Endogenous Matching in University-Industry Collaboration: Theory and Empirical Evidence from the UK

Albert Banal-Estañol
Inés Macho-Stadler
David Pérez-Castrillo

July 2013

Barcelona GSE Working Paper Series

Working Paper nº 704

Endogeneous Matching in University-Industry Collaboration:

Theory and Empirical Evidence from the UK*

Albert Banal-Estañol[†], Inés Macho-Stadler[‡] and David Pérez-Castrillo[§]

July 2013

Abstract

We develop a two-sided matching model to analyze collaboration between heterogeneous academics and firms. We predict a positive assortative matching in terms of both scientific ability and affinity for type of research, but negative assortative in terms of ability on one side and affinity in the other. In addition, the most able and most applied academics and the most able and most basic firms shall collaborate rather than stay independent. Our predictions receive strong support from the analysis of the teams of academics and firms that propose research projects to the UK's Engineering and Physical Sciences Research Council.

JEL Classification numbers: O32, I23

Keywords: matching, industry-science links, research collaborations, basic versus applied research, complementarity

*We thank Christian Fons-Rosen, Bing Guo, Ernest Miguelez and Anna Toldrà for very useful comments, as well as Eduard Clariana, Carlos Marcos and Jonás Nieto for excelent research assistance. We also thank the very constructive comments received at the Katholieke Universiteit Leuven, Universitat de Barcelona and at the workshops on "University-Industry Collaboration" in Barcelona, "Incentives and the Design of Mechanisms in Open Science" in Bordeaux and "Informal IO" in Barcelona. We gratefully acknowledge the financial support from the research grants (ECO2010-15052 and ECO2012-31962), Generalitat de Catalunya (2009SGR-169), Ramon y Cajal, and ICREA Academia. The last two authors are fellows of MOVE.

†Universitat Pompeu Fabra and Barcelona GSE; Department of Economics and Business. E-mail: albert.banalestanol@upf.edu.

‡Universitat Autonoma de Barcelona and Barcelona GSE; Department of Economics and Economic History. Email: ines.macho@uab.es.

§Universitat Autonoma de Barcelona and Barcelona GSE; Department of Economics and Economic History. Email: david.perez@uab.es.

1 Introduction

Science and innovation in modern economies often involves collaboration across institutional boundaries. Academic research groups sometimes work independently, but inter-institutional and international collaborations and coauthorships are very common (Wagner and Leydesdorff, 2005). Similarly, while some technologies are developed by one single firm, many others are developed by research joint ventures (Kamien et al., 1992). Fortunately, a substantial body of research in the economics and management literatures has identified the causes and the consequences of inter-institutional collaboration *within* institutional markets, i.e., "one-sided market" partnerships (see Katz and Martin, 1997, and Caloghirou et al., 2003, for reviews).

But, the full transformation of modern societies into knowledge- and science-based economies also requires collaboration *across* institutional markets, i.e., "two-sided market" partnerships. Business-science links through joint research, consulting or training arrangements, for example, are key channels of knowledge transfer from academia to industry according to both academics (Agrawal and Henderson, 2002) and firms (Cohen et al., 2002). As a result, university-industry collaborations are nowadays stronger and more widespread than ever before (Jensen and Thursby, 2001; Perkmann et al., 2013). Unfortunately, in spite of their tremendous importance, we know very little about which groups of which institutions engage in collaboration and which two-sided market partnerships are actually formed.

Take for example the partnership formed in 2007 by Professor Sir Colin John Humphreys of Cambridge University, who specializes in electron microscopy and analysis, and FEI, a world leading company in the production and distribution of electron microscopes. In this case, a prolific researcher of a top university, whose research is considered basic, collaborated with a research-intensive firm, heavily oriented toward basic research. We ask if this is the most common pattern: Do top academics collaborate with top firms, whereas less productive researchers collaborate with less productive firms? Do they collaborate because they have similar preferences? Do partners choose each other because of individual or institutional characteristics? Are other, less productive and more applied academics, more likely to stay independent?

This paper investigates the outcomes of the two-sided market collaboration process and, in particular, the characteristics of the resulting partnerships. We study what type of partners on each side of the market are more likely to collaborate with each other, and which characteristics affect the likelihood of collaborating, as opposed to working independently. We consider both "horizontal" and "vertical" characteristics, i.e., those related to affinity (e.g., preferences for a type of scientific research) and those related to ability (e.g., capacity to produce high-quality scientific output). We show theoretically and empirically that collaboration decisions are affected both by affinity-based and ability-based characteristics, as well as by individual and institutional characteristics.

Collaboration has costs and benefits for participants on both sides of the market. Academics claim that industry collaboration provides them with additional funds and insights (Lee, 2000; Mansfield, 1995), but it might also bias their selection of research topics and methodology (Florida and Cohen, 1999). Firms report collaborating with academics to get access to new university research and discoveries (Lee, 2000). Some of these outcomes, however, have no or little commercial value (Jensen et al., 2003). Firms are also concerned with the differences in terms of organizational and institutional structure, and with the existence of the open science culture in academia (Dasgupta and David, 1994).

Before deciding whether to collaborate, academics and firms must therefore weigh the benefits in terms of complementarities and the costs in terms of divergent interests. As a whole, collaboration allows firms to obtain better patents, more products and increased sales (Cockburn and Henderson, 1998; Cassiman and Veugelers, 2006; Zucker et al., 2002). On the academic side, collaboration has recently been linked to a higher number of academic research publications (Fabrizio and DiMinin, 2008; Azoulay et al., 2009).[1] Unfortunately, most of the empirical evidence on performance provides average effects, across all partnerships. Recent evidence, however, stresses the importance of the characteristics of the matched partners in assessing collaboration outcomes. Banal-Estañol et al. (2013), for example, show that the research projects in collaboration with firms produce more scientific output than those without them if and only if the firms in the project are research-intensive.

The rewards from collaboration might thus be highly heterogeneous and depend on own, as well as on the potential partner's, characteristics.[2] For instance, all academics, but especially those that are more research-oriented, might prefer firms that encourage their employees to publish scientific articles (Cockburn and Henderson, 1998). Similarly, all firms might prefer to collaborate with "star" academics, as their input increases firm performance (Zucker et al., 2002). Research-oriented firms and star academics, however, might not be willing or able to

[1] Agrawal and Henderson (2002), however, find no effect of the number of patents on the number of publications. Banal-Estañol et al. (2010) find an inverted U-shaped relationship between industry collaboration and academic research output.

[2] Academic researchers' individual characteristics and attitudes, as well as local group norms play a role in the collaboration decision (Louis et al., 1989). Firms' size, absorptive capacity and the adoption of open search strategies are also important factors in the firms' willingness to collaborate (Veugelers and Cassiman, 2005, Mohnen and Hoareau, 2003; Bercovitz and Feldman, 2007). Geographical proximity between the researchers' university and the firms has also been shown to be important, particularly for researchers in universities with modestly rated faculties (Audretsch and Stephan, 1996).

collaborate with all participants on the other side of the market. Given the costs and benefits of collaborating with each potential partner, how do academics and firms mutually choose each other, and which of them decide to work independently?

To understand the mechanisms at work, we build a one-to-one two-sided matching market model of academic researchers and firms developing research projects. Participants on each side of the market are heterogeneous in terms of project preferences (degree of "appliedness") and scientific ability (past publications, patents, or know-how). We allow each participant to develop a project on her/its own, or search for an appropriate partner on the other side of the market to agree on and develop a collaborative project. Our model is thus part of the recent literature on two-sided market matching models with endogenous contracts (e.g., Legros and Newman, 2002; Dam and Pérez-Castrillo, 2006; Alonso-Paulí and Pérez-Castrillo, 2012).[3]

We first derive results on the types of projects chosen, and individual investments made, by stand-alone and collaborating participants. Collaborating partners end up developing projects away from their most preferred ("ideal") type, or in areas where they are less productive, with a bias that is proportional to the relative value attached to the project by each partner. Investment in, and the value of, the project is increasing in the abilities and decreasing in the distance between the ideal types of the two partners. We make predictions on who collaborates with whom and who stays independent. If the costs due to collaboration are not too large, the most able researchers and the most able firms engage in collaboration whereas the least able stay independent. We also show that if the types of academics are generally more basic than those of firms, the most applied researchers and the most basic firms collaborate, whereas the most basic researchers and the most applied firms develop projects on their own.

With respect to the partnerships formed, our theoretical model makes three predictions. First, the matching is positive assortative in terms of scientific ability, i.e., top academics collaborate with top firms and less able academics collaborate with less able firms. This is because partner abilities are complementary: the higher the ability of the academics, the more they benefit from the higher investment of the firms with higher ability, and vice-versa. Second, the matching is also positive assortative in terms of affinity, i.e., academics with more applied bias collaborate with firms with more applied bias. The reasons, however, are different. A positive assortative matching in terms of affinity minimizes the total inefficiencies due to the distances between the ideal types of the matched partners. Appropriate (pecuniary or non-pecuniary) transfers ensure that the equilibrium matching maximizes the sum of values of all partnerships, and not necessarily the value of any particular one. Finally, we show that the matching is neg-

[3] Our paper is also related to the optimal assignment literature following Becker (1973).

ative assortative in terms of ability-affinity pairs, i.e., the higher the ability of the academics, the closer they are to their partners in terms of ideal type.

We test our theoretical predictions on the teams of academic researchers and firms that have proposed research projects to the Engineering and Physical Sciences Research Council (EPSRC), the main government agency for research funding for the engineering departments of the UK universities. The EPSRC grants are allocated to teams of academic researchers alone and also to teams that include one or more firms as industry partners. We build a unique dataset of $5,855$ projects with participants' past publications, which allow us to construct continuous measures of ability and affinity. We use the normal count and the impact-factor-weighted sum of publications as proxies for scientific ability and the proportion of publications in basic or applied journals as a proxy for the ideal type of research (Narin et al., 1976; Godin, 1996; van Looy et al., 2006).

Following Agrawal et al. (2008) and Gompers et al. (2012), we test for who partners with whom using both the formed partnerships and a set of plausible alternatives, or counterfactual pairs, constructed using exogenous characteristics. We show that the cross-partial derivative of the measures of the ability of each partner, as well as the cross-partial derivative of the measures of affinity, are positive, thus providing support for the theoretical predictions of positive matching in terms of ability and affinity, respectively. The cross-partial derivative of ability and distance of types is instead negative, thus providing support for the theoretical prediction of negative assortative matching in terms of ability-affinity pairs.

We quantify the effects by computing the marginal effects of the likelihood of forming a link of several dummy variables that capture the relative position of each partner on each side of the market. We show that pairs of academics and firms that are both (neither) at the top quartile of the distribution of abilities are 29.5% (10%) more likely to be matched, compared to those pairs in which one is in the top quartile and the other is not. Similarly, academics and firms are 33% (39.5%) more likely to be matched if they are both above (below) the median in terms of type. Although positive, the effects for being both above (below) the median in terms of ability are less strong. Our results suggest that matching occurs at the top of the distribution in terms of ability but over the whole distribution in terms of affinity.

We also assess the relative importance of the horizontal (affinity) versus the vertical (ability) characteristics, as well as the individual versus the institutional characteristics. The horizontal characteristics are relatively more important than the vertical ones, both in terms of magnitude and significance of the effects. Importantly, the characteristics at the individual level are more relevant than those at the institutional level. A top firm tends to collaborate with a top researcher, independently of where she is based; hence, top firms form links with top academic institutions only insofar as they include top researchers. This reinforces the view that the fundamental unit of collaboration is composed of individuals, not institutions (Katz and Martin, 1997).

Finally, we test for the characteristics of the academics that submitted collaborative instead of non-collaborative projects. Confirming our theoretical predictions, the most applied as well as the most able researchers are significantly more likely to propose collaborative projects. Academics that are above the median in terms of ability and those who are more applied than the median are 9.1% and 39.5% more likely to propose collaborative projects, respectively. In terms of institutional characteristics, we show that academics in larger universities, in terms of number of projects, are also more likely to submit collaborative projects. Other institutional variables, such as the scientific level of the department, are less important.

Our leading example fits most of the general properties of university-industry partnerships. First, professor Humphreys, a leading academic, collaborates with FEI, a leading firm. Both of them share the same preferences for the type of research, and as two leaders, they are matched with a partner with similar preferences. Second, as a top academic and a top, basic firm, they end up collaborating as opposed to remaining independent. But in this example, in contrast to one of our results, a basic academic does not stay independent. Finally, our results suggest that FEI collaborates with professor Humphreys, not because he is a professor at Cambridge, but because of his individual characteristics.

This paper provides the first theoretical two-sided matching model to analyze university-industry relationships. Based on 46 case-study interviews, Carayol (2003) proposes a typology of business-science collaborations and argues that firms involved in high (low) risk projects are matched with academic teams of a high (low) excellence. Agarwal and Ohyama (forthcoming) study, both theoretically and empirically, the labor market for scientists. The academic and private sectors choose among scientists who differ in their ability and preferences, and scientists choose between academia and industry. Our setup, of course, differs from theirs and includes more than two classes of participants on each side of the market.

Empirically, the paper closest to ours is Mindruta (forthcoming), who identifies ability-based characteristics as a source of complementarity in university-industry collaboration. We show that affinity-based characteristics are even more important than ability-based characteristics. Methodologically, she assumes a joint production function and estimates the interaction parameters and joint production values, following Fox's (2008) empirical strategy. Instead, we derive a joint production function from first principles, i.e., from preferences and optimization, in our theoretical section. In our empirical section, we do not attempt to estimate the value of the formed pairs so as to avoid making assumptions on joint production functions. Instead, we rely on matched and non-matched counterfactual pairs to analyze which characteristics make participants more likely to be matched.[4]

This paper is organized as follows. In Section 2, we introduce the model and obtain the theoretical results. In Section 3, we describe our dataset. In Section 4, we perform the empirical analysis. Finally, in Section 5 we conclude. Proofs are in the Appendix.

2 Theoretical model

We consider a "market" with m heterogeneous academic researchers $\mathcal{A} = \{A_1, A_2, ..., A_m\}$, and n heterogeneous firms $\mathcal{F} = \{F_1, F_2, ..., F_n\}$. Academics and firms can develop research projects on their own, labelled as "non-collaborative", or form academic-firm partnerships and agree on the terms of a "collaborative" project. That is, the matching and the project characteristics are endogenous. Due to time and other constraints, agents have capacity to work on a limited number of projects. For simplicity, we assume that each academic and firm develops only one project. That is, we model a "one-to-one two-sided matching market".[5]

2.1 Collaborative and non-collaborative projects

In this subsection, we describe the projects that would be independently developed by a given academic researcher A and firm F, and the collaborative project they would develop should they decide to form a partnership. In the next subsection, we endogenize the formation of partnerships.

Project type and preferences of the participants Projects can be of a more basic or of a more applied nature. We identify the project's level of "appliedness" (or of any other horizontal characteristic) by a parameter $x \in [0, 1]$, where $x = 0$ corresponds to the least applied project possible (the most basic one) and $x = 1$ corresponds to the most applied project possible.

We denote by x_A and x_F the most preferred (ideal) type of project of academic researcher A and firm F, and by v_A and v_F the value of a positive result from a research project of their most preferred type. A project has less value for academic A if its type x is different from

[4] Sorensen (2007) estimates a structural model based on a two-sided matching model between venture capitalists and entrepreneurs to analyze the market for venture capital investments.

[5] A two-sided matching market is one in which there are two distinct sets of agents. It is one-to-one if an agent from one side of the market can be matched only with an agent from the other side or remain unmatched. For an introduction to matching markets, see Roth and Sotomayor (1990).

x_A; the larger the distance $|x - x_A|$, the larger the loss in value. Following Pereira (2007) and Banal-Estañol et al. (2013), we model the loss in value as a "transportation cost", in the spirit of the Hotelling model.[6] That is, the value for academic A of a positive result of a project of type x is $v_A(1 - t(x - x_A)^2)$, where $t \leq 1$ denotes the transportation cost parameter. Similarly, the value for firm F of a positive result of a project of type x is $v_F(1 - t(x_F - x)^2)$.

Figure 1 represents the value of a positive result for A and F, as a function of the type of project x.

[Insert Figure 1 here]

Investment levels and outcome of projects When an academic or firm runs a project on their own, the number of positive results (or the probability of obtaining a positive result) depends on their own ability and investment. For simplicity, we assume that it is given by $\delta_A I_A$ and $\delta_F I_F$, where δ_A (resp. δ_F) represents the academic's (resp. firm's) ability, or efficiency, and I_A (or I_F) represents the academic's (resp. firm's) investment level. The parameter δ_A measures the technical and scientific level of academic A, her publications, the patents and know-how she owns, the quality of the labs she works in, etc., whereas the parameter δ_F measures the scientific level of firm F, its absorptive capacity, the level of its human capital, etc. The investment level I_A (or I_F) can be effort, time or money invested in the project. Investment has an associated cost $\frac{c}{2}I_A^2$ (resp. $\frac{c}{2}I_F^2$).

Summarizing, when an academic with characteristics (x_A, δ_A) runs a non-collaborative project of type x in which she invests I_A, her profits are

$$\pi_A(x, I_A) = v_A\left(1 - t(x - x_A)^2\right)\delta_A I_A - \frac{c}{2}I_A^2.$$

Similarly, a firm with characteristics (x_F, δ_F) that develops a non-collaborative project of type x in which it invests I_F, obtains profits

$$\pi_F(x, I_F) = v_F\left(1 - t(x_F - x)^2\right)\delta_F I_F - \frac{c}{2}I_F^2.$$

In collaborative projects, the value to each party depends on the agreed type of the project x and on the investments of both participants, I_A and I_F. For the academic, it is given by

$$\pi_A(x, I_A, I_F) = v_A\left(1 - t(x - x_A)^2\right)\lambda_A(\delta_F + \delta_A)(I_F + I_A) - \frac{C}{2}I_A^2, \tag{1}$$

[6] Pereira (2007) analyzes theoretically the characteristics of partnership agreements when informational problems are present. She shows that two different structures of partnership governance - centralized and decentralized - may optimally use the type of the project to motivate the supply of non-contractible resources. In her approach the partners are predetermined. In contrast, our paper ignores incentive issues and concentrates on analyzing which collaborative agreements will be formed.

whereas for the firm it is given by

$$\pi_F(x, I_A, I_F) = v_F \left(1 - t\left(x_F - x\right)^2\right) \lambda_F \left(\delta_F + \delta_A\right) \left(I_F + I_A\right) - \frac{C}{2} I_F^2. \tag{2}$$

Following previous results in the literature, we argue that collaboration has advantages as well as costs.[7] On the positive side, the probability or the number of positive results depends on the ability and investment of both partners, represented by $(\delta_F + \delta_A)(I_F + I_A)$. This is a simple way to introduce complementarities in investment levels. On the negative side, the fraction of the outcome that can be appropriated by the academic researcher and firm, represented by the parameters λ_A, $\lambda_F \in [0, 1]$, might add up to less than one. In addition, the cost of investing in a collaborative project can be higher than in a non-collaborative project, $C \geq c$, because of, for example, coordination costs. Our specification captures collaboration's benefits (in terms of complementarities) and costs (in terms of divergent interests and coordination difficulties).

An alternative interpretation We interpret $v_A t\left(x - x_A\right)^2$ as the academic's loss in terms of utility when developing a project of type x instead of a project of her most preferred type x_A. In this case, investment (effort) is equally productive in all types of research. An alternative interpretation of the same formulation is that x_A corresponds to the type where the academic's investment is most productive, given her expertise, and the value of a positive result is the same. That is, an investment level I_A in a project of type x_A gives $\delta_A I_A$ positive results, whereas the same investment level in a project of type x reduces the number of positive results by $t\left(x - x_A\right)^2 \delta_A I_A$. The same re-interpretation is valid for the firms' profits in non-collaborative projects as well as for both academics' and firms' profits in collaborative projects.

Optimal non-collaborative projects In a stand-alone project, academic A (resp. firm F) chooses the type of project x and the investment I_A (resp. I_F) that maximizes her (resp. its) profits. Proposition 1 describes the optimal project.

Proposition 1 *The optimal non-collaborative project $P_A^n := (x_A^n, I_A^n)$ for academic A with individual characteristics (x_A, δ_A) is*

$$x_A^n = x_A \qquad \text{and} \qquad I_A^n = \frac{1}{c} v_A \delta_A,$$

and the optimal non-collaborative project $P_F^n := (x_F^n, I_F^n)$ for firm F with individual characteristics (x_F, δ_F) is

$$x_F^n = x_F \qquad \text{and} \qquad I_F^n = \frac{1}{c} v_F \delta_F.$$

[7] See, for instance, Dasgupta and David (1994) and Veugelers and Cassiman (2005).

Not surprisingly, the optimal type of a non-collaborative project corresponds to the most preferred type. The optimal investment increases with the value associated to a positive result and the scientific ability but it is decreasing in the cost of the investment. Substituting into the profit functions, the profits of the stand-alone projects are:

$$\pi_A^n(x_A, \delta_A) = \frac{1}{2c} v_A^2 \delta_A^2 \qquad \text{and} \qquad \pi_F^n(x_F, \delta_F) = \frac{1}{2c} v_F^2 \delta_F^2.$$

Optimal collaborative projects When they develop a collaborative project, the academic and the firm need to agree on the type x of the project and the level of investments, I_A and I_F, each will devote. In addition to $P_{AF} = (x, I_A, I_F)$, they may agree on a monetary transfer.[8] The possibility to transfer profits (or utility) implies that both the academic researcher and the firm have incentives to agree on the type and the investments that maximize joint profits. In other words, the optimal project $P_{AF}^c = (x^c, I_A^c, I_F^c)$ corresponds to the vector (x, I_A, I_F) that maximizes joint profits $\Pi_{AF}(x, I_A, I_F) \equiv \pi_A(x, I_A, I_F) + \pi_F(x, I_A, I_F)$.

The next proposition provides the optimal agreement.

Proposition 2 *The optimal collaborative project $P_{AF}^c := (x^c, I_A^c, I_F^c)$ between academic A with individual characteristics (x_A, δ_A) and firm F with individual characteristics (x_F, δ_F) is*

$$x^c = \frac{\lambda_A v_A}{\lambda_A v_A + \lambda_F v_F} x_A + \frac{\lambda_F v_F}{\lambda_A v_A + \lambda_F v_F} x_F, \quad \text{and} \quad I_A^c = I_F^c = \frac{1}{C} V_{AF}\left(|x_F - x_A|\right)\left(\delta_A + \delta_F\right)$$

where

$$V_{AF}\left(|x_F - x_A|\right) \equiv \lambda_A v_A + \lambda_F v_F - \frac{\lambda_A v_A \lambda_F v_F}{(\lambda_A v_A + \lambda_F v_F)} t (x_F - x_A)^2.$$

The type of a collaborative project lies between the most preferred type of project for the academic and the most preferred for the firm. The distances between the agreed type and the ideal types, $|x^c - x_A| = \frac{\lambda_F v_F}{\lambda_A v_A + \lambda_F v_F} |x_F - x_A|$ and $|x^c - x_F| = \frac{\lambda_A v_A}{\lambda_A v_A + \lambda_F v_F} |x_A - x_F|$, depend on the relative value of the results for the participants and their capacity to appropriate them. The higher the value that the academic can appropriate from the project, the closer the type of project to her most preferred one is. Similar to the non-collaborative projects, the optimal investment levels are decreasing in the cost of investment and increasing in the partners' abilities and in the "joint value" of a result $V_{AF}\left(|x_F - x_A|\right)$, which depends on the distance $|x_F - x_A|$.

Substituting the results of Proposition 2, the total profits from an optimal collaborative project, $\Pi_{AF}^c\left(x_A, \delta_A, x_F, \delta_F\right)$ are

$$\Pi_{AF}^c = \frac{1}{C}\left(V_{AF}\left(|x_F - x_A|\right)\right)^2 \left(\delta_A + \delta_F\right)^2. \tag{3}$$

[8] We assume that the transfer, or compensation, from one partner to the other is monetary for simplicity. We later discuss the implications for the results if transfers were not possible.

Total profits Π_{AF}^{c} are increasing in the partners' abilities and decreasing in the costs of investing, as well as in the distance between the most preferred project for the academic and the firm.

2.2 The Market Equilibrium

Academic researchers and firms are heterogeneous in terms of preferences for the type of research and ability, i.e., it is possible that $x_{A_i} \neq x_{A_{i'}}$, $x_{F_j} \neq x_{F_{j'}}$, $\delta_{A_i} \neq \delta_{A_{i'}}$ and/or $\delta_{F_j} \neq \delta_{F_{j'}}$ for $A_i, A_{i'} \in \mathcal{A}$ and $F_j, F_{j'} \in \mathcal{F}$. For simplicity, we assume that academics share the same valuation and appropriability, $v_{A_i} = v_A$ and $\lambda_{A_i} = \lambda_A$ for any $A_i \in \mathcal{A}$, and similarly for the firms, $v_{F_j} = v_F$ and $\lambda_{F_j} = \lambda_F$ for any $F_j \in \mathcal{F}$. Therefore, each academic A_i is characterized by a pair (x_{A_i}, δ_{A_i}) and each firm F_j by a pair (x_{F_j}, δ_{F_j}).

Notice that the two characteristics are radically different from the point of view of the partnership. The differences in ability represent "vertical" differentiation: an academic with a higher δ_A is a better academic than an academic with a lower δ_A, and the joint profits of a collaborative project increase with δ_A. The differences in the preference for the type of research represent "horizontal" differentiation: an academic with a high x_A is neither better or worse than an academic with a low x_A. The joint profits of a collaborative project increase when the types of academic and firm are similar, not when one of them is high or low.

We proceed as follows. After providing the formal definitions and some basic properties, we investigate which characteristics of an academic make her more likely to be matched with a given firm, and vice-versa. Second, we study which characteristics of the academics and the firms make them more likely to engage in collaborative, versus non-collaborative, projects.

Formal definitions Any academic in the population of academics $\mathcal{A} = \{A_1, A_2, ..., A_m\}$ (resp. any firm in the population of firms $\mathcal{F} = \{F_1, F_2, ..., F_n\}$) either collaborates with some firm (resp. academic) or she (it) does not collaborate. We represent the collaboration decision through a function that links an academic and a firm if they collaborate and an academic (firm) with herself (itself) if she (it) conducts a non-collaborative project.

A *matching* is a function that describes which academics and firms form partnerships and with whom. Formally, a matching is a mapping μ from $\mathcal{A} \cup \mathcal{F}$ (the union of the sets of academics and firms) to $\mathcal{A} \cup \mathcal{F}$ such that (i) $\mu(A) \in \mathcal{F} \cup \{A\}$ for all $A \in \mathcal{A}$, (ii) $\mu(F) \in \mathcal{A} \cup \{F\}$ for all $F \in \mathcal{F}$, and (iii) $\mu(A) = F$ if and only if $\mu(F) = A$ for all $A \in \mathcal{A}, F \in \mathcal{F}$. If A and F develop a collaborative project then $\mu(A) = F$ and $\mu(F) = A$. If A (resp. F) develops a non-collaborative project then $\mu(A) = A$ (resp. $\mu(F) = F$).

A matching μ is *positive assortative with respect to some characteristic* y if, the partner $\mu(A_i)$

of an academic A_i with a higher y than an academic $A_{i'}$, has a higher (or equal) y than the partner $\mu(A_{i'})$ of academic $A_{i'}$. A negative assortative matching is defined in a similar manner.

An *outcome* is a pair $(\mu, \mathcal{P})$ that describes which partnerships are formed and which projects are developed. In an *equilibrium outcome*, no academic or firm can improve upon her or its current payoff. That is, no academic or firm in a collaborative project can be better off by developing a non-collaborative project and no academic and firm can be better off by quitting their current partners, if any, and forming a new partnership.

Basic properties We borrow from previous literature two basic properties of the equilibrium outcomes.[9,10] First, *in an equilibrium outcome* $(\mu, \mathcal{P})$, *all the projects are optimal.* That is, it is not possible for any academic, firm or existing academic-firm pair to design an alternative project in which they are better off. As a result, the terms of the projects in any equilibrium outcome are the ones described in Propositions 1 and 2. Once we identify the matching μ in an equilibrium outcome, the two propositions uniquely determine the characteristics of the set of projects $\mathcal{P}$.

Second, in an equilibrium outcome, *the matching* μ is *efficient*, in the sense that it maximizes total surplus, i.e., the total surplus cannot be increased by reassigning firms and academics to different partnerships. This property derives from competition in the market: as firms compete among themselves for the best academic partner, and academics compete among themselves for the best firm partner, the resulting matching maximizes the total market surplus. If this was not the case, an agent, or a pair of agents, would obtain more benefits in an alternative matching.

Partnerships formed We now focus on the properties of "matched" academics and firms, i.e., those that develop collaborative projects in equilibrium outcomes. We investigate if the matching is positive or negative assortative with respect to ability and type.

To isolate the effects of heterogeneity in ability, we consider a market where all academics have the same preferences with respect to the types of the projects, and similarly for all the firms.

Proposition 3 *Consider a market* $(\mathcal{A}, \mathcal{F})$ *where* $x_{A_i} = x_A$ *for all* $A_i \in \mathcal{A}$ *and* $x_{F_j} = x_F$ *for all* $F_j \in \mathcal{F}$. *Then, for any equilibrium outcome* $(\mu, \mathcal{P})$, *the matching is positive assortative in terms*

[9] The proofs of the properties in this section follow the same arguments as those used in previous papers (see, for instance, Dam and Pérez-Castrillo, 2006, and Alonso-Paulí and Pérez-Castrillo, 2012).

[10] In our setup, equilibrium outcomes always exist and therefore the solution concept that we use is well-defined. This follows directly from Kaneko (1982). He generalizes the assignment game proposed by Shapley and Shubik (1972) and demonstrates the non-emptiness of the core in this setup.

of ability: if $\mu(A_i) = F_j$, $\mu(A_{i'}) = F_{j'}$ and $\delta_{A_i} \geq \delta_{A_{i'}}$, then $\delta_{F_j} \geq \delta_{F_{j'}}$.

Therefore, among the collaborative partners, top academics collaborate with top firms and academics of lower ability collaborate with firms of lower ability. Indeed, following the second property above, the equilibrium matching is positive assortative if the efficient matching is positive assortative. This holds if the characteristics of each partner are complementary to each other in the sum of profits.[11] In this case, partner abilities are complementary because higher ability academics benefit more from collaborating with higher ability firms: the higher the ability of an academic, the more she benefits from the higher investment decided by a firm with higher ability, and vice-versa. Figure 2 provides a numerical example with two possible matchings, μ and μ'. Academic A_1 benefits more from collaborating with firm F_1 than academic A_2 (in the figure, $20 - 12 > 10 - 5$). Similarly, firm F_1 benefits more from collaborating with academic A_1 than firm F_2 ($20 - 10 > 10 - 7$). Hence, the efficient outcome requires the collaboration of the best academic with the best firm ($40 + 12 > 22 + 20$). As a result, matching μ is an equilibrium but μ' is not.

[Insert Figure 2 here]

To isolate the effects of heterogeneity in project preferences, we next consider a market where all academics have the same ability, and similarly for all the firms.

Proposition 4 *Consider the market $(\mathcal{A}, \mathcal{F})$ where $\delta_{A_i} = \delta_A$ for all $A_i \in \mathcal{A}$ and $\delta_{F_j} = \delta_F$ for all $F_j \in \mathcal{F}$. Then, for any equilibrium outcome $(\mu, \mathcal{P})$, the matching is positive assortative in terms of type: if $\mu(A_i) = F_j$, $\mu(A_{i'}) = F_{j'}$ and $x_{A_i} \geq x_{A_{i'}}$, then $x_{F_j} \geq x_{F_{j'}}$.*

Proposition 4 implies that, among the collaborative partners, academics with more applied interests collaborate with firms with more applied bias. Figure 3 provides a numerical example with two possible matchings μ and μ'. Matching μ' cannot be an equilibrium: although the total distance between types in the two matchings is the same, there is more dispersion in the distances in matching μ'. Therefore, since the profits are (decreasing and) concave in the distance, the sum of profits in matching μ is larger and μ is an equilibrium outcome. Notice that even if the individual profits for A_3 and F_2 are higher in μ' ($10 - 8 > 0$ for both A_3 and F_2), matching μ'

[11] In formal terms, the efficient matching is positive assortative with respect to the characteristic y if $\Pi^c_{AF}\left(y_{A_i}, y_{F_j}\right) + \Pi^c_{AF}\left(y_{A_{i'}}, y_{F_{j'}}\right) \geq \Pi^c_{AF}\left(y_{A_i}, y_{F_{j'}}\right) + \Pi^c_{AF}\left(y_{A_{i'}}, y_{F_j}\right)$ whenever $y_{A_i} \geq y_{A_{i'}}$ and $y_{Fj} \geq y_{F_{j'}}$. A sufficient condition for the inequality is that $\frac{\partial^2 \Pi^c_{AF}}{\partial y_A \partial y_F} \geq 0$. See also Legros and Newman (2002) for more general sufficient conditions for positive or negative assortative matchings.

cannot be an equilibrium matching as their individual loss in μ can be compensated by A_2 and F_3 with the additional gains they obtain ($8-3$ each of them).

[Insert Figure 3 here]

Our model also generates predictions on the interaction between the two characteristics. In an equilibrium outcome there is an inverse relationship between ability and type distance. We can say that the matching is negative assortative in terms of ability-distance pairs.[12] The next proposition states this property in terms of ability of the academics, but we would also have an equivalent one in terms of the firms.

Proposition 5 *Consider the market* $(\mathcal{A}, \mathcal{F})$ *where* $\delta_{F_j} = \delta_F$ *for all* $F_j \in \mathcal{F}$ *and* $x_{A_i} = x_A$ *for all* $A_i \in \mathcal{A}$. *Then, for any equilibrium outcome* $(\mu, \mathcal{P})$, *the matching is negative assortative in terms of the academic's ability-distance pair: if* $\mu(A_i) = F_j$, $\mu(A_{i'}) = F_{j'}$ *and* $\delta_{A_i} \geq \delta_{A_{i'}}$ *then* $\left|x_{F_j} - x_A\right| \leq \left|x_{F_{j'}} - x_A\right|$.

Proposition 5 stems from the property that the higher the ability of the academic (or the firm), the more damaging the distance to the most preferred type is.

Discussion on the effect of transfers In our setup, we assume that transfers between firms and academic researchers are possible. We recognize the possibility that a firm (an academic) can make the collaborative project more appealing to its (her) partner than other competing offers for collaboration that the partner may receive. Firms can provide additional funding to the academics. Similarly, academics can provide consulting services and/or facilitate access to their laboratories. But, what would be the equilibrium matching if transfers were not possible?[13]

In terms of the vertical dimension, the ability, the matching would still be positive assortative. All academics agree that the best partner is the firm with the highest ability, and all the firms agree that the best partner is the academic with the highest ability. Given that the best academic would like to collaborate with the best firm and the best firm would like to match with the best academic, they are necessarily matched in the equilibrium matching (otherwise, they would deviate by joining in a collaborative project). By the same argument, once the best academic

[12] As it is customary, we have formally defined a positive or negative assortative matching in terms of one characteristic that affects both sides of the market. In the next proposition, we also use the term "negatively assortative" to refer to a property that affects a characteristic of one side of the market and a different one for the other side.

[13] One-to-one matching models where monetary transfers are not possible are usually referred to as "marriage markets". They were introduced by Gale and Shapley (1962).

and firm are not available, the second best academic is certainly matched to the second best firm in an equilibrium matching and, by induction, the kth best academic is matched with the kth best firm. Therefore, without transfers, the matching is also positive assortative in ability.

However, the matching in affinity (or, in general, in any horizontal characteristic) is not necessarily positive or negative assortative if transfers are not possible. In this case, the relevant variable is the distance between the academic's and the firm's most preferred project. If the distribution of types of the academics is the same as the one of the firms (i.e. $x_{A_1} = x_{F_1}$, $x_{A_2} = x_{F_2}$, and so on), then each agent finds a partner on the other side of the market with the same type and the equilibrium matching is $\mu(x_{A_k}) = x_{F_k}$ for all k. In this example, the matching is positive assortative.

But, if the distribution of types is asymmetric, the previous result may not hold. For instance, consider a market with two academics and two firms, where $x_{A_1} < x_{A_2} = x_{F_1} < x_{F_2}$. In this case, firm F_1 is the best partner for academic A_2, who is the best partner for firm F_1. None of them can be persuaded away and, in equilibrium, $\mu(x_{A_2}) = x_{F_1}$ The other two agents collaborate between themselves and therefore $\mu(x_{A_1}) = x_{F_2}$. In this example, the matching is negative assortative. Therefore, although there is an effect that pushes toward a positive assortative matching (an academic with a high x_A prefers a firm with a high x_F), the matching is not necessarily positive assortative in types if transfers among partners are not possible.

Collaborating versus staying independent We now identify which characteristics of academics and firms make them more likely to develop collaborative projects, as opposed to non-collaborative projects. Denote $\Delta(x_A, \delta_A, x_F, \delta_F)$ the "net benefits from collaboration", i.e., the difference between the joint profits in a collaborative project and the sum of the individual profits in non-collaborative projects:

$$\Delta(x_A, \delta_A, x_F, \delta_F) \equiv \Pi^c_{AF}(x_A, \delta_A, x_F, \delta_F) - \pi^n_A(x_A, \delta_A) - \pi^n_F(x_F, \delta_F).$$

Collaboration is jointly beneficial if $\Delta(x_A, \delta_A, x_F, \delta_F) > 0$. In equilibrium, only partnerships that are jointly beneficial can form.

We first analyze the effects of ability. The ability has a positive impact on the benefits from collaboration but it also increases the profits in non-collaborative projects. Lemma 1 states the trade-off in function of the ratio between δ_F and δ_A.

Lemma 1 *The net benefits from collaboration (i) increase in δ_A if and only if*

$$\frac{\delta_F}{\delta_A} > \frac{C}{2c} \frac{v_A^2}{(V_{AF}(|x_F - x_A|))^2} - 1, \tag{4}$$

and (ii) increase in δ_F if and only if

$$\frac{\delta_A}{\delta_F} > \frac{C}{2c}\frac{v_F^2}{(V_{AF}(|x_F - x_A|))^2} - 1. \tag{5}$$

The net benefits from collaboration increase in δ_A for any δ_A if the right-hand side of inequality (4) is negative. This (sufficient condition) holds if the relative cost of investing in a collaborative versus a non-collaborative project, C/c, is small, or if the relative value of the successes in collaboration, $V_{AF}(|x_F - x_A|)/v_A$, is large. This sufficient condition for inequality (4), and a similar condition for inequality (5), are stated as Conditions 1a and 1b.

Condition 1a *The market* $(\mathcal{A}, \mathcal{F})$ *satisfies* $x_{A_i} = x_A$ and $x_{F_j} = x_F$ *for all* $A_i \in \mathcal{A}, F_j \in \mathcal{F}$ *and* $\frac{C}{2c} \leq \frac{V_{AF}(|x_F - x_A|)^2}{v_A^2}$.

Condition 1b *The market* $(\mathcal{A}, \mathcal{F})$ *satisfies* $x_{A_i} = x_A$ and $x_{F_j} = x_F$ *for all* $A_i \in \mathcal{A}, F_j \in \mathcal{F}$ *and* $\frac{C}{2c} \leq \frac{V_{AF}(|x_F - x_A|)^2}{v_F^2}$.

Proposition 6 *For any equilibrium outcome* $(\mu, \mathcal{P})$,
(i) if Condition 1a *holds, high-ability academics collaborate with firms whereas low-ability academics stay independent: if* $\mu(A_i) = A_i$ *and* $\delta_{A_{i'}} < \delta_{A_i}$, *then* $\mu(A_{i'}) = A_{i'}$, *and*
(ii) if Condition 1b *holds, high-ability firms collaborate with academics whereas low-ability firms stay independent: if* $\mu(F_j) = F_j$ *and* $\delta_{F_{j'}} < \delta_{F_j}$, *then* $\mu(F_{j'}) = F_{j'}$.

We now consider how the net benefits from collaboration depend on the agents' types. The effect is not monotone. The net benefits from collaboration increase in the academic's type as long as it is lower than the firm's type, and decrease otherwise. As shown by Lemma 2, the net benefits from collaboration decrease in the distance between the partners' types.

Lemma 2 *The net benefits from collaboration (i) increase in* x_A *if and only if* $x_A < x_F$, *and (ii) increase in* x_F *if and only if* $x_F < x_A$.

This result is obtained because the joint profits decrease with the distance between partners' types whereas the individual profits do not depend on the agents' types. Still, Lemma 2 alone does not allow us to characterize which types of academics or firms are more likely to collaborate. It only states that collaboration is more likely for similar partners. Obtaining a precise prediction on the likelihood of collaboration requires assumptions on the distribution of types of academics and firms. Conditions 2a and 2b define two intuitive distributions of types where the academics' preferences are, in general, more basic than those of the firms.

Condition 2a *The market* $(\mathcal{A}, \mathcal{F})$ *satisfies* $\delta_{A_i} = \delta_A$, $\delta_{F_j} = \delta_F$ and $x_{A_i} \leq x_{F_j}$ *for all* $A_i \in \mathcal{A}, F_j \in \mathcal{F}$.

Condition 2b *The market* $(\mathcal{A}, \mathcal{F})$ *satisfies* $\delta_{A_i} = \delta_A$, $\delta_{F_j} = \delta_F$ *for all* $A_i \in \mathcal{A}, F_j \in \mathcal{F}$ *and* $x_{A_1} \leq ... \leq x_{A_k} = x_{F_1} \leq x_{A_{k+1}} = x_{F_2} \leq ... \leq x_{A_m} = x_{F_{m-k+1}} \leq x_{F_{m-k+2}} \leq ... \leq x_{F_n}$.

Proposition 7 *If condition 2a or 2b holds then, for any equilibrium outcome* $(\mu, \mathcal{P})$*, the most applied academics and the most basic firms collaborate, whereas the most basic academics and the most applied firms stay independent: (i) if* $\mu(A_i) = A_i$*, then* $\mu(A_{i'}) = A_{i'}$ *if* $x_{A_{i'}} < x_{A_i}$ *and (ii) if* $\mu(F_j) = F_j$*, then* $\mu\left(F_{j'}\right) = F_{j'}$ *if* $x_{F_{j'}} > x_{F_j}$.

3 Data and descriptive statistics

3.1 Sample

We test our theoretical predictions on the teams of academic researchers and firms that propose research projects to the Engineering and Physical Sciences Research Council (EPSRC). The EPSRC is the main government agency for research funding for the engineering departments of the UK universities. More than half of the overall research funding of the engineering departments comes from the EPSRC. EPSRC grants are allocated to teams of academic researchers alone as well as to teams of academics and firms (teams of firms alone cannot apply for EPSRC funds). As defined by the EPSRC, "collaborative research grants" are those that involve one or more firms as industrial partners.[14] Industrial partners contribute cash or 'in-kind' services to the full economic cost of the project.

Our initial sample includes all the EPSRC project proposals with the starting year 2005, 2006 or 2007. For each project, we know the holding organization, the principal investigator (PI), the coinvestigators (if any) and the industry partners (if any). We take the projects with at least one academic researcher (not necessarily the PI) in the longitudinal data set in Banal-Estañol et al. (2010), which contains calendar information and publication data on all the academics employed at the engineering departments of the 40 major UK universities until 2007. Our final sample consists of 5,855 projects (1,912 in 2005, 1,835 in 2006 and 2,108 in 2007). As a whole, we have 2,411 unique academic researchers and 1,735 firms, which are involved in 2,057 out of the 5,855 projects. That is, 35% of the projects in our database are "collaborative" projects.

[14] Some partners in the EPSRC database are not private firms, but are university research centers and schools, large research infrastructures (e.g., the LNCC National Laboratory of Scientific Computing), government and municipal councils, public agencies, public hospitals, charities, and trade associations (e.g., the International Union of Railways). We disregard these partners as we analyze collaboration with private organizations.

The average number of researchers in each project is 2.86, and the average number of firms in the collaborative projects is 2.43.

3.2 Main variables

To build proxies for scientific ability and ideal type of each partner, we use publication information prior to the start of the project. In particular, we use the publications for the five years prior to, as well as for the starting year of, the project (because they relate to research developed and finished before the start of the project).[15] For example, if the initial year of a project is 2007, then we use the publications of the academics and the firms during the period 2002-2007.[16]

Publication data for the academics is extracted from the Thomson (formerly ISI) Web of Knowledge (WoK) database (for details, see Banal-Estañol et al., 2010). In total, our $2,411$ academic researchers published $44,399$ articles in the years 2000-2007. We follow a similar procedure for the industrial partners and we identify $201,296$ publications for the period 2000-2007 for the $1,735$ firms involved in the collaborative projects.

As a measure of scientific ability, we use both the normal count and the impact-factor-weighted sum of publications. The weights in the second measure are the Science Citation Index (SCI) Journal Impact Factors (JIF), attributed to the publishing journal in the year of publication. This measure takes into account not only the quantity but also the quality of the publications. Therefore, we use it as our main measure but report the results for the normal count as well. We refer from now on to these measures as "impact" and "count".

As a measure of preference for the type of research, we use the Patent Board classification (Narin et al., 1976), updated by Kimberley Hamilton for the National Science Foundation in 2005. Based on the cross-citation matrices, it classifies journals into four categories: (1) applied technology, (2) engineering and technological science, (3) applied and targeted basic research, and (4) basic scientific research. The first two categories are considered to be "technology" and the last two "science" (see Godin, 1996 and van Looy et al., 2006). We follow this distinction and define the "type" of a set of articles as the number of publications in the first two categories divided by the number of publications in all the four categories.

As main variables in our regressions, we use the impact and the count of publications of the

[15] We present our results using the variables that include information for six years, which characterize academic researchers and firms. However, we have also replicated all the analysis using information for two or four years and the results are very similar.

[16] Some of the academic researchers are not in our sample for the whole period, for example, because they are junior, or come from abroad. We take this into account by computing the average count and impact-factor-weighted-sum per year, using the years available (out of the six years prior to the start of the project).

whole team of academic researchers in the project, but we also report the results for the impact and the count of the PI. Similarly, we use the impact and the count of publications per year of the whole team of firms participating in the project (there is no equivalent to the PI for the firms). Similarly, we use the average type of the team of academic researchers, the type of the PI, as well as the average type of the team of firms in the project. In our empirical analysis, we refer, for simplicity, to a team of academics as an "academic" and to a team of firms as a "firm". Therefore, when we talk, for instance, about the ability of the academic in the project, we mean the ability of the team of academics in the project.

3.3 Institutional variables

We obtain information on the strength of the research developed by all the UK's engineering departments from the 2008 "Research Assessment Exercise (RAE)" results. The RAE provides aggregate information on the number of active academics in each department and their publications for the period 2001-2007. We construct variables that measure the number of top quality publications of all the engineering departments at each university, as well as the research funds they obtained (in millions of pounds), decomposed by different categories of funding (public, private and other funding). Also, using data from the Higher Education Statistics Agency, we obtain information related to the general characteristics of the universities: number of undergraduate and graduate students and the university's income and expenses (in millions of pounds). We assign to each project the information of its holding institution.

We also construct variables related to the firms' financial and employment information for the period 2005-2007, using the FAME and ORBIS databases. In particular, we compute the average (per firm in the team and per year) number of employees, turnover, tangible assets and profits before tax (in millions of pounds). We also assign to each project the one-digit Standard Industrial Classification (SIC) code of the firm(s). If the activity of the firms in the project spans several divisions, we randomly assign one of them.[17]

For the collaborative projects, we also assign a variable which measures the geographical distance among the partners. We retrieved the postal codes of the universities and the headquarters of the firms and, using an application of the UK government,[18] we computed the distance in miles between the postcode of the holding university and the UK-based firms in the project. For firms based outside the UK, we used a distance of $1,000$ miles. When there are several firms,

[17] We classify firms' activity according to the 10 US SIC division structure. For those firms for which we only have a UK SIC code, we make a simple translation from the two-digit UK SIC codes to US divisions. We assign the "artificial division" 11 to those projects for which we cannot associate a division to any of the firms.

[18] http://www.education.gov.uk/cgi-bin/inyourarea/distance.pl?

we use the minimum distance between the university and the firms.[19]

3.4 Descriptive statistics

As shown in Table 1, the academic teams in our database published 10.6 articles on average per year, with an impact of 16.9, during the six-year period prior to the initial date of the project. Not surprisingly, the count and the impact of the academics in the projects are highly correlated (0.889). The average PI published 3.6 articles per year, with an impact of 5.6, which are highly correlated with the count (0.562) and the impact (0.636) of the whole academic team. The average count and impact for the firms is 749 and $1,448$ respectively. The variance of the publications of the firms is much larger than the variance of the publications of the academics. The impact and count of the academics are positively correlated to the impact and count of the firms in the same project, 0.219 and 0.194 respectively. These correlations provide initial preliminary evidence for positive assortative matching in terms of ability.

[Insert Table 1 here]

We can define the type for the $5,519$ academic teams, $4,674$ PIs and $1,563$ teams of firms with at least one publication in a journal included in the Patent Board classification. The average type of the firms is more basic (0.579) than the average type of the academics (0.653) and the PI (0.666). This is probably because firms do not allow their employees to publish their most applied discoveries, which may be directly profitable for the firms. The correlation between the types of the academics and the firms (0.358) provides initial preliminary evidence for positive assortative matching in terms of affinity.

We also include the correlations between type and impact: -0.351 for academics, -0.396 for the PIs and, still very significant but lower, for the firms (-0.123). These correlations indicate that more applied researchers, PIs and firms publish less in journals of high impact factor. The (unreported) correlations between type and count are also negative, but the magnitudes are smaller. This suggests that more applied researchers, PIs and firms publish in journals with lower impact factor.

[19] This reflects the fact that the closest firm is the one that establishes the link with the university. We have also run all the regressions using the average distance of the firms, considering only the firms in the UK, and we obtain similar results.

3.5 Counterfactual pairs

We construct a set of "counterfactual" collaborations, i.e., collaborations that were possible but were not formed (see Agrawal et al., 2008, and Gompers et al., 2012, for a similar procedure). The set of counterfactual collaborations, when contrasted with the set of actual collaborations, enables us to assess the significance of various pair-wise characteristics in determining the likelihood of forming a partnership.

The set of counterfactuals is constructed as follows. We take the teams of academics and the teams of firms that have a collaborative project. A pair formed by any of these teams of academics and any of these teams of firms is a potential counterfactual if they do not form an actual collaboration but have a collaborative project in the same year and in the same sector with other partners. For each actual collaborative project, we select four of these potential counterfactual pairs in the following way. We randomly choose one counterfactual in which the team of researchers coincides with the one in the actual project; then one counterfactual in which the team of firms corresponds to the one in the actual project; then another one for the academics and another one for the firms. We alternate the choice to have a more balanced set of counterfactuals. We avoid repetitions so that a counterfactual pair can only appear once. We add the resulting $8,195$ counterfactual pairs to the $2,057$ actual pairs in the matching regressions.[20]

4 Empirical results

Our theoretical analysis provides predictions on the characteristics of the academics and the firms that make them more likely to collaborate with each other as well as on the characteristics that lead them to stay independent. In this section, we test these predictions.

4.1 Partnerships formed

We test our predictions on positive and negative assortative matching using the actual formed partnerships, as well as the non-formed counterfactual partnerships we have constructed (we do not use the non-collaborative projects here). For a given set of pair-wise characteristics of a collaboration, we estimate the likelihood that this collaboration is an actual rather than a counterfactual collaboration. To that purpose, we run probit regressions on the likelihood to form a partnership, using a dependent variable which has a value of 1 if the partnership is an actual pair and a value of 0 if it is a counterfactual pair.

[20] For a limited number of collaborative pairs, it is not possible to find four counterfactual pairs. This is because there are too few collaborations in that year and in that sector.

By construction of the counterfactual pairs, the individual characteristics have no impact on the likelihood of forming a partnership, as each academic and firm in an actual pair also appear in four counterfactual pairs. In all the regressions, we include year and sector fixed effects, and report robust standard errors clustered at the academic researcher level. Given that the previous literature has highlighted the role of geographical proximity in the decision to collaborate or not, we also include in all the regressions the variable that measures the geographical distance between the academics and the firms in the projects. This variable, however, turns out to be insignificant in all the regressions.

Continuous measures Our first approach to test for the presence of positive (or negative) assortative matching in one characteristic, say ability, is to regress the likelihood of being an actual match over the product of the abilities of the two partners. The coefficient associated to this joint variable represents the cross-partial derivative of the probability of being matched over the ability of the academic and that of the firm. On average, the matching is positive or negative assortative if the associated coefficient is positive or negative, respectively. To illustrate this property informally, consider a perfectly positive assortative matching in a market with three academics and three firms. That is, the academic with low (resp. medium and high) ability collaborates with the firm with low (resp. medium and high) ability. Ordering the partners' ability from low to high and setting the academics in rows and the firms in columns, the matching can be represented by a 3×3 matrix with ones in the main diagonal and zeros elsewhere. The estimate of the partial derivative with respect to the firms' abilities would be negative if evaluated for the low-ability academic, as the elements of the row are $(1, 0, 0)$. Similarly, the estimate of the partial derivative would be zero (resp. positive) if evaluated for the medium- (resp. high-) ability academic, as the elements of the row would be $(0, 1, 0)$ (resp. $(0, 0, 1)$). Therefore, the estimate of the cross-partial derivative would be positive, because the partial derivative goes from negative to zero to positive.[21] A similar argument implies that the coefficient of the cross-partial derivative is negative if the matching is negative assortative.

Table 2 shows the probit regressions over several continuous joint variables that measure (i) the ability of the academics and the firms, (ii) the type of the academics and the firms and (iii) the ability of one partner and the distance in terms of type between the two partners.

[Insert Table 2 here]

[21] Note that the partial derivative of the probability over the firm's ability would be zero on average (when averaged over all the academics' abilities). This is consistent with the property of individual characteristics having no impact on the likelihood of forming a partnership.

Columns 1 and 2 provide estimates of the cross-partial derivatives for the two measures of scientific ability, the impact and count of publications, as well as for the type. All the coefficients are positive and significant, thus providing support for the prediction of positive assortative matching in terms of ability and type (propositions 3 and 4). Therefore, the partnerships that have higher ability and more applied academics together with higher ability and more applied firms (and lower ability and less applied academics together with lower ability and less applied firms) are more likely to be an actual formed partnership, as opposed to a counterfactual non-formed partnership. As shown in column 3, we obtain similar results if we use the measures of the principal investigator instead of the ones for the entire team of academics in the project.

Columns 4 and 5 present the results of a more indirect test of assortativeness in types, which consists of estimating the effect of the distance between types. According to our theoretical model, the matching is positive assortative because it minimizes the sum of the (square of) distances between partners. This suggests that the distance between the types of the academics and the firms should lower the probability of being matched. The results in column 4 confirm this intuition: there is a strong negative effect of the distance in types on the probability of matching. Column 5 considers an adjusted measure of the distance between types. As mentioned earlier, the distribution of firms' types in our database is more basic than the distribution of the academics' types. This might be due to the difficulty for the firms' researchers to publish their most applied discoveries. This suggests that the "true" type of a firm is more applied than the one observed in our data. We therefore define a new distance variable obtained after increasing the type of all firms by a fraction of 0.1.[22] As column 5 shows, this variable has an even stronger effect than the original distance measure.

Column 6 takes into account the joint variables on the ability of one partner and the distance in terms of type between the two partners. The coefficients of the two measures are negative and significant, thus providing support for the prediction of negative assortative matching in Proposition 5. That is, a better academic (or firm) is less likely to be matched with a very different firm (academic) in terms of type probably because, as suggested by our model, better academics (firms) suffer relatively more than worse ones from collaborating with distant firms (academics). As in column 4, the coefficient of the joint variable for impact is positive and significant and that of the type distance is negative and significant. Similar (unreported) results are obtained with the adjusted measure of distance. Column 7 presents the results of the cross-effects between the impact of one partner and the type of the other. The coefficients for both variables are negative and significant, suggesting that a better academic (or firm) is less likely

[22] We have tried adding different amounts. Increases of 0.1 give the best estimates.

to be matched with a more applied firm (academic).

Discrete measures and marginal effects Table 3's columns 1 to 4 present the marginal effects of several probit regressions over dummy and other discrete variables accounting for the relative position of each agent on each side of the market. For example, the dummy variable "Both above median in impact" takes the value of 1 if both the academic and the firm are above their respective median in terms of impact.

[Insert Table 3 here]

According to the results in column 1, the pairs of academics and firms that are both (neither) above the median of their respective distribution of abilities are 1.4% (1.2%) more likely to be matched, compared to those pairs in which one is above the median and the other is not. Given that the unconditional probability of being matched is 20%, this represents 7% and 6% of the unconditional probability. Albeit positive, these effects are not significant. The effects for the types are instead very significant. The academics and the firms are 6.6% and 7.9% more likely to be matched if they are both above the median and both below the median in terms of type, respectively (33% and 39.5% in terms of unconditional probability). In column 2, we show that the impact matters when both the academic and the firm are in the first quartile (or none of them is in the first quartile) in terms of ability. When both are (respectively, none is) in the top 25% in terms of impact, the conditional probability that they match is 5.9% higher (2% higher), which in terms of the unconditional probability means 29.5% higher (10% higher). This suggests that, for the likelihood of a matching, both being or not at the top in terms of ability is more important than both being above or below the median.

To further understand the effect of the relative position of academics and firms, we divide the set of academics and the set of firms in quartiles with respect to each characteristic (impact and type), assign to all of them their quartile (from 1 to 4), and compute the pair's quartile difference, defined as the absolute value of the difference between the quartile of the academic and the quartile of the firm. Column 3 shows that for each unit change of quartile in impact, the probability of being matched decreases by a significant 1.4% (or an unconditional probability of 7%). Similarly, a unit change of quartile in type decreases the probability of being matched by 4% (or an unconditional probability of 20%).

Column 4 reports the results of a similar exercise but with a complete ranking of academics and firms in terms of ability and type. We construct two variables that measure the differences in ranking, with respect to impact and type, between the academics and the firms in the project.[23] To be able to appreciate the magnitude of the effects, we divide the ranking by one thousand. As expected, a higher distance in the ranking in impact and type leads to a lower likelihood of matching. In terms of magnitudes, an additional one-thousand units in distances of rankings of impact lowers the conditional probability of being matched by 2.2% (11% of the unconditional probability), whereas the effect in types is 10.1% (50.5% of the unconditional probability).

In sum, columns 1-4 confirm that the matching is positive assortative in both characteristics. They also suggest that the positive nature of the matching is stronger for the type, which is a horizontal characteristic, than for the impact, which is a vertical attribute. The coefficients for the joint variables in types are three to four times higher than the coefficients for the corresponding variables for impact. The numbers are meaningful because the variables reflect relative positions of the agents, which allow for the comparison of the two characteristics.

Columns 5-7 show the results of the cross-partial derivatives of the full rank of academics and firms in the two characteristics, ability and type. Column 5 shows that the joint variable of ranking in impact, as well as the one in type, increases the probability of being matched. Column 6 tests Propositions 3, 4 and 5 simultaneously. Indeed, this regression also includes the joint variables on the ability of one partner and the distance in terms of type between the two partners and the coefficients of these two variables are negative and significant. The coefficient of the joint variable of rank in types is now not significant (even if it has the right sign) possibly because the effect is taken away by the distance in rank of types included in the new variables.

As in Table 2, column 7 shows a regression combining the rank in type of one side of the market and the rank in impact of the other side. The coefficient for both variables is negative and significant, which provides support for the idea that a better academic (or firm) is less likely to be matched with a more applied firm (academic).

Individual and institutional characteristics Table 4 tests if there is positive or negative assortative matching using institutional characteristics of the universities and the firms. We are particularly interested in determining if firms select academic researchers because of their individual characteristics or because of the characteristics of the university they work for. Columns 1-6 show the regression results of several joint variables that measure the ability of the firm and the general characteristics of the university. The coefficients associated to most of these variables, including the one that uses the university's number of top quality papers shown in column

[23] The ordering treats equal numbers as average ranking. That is, if the impacts were 1, 7, 7, 20, then the associated ranks would be 1, 2.5, 2.5, 4.

1 and the university's income shown in column 6, turn out to be insignificant.[24] However, if we use the firm's impact together with the university's research or private funds, the coefficients are positive and significant, as shown in columns 2 and 4. This suggests a positive assortative matching between firms' ability and universities' funds. But, as shown in columns 3 and 5, these coefficients lose their significance if we include the joint variable that measures academics' and firms' ability. Individual characteristics turn out to be more important than institutional ones.

[Insert Table 4 here]

Similarly, column 7 shows that a joint variable that measures the academic's ability and a performance measure of the firm, its profits, is not significant. Finally, we run regressions on all the possible combinations of the institutional characteristics of the universities and the firms; they are all non-significant. Columns 8 and 9 report two such regressions.[25]

4.2 Collaborating versus staying independent

Our theoretical analysis also generates predictions on the characteristics of the academics that make them more likely to collaborate rather than stay independent. Under reasonable conditions, Propositions 6 and 7 show that, in equilibrium, the most able and the most applied academics should develop collaborative projects whereas the least able and the most basic should develop non-collaborative projects.

We test these hypotheses using the data on the academics that proposed collaborative projects as well as those that submitted non-collaborative projects (we do not use the counterfactual observations here). Unfortunately, we cannot test which firms would be more likely to conduct non-collaborative projects because in those cases they cannot apply for EPSRC funding and they are, therefore, not in our dataset. We run probit regressions on the academics' likelihood to collaborate, using a dependent variable which has a value of 1 if the academics choose to submit a collaborative project and a value of 0 if they submit a non-collaborative one. We control for year and university fixed effects, and report robust standard errors.

Table 5 shows the probit regressions over several measures of academics' ability and type. Columns 1 and 2 show that the most able as well as the most applied researchers are significantly more likely to collaborate, both if we measure ability in terms of impact and count. Columns

[24] We have run regressions using other university characteristics, such as number of active academic engineers, number of undergraduate and graduate students, and university expenses. All of the associated coefficients are insignificant. Similar results are obtained if we use the count instead of the impact as a measure of firms' ability.

[25] In addition to firms' employees and profits, we have used firms' turnover and assets and, in addition to universities' research and private funds, we have used all the other university characteristics mentioned earlier.

3 and 4 show that the results are similar if we use variables that refer to the PI instead of the team of academic researchers, although the PI's impact appears not to be significant. Column 5 shows that the academics who are above the median in terms of ability and those who are more applied than the median are 3.2% and 13.8% more likely to collaborate, respectively, than those below the median. Given that the unconditional probability of collaborating is 35%, the increases in probability are 9.1% and 39.5% in terms of the unconditional probability. In column 6, we consider the rank of the academics in terms of impact and type. The regression shows that moving up the rank in any of the two characteristics has a positive and significant effect on the probability of collaboration. With respect to the institutional measures, column 7 shows that researchers in larger universities, in terms of number of projects, are also more likely to collaborate. However, all the other university variables, including the number of top quality papers and the number of active members of staff, are not significant.

[Insert Table 5 here]

In sum, the empirical results provide strong support to the theoretical predictions of propositions 6 and 7: the likelihood of collaborating as opposed to staying independent increases with the ability and type of the researchers in the project. They also suggest that, as is the case in the matching regressions, the type of the academics has a stronger effect than their ability.

5 Conclusion

This paper analyzes university-industry collaboration as an endogenous matching problem. We first develop a two-sided market matching model of academic researchers and firms that are heterogeneous in terms of ability- and affinity-based characteristics. Our model predicts that the most able and the most applied academic researchers, and the most able and the most basic firms, prefer to develop collaborative projects, rather than stay independent and develop stand-alone projects. Among those that collaborate, we predict a positive assortative matching, both in terms of ability and affinity, i.e., more prolific academic researchers collaborate with more research-productive firms, and academics with more applied bias collaborate with firms with more applied bias. We also predict a negative assortative matching across the two characteristics: the academics with a higher ability collaborate with firms with which they have more affinity.

We verify our theoretical predictions on the teams of academic researchers and firms that propose research projects to the EPSRC. In addition, our empirical analysis shows that the affinity-based characteristics are relatively more important than those related to ability. The coefficients for the joint variables in types are three to four times higher than the coefficients for the corresponding variables for impact. Moreover, the characteristics at the individual-researcher level are more relevant than at the institutional level. Some of the institutional variables have a significant effect but lose their significance when individual variables are also included in the regressions.

We build our theoretical model to address university-industry collaborations in research projects. Our model, however, can also be used to analyze collaboration decisions in other two-sided markets, such as consultancy and contract research between universities and firms, suppliers and firms, and entrepreneurs and venture capitalists. As in our model, agents in these markets are heterogeneous along horizontal and vertical dimensions. Participants can also collaborate and, in some cases, stay independent. And, if they collaborate, partners need to make investment decisions and reach a compromise between their possibly conflicting interests.

In our setting, collaboration brings benefits in terms of complementarities but it also has costs in terms of project selection. Developing a project different from the ideal one, or in an area where one is less productive, is costly. This is consistent with our results in Banal-Estañol et al. (2013) showing that the distance in types of the academic researchers and the firms decreases the number of scientific publications of the project. In the case of other horizontal characteristics, however, heterogeneity can bring other benefits, as has been argued for the case of interdisciplinary research in academia (Derrick et al., 2011). Our model could also be adapted to take this feature into account.

This paper provides one of the rare efforts to understand collaboration across institutional markets, i.e., the two-sided market partnerships. Fortunately, we know a significant deal more about collaboration across institutions within institutional markets, i.e., the one-sided market partnerships. Many papers have studied the causes and the consequences of collaboration among academic researchers (Katz and Martin, 1997; Wagner and Leydesdorff, 2005), firms (Caloghirou et al., 2003), venture capitalists (Gompers et al., 2012), etc. A natural next step should be to study the interaction between the two, and identify if collaboration between academic researchers substitutes or complements collaboration between academic researchers and firms.

References

[1] Agarwal, R. and Ohyama, A. (forthcoming): "Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists". *Management Science.*

[2] Agrawal, A. and Henderson, R. (2002): "Putting Patents in Context: Exploring Knowledge Transfer from MIT", *Management Science* 48, 44-60.

[3] Agrawal, A., Kapur, D. and McHale, J. (2008): "How do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data", *Journal of Urban Economics* 64, 258-269.

[4] Alonso-Paulí, E. and Pérez-Castrillo, D. (2012): "Codes of Best Practice in Competitive Markets for Managers", *Economic Theory* 49 (1), 113-141.

[5] Audretsch, D.B. and Stephan, P. (1996): "Company-Scientist Locational Links: The case of Biotechnology", *The American Economic Review* 86 (3), 641-652.

[6] Azoulay, P., Ding, W. and Stuart, T. (2009): "The Impact of Academic Patenting on The Rate, Quality and Direction of (Public) Research Output", *The Journal of Industrial Economics* 57 (4), 637-676.

[7] Banal-Estañol, A., Jofre-Bonet, M. and Meissner, C. (2010): "The Impact of Industry Collaboration on Research: Evidence from Engineering Academics in the UK", W. P. City University London.

[8] Banal-Estañol, A., Macho-Stadler, I. and Pérez-Castrillo, D. (2013): "Research Output from University-Industry Collaborative Projects", *Economic Development Quarterly* 27 (1), 71-81.

[9] Becker, G.S. (1973): "A Theory of Marriage: Part I", *The Journal of Political Economy* 81, 813-846.

[10] Bercovitz, J. and Feldman, M. (2007): "Fishing Upstream: Firm Strategic Research Alliances with Universities", *Research Policy* 36 (7), 930-948.

[11] Caloghirou, Y., Ioannides, S. and Vonortas, N.S. (2003): "Research Joint Ventures", *Journal of Economic Surveys* 17, 541-570.

[12] Carayol, N. (2003): "Objectives, Agreements and Matching in Science-Industry Collaborations: Reassembling the Pieces of the Puzzle", *Research Policy* 32, 887-908.

[13] Cassiman, B. and Veugelers, R. (2006): "In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition", *Management Science* 52 (1), 68 - 82.

[14] Cohen, W.M., Nelson R.R. and Walsh J.P. (2002): "Links and Impacts: The Influence of Public Research on Industrial R&D", *Management Science* 48, 1-23.

[15] Cockburn, I. and Henderson, R. (1998): "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery", *Journal of Industrial Economics* 46 (2), 157-182.

[16] Dam, K. and Pérez-Castrillo, D. (2006): "The Principal Agent Market", *Frontiers in Economic Theory, Berkeley Electronics* 1 (Republished in BePress *Advances in Theoretical Economics*).

[17] Dasgupta, P. and David, P. (1994): "Towards a New Economics of Science", *Research Policy* 23 (5): 487-522.

[18] Derrick, E., Falk-Krzesinski, H., Roberts, M. (2011): "Facilitating Interdisciplinary Research and Education: A Practical Guide", report from the "Science on FIRE: Facilitating Interdisciplinary Research and Education" workshop of the American Association for the Advancement of Science.

[19] Fabrizio, K. and DiMinin, A. (2008): "Commercializing the Laboratory: Faculty Patenting and the Open Science Environment", *Research Policy* 37, 914-931.

[20] Florida, R. and Cohen, W.M. (1999): "Engine or Infrastructure? The University Role in Economic Development", in Branscomb, L.M., Kodama, F., Florida, R. (eds.), *Industrializing Knowledge: University-Industry Linkages in Japan and the United States*, MIT Press, London, 589-610.

[21] Fox, J. (2008), "Estimating Matching Games with Transfers," NBER W. P. No. W14382.

[22] Gale, D. and Shapley, Ll.S. (1962): "College Admissions and the Stability of Marriage", *American Mathematical Monthly* 69, 9-15.

[23] Godin, B. (1996): "The State of Science and Technology Indicators in the OECD Countries", Research Paper, Statistics Canada.

[24] Gompers, P., Mukharlyamov, V. and Xuan, Y. (2012): "The Cost of Friendship", NBER W. P.

[25] Jensen, R.A., Thursby, M.C. (2001): "Proofs and Prototypes for Sale: The Licensing of University Inventions" *The American Economic Review* 91, 240–259.

[26] Jensen, R., Thursby, J. Thursby, M. (2003): "Disclosure and Licensing of University Inventions: The Best We Can Do with the S**t We Get to Work with," *International Journal of Industrial Organization* 21 (9), 1271–1300.

[27] Kamien, M., Muller, E. and Zang, I. (1992): "Research Joint Ventures and R&D Cartels", *The American Economic Review* 82 (5), 1293-1306.

[28] Kaneko, M. (1982): "The Central Assignment Game and the Assignment Markets", *Journal of Mathematical Economics* 10, 205-32.

[29] Katz, J.S. and Martin, B.R. (1997): "What is Research Collaboration?", *Research Policy* 26, 1-18.

[30] Lee, Y.S. (2000): "The Sustainability of University-Industry Research Collaboration: An Empirical Assessment", *The Journal of Technology Transfer* 25, 111-133.

[31] Legros, P. and Newman, A. (2002): "Monotone Matching in Perfect and Imperfect Worlds", *Review of Economic Studies* 69, 925-942.

[32] Louis, K.S. , Blumenthal, D., Gluck, M.E. and Stoto, M.A. (1989): "Entrepreneurs in Academe: An Exploration of Behaviors among Life Scientists", *Administrative Science Quarterly* 34 (1), 110-131.

[33] Mansfield, E. (1995): "Academic Research Underlying Industrial Innovation: Sources, Characteristics and Financing", *Review of Economics and Statistics* 77, 55-65.

[34] Mindruta, D. (forthcoming): "Value Creation in University-Firm Research Collaborations: A Matching Approach", *Strategic Management Journal*.

[35] Mohnen, P. and Hoareau. C. (2003): "What Type of Enterprise Forges Close Links with Universities and Government Labs? Evidence from CIS 2", *Managerial and Decision Economics* 24 (2-3), 133-145.

[36] Narin, F., Pinski G. and Gee H. (1976): "Structure of the Biomedical Literature", *Journal of the American Society for Information Science*, 25-45.

[37] Pereira, I. (2007): "Business-Science Research Collaboration under Moral Hazard", W. P. Universitat Autonoma de Barcelona.

[38] Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Kitson, M., Krabel, S., Llerena, P., Lissoni, F., Salter,

A. and Sobrero, M. (2013): “Academic Engagement and Commercialization: A Review of the Literature on University-Industry Relations”, *Research Policy* 42 (2), 423-442.

[39] Roth, A. and Sotomayor, M. (1990): “*Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*”, Cambridge University Press, New York and Melbourne, Econometric Society Monographs.

[40] Shapley, Ll.S. and Shubik, M. (1972): “The Assignment Game I: The Core”, *International Journal of Game Theory* 1, 111-130.

[41] Sorensen, M. (2007): “How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital”, *The Journal of Finance* LXII (6), 2725-2762.

[42] van Looy, B., Callaert, J. and Debackere, K. (2006): “Publication and Patent Behaviour of Academic Researchers: Conflicting, Reinforcing or Merely Co-existing?”, *Research Policy* 35, 596-608.

[43] Veugelers, R. and Cassiman, B. (2005): “Cooperation between Firms and Universities. Some Empirical Evidence from Belgian Manufacturing”, *International Journal of Industrial Organization* 23, 355-379.

[44] Wagner, C.S. and Leydesdorff L. (2005): “Network Structure, Self-organization, and the Growth of International Collaboration in Science”, *Research Policy* 34 (10), 1608-1618.

[45] Zucker, L., Darby M.R. and Armstrong J.S. (2002): “Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology”, *Management Science* 48 (1), 138 - 153.

6 Appendix

Proof of Proposition 1. It follows from the arguments in the text. ■

Proof of Proposition 2. The optimal agreement P solves the following program:

$$\underset{x,I_A,I_F}{Max}\left[\pi_A(x,I_A,I_F;x_A,\delta_A)+\pi_F(x,I_A,I_F;x_F,\delta_F)\right]$$

where the expressions for $\pi_A(x,I_A,I_F;x_A,\delta_A)$ and $\pi_F(x,I_A,I_F;x_F,\delta_F)$ are in the main text. This function is concave in all its arguments. The FOC with respect to x is

$$-2\lambda_A v_A\left(\delta_A+\delta_F\right)\left(I_A+I_F\right)t\left(x-x_A\right)+2\lambda_F v_F\left(\delta_A+\delta_F\right)\left(I_A+I_F\right)t\left(x_F-x\right)=0.$$

From this condition we obtain

$$x^c = \frac{\lambda_A v_A x_A + \lambda_F v_F x_F}{\lambda_A v_A + \lambda_F v_F}.$$

The FOCs with respect to the investments are

$$\begin{aligned}
\lambda_A v_A (\delta_A + \delta_F)\left(1 - t(x - x_A)^2\right) + \lambda_F v_F (\delta_A + \delta_F)\left(1 - t(x_F - x)^2\right) - CI_A = 0 \\
\lambda_A v_A (\delta_A + \delta_F)\left(1 - t(x - x_A)^2\right) + \lambda_F v_F (\delta_A + \delta_F)\left(1 - t(x_F - x)^2\right) - CI_F = 0
\end{aligned}$$

from which the expressions for I_A^c and I_F^c in the proposition are easily derived. Note that $I_A^c = I_F^c > 0$: a sufficient condition for this to hold is $\lambda_A v_A + \lambda_F v_F - \frac{\lambda_A v_A \lambda_F v_F}{(\lambda_A v_A + \lambda_F v_F)} > 0$, or $\lambda_A^2 v_A^2 + \lambda_F^2 v_F^2 + \lambda_F \lambda_A v_F v_A > 0$, which is satisfied. ■

Proof of Proposition 3. We have computed the level of profits of the partners at the optimal project Π_{AF}^c after Proposition 2. From that expression, it easily follows that

$$\frac{\partial^2 \Pi_{AF}^c}{\partial \delta_A \partial \delta_F} = \frac{2}{C}\left(V_{AF}\left(|x_F - x_A|\right)\right)^2 > 0,$$

which implies that the matching is positive assortative (see footnote 11). ■

Proof of Proposition 4. We prove that $\frac{\partial^2 \Pi_{AF}^c}{\partial x_A \partial x_F} > 0$.

$$\frac{\partial \Pi_{AF}^c}{\partial x_F} = -\frac{2(\delta_A + \delta_F)^2}{C} \frac{\lambda_A v_A \lambda_F v_F}{(\lambda_A v_A + \lambda_F v_F)} 2t(x_F - x_A) V_{AF}\left(|x_F - x_A|\right).$$

The cross-partial derivative $\frac{\partial^2 \Pi_{AF}^c}{\partial x_F \partial x_A}$ is proportional to

$$\begin{aligned}
&V_{AF}\left(|x_F - x_A|\right) - 2\frac{\lambda_A v_A \lambda_F v_F}{(\lambda_A v_A + \lambda_F v_F)} t(x_F - x_A)^2 \\
&\qquad = \lambda_A v_A + \lambda_F v_F - \frac{3\lambda_A v_A \lambda_F v_F}{(\lambda_A v_A + \lambda_F v_F)} t(x_F - x_A)^2.
\end{aligned}$$

Therefore, $\frac{\partial^2 \Pi_{AF}^c}{\partial x_F \partial x_A} > 0$ if $\lambda_A v_A + \lambda_F v_F - \frac{3\lambda_A v_A \lambda_F v_F}{(\lambda_A v_A + \lambda_F v_F)} > 0$, that is, $\lambda_A^2 v_A^2 + \lambda_F^2 v_F^2 - \lambda_A v_A \lambda_F v_F > 0$, which always holds. ■

Proof of Proposition 5. The proposition holds if $\frac{\partial^2 \Pi_{AF}^c}{\partial \delta_A \partial d} < 0$, where $d = |x_F - x_A|$, that is, if

$$-\frac{8}{C}(\delta_A + \delta_F) td \left(V_{AF}(d)\right)^2 < 0,$$

which is satisfied. ■

Proof of Lemma 1. (i) The derivative of $\Delta(x_A, \delta_A, x_F, \delta_F)$ with respect to δ_A is

$$\frac{\partial \Delta}{\partial \delta_A}(x_A, \delta_A, x_F, \delta_F) = \frac{2(\delta_A + \delta_F)}{C}\left(V_{AF}\left(|x_F - x_A|\right)\right)^2 - \frac{1}{c}\delta_A v_A^2$$

and the lemma follows from this expression. (ii) The proof is similar. ■

Proof of Proposition 6. (i) We do the proof by contradiction. Suppose that $\mu(A_i) = A_i$ and $\mu(A_{i'}) = F_j$ with $\delta_{A_{i'}} < \delta_{A_i}$. We denote $\Pi^c_{AF} \equiv \Pi^c_{AF}(x_A, \delta_A, x_F, \delta_F)$, $\pi^n_A \equiv \pi^n_A(x_A, \delta_A)$, $\pi^n_F \equiv \pi^n_F(x_F, \delta_F)$ and $\Delta_{AF} \equiv \Delta(x_A, \delta_A, x_F, \delta_F)$ for any $A \in \mathcal{A}$ and $F \in \mathcal{F}$. The efficiency of the equilibrium matching implies $\Pi^c_{A_{i'}F_j} + \pi^n_{A_i} \geq \Pi^c_{A_iF_j} + \pi^n_{A_{i'}}$ that, given the definition of $\Delta_{A_iF_j}$, is equivalent to $\Delta_{A_{i'}F_j} \geq \Delta_{A_iF_j}$. However, $\Delta_{A_{i'}F_j} \geq \Delta_{A_iF_j}$ is not possible under hypothesis $1a$ because it implies that equation (4) holds, hence the net benefits from collaboration Δ are increasing in δ_R. (ii) The proof is similar. ■

Proof of Lemma 2. (i) The derivative of $\Delta(x_A, \delta_A, x_F, \delta_F)$ with respect to x_A coincides with the derivative of $\Pi^c_{AF}(x_A, \delta_A, x_F, \delta_F)$ with respect to x_A, which is

$$\frac{4\,(\delta_A + \delta_F)^2}{C} t\,(x_F - x_A)\,\frac{\lambda_A v_A \lambda_F v_F}{\lambda_A v_A + \lambda_F v_F} V_{AF}\left(|x_F - x_A|\right).$$

Therefore, $\frac{\partial \Delta}{\partial x_F}(x_A, \delta_A, x_F, \delta_F) > 0$ if and only if $x_F - x_A > 0$. (ii) The proof is similar. ■

Proof of Proposition 7. (i) Suppose, by contradiction, that $\mu(A_i) = A_i$, $x_{A_{i'}} < x_{A_i}$ and $\mu(A_{i'}) = F_j$ for some $F_j \in \mathcal{F}$. We use the same notations as in the proof of Proposition 6: Π^c_{AF}, Δ_{AF} and so on.

Under Hypothesis $1a$, $x_{F_j} \geq x_{A_i} > x_{A_{i'}}$, which implies (see Lemma 2) $\Delta_{A_iF_j} > \Delta_{A_{i'}F_j}$. Therefore, $\Pi^c_{A_iF_j} + \pi^n_{A_{i'}} \geq \Pi^c_{A_{i'}F_j} + \pi^n_{A_i}$, which contradicts the fact that the equilibrium matching μ must be efficient.

Under Hypothesis $1b$, we first show that $\mu(A_i) = A_i$ implies that $i < k$. Otherwise, consider the firm F_{i-k+1}, for which $x_{A_i} = x_{F_{i-k+1}}$. If $\mu(F_{i-k+1}) = A_{i''}$ for some $A_{i''}$, then for the same arguments as before, it would be more efficient that A_i is matched to F_{i-k+1} and $A_{i''}$ remains unmatched than the situation under μ. Similarly, if $\mu(F_{i-k+1}) = F_{i-k+1}$, then it would also be more efficient than $\mu(F_{i-k+1}) = A_i$ (note that the net benefits from this collaboration must be positive because they are the same as the benefits from the collaboration between $A_{i'}$ and F_j). The efficiency of μ implies that the two previous situations are not possible.

Finally, if $i < k$ then we have $x_{F_j} \geq x_{A_i} > x_{A_{i'}}$ which, by the same reasons as above, would contradict the efficiency of μ.

(ii) The proof is similar to the proof of (i). ■

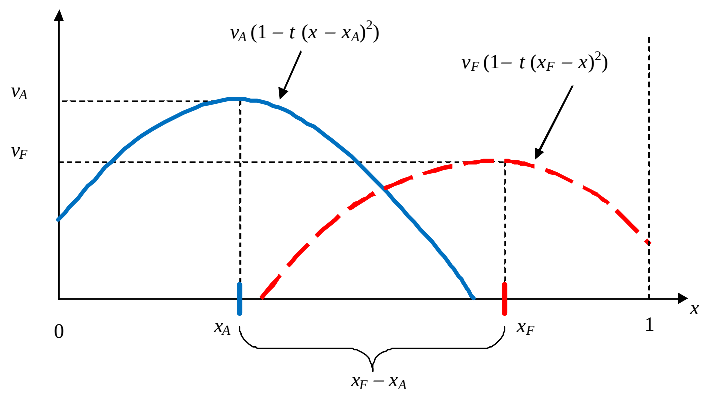

Figure 1: Value of a positive result for A and F as a function of the type of project x

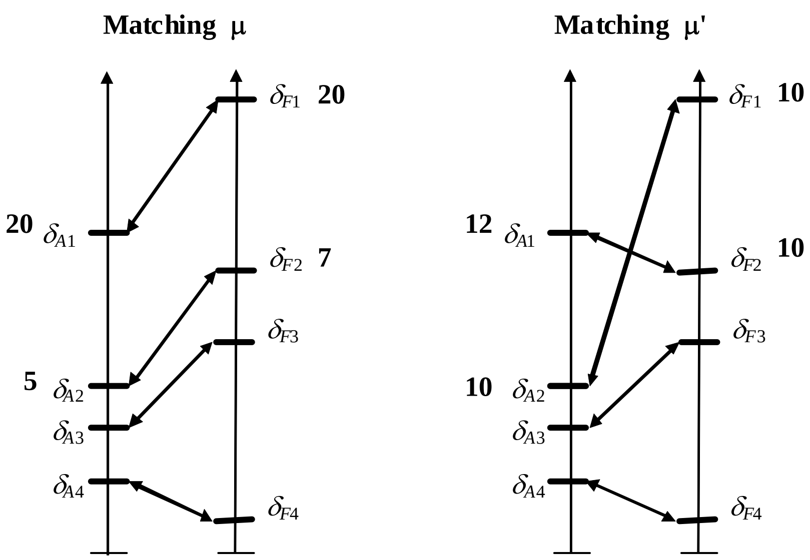

Figure 2: Matching of an equilibrium outcome in terms of ability

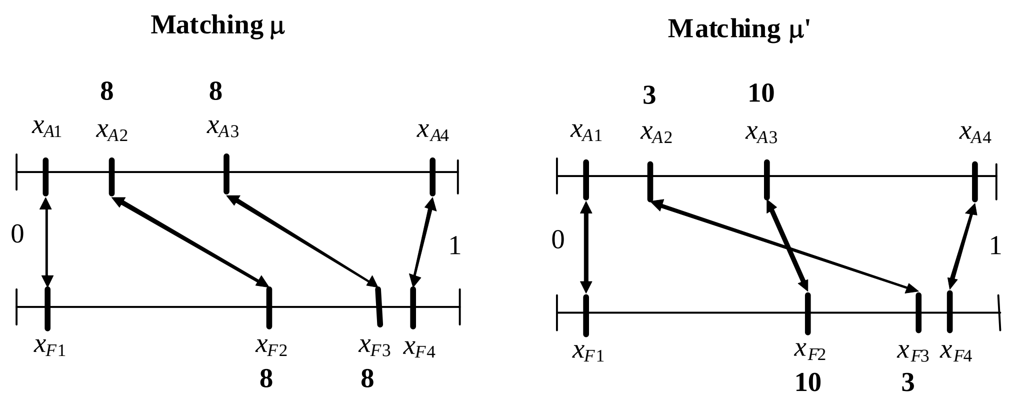

Figure 3: Matching of an equilibrium outcome in terms of type

	Observations	Mean	St dev	Corr. academics' impact	Corr. PI's count	Corr. PI's impact	Corr. firms' count	Corr. firms' impact
Academics' count (x100)	5855	0,106	0,146	0.889***	0.562***	0.502***	0.194***	0.172***
Academics' impact (x100)	5855	0,169	0,309		0.534***	0.636***	0.223***	0.219***
PI's count (x100)	5067	0,036	0,037			0.854***	0.067***	0.058***
PI's impact (x100)	5067	0,056	0,081				0.120***	0.137***
Firms' count (x1000)	2057	0,749	1,836					0.901***
Firms' impact (x1000)	2057	1,448	5,173					

	Observations	Mean	St dev	Corr. PI's type	Corr. firms' type	Corr. respect impact
Academics' type	5519	0,653	0,328	0.938***	0.358***	-0.351***
PI's type	4674	0,666	0,343		0.343***	-0.396***
Firms' type	1563	0,579	0,284			-0.123***

Table 1. Descriptive statistics

Probit linked	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Academics' impact*Firms' impact	0.027***			0.024***	0.024***	0.044***	0.046***
	[0.006]			[0.005]	[0.005]	[0.008]	[0.009]
Academics' type*Firms' type	0.163***	0.161***					0.160***
	[0.052]	[0.052]					[0.052]
Academics' count*Firms' count		0.136***					
		[0.026]					
PI' impact*Firms' impact			0.074***				
			[0.027]				
PI' type*Firms' type			0.147***				
			[0.053]				
Type distance				-0.600***		-0.463***	
				[0.068]		[0.071]	
Academics' impact*Type distance						-0.782***	
						[0.145]	
Firms' impact*Type distance						-0.033***	
						[0.009]	
Type distance adjusted					-0.654***		
					[0.072]		
Academics' impact*Firms' type							-0.420***
							[0.084]
Firms' impact*Academics' type							-0.021***
							[0.005]
Geographical distance	-0.000	-0.000	-0.000	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	-0.127***	-0.125***	-0.117***	0.174***	0.262***	0.136***	-0.110**
	[0.044]	[0.044]	[0.045]	[0.033]	[0.040]	[0.034]	[0.045]
Observations	7,323	7,323	6,426	7,323	7,323	7,323	7,323

Robust standard errors clustered at the academic researcher level in brackets. Year and sector fixed effects included in all regressions

Table 2: Probability of matching as a function of the joint characteristics of academics and firms.

Probit linked	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Both above median in impact	0.014						
	[0.009]						
Both below median in impact	0.012						
	[0.013]						
Both above median in type	0.066***	0.063***					
	[0.011]	[0.011]					
Both below median in type	0.079***	0.075***					
	[0.011]	[0.011]					
Both in top 25% in impact		0.059***					
		[0.016]					
Both in bottom 75% in impact		0.020**					
		[0.010]					
Distance of quartiles in impact			-0.014***				
			[0.005]				
Distance of quartiles in types			-0.040***				
			[0.005]				
Distance in rank of impact				-0.022**			
				[0.011]			
Distance in rank of types				-0.101***			
				[0.011]			
Rank academics' impact*Rank firms' impact					0.064***	0.075***	0.152***
					[0.016]	[0.018]	[0.021]
Rank academics' type*Rank firms' type					0.087***	0.017	0.257***
					[0.023]	[0.023]	[0.034]
Rank firms' impact*Distance in rank of types						-0.152***	
						[0.032]	
Rank academics' impact*Distance in rank of types						-0.220***	
						[0.040]	
Rank academics' type*Rank firms' impact							-0.158***
							[0.023]
Rank academics' type*Rank firms' type							-0.221***
							[0.032]
Geographical distance	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant					-0.140***	0.165***	-0.010
					[0.037]	[0.047]	[0.044]
Observations	7,323	7,323	7,323	7,323	7,323	7,323	7,323

Robust standard errors clustered at the academic researcher level in brackets. Year and sector fixed effects included in all regressions

Table 3: Probability of matching as a function of the joint characteristics of researchers and firms based on discrete variables. Columns 1 to 4 display the marginal effects.

Probit linked	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Unis' top papers*Firms' impact	0.000								
	[0.000]								
Unis' res funds*Firms' impact		0.001**	-0.000						
		[0.000]	[0.000]						
Unis' priv funds*Firms' impact				0.002**	-0.001				
				[0.001]	[0.001]				
Academics' impact*Firms' impact			0.029***		0.030***				
			[0.007]		[0.007]				
Unis' income*Firms' impact						0.006			
						[0.005]			
Academics' impact*Firms' profits							0.009		
							[0.012]		
Unis' res funds*Firms' employees								0.000	
								[0.000]	
Unis' priv funds*Firms' profits									0.002
									[0.002]
Geographical distance	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	0.005	0.006	0.008	0.006	0.008	0.008	-0.908***	-0.883***	-0.868***
	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]	[0.175]	[0.185]	[0.172]
Observations	10,178	10,178	10,108	10,178	10,108	9,187	6,464	7,719	6,503

Robust standard errors clustered at the academic researcher level in brackets. Year and sector fixed effects included in all regressions

Table 4: Probability of matching as a function of joint measures of universities/researchers and firms.

Probit collaborate	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Academics' impact	0.130**						0.263***
	[0.061]						[0.074]
Academics' type	0.635***	0.640***					0.698***
	[0.061]	[0.059]					[0.065]
Academics' count		0.515***					
		[0.123]					
PI's impact			0.092				
			[0.269]				
PI's type			0.555***	0.605***			
			[0.066]	[0.062]			
PI's count				2.278***			
				[0.554]			
Academics above median in impact					0.032**		
					[0.014]		
Academics above median in type					0.138***		
					[0.014]		
Rank academics' impact						0.100***	
						[0.024]	
Rank academics' type						0.250***	
						[0.024]	
Unis' number of projects							0.051***
							[0.013]
Unis' top papers							0.009
							[0.006]
Unis' active staff in engineering							-0.001
							[0.001]
Other university controls	No	No	No	No	No	No	Yes
University Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1.037***	-1.072***	-0.913***	-1.032***			-0.955***
	[0.181]	[0.182]	[0.187]	[0.188]			[0.139]
Observations	5,513	5,513	4,671	4,671	5,513	5,513	4,793

Robust standard errors in brackets

Table 5: Probability of collaboration as a function of the type and ability of researchers and control variables. Columns 5 and 6 display marginal effects