Micro-Foundations of R&D Alliance Formation:
The Interplay of Scientist Mobility and the Cooperative Context of Collaboration

Abstract. Alliance research emphasizes that firms can access R&D collaboration opportunities when they enjoy relational or geographic embeddedness with potential partners. But, how can firms that are not embedded with prospective partners establish alliances? We emphasize the micro-foundations of R&D alliance formation and propose that scientist mobility is an important substitutive mechanism that helps foster collaboration opportunities between firms that are poorly embedded. Specifically, we posit and show that in high-tech industries, scientist mobility is more facilitative for R&D alliance formation when potential partners lack relational ties between them or are not geographically colocated. Our findings demonstrate how incorporation of the competitive labor market context and its interplay with the cooperative context changes the insights of a fundamental research stream emphasizing the importance of the cooperative context for alliance formation.

Introduction

Prior research in R&D partner selection emphasized that firms' embeddedness along two distinct dimensions – relational and geographic – helps them access R&D alliance opportunities. Specifically, the relational view suggests that opportunities to collaborate stem from firms' prior pattern of interorganizational linkages (e.g., Ahuja 2000b; Powell et al. 1996). These linkages between potential partners are manifest as direct and indirect ties and reflect their cooperative embeddedness. Potential partners that share either direct or indirect ties enjoy familiarity with each other due to their prior linkages, so they can readily establish R&D alliances (e.g., Gulati 1995b; Gulati and Gargiulo 1999). However, these different kinds of stable relationships are often not readily available for private, high-tech new ventures (e.g., Baum et al. 2000; Shane and Cable 2002). The geographic dimension of embeddedness reflects firms' spatial proximity with potential partners and connectedness within a local institutional and cultural fabric that ensures steady access to information about each other via interpersonal communication networks (e.g., Hess 2004; Saxenian 1994). Thus, colocation of potential partners helps them discover and aggregate information about each other (e.g., Agrawal et al. 2015; Crandall et al. 2010; Hess 2004; Jaffe et al. 1993; Stigler 1961) and also promotes opportunities for establishing R&D alliances (e.g., Narula and Santangelo 2009; Reuer and Lahiri 2013). Given that R&D partner selection requires costly vetting of potential partners' intangible resources, relational and geographic embeddedness can help partners overcome informational barriers and establish linkages. However, the question that emerges from this body of research is, how can technology ventures that lack such embeddedness access opportunities for R&D collaboration?

To answer this research question, we rely on emerging research on the information-intermediation role of scientist mobility and its role in fostering interorganizational arrangements between firms (Collet and Hedström 2013; Mawdsley and Somaya 2016; Wagner and Goossen 2018). Mobile scientific personnel carry fine-grained information about firms' technological activities and are likely to be valuable for high-tech ventures to learn about external opportunities (Hess and Rothaermel 2011). The mobility of scientific personnel is valuable in transmitting information between high-tech firms (Almeida and Kogut

1999; Arrow 1962; Singh 2005; Singh and Agrawal 2011) because scientists that migrate from one firm to another are likely to be informed and have knowledge about the underlying R&D resources and activities of the firms (Dokko et al. 2009; Palomeras and Melero 2010). The informational role of mobile scientists can help us better understand the micro-foundations of alliance formation involving high-tech ventures (Salvato et al. 2017). Recent research by Wagner and Goossen (2018) has demonstrated that mobile scientists can be instrumental in facilitating R&D alliances between their current and prior employers. In our study, we build upon and extend this research by developing the idea that scientist mobility is an important micro-organizational mechanism that can offset the absence of interfirm ties and geographic colocation in a venture's cooperative context.

Our study offers several contributions to the literature. Whereas previous alliance research primarily attends to the cooperative context of collaboration by emphasizing the role of relational and geographic embeddedness (e.g., Gulati 1995b; Gulati and Gargiulo 1999; Narula and Santangelo 2009), we emphasize scientist mobility in the factor market and show that it not only influences ventures' likelihood of allying (Wagner and Goossen 2018), but it also alters the implications of interfirm ties and colocation emphasized in previous research. The fact that scientist mobility and, interfirm ties or colocation, may substitute for one another in fostering R&D alliances suggests that future research needs to attend to the competitive labor market context of collaboration rather than only the cooperative context, just as emerging research on labor market mobility needs to devote more attention to the cooperative context of collaboration and firms' embeddedness.

Our findings also replicate in a new context findings in recent research about the facilitative role of scientist mobility for R&D alliance formation (Wagner and Goossen 2018). Whereas that research focused on scientist mobility between the largest global pharmaceutical companies, we show how the mobility of scientists shapes opportunities for R&D alliances among high-tech new ventures in the biotechnology industry. Furthermore, our research reveals distinct boundary conditions that reflect partners' lack of embeddedness and shows scientist mobility is a meaningful mechanism for R&D alliance formation between partners without relational or geographic embeddedness. Our findings also demonstrate that scientist mobility is insignificant in fostering R&D collaboration when firms have prior ties or are geographically colocated.

In our empirical analyses, we consider the potential endogeneity of scientist mobility in the high-tech setting and employ novel exclusion restrictions (crime rates and weather conditions) for employee mobility, inspired by research from urban and labor economics, while employing the two-stage residual inclusion method (Rivers and Vuong 1988; Terza 2018; Terza et al. 2008). In doing so, we also demonstrate how research in urban and labor economics can be of use in addressing endogeneity in research that links labor market mobility and strategy.

Theory and Hypotheses

While R&D alliances help high-tech ventures overcome resource deficiencies and improve commercialization prospects, forming R&D alliances can be challenging for them. Lack of information about potential partners' resources make R&D alliance formation difficult, so firms may become wary about adverse selection risks (Akerlof 1970), and R&D alliances among firms can fail to occur. To overcome these information inefficiencies, firms often tap into their existing interfirm relationships or consider nearby firms for R&D collaboration (Gulati 1995b; Gulati and Gargiulo 1999; Reuer and Lahiri 2013). However, high-tech ventures often lack stable ties with other firms (Baum et al. 2000), or they find it difficult to look for partners beyond their locality (Narula and Santangelo 2009). High-tech ventures thus find it challenging to access collaborative opportunities with prospective partners and form alliances.

Interfirm movement of employees can help firms discover new information and identify new opportunities (Casper 2007; Rosenkopf and Almeida 2003). Mobility of scientific personnel transmits information between firms (Almeida and Kogut 1999; Arrow 1962; Singh 2005), and mobile scientists can serve as important information conduits for ventures as they carry information about the underlying R&D resources of the firms they worked for in the past (Palomeras and Melero 2010). They are particularly valuable for new ventures for accessing R&D alliance opportunities for several reasons.

Scientific personnel typically hold first-hand information about a firm's technological capabilities and activities (Wezel et al. 2006), thus creating opportunities for a firm to learn about the nature of resources and activities at the employees' prior employers (Almeida and Kogut 1999; Arrow 1962). Because of these informational advantages conferred through mobile scientists, managers of a venture considering an alliance are more likely to consult these scientists when deciding upon R&D collaborations. Managers become aware of the scientists' domains of expertise and employment history already at the recruitment stage, as managers are typically involved in the recruitment of knowledge workers (Rosenkopf and Almeida 2003). Evidence from the biotechnology industry demonstrates that scientists are often called upon to codify or translate scientific information for managerial use (Liebeskind et al. 1996). They also assume temporary administrative roles as R&D heads or become permanent members of executive groups (Schweizer 2005), suggesting the important role of scientists in evaluating the technological expertise and resources of prospective R&D partners. Moreover, many high-tech firms are young and of small size, with information exchange between scientific and managerial personnel occurring regularly and with some scientists even taking on executive roles (Higgins and Gulati 2006). They can therefore be an important mechanism for new ventures to discover opportunities for R&D collaboration and make partner selection decisions (Wagner and Goossen 2018).

In the hypotheses developed below, we focus on the micro-foundations of R&D alliance formation and suggest that the informational advantages stemming from scientist mobility between prospective partners are particularly valuable for ventures that are poorly embedded in a network of prior interfirm ties or are not geographically colocated with prospective alliance partners. Therefore, whereas previous research has attended to the cooperative context of collaboration by emphasizing the role of interfirm ties and colocation, below we focus on the interplay between mobility in the labor market context and the cooperative context in fostering R&D alliances for high-tech ventures.

Scientist Mobility and the Absence of Relational Ties

The value of relational ties between partners has been extensively studied in alliance research. In particular, this research used an embeddedness perspective to argue that prior collaboration ties can act as a remedy for information disadvantages and can thereby diminish the risks involved in R&D alliance formation. For example, prior alliance agreements provide a means for prospective collaborators to collect fine-grained information about each other's technological know-how as well as research and development activities (Gulati 1995a; Gulati and Gargiulo 1999). Relatedly, prior ties help firms cope with uncertainties related to the disclosure of firm-specific information such as information about intangible assets that prevail in knowledge-intensive industries, yet are also difficult to value and discern from financial statements (Chi 1994). Previous alliances therefore enable prospective partners to accumulate rich information about each other's current and future technological endeavors and develop deeper understandings about each other's resources (Vanhaverbeke et al. 2002). In the context of R&D alliance formation, prior ties offer firms critical information on intangible assets such as R&D resources, which are not easily accessible by other firms (Almeida et al. 2002; Higgins and Rodriguez 2006).

Similar advantages can be provided to firms via indirect ties that firms have through common partners. Indirect ties via shared alliance partners serve as conduits of information about the underlying R&D resources and competences of potential collaborators because a firm's partner can provide information from their interactions with their other partners (Ahuja 2000b; Gulati and Gargiulo 1999). When indirectly connected through a shared alliance partner, prospective partners each have incentives to represent themselves accurately (Powell 1990), making the collection and interpretation of R&D information during due diligence less difficult and costly. Indirect ties between prospective alliance partners can thus serve as a means for both information-gathering and information-screening (Rangan 2000). For example, firms with indirect ties may gather information about the success and failure of research efforts of prospective partners (Rogers and Kincaid 1981), but also screen, absorb and classify more information about prospective partners' R&D resources and prospects than the information processing capability of an individual firm alone would allow (Leonard-Barton 1984). The R&D-related information transferred through shared third-party ties is made credible and interpretable because the shared partner provides reliability by acting as a referee. Indirect ties thus also allow prospective partners to evaluate each other's network resources and infer whether partnering with one another could enable them to benefit from the other party's embeddedness (Gulati 1999).

Given that prior direct and indirect relationships provide firm access to significant information about prospective collaborators' R&D resources and activities, they diminish the role that information carried by mobile scientists can play for these same purposes. By contrast, absent prior relationships between firms, prospective partners can rely more upon mobile scientists to obtain information about one another's R&D resources and activities. As a consequence, the informational benefits we posit for mobile scientists, in many ways, parallel those ascribed to relational ties by previous research, and we thus expect that prior relational ties and scientist mobility will substitute for each other in fostering R&D alliances.

> **Hypothesis 1***: The positive effect of scientist mobility on the likelihood of R&D alliance formation will be stronger when prospective partners lack relational ties.*

Scientist Mobility and the Absence of Colocation

The value of spatial colocation has also been extensively studied in the research stream on alliances (e.g., Felzensztein et al. 2010; Narula and Santangelo 2009). In particular, this research has also used an embeddedness perspective to argue that physical closeness allows for frequent interactions and effective exchange of information between firms, thereby acting as a remedy for information disadvantages and the risks involved in R&D alliance formation. Being colocated engenders connectedness for firms within a local institutional and cultural fabric and ensures steady access to information about each other through interpersonal communication networks (e.g., Saxenian 1994).

Previous research in economic geography and management has highlighted the economic significance of geographic proximity for firms as it helps them easily discover each other and aggregate information in several informal ways (e.g., Agrawal et al. 2015; Crandall et al. 2010; Hess 2004; Jaffe et al. 1993; Stigler 1961). For example, firms may enjoy social and professional gatherings, such as conferences and boot camps, which serve as conduits for information exchange about technological developments and emerging R&D opportunities (e.g., Liebeskind et al. 1996; Owen-Smith and Powell 2004). Colocation enables face-to-face communication between organizational actors, which is critical to transferring tacit knowledge (Daft and Lengel 1986). Colocation may foster information exchange not only due to spatial proximity but also through the various information channels it provides, such as employees joining community groups, residing in the same neighborhoods, or participating in local industry events (e.g., Almeida and Kogut 1999; Saxenian 1990). Such information exchange is especially important in high-tech industries, where business involves a high degree of tacit knowledge. These interactions can provide firms critical information on intangible assets of prospective partners, such as R&D resources or project trajectories, which are typically not easily accessible by other firms (e.g., Almeida et al. 2002; Higgins and Rodriguez 2006). Given the informational benefits of colocation, it may be of particular value for new ventures that are usually short of track records and have limited publicly available information. Colocation thereby help them convey information about their resources and prospects to potential partners as well as learn about them to access collaboration opportunities.

The foregoing discussion suggests that geographic colocation and scientist mobility are distinct but partially redundant means for accessing information about potential R&D alliance partners. We expect that the information benefits from scientist mobility will substitute for the information benefits of colocation-enabled interactions in fostering R&D alliances. In other words, we argue that the role of scientist mobility in providing informational advantages becomes less useful for R&D alliance formation when firms are colocated. By contrast, when firms are not colocated, scientist mobility in the labor market is expected to have a greater bearing on R&D alliance formation to address the heightened information frictions that accompany greater geographic distance.

> **Hypothesis 2*:*** *The positive effect of scientist mobility on the likelihood of R&D alliance formation will be stronger when prospective partners are not colocated.*

Methodology

Data and Sample

To test these hypotheses, we focus on R&D alliances among ventures in the biotechnology industry. This industry is well recognized as a setting where access to resources that enable future development of a firm rests on alliances (Powell et al. 1996). However, discerning the value of a biotech venture as a potential alliance partner can be difficult because of its intangible resources and short business record (Stuart et al. 1999). Because mobility of scientists is also a pervasive phenomenon in the biotech industry (Casper

2007), scientists that move between firms can serve as important information intermediaries that help new ventures learn about potential R&D alliance partners. Additionally, the intensity of patenting activity in this industry is high (Roijakkers and Hagedoorn 2006), providing a means to infer scientist mobility and firms' technological resources from publicly-available patent records (Almeida and Kogut 1999).

We collected data on alliances among biotechnology ventures using Clarivate Analytics' Cortellis Deals Intelligence database during the fifteen-year period 1995-2009. This database, previously known as Thomson Reuters' Recap database, is considered to be one of the most reliable and comprehensive data sources for alliance agreements in the biotechnology industry and has been used in prior strategy research to investigate cooperative strategies of biotech ventures (Schilling 2009). We applied several sample screening criteria to the alliance data. We limited our sample to alliances that have an R&D component (Oxley 1997), that is deals that the database classifies as research, collaboration, development or co-development alliances, though deals might involve more than one type and also might involve other activities such as manufacturing and marketing. We also limited our sample to alliances between firms based in the United States in order to reduce unobserved heterogeneity from cross-border alliances. Further, we included only alliances between ventures that, five years prior to the focal alliance, had at least one patent granted by the patent office, as this allowed inferences to be drawn on scientist mobility from the patent data (Almeida and Kogut 1999).

We hand-matched each firm from the Cortellis database with the corresponding assignee in the US Patent Inventor Database (Li et al. 2014) based on firm name. We used the US Patent Inventor Database from Harvard Dataverse to identify scientist mobility events because this source contains patent data with disambiguated patent inventor names. To identify scientist mobility events, we relied on the coding procedure developed in prior research (Almeida and Kogut 1999; Corredoira and Rosenkopf 2010). For each firm in our sample, we examined the full set of patents available in the patent dataset from 1990 onwards. We chronologically traced each individual scientist's patenting history to determine instances where a scientist was employed by more than one firm over his or her patent trajectory. A mobility event was identified when a scientist was listed as an inventor in patents granted to two different firms, and the corresponding year of mobility was calculated as the halfway point between the last patenting date at the originating firm and the first patenting date at the destination firm (Singh and Agrawal 2011).

We also followed prior research in undertaking steps to minimize potential errors with respect to the identification and timing of scientist mobility events (Corredoira and Rosenkopf 2010; Singh and Agrawal 2011). Specifically, we verified that patents suggesting potential scientist mobility events were not granted to more than one firm. We also examined in detail each instance when a scientist appeared to have moved back and forth between firms, and we excluded such cases. To further reduce the possibility of misclassifying firm name changes or acquisitions as instances of mobility, we manually inspected algorithm-identified mobility events.

Since our dependent variable is the likelihood of R&D alliance formation between ventures, our analyses required sampling on realized alliances between ventures and constructing a set of corresponding counterfactuals that were potential, yet unrealized alliances. For each realized alliance deal, we constructed the set of unrealized alliance deals by considering all biotech ventures from the population of firms in the Cortellis database that fulfilled the same screening criteria used for sampling alliance deals, namely firms that (a) were performing their primary business activity in the biotechnology industry at the time of the focal realized alliance deal, (b) had their location in the US, and (c) five years prior to the focal alliance had at least one patent reported in the US Patent Inventor Database. This procedure generated a comprehensive set of unrealized alliance deals and enabled us to exploit heterogeneity in the sample for testing the hypotheses in an unbiased way when *a priori* knowledge about the likelihood of an R&D alliance in a dyad is unavailable (Gulati and Gargiulo 1999). We also made additional adjustments for acquisitions, because biotech firms are frequent takeover targets (e.g., Grigoriou and Rothaermel 2017). In particular, if one firm in the sample acquired another one, the acquired firm was dropped from the sample in the year of acquisition. After matching realized and potential alliance dyads, we obtained a resulting dataset of 234,786 dyads of which 506 are realized R&D alliances. These alliance dyads in the final sample involved a total of 751 ventures.

Finally, we obtained founding date and location information for each firm in our sample from corporate websites, Securities and Exchange Commission (SEC) filings, online platforms such as zoominfo.com and manta.com, as well as news articles.

Dependent Variable

We investigate the likelihood that two ventures in the biotechnology industry engage in an R&D alliance with each other. Therefore, our dependent variable, *R&D Alliance Formation*, is dyadic and dichotomous, and equals 1 for realized alliances and 0 for unrealized alliances between two ventures in year *t*.

Independent Variables

Our independent variable *Scientist Mobility* is also a dichotomous measure that takes on the value of 1 if a scientist mobility event occurred between the two firms in a dyad during the period of five years from *t*-5 to *t*-1, where *t* is a year of a focal alliance deal, and 0 otherwise (Corredoira and Rosenkopf 2010). We introduce the one-year lag in this variable to alleviate concerns of endogeneity in our analyses, and below we address several additional steps taken to mitigate and investigate this concern. The sample we exploit is based on unique dyads meaning that, for all observations, the specification of a firm as either the first or the second partner in a realized or unrealized alliance dyad makes no difference for any of the variables. Hence, our mobility variable also accounts for scientists' movements in either of the two possible directions in a dyad.

In Hypothesis 1, we posit that the positive effect of scientist mobility on R&D alliance formation will be more pronounced when the firms in a prospective alliance dyad lack relational ties. We thus constructed the variable *Absence of Ties* which takes on a value of 1 if firms in a dyad have no direct or indirect (i.e. shared third-party) interfirm alliances during the five years prior to the focal R&D alliance, and 0 otherwise.

In Hypothesis 2, we propose that the positive effect of scientist mobility on R&D alliance formation becomes stronger when potential R&D partners are not geographically colocated. For the variable *Absence of Colocation* we used a binary indicator variable which takes on a value of 1 if the headquarters of both prospective R&D alliance partners are not located in the same state of the United States, and 0 otherwise.

In supplemental analyses described below, we carry out a number of robustness checks for these variables such as employing alternative time lags, measurement windows, and econometric specifications.

Control Variables

We introduced a set of controls that could be related to the likelihood of R&D collaboration as well as the cooperative context of a potential collaboration. Because firms with overlapping technologies might be more likely to engage in R&D collaboration, we controlled for *Technology Similarity*, measured as the number of the two firms' overlapping technological niches, as proxied by patent classes at the three-digit level, divided by the total number of distinct technological niches of the prospective alliance partners (Mowery et al. 1996; Stuart and Podolny 1996). The measure is based on the past five years of firms' patenting activities.

Alliance activity of a firm might shape the resources to which it has access and might also convey to prospective R&D partners that its resources are in demand by other organizations (Stuart et al. 1999). Previous alliance partners have carried out evaluations of the firm's resources and capabilities (Hitt et al. 2004) and are involved in the operations of the firm (Almeida et al. 2002). In addition, prospective partners that have already engaged in collaborations might be more effective in engaging in new alliance activities (Hoang and Rothaermel 2005). We therefore incorporated the variable *Alliance Experience*, measured as the log-transformed value of the average number of alliances formed by two partners in a dyad during the period of five years prior to the focal alliance deal.

For new ventures, their prior technological achievements can convey information about their technological resources and prospects (Stuart 2000). To account for the patenting records of prospective partners as an indication of their technological achievements, we created the variable *Patents*, measured as the log-transformed value of the average number of patents granted to two firms in a dyad during the period of five year prior to the focal R&D alliance (Agarwal et al. 2009).

Further, the availability of information about a firm can also depend on its age. While the technology of a young startup might be attractive to prospective partners, it could also present greater uncertainty given its short track record (Nicholson et al. 2005; Stuart et al. 1999) or lack of reliability and accountability in its business (Hannan and Freeman 1989). We measured firm age as the number of years from firm founding until the year of a particular realized or unrealized alliance deal, and then created the variable *Age* by log-transforming the average age of the two firms in a dyad.

We incorporated different types of fixed effects to account for other sources of unobserved effects. We included *Therapeutic Area Effects* to control for the effect related to the therapeutic area of the focal R&D alliance. Further, we defined *Patent Class Effects* to control for whether firms in a dyad patented in particular three-digit technology classes that might affect R&D alliance formation (Jaffe 1989; Jaffe et al. 1993). We added *Year Effects* to capture temporal trends that might affect firms' decisions regarding R&D alliance formation. Finally, we incorporated *Firm State Effects* and *Partner State Effects* to control for attractiveness of the location in which the prospective partners reside (i.e. state of headquarters).

Empirical Approach and Identification Strategy

Unobserved heterogeneity surrounding the mobility of scientists could potentially account for our findings regarding the contingent effects of scientist mobility on R&D alliance formation. For example, certain strategic intents of a firm, for instance technological repositioning (Kogut and Zander 1992), could potentially affect both the recruitment or layoff of scientists (Tzabbar 2009) and the firm's engagement in R&D alliances (Mowery et al. 1996). We therefore accounted for the endogeneity of scientist mobility by using a two-stage residual inclusion (2SRI) estimation method, which is a particularly suitable approach for correcting endogeneity bias in nonlinear models that involve dichotomous dependent variables (Rivers and Vuong 1988; Terza 2018; Terza et al. 2008). The 2SRI involves specification of a maximum likelihood (ML) estimation method, which helps capture the nonlinearity in the dependent variables and the endogenous regressor, and thereby mitigates a specification error (Nakamura and Nakamura 1998; Wooldridge 2014). More importantly, the 2SRI involves the inclusion of residuals from first stage ML estimation of the endogenous regressor in the second stage conditional ML estimation of the dependent variable (Blundell and Powell 2004; Newey 1987; Rivers and Vuong 1988; Terza 2018). The inclusion of residuals from the first-stage regression substitutes for unobservable confounders and therefore helps better account for endogeneity of the regressors (e.g., Terza et al., 2008; Wooldridge, 2014; Geraci et al., 2018). Terza et al. (2008) have demonstrated the superiority of 2SRI in addressing endogeneity in nonlinear models over the extension of the 2SLS estimator to nonlinear models. Specifically, they showed that for nonlinear models, the two-stage least square (2SLS) estimator is not consistent, whereas the 2SRI estimator is. Terza (2018) recently also offered further guidance as well as a step-by-step protocol for the implementation of the 2SRI approach.

Following this approach, we developed a first stage logit model in which we predicted the likelihood of scientist mobility between firms in a potential alliance dyad. This model included all our independent and control variables as well as additional exclusion restrictions. Relying on research in urban and labor economics that suggests safety and weather conditions affect labor mobility (e.g., Andrienko and Guriev 2004; Malecki and Bradbury 1992; Quigley and Weinberg 1977; Rappaport 2007), we used crime rates from the FBI database and WalletHub's weather conditions as exclusion restrictions. We expected a high crime rate in the scientist's potential destination location to decrease scientist mobility to that location, whereas we expected favorable weather conditions to increase the movement of scientists to that location. As mentioned above, in our dyad-level sample, the specification of a firm as either the first or the second partner in a dyad is arbitrary and makes no difference for any of the variables. Accordingly, our scientist mobility variable captures movements of scientists in either of the two directions between prospective partners. To accommodate such a study design, we included two exclusion restriction variables referring to the differences in the crime rates and in the weather conditions in the locations of the alliance partners. To check for the strength of our exclusion restrictions, we ran the first stage with and without exclusion restrictions, and we found a significant difference in model fit when we include these variables. Crime rates and weather ranking in the scientists' potential destination location appeared strong exclusion restriction variables with high joint significance in the first stage model ($p < 0.01$) (Bascle 2008). Also, none of our exclusion restriction variables appeared correlated with the residuals associated with the dependent variable in the second-stage regression – R&D alliance formation (r = -0.001; r = - 0.002, respectively; $p > 0.20$ for both exclusion restrictions) (Semadeni et al. 2014). Further, the overidentifying restrictions test statistic suggested that we cannot reject the null that the instruments are exogenous (Cameron and Trivedi 2005). The results of the first-stage model are presented in Table 1. We obtained *Residuals* from this model and included them as an additional control variable in the second stage model predicting R&D alliance formation.

For the second-stage models of R&D alliance formation, we specified logistic regressions since our dependent variable is also of a dichotomous nature. We accounted for dyadic dependence due to the repeated occurrence of firms as well as for autocorrelation stemming from repeated occurrence of dyads in multiple years. Specifically, we estimated robust standard errors that are clustered on three dimensions including the firm dyad and the two potential alliance partners (Cameron et al. 2011). We implemented three-way clustering of standard errors using the *cluster.vcov* function in the *multiwayvcov* package in R (Graham et al. 2016).

Results

Table 1 provides a correlation matrix and descriptive statistics. As expected, a moderate correlation is present between partners' alliance experience and patents as well as between partners' average age and patents. Omitting any of these control variables, however, does not affect our findings presented below. Moreover, the maximum variance inflation factor across the models is 1.72, well below the recommended critical cutoff level of 10 (Neter et al. 1996), suggesting that there are no multicollinearity concerns evident for our models. As for the summary statistics, there are some differences across the realized and unrealized alliance dyads for the main theoretical variables. For instance, in the set of realized alliances, 7.11% of dyads experienced at least one scientist mobility event in the period of five years prior to the focal alliance, whereas in the set of unrealized alliances, only 0.38% of dyads had such instances of scientist mobility ($p < 0.01$). Relational ties between prospective alliance partners are absent in 80.43% of dyads in the set of realized alliances, and in 98.63% of dyads in the set of unrealized alliances ($p < 0.01$). Firms in our sample originate in 34 different US states, comprising 802 unique state-dyads. Figure 1 in Appendix A shows the distribution of firms across the US states. Whereas prospective collaboration partners were not colocated in 75.49% of realized alliances, the absence of colocation is higher with 81.42% of dyads in the set of unrealized alliances (p < 0.01).

**** Insert Table 1 here ****

The results of logistic regression models analyzing the probability of R&D alliance formation are presented in Table 2. Following Hoetker (2007), we report coefficient estimates, standard errors and marginal effects. Model 1 is a baseline specification including control variables only. The multivariate estimation results are in line with research which suggests that scientist mobility increases the probability of alliance formation between the two firms ($p < 0.01$) (Wagner and Goossen 2018). The marginal effect of scientist mobility on R&D alliance formation suggests that scientist mobility between firms increases the likelihood of R&D alliance formation between them by approximately 0.2% (1.0% in the full model), when variables are at sample averages, which is comparable to the marginal effects reported in Wagner and Goossen (2018). These results demonstrate the generalizability of the scientist mobility effect across various settings.

The baseline results for other variables are also in line with work on alliance formation. We find significant and negative effect for the absence of ties variable ($p < 0.01$) (e.g., Ahuja 2000b; Gulati 1995a). Moreover, when prospective partners are not colocated they are less likely to engage in an alliance ($p < 0.10$) (e.g., Reuer and Lahiri 2013). Firms with similar technological knowledge are also more likely to engage in an alliance ($p < 0.01$) (e.g., Chung et al. 2000). Also, firms experienced in interfirm collaboration are more likely to engage in future alliances ($p < 0.01$) (Stuart et al. 1999). We also note that the coefficient of residuals which we included in the model to address endogeneity is significant ($p < 0.01$), indicating that unobserved factors surrounding the mobility of scientists could also affect alliance formation. These results suggest that controlling for endogeneity is indeed important in investigating the impact of scientist mobility on firm-level outcomes (Tzabbar 2009).

**** Insert Table 2 here ****

Models 2 and 3 report the results of logistic regressions analyzing the hypothesized interaction effects, and Model 4 is the full model that contains both interaction effects estimated at once. Hypothesis 1 suggests that the positive effect of scientist mobility on R&D alliance formation will be more pronounced when prospective partners lack relational ties. In line with this hypothesis, the coefficient estimate of the interaction between *Absence of Ties* and *Scientist Mobility* in Model 2 is positive and significant ($p < 0.01$). Hypothesis 2 proposes that the positive effect of scientist mobility on R&D alliance formation will be more pronounced when the parties are not colocated. Consistent with expectations, in Model 3 the coefficient estimate of the interaction between *Absence of Colocation* and *Scientist Mobility* is also positive and significant ($p < 0.05$), thus furnishing support to Hypothesis 2. Both interaction coefficients in the full model, Model 4, yield consistent interpretations.

To further interpret the interaction effects, we plotted them graphically in Figures 1-2. In Figure 1, we depict the average marginal effects of scientist mobility with respect to relational ties between potential partners. The bars surrounding the line denote the 95% confidence intervals. In line with the regression estimates, Figure 1 illustrates that the effect of scientist mobility is stronger when prospective partners have no ties. The calculation of the corresponding average marginal effects suggests that average marginal effect of scientist mobility on the likelihood of R&D alliance formation increases eight-fold in the absence of ties. In Figure 2, we depict the average marginal effects of scientist mobility with respect to whether potential partners are colocated, indicating that the effect of scientist mobility on the likelihood of R&D alliance formation is more pronounced for prospective partners that are not colocated. The marginal effect of scientist mobility on the likelihood of R&D alliance formation increases almost three-fold for partners that are not colocated. Taken together, these results provide evidence that the effects of scientist mobility are particularly important for firms that have not developed ties or are not colocated with prospective alliance partners. Scientist mobility that is effected through labor markets and the cooperative context providing ties to other firms and proximity thus appear to be substitutive in fostering R&D alliances. Moreover, the confidence intervals indicate the strength of this substitution effect in that scientist mobility becomes insignificant in effecting R&D alliance formation for firms that have prior ties or that are colocated.

*** *Insert Figures 1-2* ***

Supplementary Analyses

We also carried out a number of supplementary analyses. We first examined whether our findings extend to alternative measures of our key explanatory variables. Whereas in the main analyses we capture prior ties using a dummy variable, we also constructed a continuous measure capturing the number of prior direct and indirect ties between the prospective alliance partners during the period of five years preceding the focal alliance. The interpretation of our main findings remained the same. Similarly, we constructed an alternative measure of absence of colocation – a dummy variable taking the value of 1 if the geographical distance between the prospective partners is more than 50 miles, and 0 otherwise. In calculating the distance between the headquarters of the two firms in the dyad, we used zip codes to determine the latitude and longitude of the firms, and we applied the great circle distance formula. Also, we created a continuous variable measuring the geographical distance in miles between the headquarter locations of prospective partners in a dyad. In both cases, the interpretation of our main findings remained the same. We also assessed whether our results extend to alternative measures for the scientist mobility variable, using time windows such as seven or ten years (e.g., Tzabbar 2009). We found that the interaction effects of scientist mobility and the absence of ties (or absence of colocation) on R&D alliance formation remain positive and significant.

In another set of analyses, we examined whether our results are robust to alternative sampling approaches featured in the literature on alliance formation (Gulati and Gargiulo 1999). To exploit heterogeneity in the sample for testing our hypotheses, our main sample was constructed by creating all the possible counterfactuals for a given realized alliance. However, we also wanted to ensure the robustness of our results to different draws of counterfactuals. For each realized alliance, we randomly selected counterfactuals to obtain different ratios between unrealized and realized dyads, such as 20:1 and 50:1. Through these analyses we obtained results consistent to our main results.

To further examine the sensitivity of the results to the estimation approach employed, we also employed the two-stage least square (2SLS) regression. The results from both the first and second stage models were consistent with those obtained by 2SRI (see Appendix B). Specifically, with respect to the exclusion restrictions in the first stage, we found that the F-statistic (Kleibergen-Paap rk Wald F statistic) for weak instruments is 26.49 and above the Stock-Yogo cut-off value at the 5% level (13.46) for weak instrument bias (Stock and Yogo, 2002). The Hansen J statistic was 0.002 ($p = 0.96$), suggesting that the overidentifying restrictions are valid. Together, these tests provided additional assurances that the instruments are valid. Further, the underidentification test statistic (Kleibergen-Paap rk LM statistic) was 14.51 and significant ($p < 0.001$), indicating that the instruments are not irrelevant. However, the parameter estimates in the 2SLS suggest a downward bias when we compare these results with estimates in our main analyses from the two-stage residual inclusion (2SRI) procedure. While our substantive interpretations do not change, this pattern is consistent with previous conclusions that 2SLS can be an incorrect specification in the context of limited dependent variable models with endogenous regressors such as ours (Hausman 1978; Rivers and Vuong 1988; Smith and Blundell 1986) and that the two-stage residual approach is a superior method to control for endogeneity in nonlinear regressions frameworks (Blundell and Powell 2004; Newey 1987; Rivers and Vuong 1988; Terza et al. 2008; Wooldridge 2014).

Finally, while in our main analyses we employed a two-stage residual inclusion approach to address the potential endogeneity of scientist mobility, in additional analyses we also utilized matching estimators (see Appendix C1 and C2). To that end, we performed treatment effect analyses based on propensity-score matching (PSM) and inverse-probability weighting (IPW) (Morgan and Winship 2015) using *teffects* in Stata. These techniques help control for the potential non-randomness of a treatment variable and obtain a quasi-experimental setup via the construction of observations, with and without the treatment, respectively, that are comparable to each other on observed covariates that potentially influence the treatment and outcome variables (Gangl 2014). In the PSM approach (using *teffects psmatch*), we first estimated a propensity score for scientist mobility which is the treatment of interest in our study, by specifying a logistic regression and using other covariates that can predict both scientist mobility and R&D alliance formation between potential partners (i.e. relational ties, colocation, technology similarity, alliance experience, patents, age, and various fixed effects). Treated and control groups were identified using nearest neighbor matching and estimated the average treatment effect of scientist mobility on R&D alliance formation. We implemented IPW (using *teffects ipw* in Stata) and used the inverse of propensity scores as weights in the second stage, and estimated the average treatment effect. The results across these models demonstrate that the average treatment effect of scientist mobility on the likelihood of R&D alliance formation is positive and significant, and is also similar to results reported in Table 3. Finally, we also employed a doubly robust estimator using inverse-probability weights with regression adjustment (IPWRA) using *teffects ipwra* in Stata, as an additional robustness check (Wooldridge, 2010). In this approach, we estimated the treatment model for scientist mobility by implementing the first stage described in the 2SRI estimation, and we used inverse-probability weights in the second stage regression along with the set of covariates explaining R&D alliance formation (Morgan and Winship 2015). More importantly, as we are interested in the effects of scientist mobility in the absence of ties between partners and when partners are not colocated, we applied the matching estimator approaches to examine the average effects of scientist mobility in subsamples based on these two conditions. We found across both matching estimator approaches that the average treatment effect of scientist mobility is positive and significant when ties are absent between partners. By contrast, the effects are insignificant when ties are present between partners, suggesting that the role of scientist mobility in alliance formation becomes redundant for partners having previous alliances with each other. Similarly, we found that across both models the average treatment effect of scientist mobility is positive and significant for firms that are not colocated, while it is insignificant for those that are colocated, suggesting that scientist mobility is a useful mechanism for R&D partner selection for partners that are geographically apart. Taken together, these results provide additional support for our theoretical expectation on the positive role that scientist mobility plays in fostering R&D alliances between firms that lack relational ties or geographic colocation.

Discussion

Theoretical Implications

Our study provides several contributions to the literatures on strategic alliances and mobility in the labor market for knowledge workers. Whereas previous alliance formation research has attended to the cooperative context of collaboration by emphasizing the role of prior organizational ties of various types (e.g., Gulati 1995b; Gulati and Gargiulo 1999) and geographic colocation (e.g., Felzensztein et al. 2010;

Narula and Santangelo 2009), we submit that scientist mobility in the factor market is an aspect of the competitive context that not only shapes firms' likelihood of engaging in R&D collaborations but also the implications of various interfirm ties and colocation. A focus on mobile scientists shows that their influence is highly interdependent with the well-known relational and geographic factors that characterize the cooperative context in which firms are embedded. Specifically, we argue and show scientist mobility might be especially important in providing collaboration opportunities for new high-tech ventures that lack extensive relational ties with other firms or are non-colocated (Eisenhardt and Schoonhoven 1996; Hagedoorn 2002). This also indicates that as firms develop and are more embedded, the importance of scientist mobility for their R&D collaboration opportunities will diminish.

Our results suggest that mobile scientists represent one important micro-foundation of R&D alliance formation between ventures in high-tech industries. We thus replicate and extend to a new domain recent research demonstrating the role of scientist mobility for alliance formation of large corporations (Wagner and Goossen 2018), and we answer recent calls for more research on micro-foundations that fundamentally revolve around the impact of individuals on strategic decisions of firms (Felin et al. 2015). More broadly, we also answer multiple calls for research taking a cross-level perspective on interorganizational relationships (Lumineau and Oliveira 2018; Mawdsley and Somaya 2016; Salvato et al. 2017) and advance this research in two important ways. First, we build upon and extend the scant research that has considered how ties between individual employees can influence ties at the interfirm level (Collet and Hedström 2013; Gulati and Westphal 1999). We thus add to the cross-level perspectives on the formation of alliances by showing how scientist mobility can play an important role in the development of inter-organizational ties (Hess and Rothaermel 2011; Wagner and Goossen 2018). Further, we show that scientist mobility and firm-level ties are alternative channels that are substitutive in facilitating interorganizational economic exchanges such as strategic alliances. Our study suggests that future research should consider the multiple ways that labor market mobility can affect other strategic transactions (e.g., acquisitions, international joint ventures, foreign direct investment, etc.).

Our study also builds upon and extends research on the implications of labor market mobility in several unique ways. Our work sheds light on an important outcome of labor market mobility between high-tech ventures that complements existing perspectives in this literature. Much of the extant research suggests that mobile scientists facilitate outcomes such as learning and knowledge transfer (Rosenkopf and Almeida 2003; Song et al. 2003). Loss of scientific human capital can therefore erode a firm's competitive advantages and can be detrimental to high-tech ventures at an important stage in their development. At the same time, to the extent that information intermediation by mobile scientists can provide information advantages and reduce risks pertaining to prospective R&D alliances, our arguments and findings point to an unexamined potential benefit of scientist mobility between firms. We therefore provide further evidence that scientist mobility not only has important implications for learning and firm's innovation strategies, as prior research has emphasized, but also strategy and specifically firms' access to opportunities for external corporate development activities for firms that are less embedded.

Moreover, recent research points to the limitations of mobility driven knowledge transfer in that it does not allow firms to access deeply buried know-how and skills in other firms. Specifically, the potential for knowledge transfer from scientist mobility may be limited due to knowledge being tacit and embedded within teams or organization structures (Palomeras and Melero 2010). Knowledge transfer can also be, highly contingent upon firm characteristics such as the firm's knowledge integration capability and innovation intensity (Herstad et al. 2015) or the ability of a scientist to collaborate with new colleagues (Tzabbar and Kehoe 2014). In line with this research, our findings suggest that scientist mobility does not necessarily obviate the need for alliances but can promote R&D collaborations under specific conditions.

Limitations and Future Research Directions

In addition to the research possibilities already noted, extensions might pursue research opportunities presented by some of the specific limitations of this study. We placed our analysis in the context of R&D collaborations between ventures in the biotechnology industry. This focus partially reflects prior research on alliance formation and the information hazards that accompany transactions in this context (Rothaermel and Boeker 2008). While we would expect our findings to generalize to other high-tech industries, ventures in other industries or at other stages of development might rely on other information conduits besides mobile scientists (e.g., board interlocks, relationships with financial intermediaries such as investment banks, etc.) and they may or may not substitute for scientist mobility. Similarly, since we investigate collaborating firms within the US, it would be worthwhile to examine alliance formation in international contexts where expatriate assignments, cross-border migration, and the return of employees to their countries of origin are common features of mobility. Such research could examine the potential limitations of mobility-driven learning and knowledge transfer as well as investigate whether mobility enhances the efficiency of R&D partnerships and other forms of collaboration.

We have investigated whether and under what conditions scientist mobility affect high-tech ventures' formation of alliances, yet we do not address the performance implications of scientist mobility for R&D alliances (Sampson 2007). In future research, it would therefore be valuable to examine whether and under what conditions the information intermediation role of mobile scientists actually reduces potential risks during the alliance formation process and improves subsequent alliance performance (Ahuja 2000a). Similarly, research on the informational advantages of mobile scientists might investigate other strategic outcomes such as innovation output, commercialization successes, and growth for individual partners and for their alliances.

In our analysis of who partners with whom in the market for R&D collaborations, we have used sets of realized alliances and unrealized deals. This research design accommodates unrealized alliances as counterfactuals rather than sampling only upon completed transactions. However, this empirical approach is limited in not being able to provide direct insights into the actual decision processes through which high-tech ventures search for R&D partners, consider a set of potential partners, evaluate certain information, or make choices between prospective partners. Although there is some evidence from prior research on how scientists affect strategic decision processes (Liebeskind et al. 1996; Schweizer 2005; Wagner and Goossen 2018), it would be interesting and important to investigate such processes in a fine-grained way using primary data from field surveys or longitudinal case studies and explore the role of executives, middle managers, scientists, legal staff, and other employees in these processes. We have employed a set of econometric techniques to address potential endogeneity concerns surrounding the relationship between scientist mobility and R&D alliance formation, but we cannot draw definitive, causal effects from observational data such as ours, so it would also be valuable to use other methodologies to investigate the implications of scientist mobility for R&D collaborations (e.g., field experiments, natural experiments, etc.).

Another empirical limitation of the study lies in the procedure for identifying the mobility events. While we followed the established procedure for tracking scientist mobility from publicly available patent data, we encourage future research that examines scientist mobility using other techniques or that considers the mobility of other employees and teams across organizations. For example, this work could examine how firms carry out acquihire strategies in different industries, and such transactions might substitute for a particular alliance or might open up new collaboration opportunities in the future. Although our data do not allow us to empirically examine this question, we were also interested more broadly in whether high-tech ventures consider hiring and partnering as substitutive choices for particular R&D projects. In interviews with executives at biotechnology firms, they indicated to us that they do not make decisions to hire versus partner, the reason being that the knowledge sought in partnerships is tightly tied to intellectual property contained in patents. It might be that in other industrial or national contexts, particularly those with weaker intellectual property regimes or in non-technology domains, firms make choices to hire versus partner for particular projects. Research in these directions will be able to investigate the impact of labor market mobility on firms' collaborative strategies and consider how the functioning of labor markets and competition for knowledge workers might sow the seeds of broader strategic opportunities, or risks, for high-tech new ventures (Eisenhardt and Schoonhoven 1996).

References

Agarwal, R., M. Ganco, R.H. Ziedonis. 2009. Reputations for toughness in patent enforcement: Implications for knowledge spillovers via inventor mobility. *Strateg. Management J.* **30**(13) 1349-1374.

Agrawal, A., C. Catalini, A. Goldfarb. 2015. Crowdfunding: Geography, social networks, and the timing of investment decisions. *J. Econ. Manag. Strategy* **24**(2) 253-274.

Ahuja, G. 2000a. Collaboration networks, structural holes, and innovation: A longitudinal study. *Admin. Sci. Quart.* **45**(3) 425-455.

Ahuja, G. 2000b. The duality of collaboration: Inducements and opportunities in the formation of interfirm linkages. *Strateg. Management J.* **21**(3) 317-343.

Akerlof, G. 1970. The market for lemons: Quality uncertainty and the market mechanism. *Quart. J. Econ.* **84**(3) 488-500.

Almeida, P., B. Kogut. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Manag. Sci.* **45**(7) 905-917.

Almeida, P., J. Song, R.M. Grant. 2002. Are firms superior to alliances and markets? An empirical test of cross-border knowledge building. *Organ. Sci.* **13**(2) 147-161.

Andrienko, Y., S. Guriev. 2004. Determinants of interregional mobility in Russia. *Econ. Transition* **12**(1) 1-27.

Arrow, K. 1962. Economic welfare and the allocation of resources for invention *The rate and direction of inventive activity: Economic and social factors*. Princeton University Press, 609-626.

Bascle, G. 2008. Controlling for endogeneity with instrumental variables in strategic management research. *Strat. Org.* **6**(3) 285-327.

Baum, J.A., T. Calabrese, B.S. Silverman. 2000. Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. *Strateg. Management J.* **21**(3) 267-294.

Blundell, R.W., J.L. Powell. 2004. Endogeneity in semiparametric binary response models. *Rev. Econ. Stud.* **71**(3) 655-679.

Cameron, A.C., J.B. Gelbach, D.L. Miller. 2011. Robust inference with multiway clustering. *J. Bus. Econ. Statist.* **29**(2) 238-249.

Cameron, A.C., P.K. Trivedi. 2005. *Microeconometrics: methods and applications*. Cambridge University Press, Cambridge, UK.

Casper, S. 2007. How do technology clusters emerge and become sustainable? Social network formation and inter-firm mobility within the San Diego biotechnology cluster. *Res. Pol.* **36**(4) 438-455.

Chi, T. 1994. Trading in strategic resources: Necessary conditions, transaction cost problems, and choice of exchange structure. *Strateg. Management J.* **15**(4) 271-290.

Chung, S.A., H. Singh, K. Lee. 2000. Complementarity, status similarity and social capital as drivers of alliance formation. *Strateg. Management J.* **21**(1) 1-22.

Collet, F., P. Hedström. 2013. Old friends and new acquaintances: Tie formation mechanisms in an interorganizational network generated by employee mobility. *Soc. Netw.* **35**(3) 288-299.

Corredoira, R.A., L. Rosenkopf. 2010. Should auld acquaintance be forgot? The reverse transfer of knowledge through mobility ties. *Strateg. Management J.* **31**(2) 159-181.

Crandall, D.J., L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, J. Kleinberg. 2010. Inferring social ties from geographic coincidences. *Proc. Natl. Acad. Sci.* **107**(52) 22436-22441.

Daft, R.L., R.H. Lengel. 1986. Organizational information requirements, media richness and structural design. *Manag. Sci.* **32**(5) 554-571.

Dokko, G., S.L. Wilk, N.P. Rothbard. 2009. Unpacking prior experience: How career history affects job performance. *Organ. Sci.* **20**(1) 51-68.

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

Felin, T., N.J. Foss, R.E. Ployhart. 2015. The microfoundations movement in strategy and organization theory. *Acad. Manag. Ann.* **9**(1) 575-632.

Felzensztein, C., E. Gimmon, S. Carter. 2010. Geographical co-location, social networks and inter-firm marketing co-operation: the case of the salmon industry. *Long Range Planning* **43**(5-6) 675-690.
Gangl, M. 2014. Matching estimators for treatment effects. H. Best, C. Wolf, eds. *The SAGE handbook of regression analysis and causal inference*. SAGE Publications, London, 251-276.
Geraci A, Fabbri D, Monfardini C (2018) Testing exogeneity of multinomial regressors in count data models: Does two-stage residual inclusion work? *Journal of Econometric Methods* 7(1).
Graham, N., M. Arai, B. Hagströmer. 2016. Multiwayvcov: Multi-way standard error clustering. R package.
Grigoriou, K., F.T. Rothaermel. 2017. Organizing for knowledge generation: Internal knowledge networks and the contingent effect of external knowledge sourcing. *Strateg. Management J.* **38**(2) 395-414.
Gulati, R. 1995a. Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances. *Acad. Management J.* **38**(1) 85-112.
Gulati, R. 1995b. Social structure and alliance formation patterns: A longitudinal analysis. *Admin. Sci. Quart.* **40**(4) 619-652.
Gulati, R. 1999. Network location and learning: The influence of network resources and firm capabilities on alliance formation. *Strateg. Management J.* **20**(5) 397-420.
Gulati, R., M. Gargiulo. 1999. Where do interorganizational networks come from? *Amer. J. Sociol.* **104**(5) 1439-1493.
Gulati, R., J.D. Westphal. 1999. Cooperative or controlling? The effects of CEO-board relations and the content of interlocks on the formation of joint ventures. *Admin. Sci. Quart.* **44**(3) 473-506.
Hagedoorn, J. 2002. Inter-firm R&D partnerships: An overview of major trends and patterns since 1960. *Res. Pol.* **31**(4) 477-492.
Hannan, M.T., J. Freeman. 1989. *Organizational ecology*. Harvard University Press, Cambridge, MA.
Hausman, J.A. 1978. Specification tests in econometrics. *Econometrica* **46**(1) 1251-1271.
Herstad, S.J., T. Sandven, B. Ebersberger. 2015. Recruitment, knowledge integration and modes of innovation. *Res. Pol.* **44**(1) 138-153.
Hess, A.M., F.T. Rothaermel. 2011. When are assets complementary? Star scientists, strategic alliances, and innovation in the pharmaceutical industry. *Strateg. Management J.* **32**(8) 895-909.
Hess, M. 2004. 'Spatial'relationships? Towards a reconceptualization of embeddedness. *Progress in Human Geography* **28**(2) 165-186.
Higgins, M.C., R. Gulati. 2006. Stacking the deck: The effects of top management backgrounds on investor decisions. *Strateg. Management J.* **27**(1) 1-25.
Higgins, M.J., D. Rodriguez. 2006. The outsourcing of R&D through acquisitions in the pharmaceutical industry. *J. Finan. Econ.* **80**(2) 351-383.
Hitt, M.A., D. Ahlstrom, M.T. Dacin, E. Levitas, L. Svobodina. 2004. The institutional effects on strategic alliance partner selection in transition economies: China vs. Russia. *Organ. Sci.* **15**(2) 173-185.
Hoang, H., F.T. Rothaermel. 2005. The effect of general and partner-specific alliance experience on joint R&D project performance. *Acad. Management J.* **48**(2) 332-345.
Hoetker, G. 2007. The use of logit and probit models in strategic management research: Critical issues. *Strateg. Management J.* **28**(4) 331-343.
Jaffe, A.B. 1989. Real effects of academic research. *Amer. Econ. Rev.* **79**(5) 957-970.
Jaffe, A.B., M. Trajtenberg, R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quart. J. Econ.* **108**(3) 577-598.
King, G., L. Zeng. 2001. Logistic regression in rare events data. *Political Analysis* **9**(2) 137-163.
Kogut, B., U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organ. Sci.* **3**(3) 383-397.
Leonard-Barton, D. 1984. Interpersonal communication patterns among Swedish and Boston-area entrepreneurs. *Res. Pol.* **13**(2) 101-114.

Li, G.-C., R. Lai, A. D'Amour, D.M. Doolin, Y. Sun, V.I. Torvik, Z.Y. Amy, L. Fleming. 2014. Disambiguation and co-authorship networks of the US patent inventor database (1975–2010). *Res. Pol.* **43**(6) 941-955.
Liebeskind, J.P., A.L. Oliver, L. Zucker, M. Brewer. 1996. Social networks, learning, and flexibility: Sourcing scientific knowledge in new biotechnology firms. *Organ. Sci.* **7**(4) 428-443.
Lumineau, F., N. Oliveira. 2018. A pluralistic perspective to overcome major blind spots in research on interorganizational relationships. *Acad. Manag. Ann.* **12**(1) 440-465.
Malecki, E.J., S.L. Bradbury. 1992. R&D facilities and professional labour: Labour force dynamics in high technology. *Reg. Stud.* **26**(2) 123-136.
Mawdsley, J.K., D. Somaya. 2016. Employee mobility and organizational outcomes an integrative conceptual framework and research agenda. *J. Manage.* **42**(1) 85-113.
Morgan, S.L., C. Winship. 2015. *Counterfactuals and causal inference*. Cambridge University Press, New York.
Mowery, D.C., J.E. Oxley, B.S. Silverman. 1996. Strategic alliances and interfirm knowledge transfer. *Strateg. Management J.* **17**(S2) 77-91.
Nakamura, A., M. Nakamura. 1998. Model specification and endogeneity. *J. Econometrics* **83**(1-2) 213-237.
Narula, R., G.D. Santangelo. 2009. Location, collocation and R&D alliances in the European ICT industry. *Res. Pol.* **38**(2) 393-403.
Neter, J., M.H. Kutner, C.J. Nachtsheim, W. Wasserman. 1996. *Applied linear statistical models*. McGraw-Hill, New York.
Newey, W.K. 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. *J. Econometrics* **36**(3) 231-250.
Nicholson, S., P.M. Danzon, J. McCullough. 2005. Biotech-pharmaceutical alliances as a signal of asset and firm quality. *J. Bus.* **78**(4) 1433-1464.
Owen-Smith, J., W.W. Powell. 2004. Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. *Organ. Sci.* **15**(1) 5-21.
Oxley, J.E. 1997. Appropriability hazards and governance in strategic alliances: A transaction cost approach. *J. Law, Econ., Organ.* **13**(2) 387-409.
Palomeras, N., E. Melero. 2010. Markets for inventors: Learning-by-hiring as a driver of mobility. *Manag. Sci.* **56**(5) 881-895.
Powell, W. 1990. Neither market nor hierarchy: Network forms of organization. B.M. Staw, L.L. Cummings, eds. *Research in Organizational Behavior*. JAI Press, Greenwich, CT, 295-336.
Powell, W.W., K.W. Koput, L. Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Admin. Sci. Quart.* **41**(1) 116-145.
Quigley, J.M., D.H. Weinberg. 1977. Intra-urban residential mobility: A review and synthesis. *Int. Reg. Sci. Rev.* **2**(1) 41-66.
Rangan, S. 2000. The problem of search and deliberation in economic action: When social networks really matter. *Acad. Management Rev.* **25**(4) 813-828.
Rappaport, J. 2007. Moving to nice weather. *Reg. Sci. Urban Econ.* **37**(3) 375-398.
Reuer, J.J., N. Lahiri. 2013. Searching for alliance partners: Effects of geographic distance on the formation of R&D collaborations. *Organ. Sci.* **25**(1) 283-298.
Rivers, D., Q.H. Vuong. 1988. Limited information estimators and exogeneity tests for simultaneous probit models. *J. Econometrics* **39**(3) 347-366.
Rogers, E.M., D.L. Kincaid. 1981. *Communication networks: Toward a new paradigm for research*. Free Press, New York.
Roijakkers, N., J. Hagedoorn. 2006. Inter-firm R&D partnering in pharmaceutical biotechnology since 1975: Trends, patterns, and networks. *Res. Pol.* **35**(3) 431-446.
Rosenkopf, L., P. Almeida. 2003. Overcoming local search through alliances and mobility. *Manag. Sci.* **49**(6) 751-766.

Rothaermel, F.T., W. Boeker. 2008. Old technology meets new technology: Complementarities, similarities, and alliance formation. *Strateg. Management J.* **29**(1) 47-77.
Salvato, C., J. Reuer, P. Battigalli. 2017. Cooperation across disciplines: A multilevel perspective on cooperative behavior in governing interfirm relations. *Acad. Manag. Ann.* **11**(2) 960-1004.
Sampson, R.C. 2007. R&D alliances and firm performance: The impact of technological diversity and alliance organization on innovation. *Acad. Management J.* **50**(2) 364-386.
Saxenian, A. 1990. Regional networks and the resurgence of Silicon Valley. *California Management Review* **33**(1) 89-112.
Saxenian, A. 1994. *Regional advantage: Culture and competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.
Schilling, M.A. 2009. Understanding the alliance data. *Strateg. Management J.* **30**(3) 233-260.
Schweizer, L. 2005. Organizational integration of acquired biotechnology companies into pharmaceutical companies: The need for a hybrid approach. *Acad. Management J.* **48**(6) 1051-1074.
Semadeni, M., M.C. Withers, T.S. Certo. 2014. The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strateg. Management J.* **35**(7) 1070-1079.
Shane, S., D. Cable. 2002. Network ties, reputation, and the financing of new ventures. *Manag. Sci.* **48**(3) 364-381.
Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Manag. Sci.* **51**(5) 756-770.
Singh, J., A. Agrawal. 2011. Recruiting for ideas: How firms exploit the prior inventions of new hires. *Manag. Sci.* **57**(1) 129-150.
Smith, R.J., R.W. Blundell. 1986. An exogeneity test for a simultaneous equation Tobit model with an application to labor supply. *Econometrica: Journal of the Econometric Society* 679-685.
Song, J., P. Almeida, G. Wu. 2003. Learning–by–hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Manag. Sci.* **49**(4) 351-365.
Stigler, G.J. 1961. The economics of information. *J. Polit. Economy* **69**(3) 213-225.
Stuart, T.E. 2000. Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strateg. Management J.* **21**(8) 791-811.
Stuart, T.E., H. Hoang, R.C. Hybels. 1999. Interorganizational endorsements and the performance of entrepreneurial ventures. *Admin. Sci. Quart.* **44**(2) 315-349.
Stuart, T.E., J.M. Podolny. 1996. Local search and the evolution of technological capabilities. *Strateg. Management J.* **17**(S1) 21-38.
Terza, J.V. 2018. Two-stage residual inclusion estimation in health services research and health economics. *Health Services Res.* **53**(3) 1890-1899.
Terza, J.V., A. Basu, P.J. Rathouz. 2008. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *J. Health Econ.* **27**(3) 531-543.
Tzabbar, D. 2009. When does scientist recruitment affect technological repositioning? *Acad. Management J.* **52**(5) 873-896.
Tzabbar, D., R.R. Kehoe. 2014. Can opportunity emerge from disarray? An examination of exploration and exploitation following star scientist turnover. *J. Manage.* **40**(2) 449-482.
Vanhaverbeke, W., G. Duysters, N. Noorderhaven. 2002. External technology sourcing through alliances or acquisitions: An analysis of the application-specific integrated circuits industry. *Organ. Sci.* **13**(6) 714-733.
Wagner, S., M. Goossen. 2018. Knowing me, knowing you: Inventor mobility and the formation of technology-oriented alliances. *Acad. Management J.* **61**(6) 2026–2052.
Wezel, F.C., G. Cattani, J.M. Pennings. 2006. Competitive implications of interfirm mobility. *Organ. Sci.* **17**(6) 691-709.
Wooldridge, J.M. 2014. Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *J. Econometrics* **182**(1) 226-234.

Table 1. Descriptive Statistics and Correlation Matrix (N = 234,786)

	Variables	Mean	Standard deviation	1.	2.	3.	4.	5.	6.	7.	8.
1.	R&D Alliance Formation	0.002	0.046								
2.	Scientist Mobility	0.004	0.063	0.050							
3.	Absence of Ties	0.985	0.117	-0.071	-0.142						
4.	Absence of Colocation	0.814	0.389	-0.007	-0.040	0.015					
5.	Technology Similarity	0.158	0.289	0.033	0.085	-0.122	-0.020				
6.	Alliance Experience	0.659	0.660	0.044	0.090	-0.213	-0.017	0.206			
7.	Patents	2.224	1.221	0.028	0.092	-0.122	-0.065	0.286	0.532		
8.	Age	2.252	0.419	0.018	0.054	-0.097	0.060	0.202	0.425	0.460	
9.	Residuals	-0.966	0.630	0.032	0.219	0.001	0.001	0.001	-0.001	0.001	-0.001

Note. Correlation coefficients with a magnitude larger than 0.005 are significant at $p < 0.05$ level.

Table 2. The First Stage Model of the Two-Stage Residual Inclusion (2SRI) Model: Logit Estimates for Scientist Mobility

Variables	Model 1
Absence of Ties	-1.354***
	(0.124)
	*-0.889****
Absence of Colocation	-1.600***
	(0.418)
	*-0.772****
Technology Similarity	1.687***
	(0.155)
	*0.663****
Alliance Experience	0.544***
	(0.120)
	*0.190****
Patents	0.697***
	(0.092)
	*0.260****
Age	0.492*
	(0.285)
	*0.178**
Weather Conditions	0.012***
	(0.003)
	*0.003****
Crime Rates	0.024***
	(0.007)
	*0.005****
Therapeutic Area Effects	53.447***
Patent Class Effects	1473.272***
Year Effects	245.189***
Firm State Effects	288.376***
Partner State Effects	306.832***
Intercept	-37.648***
	(6.465)
N	234,786
Log Likelihood	-4235.472
Chi-Squared	3695.588***

Notes. Clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.
Marginal effects, multiplied by 100 for easier interpretation, appear in italics.
Chi-Squared values for joint significance of fixed effects.

Table 3. The Second Stage Model of the Two-Stage Residual Inclusion (2SRI) Model: Logit Estimates for R&D Alliance Formation

Variables	Model 1	Model 2	Model 3	Model 4
H1: Scientist Mobility x Absence of Ties		1.743***		1.781***
		(0.398)		(0.400)
		*0.366****		*0.375****
H2: Scientist Mobility x Absence of Colocation			0.912**	0.946**
			(0.421)	(0.463)
			*0.185***	*0.198***
Scientist Mobility	0.864***	1.748***	1.015***	1.921***
	(0.237)	(0.252)	(0.252)	(0.269)
	*0.238****	*0.817****	*0.303****	*0.993****
Absence of Ties	-1.288***	-1.487***	-1.291***	-1.491***
	(0.145)	(0.140)	(0.141)	(0.139)
	*-0.406****	*-0.523****	*-0.405****	*-0.525****
Absence of Colocation	-0.271*	-0.265*	-0.344**	-0.334**
	(0.158)	(0.154)	(0.152)	(0.159)
	*-0.051**	*-0.049**	*-0.067***	*-0.067***
Technology Similarity	0.832***	0.836***	0.823***	0.830***
	(0.152)	(0.150)	(0.150)	(0.150)
	*0.148****	*0.149****	*0.146****	*0.147****
Alliance Experience	1.168***	1.149***	1.172***	1.152***
	(0.123)	(0.122)	(0.122)	(0.122)
	*0.226****	*0.241****	*0.243****	*0.242****
Patents	0.059	0.056	0.056	0.053
	(0.072)	(0.071)	(0.071)	(0.071)
	0.017	*0.017*	*0.017*	*0.018*
Age	-0.123	-0.104	-0.129	-0.111
	(0.236)	(0.235)	(0.235)	(0.235)
	-0.044	*-0.042*	*-0.044*	*-0.042*
Residuals	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
	*0.000****	*0.000****	*0.000****	*0.000****
Therapeutic Area Effects	2.832	2.417	2.996	2.388
Patent Class Effects	249.212***	247.047***	248.101***	247.540***
Year Effects	39.018***	38.239***	38.553***	38.858***
Firm State Effects	158.660***	158.283***	158.219***	158.202***
Partner State Effects	119.297***	121.040***	120.653***	120.808***
Intercept	-18.611***	-19.877***	-18.779***	-20.028***
	(1.359)	(1.652)	(1.652)	(1.108)
N	234,786	234,786	234,786	234,786
Log Likelihood	-3131.300	-3120.902	-3126.751	-3116.034
Chi-Squared	961.879***	982.676***	968.978***	988.412***

Notes. Clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.
Marginal effects, multiplied by 100 for easier interpretation, appear in italics.
Chi-Squared values for joint significance of fixed effects.

Figure 1. Marginal Effects (with 95% Confidence Intervals)
of Scientist Mobility in the Presence and Absence of Ties

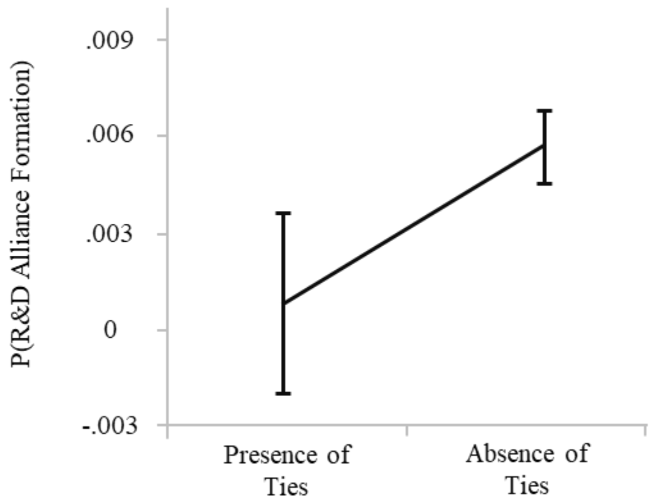
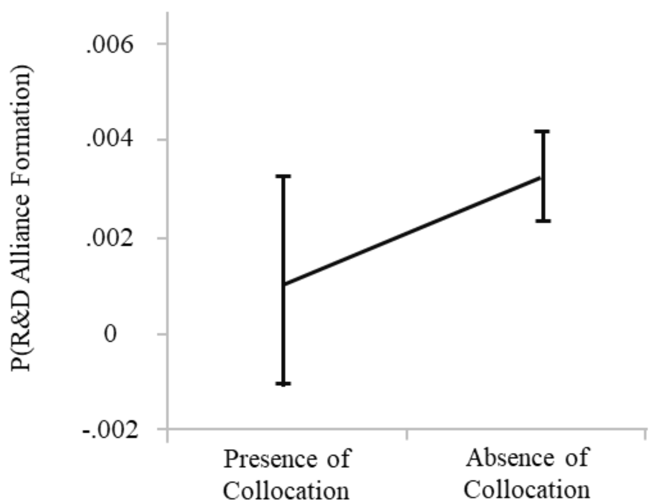
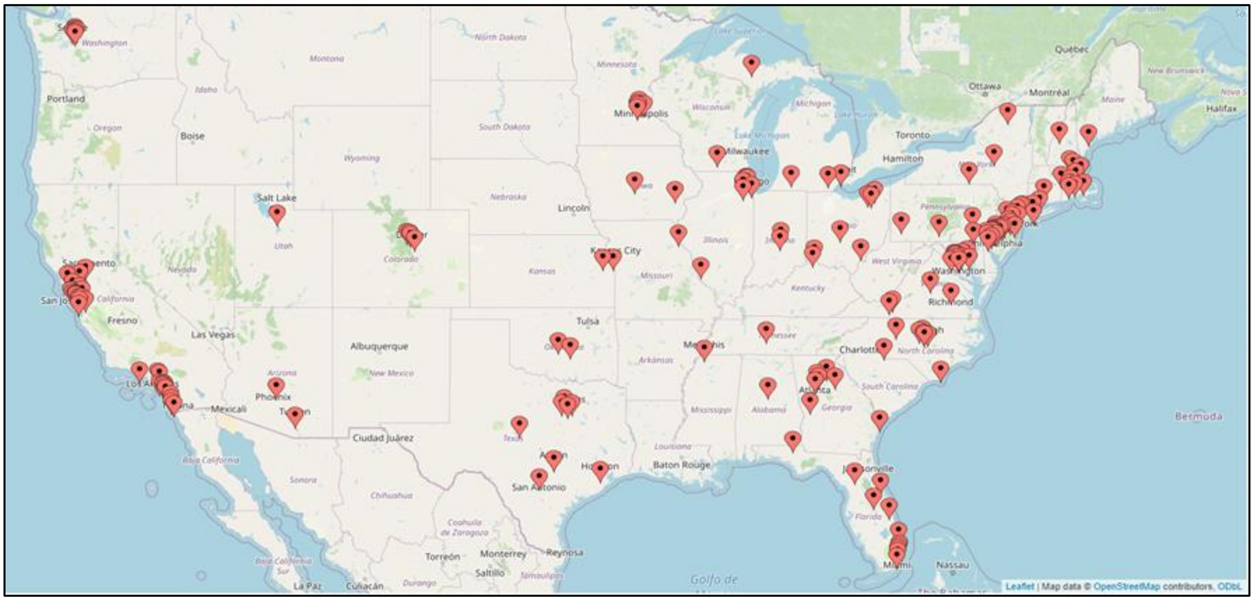


Figure 2. Marginal Effects (with 95% Confidence Intervals)
of Scientist Mobility in the Presence and Absence of Colocation



Appendix A. Geographic Distribution of Firms' Headquarters



Appendix B. Two-Stage Least Squares (2SLS) Estimation Results

Variables	OLS first Stage	2SLS (using ivreg2 in STATA)			
Scientist Mobility x Absence of Colocation				0.077^{**} (0.036)	0.071^{**} (0.035)
Scientist Mobility x Absence of Ties			0.086^{**} (0.039)		0.084^{**} (0.041)
Scientist Mobility		0.064^{**} (0.030)	0.074^{**} (0.039)	0.069^{**} (0.033)	0.083^{*} (0.049)
Weather Conditions	0.004^{***} (0.000)				
Crime Rates	-0.001^{***} (0.000)				
Absence of Ties	-0.061^{***} (0.015)	-0.018^{***} (0.004)	-0.017^{***} (0.004)	-0.019^{***} (0.004)	-0.016^{***} (0.004)
Absence of Colocation	-0.009^{***} (0.003)	-0.001^{***} (0.000)	-0.001^{***} (0.000)	-0.001^{*} (0.001)	-0.001^{*} (0.000)
Technology Similarity	0.008^{***} (0.002)	0.001^{***} (0.001)	0.002^{***} (0.000)	0.001^{***} (0.000)	0.001^{***} (0.000)
Alliance Experience	0.005^{**} (0.002)	0.003^{***} (0.001)	0.003^{***} (0.001)	0.003^{***} (0.001)	0.003^{***} (0.000)
Patents	0.003^{***} (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000) (0.000)
Age	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Therapeutic Area Effects	2.37^{*}	16.71^{*}	11.90	10.68	11.71
Patent Class Effects	6.89^{***}	17.27^{*}	17.30^{*}	17.00^{*}	17.32^{*}
Year Effects	24.37^{***}	43.59^{***}	42.81^{***}	42.92^{***}	42.89^{***}
Firm State Effects	50.86^{***}	102.00^{***}	146.38^{***}	106.71^{***}	165.50^{***}
Partner State Effects	130.60^{***}	127.74^{***}	226.14^{***}	240.06^{***}	238.26^{***}
Intercept	0.063^{***} (0.013)	0.019^{***} (0.003)	0.021^{***} (0.003)	0.020^{***} (0.003)	0.019^{***} (0.003)
N	234,786	234,786	234,786	234,786	234,786
R-squared	0.035	0.007	0.008	0.008	0.010

Notes. Clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.
Chi-Squared values for joint significance of fixed effects.

Appendix C1. Average Treatment Effects Analyses of Scientist Mobility

Panel (a)	Model 1		Model 2		Model 3	
	PSM	P > z	IPW	P > z	IPWRA	P > z
Scientist Mobility	0.089	(0.045)	0.091	(0.010)	0.094	(0.007)
SE	(0.044)		(0.035)		(0.035)	
95% CI	[0.002, 0.176]		[0.022, 0.161]		[0.025, 0.163]	

Appendix C2. Conditional Average Treatment Effects Analyses of Scientist Mobility

Panel (b)						
	Model 1		Model 2		Model 3	
	PSM	P > z	IPW	P > z	IPWRA	P > z
Absence of Ties	0.092	(0.010)	0.092	(0.010)	0.101	(0.005)
SE	(0.036)		(0.036)		(0.036)	
95% CI	[0.021, 0.163]		[0.021, 0.163]		[0.030, 0.173]	
Presence of Ties	0.004	(0.629)	0.025	(0.261)	0.028	(0.220)
SE	(0.009)		(0.022)		(0.023)	
95% CI	[-0.022, 0.013]		[-0.018, 0.068]		[-0.016, 0.073]	

Panel (c)						
	Model 1		Model 2		Model 3	
	PSM	P > z	IPW	P > z	IPWRA	P > z
Absence of Collocation	0.098	(0.015)	0.129	(0.020)	0.133	(0.008)
SE	(0.041)		(0.055)		(0.050)	
95% CI	[0.019, 0.178]		[0.020, 0.238]		[0.034, 0.232]	
Presence of Collocation	0.005	(0.339)	0.011	(0.246)	0.021	(0.422)
SE	(0.006)		(0.009)		(0.026)	
95% CI	[-0.006, 0.017]		[-0.008, 0.031]		[-0.030, 0.073]	