examine status homophily note that low status firms may not have access to high status firms who prefer to ally with each other (Rosenkopf & Padula, 2008; Stuart, 1998).

However, as Greve et al., (2013) note, a proper theoretical treatment of the issue requires the use of matching logic and consideration of outside options, since ex-ante partner selection constraints imposed due to the two-sided competition are important in ex-post outcomes in alliance formation.

The relevance of these assumptions is exemplified by the discussion across two recent papers published in the *Strategic Management Journal*. Diestre & Rajagopalan (2012) take the start-up biotechnology firm's perspective to conclude that start-ups can and do avoid pharmaceutical firms that represent greater appropriation risks and seek partners that provide them complementary resources. In response, Mason & Drakeman (2013), practicing executives in biotechnology firms, note that given competition among several thousand biotechnology firms for a partnership with one of at most 100 pharmaceutical companies, it is likely the case that pharmaceutical firms, rather than the biotechnology startups, are responsible for the observed partner selection. Accordingly, their industry experience and survey-based arguments highlight pharmaceutical firms' motives for diversification beyond in-house capabilities as more salient than biotechnology firms' concerns regarding value appropriation. Clearly, both perspectives matter, as does the fit between partner attributes, and the competition on each "side" of the partner selection for the best available partner.

Empirical Approaches Used in Extant Literature on Partner Selection in Alliances

The previous section discussed the theoretical assumptions in the current literature when examining the question of alliance formation. In this section, we turn our attention to the current empirical practice for inferring partner selection, and complementarities in partner attributes. As is evident from Table 1, binary choice models are the predominant empirical approach in addressing the question of partner selection in the alliance formation literature.

When conducted at the firm level of analysis, the empirical specifications address the likelihood of undertaking an alliance based on focal firm or environmental characteristics, but independent of partner preferences (Baum et al., 2005; Beckman et al., 2004; Geringer, 1991; Hitt et al., 2000; Hitt et al., 2004; Katila et al., 2008; Li & Rowley, 2002). Stuart (1998) discusses the challenges with this approach, and even though he theorizes from a focal firm perspective, he provides empirical tests at both the firm and dyad level of analysis. As he notes, given that alliance formation is influenced by both firm and dyadic factors, it is important to parse out the variation attributed to dyadic factors when trying to estimate the variance that can be attributed to firm characteristics. Accordingly, most of the recent research employs the dyad as the unit of analysis, even when taking a focal firm perspective (See Table 1).

At the dyadic unit of analysis, all of the studies make inferences regarding partner selection based on the signs and significance of the coefficients in binary choice model specifications. Given recent cautionary notes regarding interpretation of interaction effects in binary choice models (Hoetker, 2007; Zelner, 2009), later studies assess the probability of partner selection by examining the significance and magnitude of the interactions over the range of partner attribute values, and typically at the mean levels of control variables (e.g. Diestre & Rajagopalan, 2012; Reuer & Lahiri, 2013). However, inferences on complementarities in partner attributes have been made only on the basis of signs and significance of the interaction terms between partner attributes[3] (Ahuja et al., 2009; Chung et al., 2000; Gimeno, 2004; Gulati, 1995; Gulati & Garguilo, 1999; Hallen, 2008; Li & Rowley, 2002; Mitsuhashi & Greve, 2009; Rosenkopf & Padula, 2008; Rothearmel & Boeker, 2008; Sorenson & Stuart, 2001). In these studies, positive and significant interactions among partner attributes were deemed as representative of increasing returns, and thus fit or complementarities.

An additional issue at the dyadic unit of analysis relates to the underlying sample frame created for unrealized outcomes. Scholars have utilized two approaches when creating the choice set of potential partner combinations (See Table 1). In the unconstrained approach, the full list of dyads at the risk of alliance formation is created (Ahuja et al., 2009; Alcacer et al., 2009; Chung et al., 2000; Diestre & Rajagopalan, 2012; Dushnitsky & Shaver, 2009; Garcia-Pont & Nohria, 2002; Gimeno, 2004; Gulati, 1995; Podolny, 1994; Rosenkopf et al., 2001; Rothearmel & Boeker, 2008; Rosenkopf & Padula, 2008; Stern et al., 2013; Stuart, 1998). Typically, this list of dyads comes from all pairwise combinations of firms that are in the sample and at risk of alliance formation in the focal

[3] Studies vary on whether they directly interact partner attributes, or create a composite, pair characteristic variable that reflects these interactions.

time period. In addition to the above theoretical issue of a lack of attention to constraints to the choice set because of firms of each side competing for the best partner, this unconstrained choice set creates empirical biases and inference issues, since it impacts the baseline odds when using binary choice models. The unconstrained approach causes each firm on either side to be included in the sample for a large number of times, causing both interdependence of observations, and the problem of rare events (Gulati & Gargiulo, 1999; Hallen, 2008; Sorenson & Stuart, 2001).

Besides using model specifications that address the above problem (e.g. rare events bias correction), scholars have tried to address this empirical issue by constraining the sample choice set. Constrained approaches include the use of randomly generated sample of unrealized outcomes (Gulati & Gargiulo, 1999; Mitsuhashi & Greve, 2009), choice-based sampling (Hallen, 2008; Sorenson & Stuart, 2001), or case-control sampling (Reuer & Lahiri, 2013; Sorenson & Stuart, 2008).

As elaborated later, each of the improvements in empirical specifications highlighted above, while representing improvements to earlier strategies, are nonetheless limited to the extent that the assumptions made in binary choice models pose challenges in inference.

TWO-SIDED MATCHING MODELS: THEORY AND EMPIRICAL ESTIMATION

Our literature review in the preceding section underscored that regardless of whether scholars adopted a single or dual firm perspective, they do not explicitly account for the implications of competition, and thus sorting, among potential partners. We suggest that this abstraction has theoretical as well as empirical consequences and that the two-sided matching models have the potential to alleviate these problems. Theoretically, we note that firms enter into alliances rather than arms-length market transactions precisely because the resources and capabilities that they seek to access through alliances do not conform to "perfectly competitive" market conditions: where many buyers and sellers transact for homogenous and perfectly divisible goods. In fact, firms are in a generalized competition for scarce resources (Pfeffer & Novak, 1976), such as valuable scarce partners. In alliancing context, competition is driven by capacity constraints in the number of partners that firms can undertake at a time. The management of alliances is a complex process facing numerous challenges, including coordination costs, creating appropriate channels for sharing and transferring knowledge, and making inter-partner learning possible (Ireland, Hitt & Vaidyanath, 2002). In order to manage alliances effectively, firms need to develop special, relational capabilities and create an appropriate organizational structure (Kale, Dyer & Singh, 2002). All these constraints limit the number of alliances that firms can pursue at once. Furthermore, even if capacity constraints are not an issue, firms might decide to limit the number of partnerships for reputation reasons.

More generally, competition for partners emerges because firms on each side of the transaction are imperfectly substitutable (Peteraf, 1993). For a firm contemplating an alliance, the distribution of potentially attractive partners is highly asymmetric. Likewise, a firm will be more attractive for some partners than for others, because its attributes such as resources and capabilities will be differentially valuable in combination with attributes of different partners. Importantly, although mutual attractiveness is recognized in dyadic perspectives that emphasize fit, it is not a *sufficient* condition for alliance formation between the two of them. Rather, an alliance is formed only if either firm cannot achieve a higher level of mutual value creation with other firms in the market. This noteworthy but less studied aspect relates to the competition on each side for the "best" partner, which has consequences not only for understanding of alliance formation and performance link, but also for the nature of complementarities among alliance partners.

Two-sided matching models take into account market-level interactions among dyadic decisions. So, not only do two-sided matching models account for the preferences of both sides in the market, and explicitly incorporate complementarities among partner characteristics, but they also address the effect of competition for partners in the market. As we elaborate in the next sub-section on matching model theory, these characteristics of the matching models make it particularly suitable for the study of alliance formation, and allow future generations of research in the study of complementarities and alliance formation to address two key characteristics of the alliance market: competition and sorting. In a later sub-section on the comparison of matching and binary choice empirical estimation, we discuss why the existing estimation techniques which consider dyads in isolation are unable to explicitly incorporate the competition in the market for alliance partners.

Two-Sided Matching Models: Theoretical Underpinnings

Most of the matching models literature in economics relevant for firm alliances is in the stream that focuses on "voluntary pairing under competitive conditions" among complementary pairs (Mortensen, 1988, p. S215). Matching markets are typically characterized as consisting of two (or more) kinds of heterogeneous agents, each of whom have preferences over potential matches with agents of the other kind, and are in competition for desirable partners with agents of their own kind (Roth, 1984; Roth & Sotomayor, 1990)[4]. The quintessential example of such a partnership is marriage, which was the first social relationship studied as a matching process: as Becker (1973) elaborated, men and women (the two kinds of heterogeneous agents) have preferences regarding partner characteristics, and on each side, there is competition for desirable partners. Thus, the original matching model was also called the marriage model, since gains from a match and the mechanism of the market force competition explained not only "who marries with whom," but also complementarity and substitutions in partner characteristics (Becker, 1973).

Matching governs the formation of many relationships in real life markets. In economics, finance and sociology, and to a very limited extent in strategy, scholars have examined relationships such as school choice (Gale & Shapley, 1962), employer-employee job matching (Coleman, 1991); and a wide range of business contexts, including mergers and acquisitions (Park, 2008; Akkus, Cookson, & Hortacsu, 2012); funding of start-ups by VC investors (Sorensen, 2007); IPO underwriters (Fernando, Gatchev, & Spindt, 2005); bank lenders (Chen & Song, 2012); assignment of CEOs to firms (Tervio 2008; Gabaix & Landier, 2008, Pan, 2010); buyer-supplier relationships (Fox, 2010b; Chatain 2013; Ostrovsky, 2008); and firm-university relationships (Mindruta, 2013).

"Who partners with whom?" Matching models solve the problem of finding a suitable "assignment" (i.e. a structure of who matches with whom) that advances the common interests of agents wishing to be matched and reconciles the conflicting interests of agents of the same kind

---

[4] "Matching" has been used in two other forms, which are not discussed here. One refers to the assignment of objects to individuals (such as housing choice). Here, individual agents have preferences over objects, and objects are characterized by various degrees of scarcity. The other usage is in the context of various kinds of agents that interact through a common trading platform (Rochet & Tirole, 2003).

competing in the match formation arena. Matching may be accomplished through a centralized system of rules (such as in the matching of medical interns and residents in the United States with hospitals (Roth, 1984) or through a decentralized, often invisible process (Stovel & Fountain, 2009) where the acceptance and rejection rules are not directly observable, but the final configuration of pairing is. We focus our exposition on decentralized matching processes as they better reflect alliance formation. Common to the decentralized matching process situations is a matching problem where agents seek a mutually agreeable match from a pool of potential partners, who are also deciding whom to partner with given their own preferences and feasible choice-set. For the brevity of exposition, consider a typical upstream-downstream alliance relationship, where each firm may only enter into one alliance. If we take a matching perspective to alliance formation, whether downstream firm ***D*** chooses upstream firm ***U*** depends upon ***D's*** payoff from partnering with ***U*** relative to its payoff from partnering with any other possible upstream firms who are also willing to work with downstream firm ***D***. In turn, an upstream firm ***U's*** willingness to accept a partnership ***D*** will depend upon its own utility from the relationship as well as its "effective" choice set, i.e., the set of downstream firms willing to ally with this particular upstream firm given their own alternatives. Therefore, neither individual preferences for a particular partner, nor dyad-level mutual acceptance are sufficient conditions for an alliance to occur. Instead, each dyad-level decision must be analyzed relative to all other possibilities of partnering that each party might have to forego by committing to a particular relationship. A partnership between ***D*** and ***U*** is not feasible as long as any of the two partners is able to realize higher gains elsewhere. If upstream firm ***U*** has higher gains from partnering with a different downstream firm than ***D***, then a partnership between ***U*** and ***D*** would not only "hurt" the upstream firm, but also the firm's forgone partner, i.e. the downstream firm that would have made the ***U*** better off[5].

[5] A similar reasoning applies in situations where each partner can make multiple relationships, where partners are selected up to a "quota" which sets the upper limit in the number of ties.

Complementarities and Substitutions in Partner Attributes: The reason why both parties care about the identity of the trading partner is because transactions represent trade in heterogeneous and indivisible partner attributes, i.e. transactions in matching markets involve an "all-or-nothing trade of bundle of traits" (Mortensen, 1988) that each partner brings into a relationship. Naturally, all agents possess many different characteristics that make them more or less desirable in a relationship. Therefore, while participants in typical markets exchange fungible and homogeneous goods, participants in matching markets make trade-offs between indivisible and heterogeneous goods.

Becker (1973) predicted the patterns of sorting that could emerge in a marriage market where men and women differ along one-dimensional attribute called "type". Positive assortative matching means that highest type agents on each side of the market will pair-up, leaving the second-highest type agents to select each other and so forth. Negative or anti-assortative matching denotes the pairing of the highest type agent on one side to the lowest type agents on the other side. Positive or negative assortative matching is related to whether partner attributes represent complementarity or substitution, respectively. Formally, these terms relate to the cross-partial derivative in the joint production function (Becker, 1973).

Partner attributes are complements if the cross-partial derivative of the production function is positive, thus in equilibrium, complementary attributes will result in positive assortative matching. As noted earlier, two attributes are complements if having more of one raises the marginal value (or the incremental return) of having more of the other. For example, in Becker (1973), partner attributes such as education and intelligence in the marriage market are considered complements. Importantly, and in the firm alliances context, complementarity may occur regardless of whether the partner attributes are same or different; i.e. complementarity may arise because either (1) more of attribute A (say manufacturing capabilities) of partner ***D*** results in greater gains from having more of attribute B (say technological capabilities) of partner ***U***, or (2) more of the same attribute (say R&D capabilities as in absorptive capacity; Cohen & Levinthal, 1990) in both partners is reinforcing and results in greater gains. We note that in the alliance formation literature, scholars have used the term "complementarities" to mean only the first kind, and the term "similarity" or "compatibility" to denote the second kind (e.g. Chung et al., 2000; Rothaermel & Boeker, 2008; Stuart, 1998).

While not highlighted in the alliance formation literature, matching models also allow assessing if partner attributes are substitutes. Partner attributes are substitutes if the cross-partial derivative is negative, i.e. if having more of one decreases the marginal value of having more of the other (equivalently, two attributes are substitutes if having *less (more)* of one raises the marginal value of having *more (less)* of the other). In equilibrium, substitute attributes will have negative assortative matching. In Becker (1973), dominant and deferential personalities are more likely to marry, given that these partner attributes are substitutes. In firms partnering with university scientists, Mindruta (2013) shows negative assortative matching in partner patenting capabilities.

We note that to maintain simplicity, the above description of the theoretical underpinnings of matching models used a one-to-one matching context, and focused on observable partner attributes. However, extensions of the models permit for one-to-many and many-to-many partner matching, and also discuss matching on unobservables (Sotomayor, 1998; Fox & Yang, 2012).

Two-Sided Matching Models: Empirical Estimation

While matching model theory has a long and rich history, empirical techniques that incorporate its logic have only recently begun to be used. Largely, this is because of the difficulty of estimating models in which all decisions are interrelated and have to be modeled at once. As noted in Fox (2010b), there are two "curses of dimensionality" in the number of agents in a matching market that result in computational challenges. The first curse of dimensionality results from having to compute and check equilibria, and the second curse results from the need to estimate matching probabilities as a function of all agent attributes. In a series of papers, Fox and his colleagues have addressed these computational challenges by creating a maximum score identification approach (Fox, 2010a, 2010b; Fox & Bajari, 2010; Fox & Yang, 2012). The computational ease of using the maximum score approach has resulted in its budding use, particularly in economics and finance (Akkus et al., Cookson, & Hortacsu, 2012; Baccara et al., 2011; Chatain, 2013; Chen & Song, 2013; Levine, 2009; Mindruta, 2013; Nakajima, 2012; Pan, 2010; Yang et al., 2009).

In essence, the estimation approach builds on the idea that a configuration of observed partnerships in a market is the equilibrium outcome of a matching process, to derive a method for estimating the characteristics of the partners that drive, in interaction, the formation of partnerships in a specific empirical context. This method allows researchers to make inferences about the nature of relationships (complementarity versus substitutability) between the attributes of the partners, as well as about the relative importance of various relationships in explaining the patterns governing the sorting of partners into the observed alliances. We refer interested readers to Fox (2010a, 2010b) and Appendix A for greater details, while providing the intuitive logic in this section. To keep the description of the method tractable, we focus on one-to-one matching, and the more general many-to-many matching can be found in Fox (2010a). Continuing with the upstream-downstream terminology, we denote downstream firms by **d** and upstream firms by **u.** Firms may decide not to ally, but the estimation focuses on pair formation and does not use information on single agents.

Match Production Function: A pair (*d, u*) realizes a match value *f* (*d, u*) which is analogous to a revenue or utility function for the matched pair. The match production function *f* transforms partners' endowments (tangible and intangible skills, resources they contribute to the alliance, etc) into a joint output. The production function may contain both interaction terms between partner-specific attributes (e.g. size, age, technological capabilities), and match-specific or pair-specific attributes (e.g. geographical distance, or knowledge distance). The focus is then to estimate this function from the observable partner characteristics (See Fox and Yang (2012) for estimation based on matching on unobservables). The model thereby abstracts away from problems of adverse selection associated with unobservable attributes whose impact on the joint production and implicitly preferences over partners could make the problem empirically intractable.

Equilibrium: There are two important concepts for understanding the equilibrium conditions that lie behind the estimation. First is the idea of ***transferable utility***. Transferable utility implies that agents can transfer part of their own utility to other players. This allows for modeling situations where agents can lower prices to attract better matches or conversely, trade away match quality in order to obtain a higher share of the surplus. Under transferable utility, the rules for value appropriation, or sharing of the joint surplus generated in a match, are negotiated at the pair-level. Conversely, non-transferable utility implies the existence of a fixed rule for splitting the surplus, which is common across all pairs in the market. We believe that transferable utility is a better characterization of alliance formation, though knowledge of pie-splitting rules in a particular context may also be incorporated.[6] Second is the idea of ***pairwise stability***, an equilibrium concept specific to matching games. Stability asserts that that neither party in a match (observed alliance) has an incentive to separate and form a new match with a different partner. Technically, a match assignment is stable if: 1) no matched agent prefers the single state, and 2) for each agent in a match, the payoff from any other potential partnership is less than or equal to the share of the surplus the agent is getting from the actual match.

To summarize, firms in an alliance realize a joint match value (revenue) that is split between the two partners. A partner's payoff is the match value, net of transfer payments and other contributions made to the other party. In equilibrium, each party in an alliance receives a higher payoff from the observed partner than it could get from counterfactual partners.

Local Maximization Condition: Fox (2010a, 2010b) provides formal proof that a pairwise stable assignment with transferable utility implies the following relationship: If we take any two pairs that are matched in a pairwise stable configuration of alliances and swap the partners, the original sum of match revenues is greater than or equal to the sum of match revenues of the counterfactual pairs constructed by exchanging partners. This is a necessary condition derived from the equilibrium properties of matching models. As a consequence, the estimation technique does not require scholars to provide a formal calculation of the equilibrium outcome.[7] The inequality condition explained above does not use information on how the surplus is divided among the two alliance partners (See also Appendix A). This is a major strength of the estimator, since it permits scholars to

[6] A classic example of non-transferable utility is the matching of applicants and schools in the college admission process. Some important features of matching, such as positive assortative matching of complementary partners, are preserved with both types of utility. Transferable utility allows for more flexibility in estimation as it does not require the ex-ante specification of agents' preferences.

[7] The existence of an equilibrium assignment is discussed, among others, in Shapley and Shubik (1971) and Koopmans and Beckmann (1957) for one-to-one matching; and Sotomayor (1999) for many to many matching).

uncover complementarities in partner attributes that result in value creation, without needing full information on the contractual details that are often unavailable in real world data.

In any given market, the estimation procedure entails computing the match production function for all observed alliances and checking the inequality for all possible pairs of agents. Each time the inequality holds for a trial guess of the vector of production function parameters, the score of correct prediction increases by 1. The vector of parameters that maximizes the number of correctly predicted inequalities provides a consistent estimator of production function parameters.

Derivative Based Identification of Complementarities and Substitution in Partner Attributes

Since the match value of the production function depends on the complementarities and substitutions in partner attributes, the estimation approach uses data on equilibrium outcomes to matching models to assess whether partner attributes will positively or negative assortatively match with each other (Fox, 2010a). The intuition behind it is simple: once the vector of parameters that maximizes the number of inequalities is estimated, the cross-partial derivatives of the underlying attributes provide information about whether they are complements or substitutes. As noted above, if the interaction terms of partner attributes are positive, then the two attributes are complements; if it is negative, then the two attributes are substitutes. Further, the estimation approach also allows scholars to assess which complementarities and substitutions are more important than the others.

Some Practical Considerations: When using the maximum score estimation approach, the template described above can be accessed from Fox's website at http://www-personal.umich.edu/~jtfox/. The data should clearly identify the two sides of the market. Further, scholars need to make the choice of whether inferences will be made from single or multiple markets (e.g. each year in a panel of alliances can represent a "market"). The boundaries of the market determine the set of counter-factual pairs, for example, when multiple markets representing each year is used, the set of counter-factual pairs are determined from the firms that entered into an alliance that year. Fox (2010a) shows Monte Carlo simulations testing the finite-sample performance of the estimator. Although parameter estimates are consistent regardless of whether scholars analyze data from multiple markets (where consistency is achieved by assuming asymptotics in the number of markets) or from one big market (where consistency is achieved by assuming asymptotics in the number of pairs), the bias and the root mean-squared errors are smaller in data from multiple, smaller markets than one large market. Finally, it is important to note at the onset the following three properties of the estimation approach:

1) Only interaction terms can be estimated. The method does not provide estimates for the direct effects of the characteristics of the agents on the production function. Theoretically, this means that the "un-interacted" characteristics are valued equally by all potential partners. In other words, matching is driven entirely by the interaction between the characteristics of the partners.

2) The production function can be fully identified up to a linear transformation. Thus, a scale-normalization needs to be imposed on the vector of unknown parameters. This can be done, for example, by assuming that one of the coefficients in the production function equals 1 or minus 1. All the other unknown parameters are estimated . This is a constraint that the estimator shares with discrete choice models, since imposing location and scale normalizations on the unknown parameter vector is standard procedure in semi-parametric estimation/discrete choice models (The location parameter is usually set to zero, and scale normalization typically involves the normalization of the variance of the error terms (e.g. being set to $\pi 2/ \sqrt{6}$ in logit, and 1 in probit)).

3) Hypothesis tests require constructing the confidence intervals of the point estimates. Fox suggests random sampling without replacement from the data set. For each randomly sampled data set, a coefficient is estimated. The hypothesis that a given coefficient is positive (or negative) can be rejected at the 5% level, if 5% or less estimates of the coefficient are positive (negative).

COMPARISON OF EMPIRICAL METHODOLOGIES

We now turn to a comparison between discrete or binary choice models and the matching model methodology for assessing partner selection and complementarities in partner attributes. We first elaborate on the potential shortcomings of each based on theoretical differences, and then provide a simulation that compares the characteristics of the two models, and their relative performance, when the true underlying relationship is known.

Differences in Binary choice and Maximum score Matching Model Estimation

In estimating cross-firm complementarities in the context of alliance formation, there are salient differences between the standard binary choice model (as has been used in the mainstream management/strategy journals) and the maximum score method. We note that both methods have many other applications that we do not discuss here.

Maximum score method estimates the utility of agents in a match by specifying the mutual preferences of partners as a function of their characteristics. Logit starts from the "utility" or "value" a decision maker derives from making a choice. The researcher specifies a function that relates the observable attributes of the decision maker (here, focal firms making a choice) and the attributes of the choices (i.e. potential alliance partners) to the decision maker's utility. Error terms are treated as random and iid extreme value. With this definition of error terms and the additional assumption that the decision maker chooses the alternative that provides the greatest utility, logit model allows researchers to make probability statements about the decision maker's choice.

In both methods, interaction terms between the attributes of the partners provide the basis for making inferences about complementarity and substitutability. In both methods, interaction terms are estimated from taking the cross-partial derivatives of a "function". In matching models, the "function" is the joint utility of agents in a match. In logit, the "function" is the probability of an event occurring. Nonetheless, underlying the probability choice in logit is a utility function which presumably captures mutual preference of the partners. Notwithstanding differences in the interpretation of interaction terms (which we discuss in our simulation), the two methods take fundamentally different perspectives on estimating the utility (value) from an alliance, conditional on partners' characteristics. The differences in the assumptions have important consequences when assessing complementarities and partner selection in the context of inter-firm alliances.

As discussed in greater detail in Appendix B, logit/probit models suffer from the following shortcomings: first, logit/probit models are unable to accommodate mutual choices. The method proceeds as if the realized alliance is an independent choice made by a focal firm based on own preferences over the attributes of all possible partners in the sample. Second, logit/probit models do not take into account how preferences and alliance opportunities of other firms in the market constrain dyad-level decisions. Third, logit/probit methods and the other variations of discrete choice models suffer from the "curse of dimensionality" (Fox, 2010a, 2010b). As we explain in Appendix B, even if scholars wanted to model a matching equilibrium in the spirit of discrete choice models, this is not possible because of computational constraints. Finally, variations of logit/probit that deal with ways of selecting the potential but unrealized alliances, or address biases due to rare events, do not address these fundamental problems.

The maximum score matching model estimation technique addresses many of these issues. Nonetheless, its usage may pose some problems, which we now turn to. First, the use of matching models requires scholars to identify the two sides of the market. In many instances, this can be easily done, as when matching firms to workers, schools to students, buyers to suppliers, upstream to downstream markets, etc. However, in other instances the two sides of the market may not be cleanly identifiable. This is an important issue, since market boundaries are particularly important to delineate in matching models because they dictate the set of counterfactual pairs used for creating the inequalities. Second, even though the maximum score estimation approach has reduced the computational challenges arising from the curse of dimensionality, the approach remains computationally intensive, and it requires scholars to use computational packages such as Mathematica and Matlab, rather than conventional statistical software packages (Stata, SAS, etc.). Third, the approach generates point estimates of complementarity and substitution across partner attributes, and cannot be used to determine the *range* of values of partner attributes over which complementarities or substitution may hold (Fox, 2010a). Fourth, partner-specific attributes (i.e. the direct, un-interacted effects) cancel out in the local maximization condition and cannot be identified with the maximum score estimation technique, though Akkus, Cookson and Hortacsu (2012) suggest an extension where they use data on transfer payments between partners in a match to estimate both the interaction and the main effects.

Further, the maximum score matching estimator is unable to accommodate some thorny issues that may be present in the real world. The estimator assumes that observed alliances are the equilibrium outcome of a matching process, which may not always accurately represent reality. In particular, scholars have to give some thought on the fact that assortative matching may be thwarted by the presence of frictions in the matching market. Frictions may arise due to search costs or incomplete information about the characteristics of the potential partners. With search costs, the question is whether agents are willing to wait for the most attractive partners that would match with them. To assess whether frictions from informational asymmetries are important, scholars must first ask whether the specification of the model can be improved with better data. A more serious concern is whether some agents in the market have idiosyncratic private advantages in their information on potential partners. This might lead to situations where firms which rank low on observed attributes are chosen over ones which rank high purely based on informational concerns.

The theoretical literature on matching shows that matching remains assortative when search costs are introduced under various formulations in the model, such as when stronger complementarities exists among pairs (Shimer & Smith, 2000) or when search costs are partner-independent (Atakan, 2006). However, scholars may want to turn to the large literature on dynamic matching (or search models in labor economics) if they explicitly model the impact of search costs on match formation and dissolution. The maximum score estimator does not deal with dynamic matching, but other empirical approaches that accommodate search exist (e.g. Canals & Stern, 2001).

As it is the case with most empirical methods, the estimator's ability to deal with situations where unobservable attributes shape preferences of agents remains an open question. The method is robust to unobserved firm-effects and other unobservables that are valued similarly by agents on the same side of the market because they cancel out in the local maximization condition (please see Appendix A). Fox and Yang (2012) suggest that the method is robust even when unobserved complementarities exist between partners, as long as they are not match-specific. Scholars may want to consider carefully whether the presence of unobservable characteristics could make the problem empirically intractable. Most often, alliance formation is preceded by a "courtship" period which serves, among other things, to improve the partners' reciprocal perception of their qualities. Moreover, in many markets, uncertainty on the potential partner is often mitigated by referrals from third parties. If, on the other hand, alliances are formed under conditions of persistent uncertainty about the "true" attributes then the assortative properties of the assignment problem might be seriously undermined by adverse selection and moral hazard problems. In aggregate, with idiosyncratic informational asymmetries, we would no longer observe assortative matching.

Undoubtedly, in real life, there is some degree of uncertainty involved in any partner selection problem, and better information on a partner is often revealed only after the alliance is formed. New information could lead either to a continuation of the relationship or to match dissolution. The way matching theory takes into account initial uncertainty on the match quality is by including this consideration in models explaining the *dynamics* of the matches (Mortensen, 1988). This is however, a separate literature that the estimator does not address directly.

Simulation Runs Comparing the Empirical Methodologies

In this section, we describe the results from a simulation where we start with a known relationship of complementarities between partner attributes, generate data for these attributes, and estimate the coefficients using both the maximum score matching estimator and logistic regression. This enables us to compare the characteristics of both estimation methods, with a particular focus on the relative performance of the methods in estimating the true value of the coefficients.

Continuing with the example of upstream-downstream alliances, we assume that there are three partner attributes (x1, x2 and x3) that are relevant for value creation through alliance formation. Accordingly, we use a simple specification where the true matching production function has the following functional form:

$$F(x_u, x_d) = \beta_1 \cdot x1_u \cdot x1_d + \beta_2 \cdot x2_u \cdot x2_d + \beta_3 \cdot x3_u \cdot x3_d$$

where xi_u with i=1,..3 denote upstream firm attributes and xi_d with i=1,..3 denote downstream firm attributes. For each simulation run, we generated the values for all variables from a standard normal distribution. Choosing the variable values from a standard normal distribution allows us to abstract away from the effect of exogenous differences in the values of the partner attributes, and focus only on the endogenous value created due to complementarities in the alliance formation.

To obtain scale normalization, we set β_1=1. We chose the values of the other coefficients to be β_2=5 and β_3=3. The true coefficient values imply complementarities between partner attributes. Because the coefficient β_2 is the highest, firms match more assortatively on the second dimension than on the other dimensions. We then created 10 markets, with 50 agents on each side. Starting from the vector X=(xi_u, xi_d), i=1,2,3 of randomly distributed variables and the true payoff functions, we used a linear programming procedure in Mathematica to generate the optimal assignment in the market, i.e. the one-to-one pairing of upstream and downstream agents that maximized the sum of the payoffs in each market. This procedure ensures that the resulting assignment (matched agents) is the outcome of a one-to-one matching game (Shapley & Shubik, 1971).

Using these simulated data, we are able to mimic a typical sample of alliances where researchers observe who matches with whom and the attributes of the participants. We then constructed a dependent variable that takes 1 for a true alliance and 0 for all other possible but unrealized alliances. Counterfactual pairs were obtained from the unrealized alliances between agents on the opposite sides of the same matching market, which is the standard approach to creating counterfactuals in both the maximum score estimation, and in work using logit method.

We estimate parameters β_2 and β_3 of the match production function F(x_u, x_d) using both the maximum score estimator and logit method. We included two specifications of logit. In the *interactions only* specification, we proceed as if the researcher knows the true functional form. Here logit regression mirrored the match production function, and did not include any simple effects, since we know that the coefficient values of these terms are zero in the true relationship. In the *full* specification we proceed as if the researcher is agnostic about the true production function and we adopted the common approach in the literature: we included a constant term, each partner's attributes, and the three interaction terms. To summarize, the maximum score method estimates

$$F(x_u, x_d)= \beta_1 \cdot x1_u \cdot x1_d + \beta_2 \cdot x2_u \cdot x2_d + \beta_3 \cdot x3_u \cdot x3_d,$$

a function that represents the joint utility of partners in a match. The logit *interactions-only* specification estimates the alliance probability assuming the underlying choice utility to be:

$$F(x_u, x_d)= \beta_1 \cdot x1_u \cdot x1_d + \beta_2 \cdot x2_u \cdot x2_d + \beta_3 \cdot x3_u \cdot x3_d.$$

The logit choice probabilities in this model are given by the cumulative distribution:

$$\Phi(F(x_u, x_d)) = 1/(1+\exp(-F(x_u, x_d)).$$

The logit *full specification* provides the choice probability given by the cumulative distribution:

$$\Phi(G(x_u, x_d)) = 1/(1+\exp(G(x_u, x_d)) \text{ where}$$

$$G(x_u, x_d) = \beta_0 + \beta_1 \cdot x1_u \cdot x1_d + \beta_2 \cdot x2_u \cdot x2_d + \beta_3 \cdot x3_u \cdot x3_d + \beta_{1u}\, x1_u + \beta_{1d}\, x1_d + \beta_{2u}\, x2_u + \beta_{2u}\, x2_u + \beta_{2d}\, x2_d + \beta_{3u}\, x3_u + \beta_{3d}\, x3_{d.}$$

Our interest is to see how well logit specifications estimate the (known to us) values of $\{\beta_{1,} \beta_{2,} \beta_3\}$ because they represent the coefficients of interest in inferring complementarities between the attributes of the partners. Given the theoretical assumptions behind the two methods, we believe that the main issue for assessing complementarities with logit is that logit obtains its estimates by maximizing the utility of an individual decision maker over alternative choices, while the maximum score technique obtains its estimates by maximizing the utility of all agents in the market. Two observations are important when comparing the estimates across methods:

1) In both methods the overall scale of utility is irrelevant. However, while the overall scale of utility does not matter, the scale normalization affects the absolute values of the estimates. Therefore, to avoid making erroneous comparisons across models that have different scale normalizations, we focus on the ratios of coefficients, which are invariant to scale. We know, in addition to the true signs, that the ratios in the true model are $\beta_{2/}\ \beta_1=5$; $\beta_{3/}\ \beta_1=3$; and $\beta_{2/}\ \beta_3=1.66$.

2) While the maximum score method estimates coefficients β in the utility (or match production) function *F* directly, logit models estimate coefficients β by taking the extra step to infer choice probabilities. The coefficients of the interaction terms in the two methods have different meanings because they represent the coefficients of the cross-partial derivatives of two different functions: the match utility function *F* in the maximum score method, and the cumulative distribution functions $\Phi(F)$ and $\Phi(G)$ in logit. Thus, when researchers familiar with discrete choice models proceed to examine interaction effects, they focus on the cross-partial derivative of the expected value of the *probability* to observe an alliance event (Norton, Wang & Ai, 2004). On the contrary, our interest is in the cross-partial derivative of the *utility* function underlying match

formation. We nonetheless report the coefficients of interaction terms from the cross-partial derivatives of the probability function by doing the necessary adjustments in the calculations as suggested by Norton, Wang & Ai (2004).

We ran the above process 500 times, and thus estimated the model for 500 datasets. Tables 2-4 provide the statistics (mean, standard deviation, minimum, maximum) on the coefficients and model fit for each method, summarized over the 500 Monte Carlo runs.

[Tables 2-4 about here]

The coefficient estimates in Table 2 show that the maximum score estimation method does remarkably well in estimating the true underlying values of the coefficients. Across 500 runs, the mean value of the estimates for β_2 is 4.9899, and for β_3 is 2.9957. Thus the bias (deviation from true value) is almost zero, and the root mean error (standard deviation) is low at 0.2475. Further, the maximum score estimation also reveals that complementarities between attributes $x2_u$ and $x2_d$ is more important in value creation than complementarities between $X3_u$ and $X3_d$.

However, the coefficient estimates in Table 3 (interactions only logit regression) do not mirror the true values of the coefficients. In the interactions only model, across 500 runs the average value of β_2 / β_1 is -1.2504 (while the true value is 5) and the average value of β_3 / β_1 is -1.1459 (while the true value is 3). The bias is also very high in the full specification logit regression reported in Table 4: the average value of β_2 / β_1 is 6.6071 and the average value of β_3 / β_1 is 6.2654.

We also compared the two models with respect of measures of fit. In terms of the model's predictive power, the logit specification performs extremely well, particularly in the full model. In Table 3, the average percent of correctly predicted outcomes is 56%, and in Table 4, it is 98%. Thus, although binary choice models such as logit cannot identify the true nature of complementarities, they perform well in terms of predicting partner selections. However, if scholars are interested in inferring explanatory power of the variables in the model, and the extent of complementarities among partner attributes, matching models are the more appropriate method.

We also replicate the current empirical approach in the existing literature with regard to assessing the extent of complementarities. The current literature uses a *full specification* logit model and infers complementarities based on the effect of the interaction terms on the probability of an alliance to occur. Interpretation of an interaction term in a logit model requires computation of the change in predicted probability of alliance formation for a given change in the value of alliance partner attributes either by direct calculation of cross-partial derivatives (Hoetker, 2007; Norton, Wang & Ai, 2004) or simulation-based inferences (King, Tomz & Wittenberg, 2000; Zelner, 2009). We used the cross-partial derivatives approach and compared the effect of interactions (what has been attributed to complementarities in the existing literature) with the true matching function. As Norton, Wang and Ai (2004) note, the interaction effects will take on different values across the data range, and will be positive or negative at different data points.[8] Accordingly, in Table 5, we report the key statistics for the distribution of interaction effects over the range of data. Note that for each simulation run, there are 25,000 different interaction term values, corresponding to each observation in the data. The rows reported in Table 5 show the average, across 500 runs, of the mean, standard deviation, minimum and maximum values of the interaction effects. These results show that interaction effects vary widely in terms of the sign—thus, scholars relying on binary choice models to infer whether there are complementarities or substitution across partner characteristics may be led to believe the opposite is true. This inference is even more problematic than the actual values of the coefficients themselves, because scholars will be led to believe that the reverse relationship holds among partner attributes.

[Table 5 about here]

APPLICATION: PARTNER SELECTION IN BIO-PHARMACEUTICAL INDUSTRY

While the earlier section illustrated the performance of matching models in estimating coefficients when the true underlying model is known, in this section, we demonstrate its use, with empirical data on the bio-pharmaceutical context, to examine factors that affect partner selection and complementarities in partner attributes.

---

[8] When plotted against the predicted probability values, the interaction effect has an S-shaped curve, with negative and positive values (Norton, Wang and Ai, 2004). Thus, even for the cross partial derivative calculated interaction effects, this functional relationship creates problems in interpretation of the interaction effect's sign as demonstrating complementarities or substitution.

Research Setting

We focus on the inter-firm alliances within the context of the bio-pharmaceutical industry for a number of reasons. First, alliances are widely used by firms to access resources and capabilities across firm boundaries in this context, with scholars reporting that the industry accounts for about 20 percent of all alliances observed in high technology industries (Rothaermel & Boeker, 2008). Second, the industry context has been used by strategic management scholars to study various aspects of inter-firm alliances in general (Lane & Lubatkin, 1998; Powell, Koput & Smith-Doerr, 1996; Rothaermel & Deeds, 2004), and partner selection in particular (Alcacer et al., 2009; Diestre & Rajagopalan, 2012; Rothaermel & Boeker, 2008). Thus, the usage of the context enables us to clearly focus on the implications of using a two-sided matching model from the empirical perspective.

Data Sources and Sample Construction

Our main data source for inter-firm alliances is the *ReCap* database. Additional data on firm characteristics (e.g. age, size, patent stock, etc.) were compiled, as indicated in greater detail in the variable description section, from a variety of sources including the *BioScan Directory*, the *BioCentury Online Intelligence*, the *Bloomberg-BusinessWeek Private Firms' directory*, *Compustat, Directory of Corporate Affiliations, drugs@FDA* database*, FDA Orange Book*, *Hoover's* database, the *LexisNexis* database, *Mergent*, and the *NBER* 2006 patent database. The data on alliances span the 1994-2006 time frame, a period representing significant alliance activity to permit the estimation of the coefficients for every year, which represents the "market" within which the alliances are formed. The time period of our sample compares well to those used in other studies: Rothaermel & Boeker (2008) studied the industry during the 1998-2001 period, and Diestre & Rajagopalan (2012) used the 2002-2007 period[9]. To enable clean identification of both sides of the "market" for alliances, we continue with the upstream-downstream context by focusing on research collaborations with a licensing component, so that firms engaged in an alliance could be designated as either a licensor (upstream) or a licensee (downstream). We believe the research component of the alliance makes the matching

[9] While Alcacer et al. (2009) use data from the late eighties, their analysis is not at the dyad-year. We omit the earlier years in the bio-pharmaceutical industry given lack of sufficient yearly observations in the pre-1995 period.

approach more suitable than treating these relationships as spot transactions because firms are more likely to care about identity of their partners in research rather than in non-research licensing agreements (where the number of transacting partners might be more important). We note of course that a firm can be a licensor in any one alliance, and a licensee in another alliance. Further, the same firm could engage in multiple alliances in any one year. After carefully reading the scope of each alliance we concluded that multiple alliances in the same year by a firm constituted different transactions for different technologies (disease areas and stages of development). We thus assumed that alliance decisions were separate decisions taken by these firms, which allowed us to treat our sample as one-to-one matching (rather than one to many when one licensor has multiple licensees for the same technology). We also limit the sample to non-equity alliances (as opposed to including both equity and non-equity), so that we have a homogenous set of alliances. To summarize, there were 1,551 non-equity, research licensing alliances in the 1994-2006 period, representing a total of 1,153 firms. After accounting for missing data on some variables, the final alliance sample consists of 1,068 alliances between 774 total firms. We describe the data sources for each variable below.

Variables

The explanatory variables used in our analysis represent well-established relationships within the field of the strategic management. We note that our objective in this paper is not to hypothesize about novel factors, but rather to show how the matching model technique can be used in well researched contexts. Accordingly, we refrain from developing explicit hypotheses, and refer interested readers to studies that provide the theoretical justification for the relevant variables (Alcacer et al., 2009; Diestre & Rajagopalan, 2012; Rothaermel & Boeker, 2008). Given the focus of extant work on alliance partner selection on factors such as firm age, firm size, technological characteristics, and geographic distance, we conduct our analysis using these variables. Given few instances of repeat licensing alliances deals in the data, and in line with Alcacer et al. (2009) who also focus on licensing deals, we do not include prior history of tie formation as a separate variable.

Firm Size: We measure firm size using the (logged) number of employees in the focal year. We obtained the information regarding firm size from various databases. Given our interest in the number of employees at the year of the licensing deal (and not any year), we started with the information available at the *Directory of Corporate Affiliations* and *Compustat* historical databases. In cases where the number of employees at the year of licensing was not available, we instead searched for the number of employees at the closest available year, number of employees in 2012 or number of employees at the last year of firm activity. We used a multiple imputation technique in Stata to infer the number of employees for the year of licensing using information on firm age, number of patents and drugs, private and public status, and variables capturing M&A activity.

Firm Age: We measure firm age using years since founding year (logged). We obtained the information about a firm's founding year using multiple databases such as the *Directory of Corporate Affiliations*, the *BioScan Directory*, the *BioCentury Online Intelligence*, the *Bloomberg Businessweek Private Firms' Directory*, *Hoover's* database, *Mergent*, and news releases in the *LexisNexis* database.

Patent Stock: We measure patenting stock as the (logged) number of a firm's patents, weighted by the number of forward citations, in the three-year window prior to formation of the licensing deal, and use a ten percent depreciation to allow for more recent patents to have greater weights (Hall, et al., 2005). We use data from the *NBER 2006 patent project* for this variable. In creating patent-based measures, we account for the organizational structure of firms and their history of acquisitions in the preceding years, by incorporating information about a firm's history of acquisitions from the ReCap database and supplementing the information in the *Directory of Corporate Affiliations*.

Number of Approved Drugs: We measure this variable as the (logged) stock of a firm's FDA-approved drugs in the ten-year window prior to formation of the licensing deal (Diestre & Rajagopalan, 2012; Nerkar & Roberts, 2004). Similar to the patent stock, we use a ten percent depreciation rate. The information about a firm's approved drugs is gathered from the *FDA Orange Book* and the *drugs@FDA* database. As described above for the patent measures, we account for the organizational structure of firms and their history of acquisitions in the preceding years.

Geographic Distance: This is a pair-specific characteristic measured as a dummy variable equal one if the two alliance partners are located in the same state (for US firms) or the same country (for foreign firm), and zero otherwise. The data stems from multiple databases, including the *Directory of*

Corporate Affiliations, the *BioScan Directory,* the *BioCentury Online Intelligence, the Bloomberg BusinessWeek Private Firms' Directory, Hoover's* database*, Mergent,* and news releases in the *LexisNexis* database.

Technological Overlap: Based on the patent portfolio of each firm, we create a pair characteristic of the technological overlap, as measured by the number of shared international patent class (IPC) codes in which both firms have patenting activity (Diestre & Rajagopalan, 2012). In alternative specifications, we used the measures in Rothaemel and Boeker (2008) to find similar results.

Table 6 provides the descriptive statistics and the correlation matrix.

Model Specification

As described in greater detail above, we estimate the coefficients using a maximum score estimation approach developed by Fox (2010a, 2010b). We define each year as a separate market for alliance so that the set of counterfactual partners come from the list of firms that had alliance activity within that year. We use yearly markets rather than treating the entire set of alliances as one big market given that the bias and the root mean-squared errors are smaller in data from multiple, smaller markets than one large market (Fox, 2010a). The average number of alliances across the 13 markets is 82. In the earlier years of our sample we observe less alliances (with a minimum of 14 in 1994 but more than 40 in each year after 1995) and a maximum of 162 in 2006. In order to create confidence intervals, we used 200 random samples of 4 markets at a time.

Results

Table 7 provides the estimated sign of the baseline and the estimated values of coefficients and the 95 percent confidence intervals. As noted above, the maximum score approach estimates only the interaction terms, under the assumption that the "un-interacted" characteristics are valued equally by all potential partners, and that matching is driven entirely by the interaction between the characteristics of the partners. A positive coefficient suggests that the two attributes are complements, while a negative coefficient suggests that the two variables are substitutes. We include multiplication of firm-level attributes in order to assess complementarity. Since we need scale normalization by assuming that one of the coefficients in the production function equals 1, we set the coefficient for firm size as our baseline, and compare the other coefficients to this base line.

[Tables 6-7 about here]

Our results indicate substitution in terms of age and complementarities in drug development capabilities. We thus infer that experience in one firm substitutes for lack thereof in another, given that younger firms match with older firms. There is a positive assortative matching in drug development capabilities; higher values of technological capabilities in one firm increases the marginal value of the technological capabilities of the other. Coefficient estimates allow us to compare the increase in the match revenue caused by one unit change in the interacted characteristics relative to the increase in the match revenue caused by one unit change in the baseline relationship (i.e. the relationship that represents the scale of the payoff function). In this example, a one unit increase in the baseline is given by a one percent increase in the size of the firms in a relationship. Therefore, the marginal impact of one percent increase in the cumulated number of partners' FDA-approved drugs over a decade is seventeen times higher than the marginal impact on the match surplus of a one percent increase in the number of employees.

DISCUSSION AND CONCLUSIONS

Complementarities in partner attributes represent an important question in the strategic alliances literature, since value creation in alliances is contingent on them. Matching models are particularly conducive to the study of "who partners with whom," since alliances represent the transaction of indivisible and heterogeneous goods when the identity (i.e. partner attributes) of the agents of both sides of the market matter, and the relatively small numbers of agents imply that there is competition on each side of the market for the "best" available partner. While matching models traditionally suffered from the curse of dimensionality, recent econometric advancements have significantly reduced the computational burden of estimating these models.

In this paper, we discussed how strategic management scholars may benefit from using these methods, which are gaining traction in economics and finance. Theoretically, matching models are more representative of alliance contexts since they not only explicitly account for both partners' preferences but also for the competition in each side for the "best partner." The latter permits scholars to infer that realized matches result not only because of true preferences, but also because of market constraints. The empirical comparison of the maximum score matching model estimator with the dominant binary choice model reveals that while the latter may be appropriate for predictive purposes, there are several limitations in its use for assessing complementarity in partner attributes. Accordingly, we hope that our paper has provided the basis for a future generation of research that examines the relationships using alternative methodologies. In doing so, our work may help scholars improve the answers, based on reliable research design. In particular, the use of matching models is in line with recent calls for identification, since they address selection and sorting effects, and permit scholars to distinguish the role of the distribution of exogenous partner characteristics from the role of the alliance as a "match production function." Thus, by assessing complementarities and substitutions of partner attributes, matching models permit better identification of the value creation through strategic alliances.

Our assertion in this paper rests on the assumption that even if scholars might not be willing to treat alliance formation as a matching game, there is a consensus in the literature that firms typically do not have the same choice set available. Further, in their most general forms, discrete choice models do allow astute scholars who have a theoretical prior on the distribution of choice set to integrate these complex structures and obtain estimates that accurately capture the decision making process. Our main claim is that it would be difficult, both theoretically and empirically, to specify ex-ante the structure of choice set for each firm on each side of the market. Therefore, we highlight the maximum score matching model estimation technique as we believe it builds on a theoretical perspective that captures the essence of the alliance formation process and it is flexible enough to be implemented by scholars without the need to build and solve complicated formal models. Ultimately, however, all models are approximations to reality and scholars need to make assumptions that they see appropriate for studying the context of interest.

Table 1: Empirical Specifications in Scholarly Work on Partner Selection

Scholarly work	Unit of Analysis	Choice Set for Unrealized Outcomes	Econometric specification	Inference for complementarity of partner attributes
A: Partner preferences matter = NO, Partner fit matters= NO				
Baum et al., (2005); Beckman et al., (2004); Katila et al., (2008); Stuart (1998)	Focal firm	Not applicable	random effects negative binomial, Poisson or GLS; GEE logit	No
Li, et al. (2008); Stern et al. (2013); Stuart (1998)	Dyad	Unconstrained	multinomial logit; probit; discrete time event history	No
Geringer (1991); Hitt et al., (2000); Hitt et al., (2004)	Focal firm	Not applicable	hierarchical linear model or correlation analysis	No
B: Partner preferences matter = NO, Partner fit matters= YES				
Shah & Swimanathan (2008)	Focal firm (in an experiment)	Not applicable	t-tests	No
Li & Rowley (2002)	Focal firm	Not applicable	logit	signs and significance of (a) pair specific variables and (b) interaction terms
Alcacer et al., (2009); Diestre & Rajagopalan (2012)	Dyad	Unconstrained	Conditional or GEE logit	No
Dushnitsky & Shaver (2009); Podolny (1994)	Dyad	Unconstrained	logit	signs and significance of pair specific variables
Hallen (2008); Sorenson & Stuart (2001)	Dyad	Constrained	rare events logit	signs and significance of (a) pair specific variables and (b) interaction terms
C: Partner preferences matter = YES, Partner fit matters= YES				
Garcia-Pont & Nohria (2002); Rosenkopf et al., (2001);)	Dyad	Unconstrained	logit;probit; negative binomial; competing risks	No
Ahuja et al., (2009); Chung et al., (2000); Gimeno (2004); Gulati (1995); Rosenkopf & Padula (2008); Rothearmel & Boeker (2008)	Dyad	Unconstrained	logit;probit; negative binomial; competing risks	signs and significance of (a) pair specific variables and (b) interaction terms
Reuer & Lahiri (2013); Sorenson & Stuart (2008)	Dyad	Constrained	logit or conditional logit, probit or random effects probit	No
Gulati & Garguilo (1999); Mitsuhashi & Greve (2009);	Dyad	Constrained	logit or conditional logit, probit or random effects probit	signs and significance of (a) pair specific variables and (b) interaction terms

There are no studies in the category of yes for partner preferences, and no for partner fit.

Table 2: Distribution of Key Statistics for Maximum Score Estimation

Attributes	Mean	Std. dev.	Min	Max
$X1_u$*$X1_d$	Fixed to 1	…	…	…
$X2_u$*$X2_d$	4.989928	.2474937	3.588092	6.269969
$X3_u$*$X3_d$	2.995729	.1541879	2.219962	3.669952
β_2/β_1 (true = 5)	4.989928	.2474937	3.588092	6.269969
β_3/β_1 (true = 3)	2.995729	.1541879	2.219962	3.669952
β_2/β_3 (true = 1.66)	1.667645	.0781835	1.370278	2.107691
% correctly predicted inequalities	46	179	44	48.5

* Simulation over 500 runs

Table 3: Distribution of Key Statistics for Logit Regression

(Only Interaction Terms)

Attributes	Mean	Std. dev.	Min	Max
$X1_u$*$X1_d$ (coefficient β_1)	.0590861	.006342	.038259	.0795019
$X2_u$*$X2_d$ (coefficient β_2)	-.073065	.013264	-.110059	-.009017
$X3_u$*$X3_d$ (coefficient β_3)	-.066947	.013134	-.118266	-.020055
β_2/β_1 (true = 5)	-1.25041	.26487	-2.87393	-.13274
β_3/β_1 (true = 3)	-1.14596	.255790	-2.01349	-.405727
β_2/β_3 (true = 1.66)	1.140683	.367626	.1362432	5.084715
Log-likelihood ratio	-16607.4	151.205	-16924.5	-16252.9
% correctly predicted outcomes	55.86	1.14468	53.15	58.79

* Simulation over 500 runs

Table 4: Distribution of Key Statistics for Logit Regression

(Simple and Interaction Terms)

True value	Mean	Std. dev.	Min	Max
Constant	4.33177	.053911	4.208925	4.502775
X1u	.288079	.025422	.213495	.361684
X1d	.183048	.021041	.130728	.254239
X2u	.001713	.051071	-.15509	.169813
X2d	-.00075	.05263	-.14839	.164317
X3u	-.00195	.047941	-.14674	.129519
X3d	-.00072	.049594	-.18369	.157618
$X1_u*X1_d$ (coefficient β_1)	-.11581	.008463	-.15603	-.09109
$X2_u*X2_d$ (coefficient β_2)	-.76192	.046747	-.91398	-.61851
$X3_u*X3_d$ (coefficient β_3)	-.72231	.042725	-.84639	-.61078
β_2/β_1 (true = 5)	6.60716	.550027	5.016406	8.412267
β_3/β_1 (true = 3)	6.265471	.534813	4.644454	8.556417
β_2/β_3 (true = 1.66)	1.05771	.079905	.799011	1.293881
Pseudo-R^2	.20421	.005746	.190652	.216105
Log-likelihood ratio	-1907.562	18.990	-1952.995	-1859.64
% correctly predicted outcomes	98.03	.0262	97.99	98.11

* Simulation over 500 runs

Table 5: Distribution of Key Statistics for Interaction Terms Effects for Logit Regressions

(Simple and Interaction Terms, Norton et al Approach)

True value	Mean	Std. dev.	Min	Max
$X2_u*X2_d$ (interaction effect, mean)	.01744	.001235	.01364	.02028
$X2_u*X2_d$ (interaction effect, max)	.47153	.044548	.35742	.57480
$X2_u*X2_d$ (interaction effect, min)	-.26852	.062349	-.39777	-.14517
$X2_u*X2_d$ (interaction effect, std. dev.)	.037517	.002925	.028161	.044590
$X3_u*X3_d$ (interaction effect, mean)	.016343	.001270	.014001	.020145
$X3_u*X3_d$ (interaction effect, max)	.430740	.042230	.358891	.540529
$X3_u*X3_d$ (interaction effect, min)	-.24146	.067230	-.38604	-.09939
$X3_u*X3_d$ (interaction effect, std. dev.)	.034457	.003100	.028419	.043394

* Simulation over 500 runs

Table 6: Descriptive Statistics for Bio-Pharmaceutical Alliances

		Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9	10
1	Licensee Size	8.57	2.75	0.69	13.18	1									
2	Licensee Age	3.44	1.28	0.00	5.82	0.64	1								
3	Licensee Patents	4.34	8.63	-9.21	17.47	0.47	0.24	1							
4	Licensee Drugs	-3.61	6.29	-9.21	6.08	0.62	0.37	0.47	1						
5	Licensor Size	4.77	1.75	0.00	12.73	0.01	0.00	0.00	0.00	1					
6	Licensor Age	2.05	0.79	0.00	5.33	-0.01	-0.00	-0.00	0.02	0.49	1				
7	Licensor Patents	0.62	7.25	-9.21	17.87	-0.01	-0.04	0.06	-0.02	0.32	0.21	1			
8	Licensor Drugs	-8.89	1.85	-9.21	5.56	-0.04	-0.02	-0.01	0.01	0.36	0.33	0.16	1		
9	Geographic Distance	0.60	0.39	0.00	1.00	0.09	0.11	0.00	0.07	0.00	-0.03	0.03	0.02	1	
10	Technological Overlap	2.06	3.29	0.00	23.00	0.26	0.16	0.37	0.19	0.27	0.26	0.50	0.11	0.03	1

Table 7: Matching Model Estimations of Complementarities in Partner Attributes

	Relationship	Coefficients	Confidence Interval
Baseline	Licensee Size * Licensor Size	1	
	Licensee Age * Licensor Age	**-3.85**	(-15.50 ; -1.52)
	Licensee Patents * Licensor Patents	-0.10	(-3.91 ; 0.46)
	Licensee Drugs * Licensor Drugs	**17.35**	(7.85 ; 57.07)
	Geographic Distance	-37.58	(-59.23 ; 86.76)
	Technological Overlap	3.92	(-10.89 ; 8.49)

References

Agarwal R, Croson R, Mahoney JT. 2010. The role of incentives and communication in strategic alliances: an experimental investigation. *Strategic Management Journal* **31**(4): 413–437.

Ahuja G. 2000. Collaboration networks, structural holes, and innovation: a longitudinal study. *Administrative Science Quarterly* **45**(3): 425–455.

Ahuja G, Polidoro F Jr., Mitchell W. 2009. Structural homophily or social asymmetry? The formation of alliances by poorly embedded firms. *Strategic Management Journal* **30**(9): 941–958.

Akkus O, Cookson T, Hortacsu A. 2012. The Determinants of Bank Mergers: A Revealed Preference Analysis. Working paper, University of Chicago.

Alcacer, J. Cantwell. J., Gittelman, M. 2009. Are licensing markets local? An analysis of the geography of vertical licensing agreements in Bio-Pharmaceuticals. *Working paper*, Harvard Business School.

Atakan AE. 2006. Assortative matching with explicit search costs. *Econometrica* **74**(3): 667-681.

Baccara M, Imrohoroglu A, Wilson A, Yariv A. 2012. A field study on matching with network externalities. *American Economic Review*: **102** (5): 1773-1804.

Baum J, Rowley TJ, Shipilov AV, Chuang YT. 2005. Dancing with strangers: aspiration performance and the search for underwriting syndicate partners. *Administrative Science Quarterly* **50**(4): 536–575.

Becker GS. 1973. A theory of marriage: part I. *Journal of Political Economy* **81**(4): 813–846.

Beckman CM, Haunschild PR, Phillips DJ. 2004. Friends or strangers? firm-specific uncertainty, market uncertainty, and network partner Selection. *Organization Science* **15**(3): 259–275.

Chatain O. 2013. Estimating Value Creation from Revealed Preferences: Application to Value-Based Strategy. Working Paper, University of Pennsylvania

Chen J, Song K. 2013. Two-sided matching in the loan market. *International Journal of Industrial Organization* **31**(2): 145-152.

Choo E, Siow A. 2006. Who Marries Whom and Why. *Journal of Political Economy* **114**(1): 175–201.

Chung SA, Singh H, Lee K. 2000. Complementarity, status similarity and social capital as drivers of alliance formation. *Strategic Management Journal* **21**(1): 1–22.

Cohen WM, Levinthal DA. 1990. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* **35**(1): 128–152.

Cockburn I, Griliches Z. 1988. Industry effects and appropriability measures in the stock market's valuation of R&D and patents. *The American Economic Review* **78**(2): 419–423.

Coleman(1991) Matching processes in the Labor Market. Acta Sociologica, 34:3-12.

Das TK, Teng B-S. 1998. Between Trust and Control: Developing Confidence in Partner Cooperation in Alliances. *The Academy of Management Review* **23**(3): 491–512.

Diestre L, Rajagopalan N. 2012. Are all 'sharks' dangerous? New biotechnology ventures and partner selection in R&D alliances. *Strategic Management Journal* **33**(10): 1115–1134.

Dushnitsky G, Shaver JM. 2009. Limitations to interorganizational knowledge acquisition: the paradox of corporate venture capital. *Strategic Management Journal* **30**(10): 1045–1064.

Eisenhardt, K.M., Schoonhoven, C.B. 1996. Resource-based view of strategic alliance formation: strategic and social effects in entrepreneurial firms. *Organization Science*, **7**(2): 136-150.

Fernando CS, Gatchev VA, Spindt PA. 2005. Wanna dance? how firms and underwriters choose each other. *The Journal of Finance* **60**(5): 2437–2469.

Fox JT. 2010a. Identification in matching games. *Quantitative Economics* **1**(2): 203–254.

Fox JT. 2010b. Estimating matching games with transfers. NBER working paper 14382, National Bureau of Economic Research, Cambridge, MA. Available at: http://www.nber.org/papers/w14382.

Fox JT, Bajari P. 2013. Measuring the efficiency of an FCC spectrum auction. *American Economic Journal: Microeconomics* **5**(1): 100–146.

Fox JT, Yang C. 2012. Unobserved heterogeneity in matching games. NBER working paper 18168, National Bureau of Economic Research, Cambridge, MA. Available at: http://www.nber.org/papers/w18168.

Gabaix X, Landier A. 2008. Why Has CEO Pay Increased So Much? *Quarterly Journal of Economics* **123**(1): 49–100.

Gale D, Shapley LS. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly* **69**(1): 9–15.

Garcia-Pont C, Nohria N. 2002. Local versus global mimetism: the dynamics of alliance formation in the automobile industry. *Strategic Management Journal* **23**(4): 307–321.

Geringer JM. 1991. Strategic determinants of partner selection criteria in international joint ventures. *Journal of International Business Studies* **22**(1): 41–62.

Gimeno J. 2004. Competition within and between networks: the contingent effect of competitive embeddedness on alliance formation. *The Academy of Management Journal* **47**(6): 820–842.

Gittelman M. 2013. Serendipity, skills and learning across space: The effect of general and specific knowledge on the geographic and technological proximity of inter-firm alliances. Working paper, Rutgers Business School.

Greve HR, Mitsuhashi H, Baum JAC. 2013. Greener pastures: outside options and strategic alliance withdrawal. *Organization Science* **24**(1): 79–98.

Gulati R. 1995. Social structure and alliance formation patterns: a longitudinal analysis. *Administrative Science Quarterly* **40**(4): 619–652.

Gulati R, Gargiulo M. 1999. Where do interorganizational networks come from? *American Journal of Sociology* **104**(5): 1439–1493.

Gulati R, Nohria N, Zaheer A. 2000. Strategic networks. *Strategic Management Journal* **21**(3): 203–215.Hall BH, Jaffe A, Trajtenberg M. 2005. Market value and patent citations. *The Rand Journal of Economics* **36**(1): 16–38.

Hallen BL. 2008. The causes and consequences of the initial network positions of new organizations: from whom do entrepreneurs receive investments? *Administrative Science Quarterly* **53**(4): 685–718.

Hamel G. 1991. Competition for competence and inter-partner learning within international strategic alliances. *Strategic Management Journal* **12**(S1): 83–103.

Hitt MA, Ahlstrom D, Dacin MT, Levitas E, Svobodina L. 2004. The institutional effects on strategic alliance partner selection in transition economies: china vs. russia. *Organization Science* **15**(2): 173–185.

Hitt MA, Dacin MT, Levitas E, Arregle J-L, Borza A. 2000. Partner selection in emerging and

developed market contexts: resource-based and organizational learning perspectives. *The Academy of Management Journal* **43**(3): 449–467.
Hoetker G. 2007. The use of logit and probit models in strategic management research: critical issues. *Strategic Management Journal* **28**(4): 331–343.
Ireland RD, Hitt M, Vaidyanath D. 2002. Alliance management as a source of competitive advantage. *Journal of Management* **28**(3): 413–446.
Kale P, Dyer JH, Singh H. 2002. Alliance capability, stock market response, and long-term alliance success: the role of the alliance function. *Strategic Management Journal* **23**(8): 747–767.
Katila R, Rosenberger JD, Eisenhardt KM. 2008. Swimming with sharks: technology ventures, defense mechanisms and corporate relationships. *Administrative Science Quarterly* **53**(2): 295–332.
Khanna T, Gulati R, Nohria N. 1998. The dynamics of learning alliances: competition, cooperation, and relative scope. *Strategic Management Journal* **19**(3): 193–210.
Kim JW, Higgins MC. 2007. Where do alliances come from? *Research Policy* **36**(4): 499–514.
King G, Tomz M, Wittenberg J. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2): 347–61.
Koopmans TC, Beckmann M. 1957. Assignment problems and the location of economic activities. *Econometrica* **25**(1): 53–76.
Lane PJ, Lubatkin M. 1998. Relative absorptive capacity and interorganizational learning*Strategic Management Journal* 19(5): 461-477.
Lavie D. 2007. Alliance portfolios and firm performance: a study of value creation and appropriation in the U.S. software industry. *Strategic Management Journal* **28**(12): 1187–1212.
Levine AA. 2009. Licensing and scale economies in the biotechnology pharmaceutical industry. Working paper, Harvard University.
Li D, Eden L, Hitt MA, Ireland RD. 2008. Friends, acquaintances, or strangers? partner selection in R&D alliances. *Academy of Management Journal* **51**(2): 315–334.
Li S, Rowley TJ. 2002. Inertia and evaluation mechanisms in interorganizational partner selection: syndicate formation among U.S. investment banks. *The Academy of Management Journal* **45**(6): 1104–1119.
Mason R, Drakeman DL. 2013. Fishing for sharks: Partner selection in biopharmaceutical R&D alliances. *Strategic Management Journal* (forthcoming)
Milgrom P, Roberts J. 1995. The economics of modern manufacturing: reply. *The American Economic Review* **85**(4): 997–999.
Mindruta D. 2013. Value creation in university-firm research collaborations: a matching approach. *Strategic Management Journal* **34**(6): 644–665.
Mitsuhashi H, Greve HR. 2009. A matching theory of alliance formation and organizational success: complementarity and compatibility. *Academy of Management Journal* **52**(5): 975–995.
Mortensen DT. 1978. Specific capital and labor turnover. *The Bell Journal of Economics* **9**(2): 572–586.
Mortensen DT. 1988. Matching: finding a partner for life or otherwise. *American Journal of Sociology* **94**: S215-S240.
Mowery DC, Oxley JE, Silverman BS. 1996. Strategic alliances and interfirm knowledge transfer. *Strategic Management Journal*, Winter Special Issue **17**: 77–91.

Nerkar A, Roberts PW. 2004. Technological and product-market experience and the success of new product introductions in the pharmaceutical industry. *Strategic Management Journal* **25**(8/9): 779–799.

Nohria N, Garcia-Pont C. 1991. Global strategic linkages and industry structure. *Strategic Management Journal,* Summer Special Issue **12**:105–124.

Norton EC, Wang H, Ai C. 2004. Computing interaction effects and etandard errors in logit and probit models. *The Stata Journal* **4**(2): 154-167.

Ostrovsky M. 2008. Stability in Supply Chain Networks. *American Economic Review* **98**: 897–923.

Oxley JE. 1997. Appropriability hazards and governance in strategic alliances: a transaction cost approach. *Journal of Law, Economics, & Organization* **13**(2): 387–409.

Pan Y. 2010. The determinants and impact of executive-firm matches. Working Paper, University of Minnesota.

Pfeffer J, Nowak P. 1976. Joint ventures and interorganizational interdependence. *Administrative Science Quarterly* **21**(3): 398–418.

Podolny JM. 1994. Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly* **39**(3): 458–483.

Powell WW, Koput KW, Smith-Doerr L. 1996. Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. *Administrative Science Quarterly* **41**(1): 116–145.

Reuer JJ, Lahiri N. 2013. Searching for alliance partners: effects of geographic distance on the formation of R&D collaborations. *Organization Science* (forthcoming).

Rochet, JC., Tirole, J. 2003. Platform competition in two-sided markets. Journal of European Economic Association, **1**(4): 990-1029.

Rosenkopf L, Almeida P. 2003. Overcoming local search through alliances and mobility. *Management Science* **49**(6): 751–766.

Rosenkopf L, Metiu A, George VP. 2001. From the bottom up? technical committee activity and alliance formation. *Administrative Science Quarterly*. **46**(4): 748–772.

Rosenkopf L, Padula G. 2008. Investigating the microstructure of network evolution: alliance formation in the mobile communications industry. *Organization Science* **19**(5): 669–687.

Roth AE. 1984. Stability and Polarization of Interests in Job Matching. *Econometrica* **52**(1): 47–58.

Roth AE, Sotomayor MAO. 1990. *Two-Sided Matching*. Cambridge University Press: New York, NY.

Rothaermel FT, Boeker W. 2008. Old technology meets new technology: complementarities, similarities, and alliance formation. *Strategic Management Journal* **29**(1): 47–77.

Rothaermel FT, Deeds DL. 2004. Exploration and exploitation alliances in biotechnology: a system of new product development. *Strategic Management Journal* **25**(3): 201–221.

Sampson RC. 2007. R&D alliances and firm performance: the impact of technological diversity and alliance organization on innovation. *The Academy of Management Journal* **50**(2): 364–386.

Shah RH, Swaminathan V. 2008. Factors influencing partner selection in strategic alliances: the moderating role of alliance context. *Strategic Management Journal* **29**(5): 471–494.

Shan W, Hamilton W. 1991. Country-specific advantage and international cooperation. *Strategic Management Journal* **12**(6): 419–432.

Shapley LS, Shubik M. 1971. The assignment game I: The core. *International Journal of Game Theory* **1**(1): 111–130.

Shimer R, Smith L. 2000. Assortative matching and search. *Econometrica* **68**(2): 343–369.
Sørensen M. 2007. How smart is smart money? A two-sided matching model of venture capital. *Journal of Finance* **62**(6): 2725–2762.
Sorenson O, Stuart TE. 2001. Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology* **106**(6): 1546–1588.
Sorenson O, Stuart TE. 2008. Bringing the context back in: settings and the search for syndicate partners in venture capital investment networks. *Administrative Science Quarterly* **53**(2): 266–294.
Sotomayor M. 1999. The lattice structure of the set of stable outcomes of the multiple partners assignment game. *International Journal of Game Theory* **28**:567-58.
Sotomayor M. 2004. Implementation in the many-to-many matching market. *Games and Economic Behavior* **46**(1): 199–212.
Stern I, Dukerich JM, Zajac E. 2013. Unmixed signals: how reputation and status affect alliance formation. *Strategic Management Journal* (forthcoming).
Stovel K, Fountain C. 2009. The Social Dynamics of Matching Processes. The Handbook of Analytic Sociology. Oxford University Press.
Stuart TE. 1998. Network Positions and Propensities to Collaborate: An Investigation of Strategic Alliance Formation in a High-Technology Industry. *Administrative Science Quarterly* **43**(3): 668–698.
Tervio M. 2008. The Difference That CEOs Make: An Assignment Model Approach. *American Economic Review* **98**(3): 642-668.
Norton, E, Wang, H., Ai, C. 2004. Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, 4(2): 154-167.
Train K. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press: Cambridge, MA.
Wang L, Zajac EJ. 2007. Alliance or acquisition? A dyadic perspective on interfirm resource combinations. *Strategic Management Journal* **28**(13): 1291–1317.
Yang Y, Shi M, Goldfarb A. 2009. Estimating the value of brand alliances in professional team sports. *Marketing Science* **28**(6): 1095–1111.
Zelner BA. 2009. Using simulation to interpret results from logit, probit, and other nonlinear models. *Strategic Management Journal* **30**(12): 1335–1348.

A Two-sided Matching Approach for Partner Selection and Assessing Complementarities in Inter-firm Alliances

APPENDIX A

Overview of Maximum Score Estimator and Implication for Inference of Complementarities in Alliance Formation

We present here what we consider to be the most relevant mathematical formulae behind the maximum score estimator. To make the exposition tractable, we focus on one-to-one matching. For a more comprehensive description of the estimator the reader should consult Fox (2010a, 2010b). Let $\boldsymbol{N_h}$ be the total number of alliances in a market and $\boldsymbol{h}$ a market index which takes values from 1 to M, where M is the total number of observed markets. We denote downstream firms by $\mathbf{d}_i$, i= 1, ... $\boldsymbol{N_h}$ and upstream firms by $\mathbf{u}_j$, **j**= 1, … $\boldsymbol{N_h}$. A partnership is characterized by a "production function" f that relates total output Z to different inputs:

$$Z_{uidi}=f(u_i, d_i \mid \beta) = \beta\ [X_{ui} * X_{di}] + \varepsilon_{uidi} \tag{1}$$

where X_{ui} represents the vector of upstream firms' characteristics and X_{di} the vector of downstream firms' characteristics and ε_{uidi} match error term. The purpose of the method is to estimate parameters $\boldsymbol{\beta}$ that describe the relationship between the interacted attributes of the partners.

As we explained in the main text, the mechanism behind the estimation is driven by the equilibrium properties of a matching game. These properties can be summarized as follows:

Let V denote the payoff that a firm receives from a match and t the transfer (payments and other contributions) it has to make to the other alliance partner. For the brevity of exposition we will assume that downstream firms make a transfer towards upstream firms t_{ud} (transfers can be positive or negative). Thus, the downstream firm d_i's net payoff from a match with u_i is $V_d\ (u_i, d_i)$-t_{uidi} and the upstream firm u_i's net payoff from the match is V_u (ui, di) + $t_{uidi.}$. The match surplus $f(u_i, d_i)$ is the summation of the individual (pre-transfer) payoffs that firms receive in a match:

$$f(u_i, d_i)= V_u\ (u_i, d_i) + V_d\ (u_i, d_i) \tag{2}$$

The equilibrium condition (pairwise stability) in a matching market states that none of the matched firms has an incentive to unilaterally deviate, that is, to form a relationship with another partner in the market.

Consider two actual matches (u_i, d_i) and (u_j, dj). Here, upstream uj is a counterfactual partner for downstream firm d_i and downstream dj firm is a counterfactual partner for upstream u_i. Because the downstream firm d_i maximizes its payoffs across potential upstream partners, the net payoff of downstream d_i from the actual match with u_i should be greater than d_i's net payoff from an alliance with u_j:

$$V_d\,(u_i, d_i) - t_{uidi} > V_d\,(u_j, d_i) - t_{ujdi} \qquad (3)$$

The same is true for u_i: its payoff from the actual match with downstream firm d_i should be greater than its payoff from a partnership with another downstream firm d_j:

$$V_u\,(u_i, d_i) + t_{uidi} > V_u\,(u_i, d_j) + t_{uidj} \qquad (4)$$

In interpreting relationship (4), one should think of t_{uidj} as the maximum transfer that downstream d_j is willing to pay to upstream firm u_i when firm d_j is contemplating to switch from its current partner (which is u_j) to u_i. The key idea captured by the equilibrium condition is that even if firm d_j offers to pay this maximum transfer to upstream firm u_i, firm u_i still prefers to pair up with d_i and receive t_{uidi} instead of pairing up with d_j and receiving t_{uidj} .

By summing up inequalities (3) and (4) we obtain the following inequality:

$$V_d\,(u_i, d_i) + V_u\,(u_i, d_i) > V_d\,(u_j, d_i) - t_{ujdi} + V_u\,(u_i, d_j) + t_{uidj} \qquad (5)$$

But note that relationship (5) should hold also for the match (u_j, dj). Thus, we obtain:

$$V_d\,(u_j, d_j) + V_u\,(u_j, d_j) > V_d\,(u_i, d_j) - t_{uidj} + V_u\,(u_j, d_i) + t_{ujdi} \qquad (6)$$

Inequalities (5) and (6) together lead (after rearranging terms) to the following relationship:

$$V_d\,(u_i, d_i) + V_u\,(u_i, d_i) + V_d\,(u_j, d_j) + V_u\,(u_j, d_j) > V_d\,(u_i, d_j) + V_u\,(u_i, d_j) + V_d\,(u_j, d_i) + V_u\,(u_j, d_i) \quad (7)$$

As stated in (2), the sum of payoffs V_d and V_u for match partners equals the match output produced by the two partners. Therefore, by applying (2) for the matched pairs (u_i, d_i) and (u_j, dj) and for the counterfactual pairs (u_i, d_j) and (u_j, d_i) we obtain directly from (7) the *local maximization condition*:

$f(u_i, d_i \mid \beta) + f(u_j, d_j \mid \beta) > f(u_i, d_j \mid \beta) + f(u_j, d_i \mid \beta)$ (8)

In the final relationship (8), all transfer payments *t* cancel out, which means that scholars do not have to observe transfer payments in order to estimate the match revenue function. However, underneath the local maximization condition is the intuition (stemming from the equilibrium condition) that each of the alliance partners obtains greater net payoffs from their actual match than elsewhere. Therefore, at market level, the observed matches are the ones that maximize the sum of joint outputs, and not the ones where some agents get better matches or higher surplus by exercising higher bargaining power. It is in this sense that matching models are about value creation and not about value capture. Once the market-level matching has formed, the division of the output is determined through a Nash-bargaining game at pair-level. It is possible that a unique matching equilibrium is supported by multiple ways of sharing the surplus among partners.

The estimation requires checking the location maximization conditions for all possible combinations of observed and counterfactual pairs in all markets observed in the sample by the scholar. To estimate the unknown parameters **β,** the method maximizes the following objective function:

$$\mathbf{Q(\beta) = \frac{1}{M} * \sum_{h=1}^{M} \sum_{1<i,j<N_h} 1[f\,(u^h_i, d^h_i \mid \beta) + f\,(u^h_j, d^h_j \mid \beta) > f\,(u^h_i, d^h_j \mid \beta) + f\,(u^h_j, d^h_i \mid \beta)]} \quad \mathbf{(9)}$$

where 1[˙] denotes an indicator function which takes a value of 1 if the expression between parenthesis is true and 0 otherwise, and the index *h*=1,..M denotes that a pair of {upstream, downstream} firms is situated in market *h*.

The method is semi-parametric, in the sense that it only requires the specification of a production function but it does not make any assumptions about distribution of the error term. This is possible because local maximization inequality does not impose constraints on the error term. Furthermore, Fox (2010a) proves that the local maximization condition produces a consistent estimator of unknown parameters **β** as long as the model satisfies Manski (1975)'s rank order property. A proof and a full discussion of the technical aspects related to the rank order condition can be found in Fox (2010a). In a nutshell, the rank order property states that if firms prefer to be in the configuration A of matches $\{(u_i, d_i), (u_j, d_j)\}$ as opposed to being in the configuration B of pairings $\{(u_i, d_j), (u_j, d_i)\}$, then the probability (to the econometrician) of observing configuration A is higher than the probability of observing B, after the error terms are drawn.

Note that the objective function in (9) is not smooth and thus, numerical techniques have to be employed in order to obtain parameters that maximize this function. Fox (2010b) recommends a global optimization routine known as differential evolution. This optimization method is built into packages such as Mathematica and MATLAB.

APPENDIX B

Overview of discrete choice models and implications for estimating partner choice in alliance formation

Discrete choice methods model the payoff of a firm to ally with another firm as a function of the characteristics of the firm making the choice, the characteristics of their potential partners (chosen or not) and some interaction terms between the attributes of the two sides. The estimates are derived by imposing the condition that a firm's payoff from allying with the chosen partner is greater than the payoff from allying with any other partner in a choice set.[10] This formulation is used to make probabilistic statements about the decision maker's choice of a partner. Scholars then proceed to make inferences on the probability that firms of certain attributes choose partners of certain attributes. We elaborate on these points below.

We use the terms downstream firm and upstream firm to indicate the two sides of the market. Assuming N alliances, each firm on each side of the market chooses the partner that will maximize its payoffs W. For each downstream firm d_i the payoff from choosing upstream firm u_j over all other possible choices must satisfy the condition that:

$$W_{ij} > W_{ik} \text{ for all } j \neq k \qquad i,j,k=1,\ldots N \qquad (1)$$

[10] We follow the standard description of discrete choice models, which is consistent with the assumption of payoff-maximizing behavior of the agent making a choice. As Train (2002) states, although this is the typical way in which models are presented, discrete-choice models can also be used to represent decision making that does not entail maximization. In other words, the model can be seen as simply describing a relationship between some explanatory variables and a choice outcome, where the scholar is not cognizant of how the choice was made. Nonetheless, even if we assume away the payoff maximization derivation, typical logit models are not able to account for the bilateral aspect of partner choice in alliance formation.

The payoff W_{ij} is not observed by the researcher (only the decision maker knows it). However, the researcher can specify an observable part O_{ij} which is a function of the attributes X_{ij} of the downstream firm d_i, the upstream firm u_j and both partners in interaction, and a pair-specific error term $\varepsilon_{ij.}$

$$W_{ij} = O_{ij} + \varepsilon_{ij} = \beta X_{ij} + \varepsilon_{ij} \quad (2)$$

The aim is to estimate parameters β of the model. This is done by the researcher making probability statements about the decision maker's choice. The probability that downstream firm d_i chooses upstream firm u_j among other alternatives is given by the following formula:

$$P_{ij}=\text{Prob}(W_{ij} > W_{ikj} \text{ for all } j \neq k) = \text{Prob}(O_{ij} + \varepsilon_{ij} > O_{ij} + \varepsilon_{ij} \text{ for all } j \neq k)=$$

$$\oint \mathrm{I}(O_{ij} + \varepsilon_{ij} > O_{ik} + \varepsilon_{ik} \text{ for all } j \neq k)\, \varphi(\varepsilon i)\ d\,\varepsilon_{i,} \quad (3)$$

where I is an indicator of whether the statement in parenthesis is true, ε_i is a vector of errors from all pairing choices of downstream firm d_i over upstream firms ($\varepsilon_{i1,}$ $\varepsilon_{i2\ ,}$ … , ε_{iN}), and $\varphi(\varepsilon_i)$ is the probability density function of ε_i. Thus, the probability is given by a N-dimensional integral over the N errors from all alternative partnerships available to downstream firm d_i. Different discrete choice models rely on different specifications of the probability density function. Logit assumes that $\varepsilon_{i1,}$ $\varepsilon_{i2\ ,}$ … , ε_{iN} are distributed i.i.d. extreme value. Probit assumes that φ(·) is multivariate normal.

To obtain an estimate of the unknown β coefficients, scholars need to calculate the probability of each downstream firm in the sample choosing the upstream firm that it was actually observed to choose. This is done by calculating a Likelihood function L(β), which is the joint density of all agents' choices in the sample[11]. The estimates of β are those that maximize the log-Likelihood function.

Various assumptions are needed to implement discrete choice methods (Train, 2002)[12]. A detailed discussion of all assumptions is beyond the scope of this paper, but we emphasize some

[11] Note that while our discussion so far referred to a joint probability density of the error term as a function of the data conditioned on the parameters, in the likelihood function, the joint density is a function of the unknown parameters, conditioned on the data.

[12] For example, both logit and probit assume a specific distribution of the error terms. Logit specification has a closed form functional form and the joint distribution is easier to calculate. Probit assumes a multivariate normal distribution that needs to be evaluated numerically through simulation. Picking the right distribution of errors is not an innocuous task. The two distributional forms differ not only on the difficulty to estimate the multidimensional joint density function of the errors, but also on subtle aspects related to choice behavior. Train (2002) describes how logit and probit deal with taste variation, substitution patterns and repeated choices. Particularly troublesome is the assumption of independence of irrelevant alternatives (IIA) in logit which overestimates the probability of choosing among equivalent alternatives.

aspects that are of concern when choices are made within a matching process. The most concerning of assumptions rest on the implicit idea that each decision maker can choose his or her most preferred partner(s) independently of: (a) the preferences of the partner(s), and (b) the choices of other decision makers in the market. These assumptions can be noticed in relationships (1) and (2) which indicate that standard discrete choice models compare the payoff of an agent from the observed partnership with the payoff from alternative partnerships. This is an adequate representation of many choice situations (e.g. firm choice of technologies, locations, or product markets). However, we believe it is too restrictive given that the choice set is composed of agents who also have preferences over whom to partner with. As Appendix A indicates, while logit models maximize the utility of individual agents, matching models formally characterize the observed assignment as the equilibrium outcome of a process that maximizes the sum of utilities of all pairs in the market.

The literature taking a dyadic approach to alliance formation has long criticized assumption (a). That one agent's choice is constrained by the decisions made by all other agents in the market has been less discussed in the literature. Our literature review corroborates this conclusion. Consider a situation where a ranking of downstream firms and upstream firms in terms of their marketing and manufacturing capabilities, respectively, exists in the market. Assume also that a downstream firm with better marketing capabilities has higher returns from partnering with an upstream firm with superior manufacturing capabilities (i.e. marketing and manufacturing capabilities are complements in alliance formation). Matching predicts a positive assortative outcome where we observe an alliance between the highest ranked downstream-upstream pair, another partnership between the second ranked firms and so forth to a match between the lowest ranked downstream and upstream firms. Some deviations from this outcome will always be observed in real datasets, but in expectation, most agents will match assortatively. Because the highest ranked firms create more value together than partnering with any other firms in the market, these two firms will not be part of the choice set of the remaining downstream and upstream firms seeking to form a match in the same market. Given that the best upstream firm has already formed a partnership with the best downstream firm, the choice set of the second best downstream firm is now smaller than the choice set of the most highly ranked downstream firm.[13] This leads to the obvious conjecture that the

[13] This is not strictly true in one-to-many or many-to-many matching, where the highest ranked upstream firm might have a bigger quota (i.e. it can work with multiple downstream firms) than in a typical one-to-one matching. However,

choice set of the higher ranked firms will differ greatly from the choice set of lower ranked firms. The same reasoning applies to all firms in the market, but it would be difficult to specify ex-ante a structure on the firm-specific constraints.

We are not aware of work within the discrete choice models tradition that deals with assumptions (a) and (b) above, at least not without resorting to formal modeling or context-specific solutions[14]. In the mainstream strategy literature, work within the dyadic perspective has rightfully acknowledged that alliance formation is subject to mutual choice but it has neither proposed an estimation method to account for this feature of the alliance process, nor acknowledged the further complication of inter-related choice constraints among all dyads in the market (an exception is Mitsuhashi & Greve, 2009). It is important to notice that while scholars have suggested choice-based or case-control sampling method as a way of dealing with the variation in the choice sets available to decision makers, sampling from correlated alternatives does not address the more fundamental aspects of choice constraints discussed here. Therefore, the current approaches that resort to removing alternatives which are rarely available, or to randomly sampling a number of unrealized choices, or more generally, to imposing a rule for picking the counterfactual partners do not address the fundamental problems raised (and solved) here in two sided matching models.

Our discussion should not lead to the conclusion that the underlying features of discrete choice models are theoretically unable to accommodate complex situations such as matching. One way to apply these models is to extend the reasoning from the single-agent choices to the configuration-level choices (i.e. the structure of pairings in a market, more formally known as "assignment"). Accordingly, scholars might proceed to evaluate the probability distribution over assignments given characteristics X of downstream firms and upstream firms. This formulation requires scholars to model why the observed assignment (and not any other alternative configurations) has emerged in the data. In a one-to-one matching market of N firms on each side, this approach entails evaluating an integral over the unobserved error terms in the market. This

---

there will always be restrictions for the lower ranked downstream firm once the higher ranked upstream firms have filled their quota.

[14] For example Choo and Siow(2006) restrain the number of covariates by using aggregate data of different types of men and women in a marriage market. Such an approach can only be followed if agents on each side of the market could be grouped into large, observationally equivalent classes. Other authors have used Bayesian estimation to evaluate the likelihood function (Sorensen, 2007), or the simulated method of moments (Boyd, Lankford, Loeb & Wyckoff, 2003; Gordon & Knight, 2009). However, simulation methods also need to enforce various restrictions to overcome the problem high dimensional integrals. For example, Gordon & Knight (2009) and Sorensen (2007) impose constraints on the preferences of the agents and Boyd et al (2003) impose limitations in the size of the market.

integral will be of dimension equal to all potential assignments N!. Furthermore, because the integral is not of closed-form expression, it must be evaluated numerically. Current optimization routines, however, typically fail to perform such a task even for sample sizes in the lower range of the alliance datasets we typically observe in the literature[15].

To understand the complexity of the estimation problem, consider a two-sided market with 3 downstream firms {d1, d2, d3} and 3 upstream firms {u1, u2, u3}. To simplify the notation, we will use ij to denote a partnership between d_i and u_j. There are 3^2=9 possible partnerships: {11}, {12}, {13, {21}, {22}, {23}, {31}, {32}, {33} and 3!=6 possible configurations of combinations between downstream firms and upstream firms:

{11, 22, 33}, {11, 23, 32}, {13, 22, 31}, {12, 21, 33}, {12, 23, 31}, {13, 21, 32}.

Assume we observe {11, 22, 33}. Standard logit/probit models take the perspective of one side of the market and evaluate the probability of a pair to occur. Let's take the match between downstream firm d1 and upstream firm u1. According to (1), it must be true in the logit/probit model that the payoff of downstream firm d1 from matching with upstream firm u1 is higher than the payoff from alternative partnerships: $P_{11} > P_{12}$ and $P_{11} > P_{13}$. Scholars then have to calculate the probability of pair {11} conditional on the observable characteristics of the downstream firm and upstream firm. This requires the evaluation of a N=3 -dimensional integral which represents the density function of the error vector (ε_{11}, ε_{12} , ε_{13}). The procedure is repeated for all remaining downstream firms, here d2 and d3. Note that this procedure evaluates the probability of a match only from the perspective of one agent at a time. The calculation of downstream firm d1's probability to match with upstream firm u1 neither takes into account u1's payoff from the match, nor the fact that the true configuration in the market is {11, 22, 33} (as opposed to {11, 23, 32}, a possible configuration that also contains pair {11}).

As we pointed out, the best way to proceed would be to calculate the probability of the equilibrium assignment {11, 22, 33} to occur, given the observed characteristics of downstream firms and upstream firms in the market. In our example, this approach requires evaluating a 3!–dimensional integral, but the procedure is not feasible once we try to apply it to typical datasets.

[15] As Fox (2010a) famously asserts, markets of only 100 pairs lead to 100! possible assignments, a number that is higher than the number of atoms in the universe.