Strategic Management Journal
Strat. Mgmt. J., **34**: 644–665 (2013)
Published online EarlyView in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/smj.2036
Received 16 July 2009; Final revision received 2 December 2012

VALUE CREATION IN UNIVERSITY-FIRM RESEARCH COLLABORATIONS: A MATCHING APPROACH

DENISA MINDRUTA*
HEC-Paris School of Management, Jouy-en-Josas, France

University-based technological opportunities are often exploited through joint corporate and academic entrepreneurship activities such as university–industry research collaborations. This paper explores the partner attributes that drive the matching of academic scientists and firms involved in these relationships. The paper models the formation of firm–faculty partnership as an endogenous selection process driven by synergy between partners' knowledge-creation capabilities. The main findings indicate that faculty–firm matching is multidimensional: firms and scientists complement each other in publishing capabilities but substitute each other in patenting skills. Furthermore, firms and scientists with specialized knowledge create more value by teaming with more knowledge-diversified partners. The paper contributes to the literature on university–industry knowledge transfer and, more generally, to the literature on alliance formation.

INTRODUCTION

Universities have always been an important source of new knowledge, and long-standing evidence shows that academic research has a significant impact on the productivity of private sector research and development (R&D) (e.g., Cohen, Nelson, and Walsh, 2002). Spillovers from published academic research and informal collaboration are influential channels of knowledge transfer, but recent studies suggest that firms also seek out new scientific knowledge and technological opportunities by engaging in corporate entrepreneurship activities such as licensing academic inventions and collaborative research (Thursby and Thursby, 2002). Of particular interest for the management field is evidence showing that firms with direct research ties to universities significantly increase their innovative performance (e.g., Cockburn and Henderson, 1998; Zucker, Darby, and Armstrong, 2002; Belderbos, Carree, and Lokshin, 2004; Fabrizio, 2009).

While the economic benefits of collaboration are well documented in the literature, the sources of value creation in university–industry research alliances remain underexplored. Much of the empirical work on alliances performance links an innovation output to an indicator measuring the presence of these relationships, while a few studies go deeper to examine the outcome of collaboration in relation to partner characteristics. This work typically regresses performance on the observed attributes of firms, scientists, or universities. However, this methodology is valid only when (1) all attributes of the partners that influence performance are included in the model and measured without errors, *or* when (2) alliance formation is a random process. The first condition is hardly ever satisfied in empirical settings. In strategy research, where great emphasis is placed on strategic, that is, *nonrandom*, choices (Shaver, 1998), the second condition is particularly problematic.

Keywords: matching; complementarity; endogeneity; value creation; university–industry alliances
*Correspondence to: Denisa Mindruta, HEC-Paris School of Management, 1 rue de la Liberation, Jouy-en-Josas, 78350 France. E-mail: mindruta@hec.fr

At the core of this study is the idea that partnership formation is a strategic decision. If partner choice is guided by preferences about whom to ally with, and motivated by the expectation that collaboration with certain partners will lead to superior performance, the process of partner selection is neither random nor independent of performance considerations. Put succinctly, if there are incentives whereby certain types of partners end up allying (i.e., they 'match' according to certain preferences), then partner attributes must be endogenous choice variables in regressing alliance performance on the characteristics of the partners. As a result, empirical models that do not account for the partner selection process will produce biased estimates and even skewed normative conclusions. These econometric concerns have been acknowledged in the alliance literature, but matching as a cause of endogeneity remains poorly understood.

To illustrate the importance of matching and its implications, consider Zucker *et al.*'s (1998, 2002) finding that partnerships with 'star' university scientists have a higher impact on firms' innovation performance than partnerships with non-star faculty. By concluding that the differential productivity of firm–star alliances is due to the added value university scientists (professors) contribute to the innovation process, these studies suggest that firms aspiring to increase performance should pursue research ties with top faculty. This conclusion raises several questions. Firms may prefer teaming-up with 'stars,' but are such collaborations equally attractive to scientists? Could scientists be equally productive no matter what firm or firms they collaborate with? Further, if success *is* a result of partnering with 'stars,' why would firms partner with 'non-star' faculty?

The existing literature has also accumulated compelling evidence that scientists value firms based on multiple factors, such as the strength of internal research, technological expertise, and organizational culture. As this paper will show, if scientists care about the quality of the firm, and firms care about scientists' research excellence, the interaction of these preferences will lead to a sorting of partners across alliances, causing 'better' scientists to work with 'better' firms. The endogeneity arises when top faculty teams with firms whose qualities *reinforce* their expertise and effort in innovation (i.e., scientists and firms match on complementary attributes) but the regression model does not include all relevant firm – quality dimensions. The error term then captures the effect of the omitted firm attributes and becomes correlated with faculty research status, overestimating star scientists' contribution to innovation. Conversely, when partners match on attributes that act as substitutes in innovation, the omitted attributes of one party will bias downward the estimates of the observed characteristics of the other party.[1]

The above example illustrates a large class of situations in which partner choice affects performance estimates. Researchers typically attempt to control for the confounding effects of partner selection through instrumental variables or experimental design. In contrast, this paper argues that faculty – firm partnering reveals important information about the sources of value creation driving cross-organizational alliances and, ultimately, their differential success in innovation. It introduces a theoretical model and a methodology that examine alliance formation as an endogenous matching process in which firms and scientists strategically select each other based on attributes that are important for knowledge creation. The model predicts positive (i.e., top-down) sorting of scientists and firms on attributes that are complements in the innovation process, and negative sorting on attributes that are substitutes.

I apply this approach to a proprietary database that contains the complete set of 447 discovery-type collaborations between university scientists at a top U.S. medical school and their industry partners during 1995–2004. Employing recent developments in econometrics, I estimate the model using a maximum score estimator (Fox, 2010). Parameters are identified based on mathematical inequalities that compare the output created by observed and counterfactual matches. These inequalities are derived from the equilibrium outcome of the matching model, which predicts that when partner choice occurs under rivalry constraints, the observed matches should create

[1] In general, the bias depends on the relationship between the included and omitted variables and the impact of the omitted variables on performance. In the example discussed here, the omitted variables measuring firms of higher quality have, most likely, a positive effect on innovation. Thus, the direction of the bias rests on the relationship between the omitted firm variables and scientists' status. In the case of complementarity between firms and top scientists, the omitted attributes of the firms will be positively related with scientists' status, leading to an upward bias impact of 'stars' on performance. In the case of substitution, attributes of the partners will be negatively related, and the impact of 'stars' on performance will be biased downward.

(weakly) higher value than counterfactual partnerships. The results show that firm – faculty research alliances create more value when partners complement each other in publishing capabilities and substitute each other for lack of patenting capabilities. There is also evidence of complementarity between knowledge-diversified and knowledge-specialized partners.

This paper extends the prior literature in several ways. The theoretical approach, rooted in the management theories that recognize heterogeneity in capabilities as a central feature of firms (e.g., Helfat *et al*., 2007), builds upon the conceptualization of two-sided matching markets to include strategic aspects of partner choice (Mortensen, 1988). Unlike previous work that focused on firms *or* universities, this paper considers alliance formation from the perspective of both partners. Furthermore, the analysis of partner choice goes beyond considerations of mutual attractiveness at the dyad level and takes into account the competition between participants on the same side of the market to ally with the most desirable exchange partners on the other side of the market.

This approach generalizes previous findings and explains empirical regularities that other models fail to clarify. While prior literature has emphasized the value-added contribution of star faculty to the innovativeness of partner firms, this paper draws attention to partner sorting in the market for research. The multidimensional sorting model captures the differences between participants' preferences and ability to attract partners more than other simpler, univariate models. Within this perspective, firm collaboration with non-star faculty is not a simple story of top-down sorting but a more nuanced reflection of heterogeneity in preferences and synergy types (complementarity vs. substitution) in innovation.[2]

THEORETICAL BACKGROUND

This paper examines how new knowledge is created when actors deliberately form new ties across organizations or sectors. Thus, although linkages between universities and industry can take multiple forms, here I focus specifically on formal research collaboration between university scientists and firms intended for scientific discovery.

Main concepts

Two key concepts are *complementarity* and *substitutability*. The mathematical definitions of these terms stipulate that two inputs are complements in production if the cross-partial derivative of the production function with respect to the two inputs is positive, and substitutes if the cross-partial derivative is negative.[3]

For this study, the research alliance 'production function' can be framed as a knowledge output whose inputs are partner attributes. Indeed, the objective of a research alliance is to create value (in the form of new knowledge or cross-learning among partners) beyond what scientists and firms would yield separately. The joint output depends on the combination of partner attributes, which may include strengths building upon strengths or the strengths of one partner substituting for the other partner's shortcomings. Conceptualizing the production function in this way implies that two attributes are complements in knowledge creation if having a higher level of one raises the return from having a higher level of the other. Conversely, two attributes are substitutes if having a higher level of one decreases the marginal value of having a higher level of the other.

The term *market for research* refers collectively to all transactions in which university scientists and firms agree upon joint research activities leading to the creation of new knowledge and technologies, without changing organizational boundaries or asset ownership. This study also advances the idea that the market for research is a *two-sided matching market*. *Two-sidedness* indicates that agents involved in a transaction belong

[2] For example, the finding of negative sorting on patenting skills predicts, *ceteris paribus*, that partnerships between firms with higher patenting activity and 'non-star' patenting scientists (i.e., scientists with fewer or no patents) are better matches than alliances between partners that both patent a lot. Thus, these partnerships are not the result of firms not having access to 'star patenting' scientists, but they reflect preferences for partner choice on dimensions that are substitutes in innovation.

[3] These terms have been used with many other meanings in the strategy literature (e.g., complementarity has been equated with compatibility, non- similarity, or non-overlap), but this paper employs the formal economic definition of these concepts. The formal definition has been applied to the study of organizational decisions related to firm boundaries (Parmigiani and Mitchell, 2009), the relationship among various organizational practices (Milgrom and Roberts, 1995), and among firm internal resources and R&D alliances (Arora and Gambardella, 1994; Cassiman and Veugelers, 2006).

DOI: 10.1002/smj

to one of two disjoint sets (here, firms and universities). By contrast, in a typical commodity market, an agent might be a buyer at some price and a seller at another price. *Matching* reflects the bilateral nature of exchange. For example, in labor markets, an employee works for a firm, and the firm employs the worker. As Roth and Sotomayor (1990:1) explain, this is in contrast with the markets for goods, 'in which someone may come to market with a truck full of wheat, and return home with a new tractor, even though the buyer of wheat doesn't sell tractors, and the seller of tractors didn't buy any wheat.' The concept of a two-sided matching market captures the theoretical underpinnings of the markets involving exchanges of heterogeneous and indivisible goods (Mortensen, 1988). In this regard, university – industry partnerships are a typical example. The inputs each partner brings to the relationship are highly differentiated: the expertise of university scientists is not a commodity, and neither are the capabilities of individual firms. Moreover, in this particular context, an exchange involves an indivisible *bundle* of distinctive characteristics that parties on one side of the market offer to parties on the other side. These two aspects—heterogeneity and indivisibility—lead to an important theoretical distinction. While the identity of agents who trade in commodity markets does not matter, it is important to understand whether agents in matching markets trade with the right partners. The critical question in these markets is *who trades with whom*.

A matching model of alliance formation

The following three features of university – industry alliances define the matching process:

(1) University-industry research alliances are voluntary relationships that form with the expectation of mutual gains. One of the university scientists interviewed for this paper described his decision to collaborate with industry in the following way:

> *[. . .] because I only have so much time, I'm not going to spend my time on something that I don't think is going to be productive. And I certainly will not have students and people in my group spend time on something that's not productive. So, I think that's where it comes down to. There is the mutual interest and then there have to be mutual benefits. And, if those aren't there, then there is no reason to do it.*

Likewise, as Stephan notes: 'firms can ill afford to fund research that has little promise of (eventually) relating to the company's objectives' (1996: 1211). Most commonly, research projects involving university partners go through a competitive selection process along with internal projects. Thus, it is reasonable to believe that only projects with justifiable benefits to the firm will be approved.

(2) The value of innovation generated through collaboration is determined, at least in expectation, by the identity of the university scientists and firms involved in a relationship. Consequently, both university scientists and firms have preferences about with whom they collaborate. In discovery-type collaborations, preferences are driven by the belief that joint work with partners of a certain type (such as those with high research productivity and expertise in a particular scientific field) will have a higher probability of generating new knowledge.

(3) Alliance partners are restricted in the number of collaborations they can undertake at one time. On the academic side, time constraints are one of the main reasons for a careful screening of the number and quality of industry partners. Moreover, research contracts often impose confidentiality conditions that prevent scientists from collaborating with multiple companies simultaneously (Kenney, 1986). Firms have a better ability to manage multiple contracts with external parties, up to the point where over-embeddedness becomes a liability (Uzzi, 1996). The management of alliances is a complex process facing numerous challenges, such as managing coordination among partners and creating an appropriate organizational structure for inter-partner learning and knowledge transfer. These constraints limit the number of simultaneous alliances firms can pursue.

Taken together, preferences for potential partners and restrictions on the number of collaborations suggest that, in the market for research, agents on each side of the market will compete to ally with the most desirable partner on the other side of the market. While prior literature has examined interfirm competition in innovation races, and scientists' competition for funding and academic prestige (Stern, 1995), firms' (respectively, scientists') competition for teaming up with the 'best' alliance partner has been largely ignored. This type of rivalry is extremely important in matching markets, where 'transactions' involve differentiated bundles of indivisible resources in limited supply.

From a theory standpoint, firm – scientist alliance formation belongs to a more general class of models (known as *assignment games*) that address the formation of voluntary partnerships among complementary agents under conditions of rivalry (Mortensen, 1988).[4] Economists have typically studied matching in marriage and labor markets. More recently, scholars have begun to apply these models to the study of various business relationships, including those between firms and initial public offering underwriters (Fernando, Gatchev, and Spindt, 2005), start-ups and venture capital investors (Sorensen, 2007), buyer–supplier contracting (Fox, 2009; Chatain, 2010), or the assignment of resources to firms (Lippman and Rumelt, 2003).

Matching theory explains what happens when all features of alliance formation—voluntary collaboration, two-sided decision making, and competition for better partners—are considered in interaction. From this perspective, partnership formation is a two-sided process. Each agent on one side of the market evaluates, and is simultaneously evaluated by, potential partners on the other side of the market. At the dyad level, and in isolation from the rest of the market, the decision rule is simple: a partnership occurs when both partners expect that collaborating will create more value than remaining single. However, under rivalry for partners and preferences on both sides, synergy at the dyad level is no longer a sufficient condition for an alliance if *any* of the partners could create a higher surplus with another agent on the other side of the market. At the market level, dyad-level decisions interact to constrain each other. On the one hand, the likelihood that an agent (firm or scientist) teams with his or her preferred partner is influenced by the existence of other agents on the same side of the market wishing to ally with the same partner. On the other hand, the likelihood of teaming with the preferred partner is also influenced by the alternative opportunities the partner might have to forego by entering a particular deal. In short, the decision of two individual agents to collaborate depends not only upon their preferences but also on their effective choice set, which is constrained by the decisions made by all other agents in the market. Ultimately, it is the interaction of preferences and competition at the market level that leads to the final matching.

The core theoretical properties of matching, in which the attributes of the partners are complements or substitutes in the match production function, were introduced to the literature by Becker (1973).[5] Becker's results lead to the following prediction: *ceteris paribus*, if two inputs or attributes (capabilities of a scientist and a firm, respectively) are complements in knowledge production, in equilibrium, firms and scientists scoring highest on these dimensions will work together, leaving firms and scientists scoring second to select each other, and so forth until the least-endowed agents select each other. Conversely, if two attributes are substitutes, then firms (scientists) scoring high on one attribute will match with scientists (firms) scoring low on the other attribute.[6]

The set of theorems proven by Becker (1973) constitute the basis for the empirical matching model. Drawing from innovation literature, I examine the matching of scientists and firms along three key dimensions in collaborative research: publishing capability, patenting capability, and the degree of knowledge specialization (focused vs. broad).

[4] An outcome in assignment games consists of an assignment (i.e., the identity of the matching parties) and a vector of payoffs (i.e., a division of surplus between the members of a matched pair). The equilibrium concept is that of *pairwise stability*, a notion that captures the idea that none of the partners to a match have an incentive to separate and team with other partners. *Optimality* means that unmatched agents could not form a partnership and make one better off without making the other worse off. Assignment games study questions such as which partnerships will form given the agents' preferences and the rules governing the market. These models are appropriate for studying the *ex ante* process of partnership formation. In contrast, prisoner's dilemma and assurance games are appropriate for studying the *ex post* dynamics of alliances, where partners decide how much to contribute to joint activity (Agarwal, Croson, and Mahoney, 2010).

[5] Becker considers the unidimensional case (one attribute on each side) and shows that if two inputs/attributes are complements (substitutes) in production, then positive (negative) assortative matching is *stable* and *optimal*.

[6] Some partnerships might look 'suboptimal' from the perspective of the dyad, such as, for example, alliances between two less-endowed agents (as in the case of positive assortative matching), or alliances between one highly endowed agent and one less-endowed agent (as in the case of negative assortative matching). However, these partnerships are 'optimal' when all linkages in the market are taken into account.

Hypotheses

As in most fields of modern science, biomedical research considered in this study is simultaneously motivated by scientific interest and considerations of usefulness and applications (Stephan, 2010). In this context (typically referred to as Pasteur's Quadrant) discoveries have *dual* scientific and applied/commercial value, providing opportunities for inventors to disclose their ideas through scientific publications, patented inventions, or both (Murray and Stern, 2007). As I discuss below, the choice of disclosure regime sends differing signals about inventors' research abilities and motivations (Hicks, 1995; Anton and Yao, 2003), and plays an important role in facilitating knowledge exchanges among academics and firms (Hellmann, 2007; Gans and Stern, 2010).

The norms of 'open science' reward priority in advancing novel ideas and provide incentives for full disclosure of scientific findings in the public domain. The scientific merit of a discovery is decided by the research community through the peer-review process and citations. Publishing is the hallmark of academia, but articles published by corporations have increased significantly in the last decades (Stephan, 2010). A salient motivation for corporate publishing is firms' desire to participate in knowledge exchanges within the broader scientific community (Hicks, 1995; Henderson and Cockburn, 1996). These publications serve two key roles. First, they indicate a firm's willingness to abide by the norms of reciprocation and full information disclosure (Murray, 2011). Second, they signal a firm's research quality (including the existence of tacit knowledge and unpublishable resources) and thus, provide the necessary 'credentials' for the firm to find exchange partners in the scientific community (Hicks, 1995).

The norms of 'commercial science' reward priority in producing ideas that are novel, useful and nonobvious. The choice of patent disclosure regime indicates that the inventor seeks to secure monopoly rights on ideas and will follow the logic of commercial practices in exchanges involving these ideas (Murray, 2011). Like publishing, patenting has a strong signaling value (Anton and Yao, 2003) and is practiced both in the industry and academia. Patents signal to other market participants the strength of a firm's proprietary knowledge and technical expertise in a field (Ahuja, 2000). In academia, patents indicate that a scientist's ideas have useful marketable applications, and are often perceived as a 'symbol of commercial savvy' (Murray, 2011).

Building on this literature, I refer to *publishing* and *patenting capabilities* to denote the ways firms and university scientists deploy their knowledge capital to signal expertise and preferences for the 'open,' and 'commercial science' regime, respectively.[7,8] By patenting or publishing in a narrower or a broader territory of the knowledge space, firms and scientists also convey the breadth of their knowledge base. Specialization implies that an innovator's knowledge is concentrated in a narrow territory of the knowledge space, and diversity or 'breadth' of knowledge indicates familiarity with a broad range of scientific or technological components.

Patenting and publishing capabilities

A number of findings in the innovation literature are consistent with the idea that corporate publishing is driven by firms' desire to participate in networks of scientific knowledge. In particular, firms that provide 'pro-publication' incentives and promote scientists based on their scientific standing are more likely to collaborate with universities and to rely on academic research in their innovation activities (Cockburn and Henderson, 1998). In contrast with prior studies that looked exclusively at the relationship between corporate publishing and firms' *propensity* and *degree* of collaboration with universities, I focus here on the *quality* of firms and faculty involved in these relationships. Specifically, if publications signal

[7] Patenting and publishing take different roles in different markets (labor, capital, product, or ideas market) and the variation in patent and publication records of firms and faculty reflects more than differences in capabilities. For example, firms might patent aggressively when they engage in 'patent blocking' and patent races, or seek to improve their bargaining power (e.g., Hall and Ziedonis, 2001). In academia, patenting might be driven by strategic and reputational calculations, as well as social influences (Murray, 2011). I focus on knowledge-related interpretations because participants in the market for research use observable patent and publication outputs to make inferences about potential partners' expertise.

[8] To my knowledge, Murray and coauthors (e.g., Gans, Murray, and Stern (2011) and Murray (2011)) were the first to systematically use the terms 'open' vs. 'commercial science' to emphasize that the main difference between publishing and patenting lies neither in the type of knowledge disclosed (basic vs. applied), nor in the locus of production (academia vs. industry), but rather in the institutional norms of the two disclosure regimes.

DOI: 10.1002/smj

scientific merit, I propose to examine the extent to which firm – faculty alliances are driven by complementarity between partners' publishing capabilities, such that firms and faculty select each other assortatively based on their ability to generate high quality scientific work. Survey studies reinforce the idea that firms aim to partner with university scientists with outstanding scientific records (Blumenthal *et al*., 1986). Particularly in basic research collaboration, firms perceive university scientists' research quality as being more important than their geographic proximity or other factors (Mansfield and Lee, 1996; Audretsch and Stephan, 1996). Preliminary interviews for this study also indicated that faculty value firms with superior records of publication. Thus, I propose to test:

Hypothesis 1: Ceteris paribus, firms' and university scientists' publication capabilities are complements in value creation.

A long line of research in the innovation literature views academic science and technological innovation as complementary and coevolving. Numerous examples of key discoveries corroborate the idea that technological development happens faster in areas with strong science, while scientific progress emerges as a response to the challenges posed by new technologies (Nelson, 2004). Furthermore, large-scale empirical studies done at high levels of aggregation have found a complementary relationship between public and industrial R&D (David, Hall, and Toole, 2000). Scholars have also argued that due to the research cost structure and the nature of incentives in universities and firms, it is more appropriate to have the earlier stages of a research process done by universities, and the later, more applied and more commercially viable innovation stages, in the industry (Aghion, Dewatripont, and Stein, 2008; Lacetera, 2009). These arguments suggest that alliances between science faculty and firms are driven by partners' motivation to build on each other's strengths in basic and applied research. Thus, university scientists with a better record of scientific publications should be more productive when teaming with firms with a better record of patenting, and vice-versa. Very few studies exist at the alliance level, but those that do tend to support this perspective. Orsenigo (1989) argues that, in the early years of the biotechnology industry, prominent scientists preferred to collaborate with firms with higher numbers of patents. Studying a broader range of industries in the mid 1990s, Mansfield and Lee (1996) found that university – industry collaboration was driven by the synergy between university scientists' research and firms' efforts to create new products and processes. I therefore propose the following hypothesis:

Hypothesis 2: Ceteris paribus, firms' patenting capabilities and university scientists' publication capabilities are complements in value creation.

Prior literature tells us little about whether the relationship between the patenting capabilities of faculty and firms are complements or substitutes in innovation. Recent work has identified a variety of individual and contextual factors that prompt university scientists to disclose inventions and become further involved in their commercial development (Bercovitz and Feldman, 2008; Stuart and Ding, 2006). However, although university scientists might occasionally disclose inventions in an attempt to seize immediate opportunities (Azoulay, Ding, and Stuart, 2007), they need to build a track record of disclosures to credibly signal expertise in converting scientific ideas to useful, marketable applications. For example, Elfenbein (2007) showed that university scientists with a higher patenting reputation were more likely to have their inventions licensed (although not necessarily for higher proceeds) than first-time patentees.

A question of interest here is which firms are better positioned to harness university scientists' patenting expertise? Two opposite scenarios are conceivable. On the one hand, firms with superior patenting capabilities could contribute with technological skills and resources, accelerating scientists' inventions to commercialization. If scientists interested in the commercial prospects of their ideas find collaboration with these firms valuable, we should observe the following relationship:

Hypothesis 3a: Ceteris paribus, patenting capabilities of firms and university scientists are complements in value creation.

On the other hand, firms with a better in-house stock of patents also face higher opportunity costs

DOI: 10.1002/smj

from working on inventions generated externally, particularly when those inventions are in their early stages, as is typical with inventions originating from universities (Jensen and Thursby, 2001). On the contrary, firms with relatively lower technical capital have lower opportunity costs from working on external ideas, and are more willing to take the risks and exert the effort necessary to develop early-stage inventions (Lowe, 2001). I argue that these aspects might be important for university scientists interested in the commercial prospects of their ideas. The more technologically astute the scientist is, the more likely it is that the research done with external partners is close to the scientist's core agenda, and the more likely it is that the project's success depends more on the scientist's knowledge than on his partner's patenting skills. Under this scenario, we should observe collaboration between university scientists with patenting expertise and firms with a shorter history of patenting.

A substitution relationship between partners' patenting capabilities is only credible if the parallel story is credible, too: that scientists with lower or no patenting skills will team up with firms with higher patenting capabilities. Indeed, for a university scientist with little experience in capitalizing on the technological opportunities of scientific research, partnering with a firm with higher expertise in patenting makes it possible to learn how early-stage ideas develop into marketable innovations. Conversely, firms with an established patenting record are well positioned to utilize scientific information toward more applied goals (Arora and Gambardella, 1994). By deciding to do joint research with universities, these firms effectively 'buy an option' in early-stage research ideas that might later be able to be commercialized. For these firms with high patenting capabilities, it is less important that a prospective faculty partner has patenting expertise. Thus,

Hypothesis 3b: Ceteris paribus, firms' and university scientists' patenting capabilities are substitutes in value creation.

Knowledge breadth and specialization

Knowledge attributes—*breadth* and *specialization*—represent another set of potential sources of complementarities in research. From the perspective of a party endowed with a diversified knowledge base, there is value in partnering with a specialized external collaborator who can fill particular knowledge gaps. Equally, a specialized agent will seek to ally with a more diversified external partner who will envision new applications of the agent's knowledge. While it is difficult to construct irrefutable arguments against alternative combinations (between two broad or two specialized partners, for example), numerous arguments suggest such combinations are less likely to systematically drive the formation of firm – scientist alliances. The rationale behind this assertion stems from the mechanisms through which breadth and specialization contribute to the discovery process.

Consider firms with a diversified knowledge base. Their innovation performance stems from their ability to capitalize on economies of scope from internal know-how and external knowledge spillovers (Henderson and Cockburn, 1996). However, capabilities and resource allocation priorities that enhance the development of breadth tend to conflict with those required to develop depth (Wang and Tunzelmann, 2000). Diversification enables firms to absorb and recombine outside know-how. Thus, firms keep up with the latest advances in a field through alliances with specialized suppliers rather than through internal development (Brusoni, Prencipe, and Pavitt, 2001).

Likewise, for a specialized university scientist (i.e., an 'expert' in a field), teaming with a firm with a broader knowledge base constitutes a productive opportunity. Because diversified firms already have mechanisms to internalize focused scientific information, university scientists' interactions with diversified firms will require less effort. More importantly, formal work on downsides of specialization has shown that reliance on multidisciplinary teamwork helps innovators with narrower expertise overcome the limitations of specialization (Jones, 2009). Thus, exposure to diverse knowledge through teamwork or alliances tends to benefit specialized inventors. However, assembling teams and capitalizing on diverse expertise can be compromised by communication and coordination costs (Dahlin, Weingart, and Hinds, 2005). Faced with the trade-off between forming a partnership with multiple specialized partners and a firm with a diverse array of knowledge, narrowly specialized scientists more likely prefer the latter. These arguments imply

DOI: 10.1002/smj

cross-fertilization in innovation between specialized scientists and firms with a diverse knowledge base. Hence:

Hypothesis 4a: Ceteris paribus, a complementarity relationship exists between firms with a diversified knowledge base and university scientists with focused expertise.

The complementarity between specialized firms and faculty with broad expertise becomes evident when we examine these agents' motivations for entering a research alliance. To reduce the risk of technological exhaustion, specialized firms must pursue various boundary-spanning activities (Rosenkopf and Almeida, 2003). Prior work has shown that building alliances and relying on university research are methods of exploration (Bercovitz and Feldman, 2007). Here, I propose that specialized firms have higher chances of successful exploration when they partner with scientists whose expertise is broad rather than narrowly specialized. This argument builds on prior research showing that scientific knowledge increases the rate of invention by: (a) helping innovators identify useful directions of search, (b) preventing innovators from wasting time with dead ends, and (c) encouraging innovators in the face of failure (Fleming and Sorenson, 2004). The returns from working with broader partners will be higher when at least the first two mechanisms are considered. A partner with broader expertise provides immediate access to multiple potentially relevant pieces of information. This enables firms to become aware of possible linkages among various knowledge domains and helps them more rapidly identify ways to leverage their knowledge. In addition, broader scientists can deploy a larger set of information-processing filters, which in turn reduce the uncertainty around possible applications of firms' knowledge and eliminate inefficient experimentation (Nelson, 2004).

I also conjecture that broader scientists are attracted to collaborations with specialist firms that could enrich the scientist's understanding of a particular scientific area. Working with diversified firms is less likely to bring higher benefits to these scientists. Diversified firms and scientists excel at recombination and new knowledge integration, and, if one alliance partner possesses such skills, it is less necessary for the other partner to contribute the same skill set. As noted before, potential benefits from increasing expertise diversity are subject to diminishing returns. Thus, I propose:

Hypothesis 4b: Ceteris paribus, a complementarity relationship exists between specialized firms and university scientists with broader expertise.

DATA AND METHOD

Because research contracts between firms and scientists are typically considered confidential by both parties, prior researchers have used coauthorship as a proxy for collaboration. However, many joint contracts do not result in coauthorship; coauthorship may also be a result of informal ties between scientists. I overcame these problems by using first-hand information from the Grants and Contracts Office of a top-ranked, private U.S. medical school located on the East Coast. The sample contains the full set of faculty – industry research contracts from 1995 (the earliest date for which this information was available) until 2004. The office released the following information: the name of the principal investigator, the name of the firm, the project title, the start and end date of the contract, and the award (contract) money. I supplemented this primary information with numerous secondary sources detailed in the 'variables' section.

This empirical context is suitable for studying research collaboration between faculty and firms for multiple reasons. Medical schools have a long history of collaboration with industry, and research done in medical domains has been shown to have a strong impact on industrial R&D (Cohen *et al*., 2002). University scientists in medical schools patent more and are more involved in technology transfer than those in other fields equally important for the private sector. The medical school studied here began collecting systematic data on research contracts much earlier than other universities, allowing me to study a larger sample of collaborations over a longer time span. Furthermore, all contracts in the sample reflect a clear intent to generate innovation (i.e., they are referred to as discovery-type collaborations).

The sample consists of 447 contracts between 238 firms and 217 university scientists. The number of contracts per year varied from 32 to 54. Almost 25 percent of the scientists and firms collaborated with more than one partner during the observed decade, but there were no instances of multiple scientists (or firms) working simultaneously with a firm (scientist) on the *same project*. This data aspect implies a 'one-to-one' matching of firms and scientists.[9]

Variables

I collected firms' stock of publications during a 10-year window (up to the collaboration date) and scientists' lifetime publications from the ISI Science Citation Index. I chose a 10-year window to allow a sufficiently long time for firms' publication output to translate into reputational effects. I weighted all publications by the number of citations they received in the Web of Science database (as of April 2010). As is typical in innovation studies, I applied a yearly depreciation rate of 10 percent.

I collected firms' and scientists' patents over a six-year window prior to the contract time. I controlled for the variation in patent quality by weighting patents by citations received as of June 2010, with a 10 percent annual depreciation rate.[10]

I measured the *breadth of a firm's knowledge base* using an entropy index of diversification based on the distribution of a firm's patents in technology domains defined by the inventive International Patent Classification Reform (IPCR) patent classes. The formula for this index is: $\sum_{1}^{N}(f_i^* \ln(1/f_i))$, where N represents the total number of technological domains (here, inventive IPCR) in which a firm has been patenting and f_i represents the fraction of patents in *ith* category. The index takes higher values as firm knowledge becomes more diversified.

The *knowledge-diversification index* of the university scientists is similarly defined, but technological domains were replaced with scientific domains. I identified the scientific domains from the MEDLINE database of the National Library of Medicine (NLM). NLM specialists classify each article according to a scheme known as the Medical Subject Headings (MESH) thesaurus. The classification of articles based on the MESH thesaurus insures consistency in classifying scientists' expertise. The MESH thesaurus has a hierarchical structure composed of descriptors (main headings), qualifiers (subheadings), and supplementary concept records. Descriptors indicate the main themes discussed in the article, and qualifiers indicate a specific aspect of the descriptor. I treated a combination of MESH descriptors and MESH qualifiers as constituting a scientific domain.[11]

Controls

I included several control variables that might influence partner choice and the matching of firms and scientists. Firm size (measured by the number of employees), as an indicator of resources, and firm age (years since founding), as an indicator of experience, are likely to play a role in alliance formation. I also included the size of the lab (measured by the number of people associated with the principal investigator's lab) and academic age (the number of years since the scientist obtained his or her highest degree). I controlled for scientists' academic titles by adding dummy variables for associate professor, full professor, and chairman/chief (the excluded category was assistant professor). In a robustness check analysis, I also included a dummy variable for biotechnology firms (the excluded categories

[9] Both the theory and the estimation method can be used to study 'one-to-many' and 'many-to-many' matching. The assumption here is that projects are independent. Thus, treating alliance formation as 'one-to-one matching' is a better model for the current data and not the result of theoretical or methodological constraints. In contrast, a 'one-to-many matching' would imply, for example, that one firm collaborates with multiple scientists on the same project. In this case, the choice of a particular university scientist is dependent on the choice of other faculty, and the production function needs to be modified to include this aspect. There are no instances of such collaborations in my sample. For a typical example of 'many-to-many matching', see Fox (2009).

[10] Corporate publications and patents were aggregated at the parent level after carefully considering the corporate structure of the firm during the relevant time window. I took into account mergers and acquisitions, divestitures, and name changes. I also performed several sensitivity analyses, and the final results are very similar even if I use an alternative time window for patents (e.g., five or four years) or publications (e.g., nine or eight years).

[11] For example, 'coronary vessels' and 'cardiac surgical procedures' are descriptors, while 'injuries' and 'adverse effects' are subheadings. Accordingly, the combinations 'coronary vessels/injuries' and 'cardiac surgical procedures/adverse effects' represent two different subject areas. I eliminated all MESH categories that do not reflect substantive scientific information (e.g., the type of research support, the publication format, the geographic location, etc.) The complete list of descriptors can be accessed from the NLM.

were pharmaceutical firms and a small number of medical device firms). Table 1 describes variables and their sources. Table 2 explains the relationship between variables and hypotheses.

Model specification and estimation

The challenge in estimating a matching model arises from the fact that partner choice depends on all other relationships formed in the market. Thus, there are as many endogenous variables as the observed matches. When alliances are treated as independent decisions, as in standard discrete-choice models, the choice probability that an agent chooses a particular partner can be factored into a product of low-dimensional integrals. In a matching situation, the choice probability is a multidimensional integral over the joint density of error terms. The dimension of this integral is equal to all possible arrangements of pairings in the market. For example, in one-to-one matching involving 100 agents on each side, the dimension is $100! = 9.33 * 10^{15}$. Integrating over such a high-dimensional integral is currently not feasible.

To overcome these problems, I estimated the model using a semiparametric estimator developed by Fox (2010), which generalizes Becker (1973)'s unidimensional analysis to the multidimensional case. The estimator is semiparametric in the sense that it requires the specification of a production function but does not impose a distribution on the error terms. The method relies on the 'local production maximization' condition, which Fox (2010) proved is satisfied whenever the observed matches are the equilibrium outcome of an assignment game. The local production maximization condition states that the output created by any two observed matches is greater than the output created by counterfactual matches formed from an exchange of partners. [12]

To illustrate how the method works, I denote by **(F, S)** a firm – scientist pair in the sample, and by $\mathbf{g}\ (\boldsymbol{F}, \boldsymbol{S}\,|\boldsymbol{\beta})$ the output created by the pair working together. The production function that describes the match output takes the following form:

$$\mathbf{g}\left(\mathbf{F},\mathbf{S}|\beta\right) = \beta\ \left[\mathbf{X_F} * \mathbf{Y_S}\right] + \varepsilon_{\mathbf{SF}} \tag{1}$$

where $\mathbf{X_F}$ represents the vector of firms' characteristics, $\mathbf{Y_S}$ the vector of scientists' characteristics, $\boldsymbol{\varepsilon}_{\mathbf{SF}}$ is the error term, and $\boldsymbol{\beta}$ is the vector of parameters to be estimated.

I consider that each of the 10 years in the sample represents a separate matching market. For any two observed collaborations in a market, $(\boldsymbol{F_i}, \boldsymbol{S_i})$ and $(\boldsymbol{F_j}, \boldsymbol{S_j})$, the corresponding match output is $\mathbf{g}\ (\boldsymbol{F_i}, \boldsymbol{S_i}|\boldsymbol{\beta})$ and, respectively, $\mathbf{g}\ (\boldsymbol{F_j}, \boldsymbol{S_j}|\boldsymbol{\beta})$. The counterfactual pairings are obtained by switching partners such that $\mathbf{F_i}$ pairs with $\mathbf{S_j}$ to generate $\boldsymbol{g}\ (\boldsymbol{F_i}, \boldsymbol{S_j}|\boldsymbol{\beta})$ and $\boldsymbol{F_j}$ pairs with $\boldsymbol{S_i}$ to generate $\mathbf{g}\ (\boldsymbol{F_j}, \boldsymbol{S_i}|\boldsymbol{\beta})$. The local production maximization implies that the following inequality is true:

$$\mathbf{g}\left(\mathbf{F_i},\mathbf{S_i}|\beta\right) + \mathbf{g}\left(\mathbf{F_j},\mathbf{S_j}|\beta\right) > \mathbf{g}\left(\mathbf{F_i},\mathbf{S_j}|\beta\right) + \mathbf{g}\left(\mathbf{F_j},\mathbf{S_i}|\beta\right) \tag{2}$$

The estimation procedure requires checking the local production maximization inequality (1) for all 9,996 possible combinations of firm–scientist in the sample and the corresponding counterfactuals. The parameters $\boldsymbol{\beta}$ are estimated by choosing the values that predict the highest number of inequalities. Formally, the estimator maximizes the following objective function:

$$\mathbf{Q}\left(\beta\right) = \frac{1}{10} * \sum_{\mathbf{h}=1}^{10} \sum_{0<\mathbf{i},\mathbf{j}<N_h} 1[\mathbf{g}(\mathbf{F^{h_i}},\mathbf{S^{h_i}}|\beta) + \mathbf{g}\left(\mathbf{F^{h_j}},\mathbf{S^{h_j}}|\beta\right) > \mathbf{g}\left(\mathbf{F^{h_i}},\mathbf{S^{h_j}}|\beta\right) + \mathbf{g}\left(\mathbf{F^{h_j}},\mathbf{S^{h_i}}|\beta\right)]$$

where $1[\ \cdot\ ]$ denotes an indicator function which takes a value of 1 if the expression between parentheses is true and 0 otherwise, $\boldsymbol{N_h}$ is the number of pairs in each individual market, and $\boldsymbol{h}$ is a market index that here takes values from 1 to 10. [13]

[12] Although the existence and the uniqueness of the equilibrium in one-to-one matching games are well known, a general analytical solution for characterizing the matching of agents with multidimensional attributes does not exist in the literature. However, because the local production maximization is a *necessary* condition, the estimator does not require computing an equilibrium solution to a matching game (Fox, 2010).

[13] Note that the production function $\mathbf{g}\ (\mathbf{F}, \mathbf{S}|\boldsymbol{\beta})$ and the objective function $\mathbf{Q}(\boldsymbol{\beta})$ contain only interaction terms. Production levels are also affected by the 'main' effects of X_F and Y_S. However, the main effects cancel out in the local maximization inequality (2). This property is consistent with the theoretical model, which asserts that matching is driven entirely by

DOI: 10.1002/smj

Table 1. Variable definition

Firms	
Size (FIRM SIZE)	Number of employees at contract year. *Source*: Factiva; Corporate Affiliations; Thomson One.
Publishing capabilities (FIRM PUB CAP)	Ten-year stock of publications weighted by the number of citations (as of April 2010) and a 10% yearly depreciation rate. *Source*: ISI Web of Science combined with Corporate Affiliations to account for organizational structure at the publication time. Note that ISI is the only collection of medical journals that contains information on authors' institutional affiliations during the sample time frame. However, ISI provides this information only for the first 10 authors. Thus, this measure provides a conservative estimation of firms' publications.
Patenting capabilities (FIRM PAT CAP)	Six-year stock of patents weighted by citations (as of June 2010) and a 10% yearly depreciation rate. *Source*: Delphion combined with Corporate Affiliations to account for organizational structure at the patent application date.
Knowledge breadth (FIRM BREADTH)	Entropy measure of knowledge diversification based on inventive IPCR patent classes. Higher values indicate higher degree of knowledge diversity. I used the full range of inventive IPCR patent classes, not just the main patent class. *Source*: Firms' patent stock (see above).
Age (FIRM AGE)	Number of years since founding. *Source*: Corporate Affiliations; Factiva; Companies' Web pages.
Biotechnology firm (BIOTECH)	Coded 1 if the firm's main activity was in the biotechnology industry. *Source*: Corporate Affiliations; Factiva; Companies' Web pages.
University scientists	
Lab size (LAB SIZE)	Number of people associated with principal investigator's lab. I obtained accurate information on the lab size for 40% of the sample. For the remaining scientists, I estimated the lab size from corroborating various sources (e.g., faculty Web sites, local news articles) and, in particular, from carefully identifying the investigator's coauthors at the same university during a three-year window prior to the contract date. Although not perfectly measured, this variable captures the variation in the human capital resources that principal investigators could mobilize for a project. *Source*: Elsevier's Scopus; ISI Web of Science; Scientists' Web pages; Lexis-Nexis
Publishing capabilities (PI PUB CAP)	Cumulative number of publications weighted by the number of citations (as of April 2010). A 10% yearly depreciation rate has been applied. *Source*: ISI Science Citation Index Expanded.
Patenting capabilities (PI PAT CAP)	Number of patents and patent applications in which the scientist appears as inventor cumulated over a six-year window prior to collaboration and weighted by patent citations (self-cites excluded) as of June 2010. A 10% yearly depreciation rate has been applied. *Source*: Delphion.
Knowledge breadth (PI BREADTH)	Entropy measure of diversification based on MESH descriptor/qualifier keyword combination. *Source*: Medline and Medical Subject Headings thesaurus of the National Library of Medicine.
Academic age (PI AGE)	Number of years since the highest degree was obtained. *Source*: University scientists' CVs; university registrar.
Academic title	Dummy variables for each title category: assistant; associate; professor; chairman/chief. *Source*: University scientists' CVs; university registrar.

Table 2. Variable measurement and empirical predictions

Hypothesis	Firm side	Scientist side	Theoretical prediction	Relationship between variable measurement and expected sign	Expected sign
Hypothesis 1	Publishing capabilities	Publishing capabilities	*Complementarity*	High values on both variables indicate superior publishing capabilities. A positive sign is consistent with a complementary relationship.	+
Hypothesis 2	Patenting capabilities	Publishing capabilities	*Complementarity*	High values indicate superior patenting, respectively, publishing capabilities. A positive sign is consistent with a complementary relationship.	+
Hypothesis 3a	Patenting capabilities	Patenting capabilities	*Complementarity*	High values on both variables indicate superior patenting capabilities. A positive sign implies complementarity.	+
Hypothesis 3b	Patenting capabilities	Patenting capabilities	*Substitution*	High values on both variables indicate superior patenting capabilities. Hence, a negative sign (where high levels on one side of the market are associated with lower levels on the other side of the market) implies substitution in patenting capabilities.	−
Hypothesis 4a	Knowledge breadth	Knowledge specialization	*Complementarity*	On both sides, the entropy-based 'knowledge diversification' measure increases with the diversity of knowledge, such that high values imply 'breadth' or 'diverse knowledge' and lower values indicate 'specialization' or 'focused expertise.' Theoretically, complementarity is expected between *breadth* (i.e., *high* values of the 'knowledge diversification' index) on the firm side and *specialization (i.e., low* values of the 'knowledge diversification' index) on the scientist side. Hence, when the direction of measurement of the 'knowledge diversification' variables is taken into account, a negative sign indicates complementarity.	−
Hypothesis 4b	Knowledge specialization	Knowledge breadth	*Complementarity*	From a theoretical perspective, complementarity is expected between *specialization (i.e., low* values of the 'knowledge diversification' index) on the firm side and *breadth* (i.e., *high* values of the 'knowledge diversification' index) on the scientist side. Thus, a negative sign is consistent with the predicted complementarity relationship.	−

DOI: 10.1002/smj

Each time the inequality (2) holds for a trial guess of the vector of parameters $\boldsymbol{\beta}$, the score of correct prediction increases by 1. The vector of parameters that yields the highest score of correct predictions $\mathbf{Q}(\boldsymbol{\beta})$ provides a consistent estimator of the parameters $\boldsymbol{\beta}$ (Fox, 2010).

Note that adding a constant to the match output or multiplying it by a positive constant will not affect the inequality (2). To take account of this aspect, the scale of the production function must be normalized. I normalized the scale of the match output by setting to $|1|$ the coefficient $\boldsymbol{\beta}_0$ of the interaction term between two control variables: the size of the firm and the lab size.[14] As Stephan (2010) points out, biomedical research relies increasingly on sophisticated equipment and expensive material, which makes access to resources a necessary condition for doing research in this field. In this context, a partner of larger size is more valuable because it provides the possibility to undertake more complex research projects at lower costs. In addition, a larger partner gives access to a broader network of people. I conjectured that these aspects are important on both sides of the market. Thus, I chose the interaction term $|1|$***FirmSize * LabSize** to play the role of a 'benchmark' relationship, relative to which I assessed the magnitude of the other relationships in the model. Following a standard approach in the literature (e.g. Fox, 2009; Yang, Shi, Goldfarb, 2009), I ran separate analyses for both $+1$ and -1 and I chose the sign that produced a higher value for the objective function.

Hypothesis testing requires subsampling, a procedure that involves random sampling without replacement and generating 95 percent confidence intervals for the estimated coefficients. The reported confidence intervals were obtained from 200 random samples of four markets at a time. The online appendix offers a more detailed description of the method.

the *interaction* between partners' attributes, while the 'un-interacted' characteristics cancel out because all potential partners value them equally.

[14] All choice models impose a scale normalization of the utility function. Typically, the variance of the error terms is set to a number chosen for convenience (e.g., the variance of the error terms is set to $\pi^2/6$ in logit models and to 1 in probit models). A standard way to normalize the scale in a maximum score estimator is to set one coefficient to be equal to $|1|$.

RESULTS AND DISCUSSION

Table 3 provides the descriptive statistics and the correlation matrix. All of the key variables have a substantial amount of variation. While prior research has focused on the contribution of public research for start-ups and large, established firms (see Cohen *et al.*, 2002), current data show that university – industry research collaborations involve firms that vary considerably in age, size, and publication and patenting activity.

On average, firms in the sample published 230 publications and applied for 82 patents every year.[15] Faculty published on average 3.8 articles yearly and 70 articles over a career. Faculty in the top decile of the distribution had more than seven patents, and 42 percent of contracts involved a scientist with at least one patent application. The sample mean is equal to two faculty patents.

Because the interpretation of results is not meaningful when variables have different units of measurement and differ greatly in scale magnitude, I standardized the variables using the z-score transformation. Table 4 presents the estimated sign of the baseline and the estimated values of $\boldsymbol{\beta}$ coefficients and their 95 percent confidence intervals.

Model 1 focused on the relationships predicted in Hypotheses 1–4, controlling for several other determinants of matching. As the correlation matrix showed, larger firms tended to generate patents and publications of better quality. Thus, I controlled for potential confounding effects of firm size by including the interaction terms between firm size and scientists' capabilities. Likewise, I controlled for lab size effects by interacting this variable with firms' capabilities. Because experience and seniority in an industry/field are likely to matter in alliance formation, I added in the model the interaction between firm age and academic age. However, the variable *academic age* might reflect both the experience and the academic status of the scientist. To disentangle these effects, I included separately the interaction between firm age and scientists' title. Lastly, to separate the effects of firm age from those of firm size, I also interacted firm size with scientists' age and title variables.

[15] To make the comparison easier, I discuss here yearly averages, while Table 3 reports the statistics for the *stock* of patents and publications as described in Table 1.

Table 3. Descriptive statistics and correlations

	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Firms																		
1. Size	23110.73	34486.16	1	156400	1													
2. Publishing capabilities	37835.48	57591.10	0	221509	0.75	1												
3. Patenting capabilities	5041.54	8551.80	0	65582	0.27	0.29	1											
4. Knowledge breadth	4.41	2.09	0	7.19	0.61	0.61	0.43	1										
5. Age	47	46.93	1	155	0.53	0.31	0.20	0.61	1									
6. Biotechnology firm	0.26	0.43	0	1	−0.37	−0.29	−0.31	−0.52	−0.45	1								
University Scientists																		
7. Lab size	17.40	13	1	41	−0.10	−0.11	−0.08	−0.06	−0.11	0.04	1							
8. Publishing capabilities	1219.06	1383.27	1	12671.43	0.06	0.04	−0.01	0.08	0.02	−0.05	0.54	1						
9. Patenting capabilities	69.70	382.89	0	4382.99	−0.09	−0.09	−0.09	−0.19	−0.11	0.16	0.04	0.09	1					
10. Knowledge breadth	4.75	0.52	2.20	5.79	−0.10	−0.07	−0.12	−0.09	−0.14	0.09	0.58	0.37	0.05	1				
11. Academic age	17.12	7.35	3	53	−0.09	−0.12	−0.13	−0.10	−0.10	0.07	0.40	0.19	0.12	0.50	1			
12. Associate professor	0.29	0.45	0	1	0.07	0.06	0.12	0.08	0.08	−0.06	−0.07	−0.05	−0.09	0.04	−0.06	1		
13. Full professor	0.23	0.42	0	1	−0.13	−0.14	−0.14	−0.10	−0.07	0.06	0.44	0.18	0.16	0.26	0.39	−0.36	1	
14. Chair/chief	0.12	0.32	0	1	0.02	0.03	−0.02	0.04	−0.04	0.05	0.27	0.40	−0.01	0.31	0.35	−0.23	−0.20	1

Table 4. Point estimates and 95% confidence intervals

	Relationship	Model 1	Model 2	Z-score transformation Model 3	Model 4	Model 5	Log transformation Model 6	Model 7
Baseline	FIRM SIZE *	1	−1	−1	1	1	1	1
	LAB SIZE							
Hypothesis 1	*FIRM PUB CAP* *	**9.22**	...	...	**3.56**	**6.78**	**3.35**	**4.03**
	PI PUB CAP	(5.33; 11.97)			(2.97; 7.10)	(3.00; 9.95)	(1.25; 4.72)	(1.66; 5.77)
Hypothesis 2	*FIRM PAT CAP**	−1.01	...	...	−1.02	−0.87	−1.04	−1.26
	PI PUB CAP	(−3.18; 2.17)			(−2.96; 3.21)	(−3.03; 3.88)	(−5.60; 1.02)	(−6.49; 2.95)
Hypothesis 3	*FIRM PAT CAP**	**−15.91**	...	...	**−7.23**	**−4.61**	**−2.83**	**−2.66**
	PI PAT CAP	(−18.78; −10.36)			(−11.35; −5.32)	(−9.71;−2.18)	(−5.42; −1.07)	(−5.49; −1.43)
Hypothesis 4	*FIRM BREADTH**	**−4.78**	...	...	**−1.93**	**−2.89**	**−4.24**	**−13.67**
	PI BREADTH	(−6.65; −1.24)			(−4.91; −0.84)	(−5.12; −0.71)	(−10.98; −3.46)	(−16.88; −7.79)
5	FIRM SIZE*	**5.26**	...	**9.72**	**3.97**	**6.61**	0.17	**1.35**
	PI PUB CAP	(3.21; 9.15)		(3.05; 12.30)	(1.64; 5.86)	(2.75; 9.36)	(−0.66; 3.24)	(0.92; 3.02)
6	FIRM SIZE*	**−21.54**	...	**−12.85**	**−9.37**	**−3.21**	−1.02	−0.45
	PI PAT CAP	(−23.85; −7.25)		(−18.81; −4.23)	(−11.47; −5.11)	(−6.95; −1.43)	(−3.19; 0.99)	(−4.29; 1.53)
7	FIRM SIZE*	−0.31	...	**−6.09**	−0.35	**−5.86**	**−19.74**	**−10.25**
	PI BREADTH	(−4.40; 1.15)		(−7.54; −2.18)	(−3.28; 1.63)	(−8.17; −2.27)	(−24.40; −15.71)	(−16.68; −7.83)
8	*FIRM PUB CAP* *	**−3.66**	...	**−1.25**	**−3.94**	**−7.48**	**−2.23**	**−2.64**
	LAB SIZE	(−6.45; −1.16)		(−6.45; −0.37)	(−7.86; −1.43)	(−10.60; −3.54)	(−5.88; −1.05)	(−6.76; −1.15)
9	*FIRM PAT CAP* *	−0.78	...	0.23	−0.64	−0.32	1.46	1.01
	LAB SIZE	(−2.59; 4.06)		(−4.77; 2. 57)	(−5.77; 2.41)	(−3.15; 4.16)	(−4.48; 3.80)	(−3.65;5.16)
10	*FIRM BREADTH**	−0.24	...	0.10	0.68	0.85	**8.92**	**6.61**
	LAB SIZE	(−2.40; 5.73)		(−3.30; 4.81)	(−3.60; 5.94)	(−2. 55; 5.48)	(2.55; 9.28)	(1.73; 8.33)
11	FIRM AGE *	**−0.93**	0.04	−0.60	−0.63	**−0.89**	**−2.86**	**−5.89**
	PI AGE	(−3.92; −0.08)	(−5.90; 2.87)	(−4.63; 2.39)	(−3.90; 1.11)	(−4.20; −0.12)	(−5.90; −0.53)	(−9.33; −2.98)
12	FIRM SIZE *	**−1.11**	−0.24	**−1.36**	**−0.98**	−0.60	**−5.06**	**−3.06**
	PI AGE	(−4.07; −0.52)	(−4.78; 3.49)	(−3.10; −1.06)	(−4.01; −0.50)	(−3.27; 0.29)	(−10.77; −2.78)	(−6.50; −0.75)
13	FIRM AGE *	1.77	0.56	1.98	3.97	3.46	2.54	2.60
	ASSOCIATE	(−0.84; 4.15)	(−1.64; 4.22)	(−4.39; 3.25)	(−0.53; 5.58)	(−0.49; 5.83)	(−0.93; 5.36)	(−3.30; 5.27)

DOI: 10.1002/smj

Table 4. *(Continued)*

	Relationship	Z-score transformation Model 1	Model 2	Model 3	Model 4	Model 5	Log transformation Model 6	Model 7
14	FIRM AGE *	**5.84**	0.30	0.63	**1.72**	**2.63**	**11.51**	**3.90**
	PROFESSOR	(1.13; 6.42)	(−0.84; 4.87)	(−4.64; 2.81)	(0.74; 4.63)	(1.40; 6.14)	(9.65; 17.56)	(1.34; 8.56)
15	FIRM AGE *	**−4.87**	−0.65	**−1.84**	**−4.15**	**−6.33**	**−9.00**	**−8.00**
	CHIEF	(−7.58; −2.21)	(−6.70; 0.26)	(−6.90; −1.39)	(−8.42; −2.22)	(−8.90; −3.93)	(−16.26; −4.64)	(−13.59; −5.23)
16	FIRM SIZE *	−0.02	0.35	0.69	1.86	0.82	4.24	−0.47
	ASSOCIATE	(−2.67; 3.50)	(−5.88; 3.11)	(−3.87; 3.83)	(−4.24; 3.13)	(−3.48; 2.40)	(−2.37; 6.41)	(−4.81; 5.31)
17	FIRM SIZE *	**−17.01**	**−2.12**	**−8.64**	**−3.73**	**−8.85**	**−5.45**	**−4.11**
	PROFESSOR	(−19.77; −11.34)	(−6.96; −1.50)	(−13.52; −4.96)	(−7.13; −1.76)	(−10.48, −3.55)	(−8.89; −2.57)	(−9.82; −2.63)
18	FIRM SIZE*	**5.78**	**2.01**	**3.26**	**7.53**	**6.41**	**4.92**	**3.13**
	CHIEF	(3.30; 8.12)	(1.89; 7.28)	(0.92; 5.32)	(4.34; 9.85)	(2.80; 8.33)	(1.76; 7.08)	(1.17; 6.55)
19	FIRM AGE*	...	...	...	0.60	...	0.81	...
	PI PUB CAP				(−2.30; 3.19)		(−4.30; 3.23)	
20	FIRM AGE*	...	...	...	−0.42	...	−2.56	...
	PI PAT CAP				(−4.20; 3.62)		(−5.06; 2.47)	
21	FIRM AGE*	...	...	...	**−3.52**	...	**−23.72**	...
	PI BREADTH				(−8.30, −1.92)		(−27.31; −16.63)	
22	BIOTECH*	...	...	...	...	−1.76		−6.36
	PI PUB CAP					(−4.26; 2.85)		(−8.19; 3.09)
23	BIOTECH *	...	...	...	...	**2.33**		**11.99**
	PI PAT CAP					(1.26; 6.79)		(5.52; 14.13)
24	BIOTECH *	...	...	...	...	**2.52**		**8.80**
	PI BREADTH					(1.49; 5.25)		(6.45; 12.30)
	% Inequalities predicted	74%	59%	61%	75%	76.5%	73.5%	74%

Note: *Parameters in* ***bold*** *are significantly different than zero*. Note that confidence intervals do not have to be centered at the point estimate (see Santiago and Fox (2008)).
To estimate the Model, I adapted Santiago and Fox's (2008) template in Mathematica software for the particular production functions that are tested here. For optimization, I used the standard differential evolution procedure in Mathematica 7.0.

DOI: 10.1002/smj

The results of the multivariate matching analysis strongly corroborate the relationships hypothesized in Hypotheses1, 3b, 4a, and 4b. As predicted in Hypothesis 1, firms' and university scientists' publishing capabilities are complements in innovation. As conjectured in Hypothesis 3b, there is a negative relationship between partners' patenting capabilities, consistent with the idea that patenting capabilities are substitutes. The fourth interaction term shows that firms and university scientists whose research is concentrated in fewer scientific domains create more value by teaming up with more knowledge-diversified partners. An unexpected result is the negative relationship between firms' patenting capabilities and university scientists' publishing capabilities, while Hypothesis 2 predicted a positive relationship. This finding echoes Gittelman and Kogut (2003), who found that biotechnology firms' high-impact patents did not build on important scientific papers, a result the authors attributed to the fact that valuable inventions and valuable scientific knowledge follow different and conflicting evolutionary selection 'logics.' While Gittelman and Kogut's study was based on patents and publications by biotechnology firms, their arguments offer a possible explanation for the negative relationship between firms with a higher stock of important patents and university scientists with a higher stock of important publications. However, I cannot rule out a simpler argument that the operationalization of variables did not capture all relevant aspects for the division of innovation labor between firms and scientists.

The inclusion of various controls in the equation revealed interesting aspects of matching. All other things being equal, faculty with a higher stock of important patents create higher synergies by partnering with firms of smaller size. This result reinforces the arguments behind the finding that scientists' and firms' patenting capabilities are substitutes. Contrarily, university scientists with a better publication record create higher synergies with more resourceful firms. This finding echoes a longstanding view in the literature that university – industry collaboration is driven by scientists' need to obtain resources from industry partners. However, this result also shows that, all else equal, access to firms with more resources is given to (and more valued by) prominent scientists.

We can also observe a negative relationship between lab size and firms' publishing capabilities, suggesting that firms with higher publication capabilities get more rewards from collaborating with individual scientists or smaller teams. Likewise, it shows that firms' publishing capabilities are more important for smaller labs than for larger ones.

The relationship between firm age and academic age is negative, an indication that more experienced partners are substitutes for less experienced partners. Other interesting effects can be observed when the role of firm experience (age) and resources (size) are examined in interaction with faculty academic title. In particular, both larger firms and younger firms create higher synergies by partnering with university scientists of high standing, such as chairmen and chiefs of various divisions. Thus, both large and start-up firms compete for higher status faculty. These findings suggest that on the one hand, university scientists who can manage larger-scale projects match with firms that can provide an appropriate level of resources; on the other hand, university scientists with higher standing in administration also tend to match with younger firms, which are riskier but potentially more rewarding partners.

The estimated sign of the baseline is positive, indicating complementarity between partners' resources. In terms of economic significance, the baseline represents the marginal effect on the innovation output created by increasing the size of the firm and the lab by one standard deviation. In Model (1), the baseline effect on the production function was normalized to $+1$ and it represents the scale to which we can interpret the magnitude of other coefficients. The results show that increasing the publication capabilities of both partners by one unit (i.e., one standard deviation in the stock of quality-adjusted publications), has a nine times higher effect on the innovation output than the baseline. Further, in absolute value, the marginal impact of a one standard deviation change in partners' patenting capabilities is 15 times more important than the baseline, and the marginal effect of a unit change in knowledge diversity is four times more important. [16]

[16] Because of the substitution relationship, the impact is positive in absolute terms when a one unit *increase* in patenting capabilities of one partner is associated with a one unit *decrease* in patenting capabilities of the other partner. The same reasoning applies for knowledge diversity: in absolute terms, a one unit increase in knowledge diversity on one side of the market, together with a one unit decrease in knowledge diversity on

As well as being significant, the variables measuring partners' capabilities explain a fair amount of the variation in the data. Model (1) predicts 74 percent of inequalities.[17] For comparison, Model (2) shows the results of an equation that has only control variables; few coefficients were significant and the model explained 59 percent inequalities. Model (3) contained all relationships except those predicted by the main hypotheses and explained 61 percent inequalities.

Robustness checks

I performed several robustness checks to test the sensitivity of results to alternative model specifications and variable transformations. Models (4) and (5) test whether the main relationships predicted in Hypothesis 1 to Hypothesis 4 remain robust when firm age and industry effects are taken into account. In relative terms, the impact of the main relationships on the alliance output is smaller in these specifications, but previous results remain significant and of the same sign. The analysis in Model (4) shows that firm age does not play an important role when considered in combination with scientists' publishing and patenting capabilities, but younger firms and more knowledge-diversified scientists create higher synergies together. In Model (5), the results remain robust to the inclusion of a dummy variable for biotechnology firms. Interestingly, scientists with higher patenting capabilities and those with a more diversified knowledge base create higher synergies by collaborating with biotechnology firms. However, the relationship between faculty publishing capabilities and biotech dummy is not significant. Finally, Models (6) and (7) in Table 4 show that results of the empirical test are robust to the log transformation of variables.[18]

Limitations and future research

The results of the matching analysis presented here are based on several assumptions. The model discounts search costs (e.g., the cost of finding a partner), the potential initial uncertainty about the match value (e.g., adverse selection), and other types of market frictions, such as institutional policies of technology transfer that might impede the formation of public–private partnerships. These factors should be taken into account for a more comprehensive approach to match formation. For example, with high search costs, geographical location might be an important consideration in partner choice. Because collaboration contracts in this paper are from one university, all scientists in the sample are equally located from the firms' point of view. Data from more universities would allow us to examine the role of geography in university – industry partnerships. Likewise, data from universities with different technology transfer policies would allow researchers to determine if frictions created by technology transfer terms (e.g., rigid rules on intellectual property rights) might preclude some matches from happening.

This paper was based on the premise that the objective of a research alliance is to generate new knowledge. Future work could examine if considerations other than research productivity, such as perceptions of social status or affiliation with a particular social network, provide additional explanatory power or change the understanding of matching observed here.

The theoretical model presented in this paper addressed questions related to value creation and abstracted away the costs of governance, value appropriation, and partnership evolution. The connection between matching and these important aspects of alliances creates an interesting opportunity for additional research.

---

the other side of the market has a four times higher impact on the alliance output than the baseline.

[17] In a market with k agents on each side, the maximum number of inequalities predicted by a random configuration where k-2 pairs are identical to ones in the 'true' model is: **(k-2) (k-3)/k(k-1)**. For the 10 markets in the sample, we obtain (by multiplying the corresponding number of inequalities for each market) 0.386 or 38.6 percent inequalities. However, the expected value for such arrangement to occur is equal to **(k-2)(k-3)/k(k-1)*(1/(k!-1))**, which is close to 0 in the data.

[18] A small constant of 0.001 was added before logging to insure that values equal to 0 remained in the estimation. The constant shifts the distribution to the right and the log transformation minimizes the importance of extreme values. The baseline now represents the marginal effect on the innovation output generated by increasing the size of the firm and the lab size by one percent. Because of the log transformation, the interpretation of coefficients is in terms of 'percentage change.' For example, Model (7) in Table 4 indicates that relative to one percent increase in firm size and lab size, a one percent increase in publishing capabilities of both partners (measured on the new scale) yields a four percent increase in the alliance output. Additional analyses (available upon request) show that results are not driven by the presence in the sample of firms and scientists without publications or patents.

Contribution to the literature of university–industry knowledge transfer

This study made a first step toward unpacking the dimensions of firm – scientist matching in research collaboration. The empirical analysis showed that research partners complement each other in publishing capabilities, but substitute each other in patenting capabilities. Knowledge diversity also matters in partner choice, as shown by the complementary relationship between knowledge specialization and knowledge diversification.

These findings are interesting both in the context of university – industry knowledge transfer and, more generally, for the literature on alliance partner selection. The results corroborate previous work that found a performance-enhancing effect for firms connected with university scientists with an outstanding research record. However, this study contributes to a better understanding of the underlying causes of these effects. While previous work has focused the value-added contribution of 'stars' in the innovation process, I show that, through the *ex ante* matching process, 'star' scientists team up with firms that also have outstanding publication records. Further, although publishing capabilities are important in the market for research, I show that firm – scientist alliance formation is driven by a more complex set of considerations, notably, by anti-assortative matching based on patenting capabilities.

Contribution to the literature on alliance formation

Despite the importance of matching in partner choice, theoretical and empirical work in this area is in an early stage. This paper argued that the drivers of alliance performance cannot be fully understood without taking into account the matching process that precedes the formation of research alliances, and it provided a rigorous approach for examining partnership formation. The paper introduced a theoretical model that explicitly deals with the endogeneity problem created by the self-selection of ties in specific partnerships. The model demonstrated that anticipating synergistic gains and competing to ally with better partners leads to sorting in the market, which, in turn, explains why certain firm – scientist alliances create more value and enjoy higher innovation performance. The empirical analysis followed the theoretical model closely and focused on identifying the underlying sorting patterns created by complementarity and substitutability of partner attributes. Both the theoretical framework and the methodology can be applied to settings other than knowledge creation.

ACKNOWLEDGEMENTS

This paper has significantly benefited from comments and suggestions by the members of my dissertation committee, Rajshree Agarwal, David Audretsch, Janet Bercovitz, Glenn Hoetker, and Joseph Mahoney. I am also thankful to Editor Will Mitchell and three anonymous referees for their suggestions for improvement during the reviewing process. Several people have provided useful comments at various stages of this project, in particular, Gautam Ahuja, Thomas Astebro, Pierre Azoulay, Dan Elfenbein, Alfonso Gambardella, Bernard Garrette, Shane Greenstein, Nicola Lacetera, Kristina McElheran, Deepak Somaya, and Scott Stern. I am grateful to Janet Bercovitz for generously sharing data for this project. The Ewing Marion Kauffman Foundation and the Academy for Entrepreneurial Leadership at the University of Illinois at Urbana-Champaign have provided invaluable financial support. Any remaining errors are my own.

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.

Aghion P, Dewatripont M, Stein JC. 2008. Academic freedom, private-sector focus, and the process of innovation. *RAND Journal of Economics* **39**(3): 617–635.

Ahuja G. 2000. The duality of collaboration: inducements and opportunities in the formation of interfirm linkages. *Strategic Management Journal, March Special Issue* **2**: 317–343.

Anton JJ, Yao DA. 2003. Patents, invalidity, and the strategic transmission of the enabling information. *Journal of Economics and Management Strategy* **12**(2): 151–178.

Arora A, Gambardella A. 1994. Explaining technological information and utilizing it: scientific knowledge, capacity, and external linkages in biotechnology. *Journal of Economic Behavior and Organization* **24**(1): 91–114.

Audretsch DB, Stephan P. 1996. Company-scientist locational links: the case of biotechnology. *American Economic Review* **86**(3): 641–652.

Azoulay P, Ding W, Stuart T. 2007. The determinants of faculty patenting behavior: demographics or opportunities? *Journal of Economic Behavior and Organization* **63**(4): 599–623.

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

Belderbos R, Carree M, Lokshin B. 2004. Cooperative R&D and firm performance. *Research Policy* **33**(10): 1477–92.

Bercovitz J, Feldman M. 2007. Fishing upstream: firm strategic research alliances with universities. *Research Policy* **36**(7): 930–948.

Bercovitz J, Feldman M. 2008. Academic entrepreneurs: organizational change at the individual level. *Organization Science* **19**(1): 69–89.

Blumenthal D, Gluck M, Louis KS, Stoto MA, Wise D. 1986. University-industry research relationships in biotechnology: implications for the university. *Science* **232**(4756): 1361–1366.

Brusoni S, Prencipe A, Pavitt K. 2001. Knowledge specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make? *Administrative Science Quarterly* **46**(4): 597–621.

Cassiman B, Veugelers R. 2006. In search of complementarity in innovation strategy: internal R& D and external knowledge acquisition. *Management Science* **52**(1): 68–82.

Chatain O. 2010.Where do client-specific scope economies come from? A revealed preference analysis. Working paper, Wharton Business School, University of Pennsylvania, Philadelphia, PA.

Cockburn I, Henderson R.1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *Journal of Industrial Economics* **46**(2): 157–182

Cohen WM, Nelson RR, Walsh JP. 2002. Links and impacts: the influence of public research on industrial R&D. *Management Science* **48**(1): 1–23.

Dahlin KB, Weingart LR, Hinds PJ. 2005. Team diversity and information use. *Academy of Management Journal* **48**(6): 1107–1123.

David PA, Hall BH, Toole A. 2000. Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy* **29**(4-5): 497–529.

Elfenbein D. 2007. Publications, patents, and the market for university inventions. *Journal of Economic Behavior and Organization* **63**(4): 688–715.

Fabrizio KR. 2009. Absorptive capacity and the search for innovation. *Research Policy* **38**(2): 255–267.

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

Fleming L, Sorenson O. 2004. Science as a map in technological search. *Strategic Management Journal, August-September Special Issue* **25**: 909–928.

Fox J. 2009. Estimating matching games with transfers. NBER working Paper 14382. National Bureau of Economic Research, Cambridge, MA. Available at: http://www.nber.org/papers/w14382 (21 February 2012)

Fox J. 2010. Identification in matching games. *Quantitative Economics* **1**: 203–254.

Gans JS, Stern S. 2010. Is there a market for ideas? *Industrial and Corporate Change* **19**(3): 805–837.

Gans JS, Murray F, Stern S. 2011. Contracting over the disclosure of scientific knowledge: intellectual property and academic publication. Working paper series. Available at SSRN: http://ssrn.com/abstract = 1559871 (21 February 2012)

Gittelman M, Kogut B. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* **49**(4): 366–382.

Hall BH, Ziedonis RH. 2001. The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979–1995. *RAND Journal of Economics* **32**(1): 101–128.

Helfat CE, Finkelstein S, Mitchell W, Peteraf MA, Singh H, Teece, DJ, Winter SG. 2007. *Dynamic Capabilities: Understanding Strategic Change in Organizations*, Blackwell Publishing, Oxford, UK.

Hellmann T. 2007. The role of patents for bridging the science to market gap. *Journal of Economic Behavior and Organization* **63**(4): 624–647.

Henderson R, Cockburn I. 1996. Scale, scope, and spillovers: the determinants of research productivity in drug discovery. *RAND Journal of Economics* **27**(1): 32–59.

Hicks D. 1995. Published papers, tacit competencies, and corporate management of the public/private character of knowledge. *Industrial and Corporate Change* **4**(2): 401–424.

Jensen R, Thursby M. 2001. Proofs and prototypes for sale: the licensing of university inventions. *American Economic Review* **91**(1): 240–259.

Jones B. 2009. The burden of knowledge and the death of the Renaissance man: is innovation getting harder? *Review of Economic Studies* **76**(1): 283–317.

Kenney M. 1986. *Biotechnology: The University Industrial Complex*. Yale University Press: New Haven, CT.

Lacetera N. 2009. Different missions and commitment power in R&D organization: theory and evidence on industry-university relations. *Organization Science* **20**(3): 565–582.

Lippman SA, Rumelt RP. 2003. A bargaining perspective on resource advantage. *Strategic Management Journal* **24**(11): 1069–1086.

Lowe RA. 2001. The role and experience of start-ups in commercializing university inventions. In *Entrepreneurial Inputs and Outcomes*, Libecap G (ed). JAI Press: Amsterdam, the Netherlands; 189–222.

Mansfield E, Lee JY. 1996. The modern university: contributor to industrial innovation and recipient of industrial R&D support. *Research Policy* **25**(7): 1047–1058.

Milgrom P, Roberts J. 1995. Complementarities and fit: strategy, structure, and organizational change in manufacturing. *Journal of Accounting and Economics* **19**(2): 179–208.

Mortensen DT. 1988. Matching: finding a partner for life or otherwise. *American Journal of Sociology* **94**: S215–S240.

Murray F. 2011. The oncomouse that roared: hybrid exchange strategies as a source of distinction at the boundary of overlapping institutions. *American Journal of Sociology* **116**(2): 341–88.

Murray F, Stern S. 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *Journal of Economic Behavior and Organization* **63**(4): 648–687.

Nelson RR. 2004. The market economy, and the scientific commons. *Research Policy* **33**(3): 455–471.

Orsenigo L. 1989. *The Emergence of Biotechnology*. Pinter Publishing: London, UK.

Parmigiani A, Mitchell W. 2009. Complementarity, capabilities, and the boundaries of the firm: the impact of within-firm and interfirm expertise on concurrent sourcing of complementary components. *Strategic Management Journal* **30**(10): 1065–1091.

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

Roth AE, Sotomayor M. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press: Cambridge, MA.

Santiago D, Fox JT. 2008. A toolkit for matching maximum score estimation and point and set identified sub-sampling inference. Available at: http://www-personal.umich.edu/~jtfox/ (21 February 2012)

Shaver JM. 1998. Accounting for endogeneity when assessing strategy performance: does entry mode choice affect FDI survival? *Management Science* **44**(4): 571–585.

Sorensen M. 2007. How smart is smart money? A two-sided matching model of venture capital. *Journal of Finance* **62**(6): 2725–2762.

Stephan P. 1996. The economics of science. *Journal of Economic Literature* **34**(3): 1199–1235.

Stephan P. 2010. Economics of science. In *Handbook of Economics and Innovation* (Volume **1**), Hall BH, Rosenberg N (eds). Elsevier: Amsterdam, the Netherlands; 217–274.

Stern S.1995. Incentives and focus in university and industrial research: the case of synthetic insulin. In *Sources of Medical Technology: Universities and Industry (Medical Innovation at the Crossroads)* Rosenberg N, Gelijns AC, Dawkins H (eds). National Academy Press: Washington, D.C.; 157–187.

Stuart TE, Ding W. 2006. When do scientists become entrepreneurs? The social structural antecedents of commercial activity in the academic life sciences. *American Sociological Journal* **112**(1): 97–14.

Thursby JG, Thursby MC. 2002.Who is selling the ivory tower? Sources of growth in university licensing. *Management Science* **48**(1): 90–104.

Uzzi B.1996.The sources and consequences of embeddedness for the economic performance of organizations. *American Sociological Review* **61**(4): 674–698.

Wang Q, von Tunzelmann N. 2000. Complexity and the functions of the firm: breadth and depth. *Research Policy* **29**(7/8): 805–818.

Yang Y, Shi M, Goldfarb A. 2009. Estimating the value of brand alliances in professional team sports. *Marketing Science* **28**(6): 1095–1111.

Zucker LG, Darby MR, Armstrong J. 1998. Geographically localized knowledge: spillovers or markets? *Economic Inquiry* **36**(1): 65–86.

Zucker LG, Darby MR, Armstrong J. 2002. Commercializing knowledge: university science, knowledge capture, and firm performance in biotechnology. *Management Science* **48**(1): 138–153.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix

DOI: 10.1002/smj