MANAGEMENT SCIENCE
Vol. 57, No. 1, January 2011, pp. 129–150
ISSN 0025-1909 | EISSN 1526-5501 | 11 | 5701 | 0129

DOI 10.1287/mnsc.1100.1253

Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires

Jasjit Singh
INSEAD, Singapore 138676, jasjit.singh@insead.edu

Ajay Agrawal
Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada; and
National Bureau of Economic Research, Cambridge, Massachusetts 02138,
ajay.agrawal@rotman.utoronto.ca

When firms recruit inventors, they acquire not only the use of their skills but also enhanced access to their stock of ideas. But do hiring firms actually increase their use of new recruits' prior inventions? Our estimates suggest they do, quite significantly in fact, by approximately 219% on average. However, this does not necessarily reflect widespread "learning by hiring." In fact, we estimate that a recruit's exploitation of her own prior ideas accounts for almost half of the above effect, with much of the diffusion to others being limited to the recruit's immediate collaborative network. Furthermore, although one might expect the recruit's role to diminish rapidly as her tacit knowledge diffuses across her new firm, our estimates indicate that her importance is surprisingly persistent over time. We base these findings on an empirical strategy that exploits the variation over time in hiring firms' citations to the recruits' premove patents. Specifically, we employ a difference-in-differences approach to compare premove versus postmove citation rates for the recruits' prior patents and corresponding matched-pair control patents. Our methodology has three benefits compared to previous studies that also examine the link between labor mobility and knowledge flow: (1) it does not suffer from the upward bias inherent in the conventional cross-sectional comparison, (2) it generates results that are robust to a more stringently matched control sample, and (3) it enables a temporal examination of knowledge flow patterns.

Key words: inventor mobility; access to ideas; knowledge spillovers; learning by hiring; difference in differences; coarsened exact matching; collaborative networks; patent citations
History: Received August 28, 2009; accepted August 4, 2010, by Kamalini Ramdas, entrepreneurship and innovation. Published online in *Articles in Advance* November 19, 2010.

1. Introduction

The link between recruiting inventors and using their stock of prior ideas is important, yet we know surprisingly little about it. There is a broad consensus that innovation is central to many firms' strategies and the basis for their competitive advantage. Yet a firm's past experience and extant stock of knowledge constrains innovation, making the innovation trajectory highly path dependent (Nelson and Winter 1982, Dosi 1988). Firms need to balance this natural tendency toward exploitation of familiar knowledge with deliberate mechanisms that facilitate exploration of distant knowledge (March 1991), especially because combining ideas drawn from different sources is often important for innovation success (Weitzman 1998, Fleming 2001, Chesbrough 2003, Singh and Fleming 2010). Recruiting an individual from outside the organization may enhance a firm's access to external ideas and thus better enable it to complement the exploitation of native ideas with the exploration of foreign ideas. But how much do firms really increase their use of a new recruit's stock of prior ideas? To the extent they do, how do they do it? Furthermore, to what extent do firms really exhibit "learning by hiring"? Finally, given the temporal nature of diffusion, how does this process evolve over time? We set out to address these questions.[1]

Scholars from a variety of schools of thought have suggested that interfirm mobility could be a key mechanism driving diffusion of ideas across firms. Noting that knowledge spillovers through mobility can take place despite the source firm's use of legal measures to prevent them, economist Arrow (1962, p. 615) remarks: "Mobility of personnel among firms provides a way of spreading information. Legally imposed property rights can provide only a partial barrier, since there are obviously enormous difficulties in defining in any sharp way an item of information and differentiating it from other similar sounding items." Proponents of institutional theory,

[1] Our focus is on an individual moving from one firm to another *existing* firm. A related literature emphasizes knowledge transfer through mobility in the context of new firm formation (Agarwal et al. 2004, Gompers et al. 2005, Klepper and Sleeper 2005).

such as DiMaggio and Powell (1983), suggest that interfirm movement of personnel is a particularly important mechanism through which innovations diffuse among competitors in an industry. Analogously, the resource-based view of the firm acknowledges that recruitment from outside can enable firms to bypass constraints on growth imposed by relying solely on internally grown resources and capabilities (Penrose 1959, Barney 1991).

At an aggregate level, scholars have also linked regional economic growth to enhanced access to ideas afforded by interfirm mobility. In her monograph on regional advantage, Saxenian (1994, pp. 34–37) characterizes the relationship between recruiting and access to ideas as central to explaining the exceptional economic growth of northern California: "Silicon Valley was quickly distinguished by unusually high levels of job hopping. During the 1970s, average annual employee turnover exceeded 35 percent in local electronics firms and was as high as 59 percent in small firms.... Early efforts to take legal action against departed employees proved inconclusive or protracted, and most firms came to accept high turnover as a cost of business in the region.... This decentralized and fluid environment accelerated the diffusion of technological capabilities and know-how within the region." Several studies that examine the effects of restrictions on interfirm mobility due to noncompete covenants find support for this view (Franco and Mitchell 2008, Marx et al. 2009, Samila and Sorenson 2011).

While the above research *assumes* a link between mobility and knowledge flow, several studies explicitly estimate aspects of this relationship, particularly in the context of mobile inventors. In one of the first such studies, Almeida and Kogut (1999) show that locations with greater intraregional labor mobility between firms tend to have more localized knowledge flows. In another study, Song et al. (2003) illustrate that mobile inventors build upon ideas from their previous firm more often than do other inventors at the hiring firm. In yet another influential article, Rosenkopf and Almeida (2003) examine firm pairs and show that dyads that experience more labor mobility between them also demonstrate greater subsequent knowledge flow. These pioneering studies, all based on a methodology employing patent data for examining the mobility–knowledge flow relationship, have inspired a large field of research using similar data and methods to sharpen our understanding of various aspects of this relationship (e.g., Agrawal et al. 2006, Singh 2007, Oettl and Agrawal 2008, Agarwal et al. 2009, Corredoira and Rosenkopf 2010).

Although the resulting literature has advanced the field significantly, it remains subject to two related limitations associated with drawing causal inferences from cross-sectional data on inventor mobility and knowledge use. The first limitation concerns unobserved heterogeneity. The destination firm is more likely to use an idea of inherently higher quality or greater firm-specific relevance, irrespective of whether it hires the inventor. However, inventors of such ideas might also have a different likelihood of being recruited (Hoisl 2007), thus presenting a selection problem wherein an observed cross-sectional correlation between mobility and knowledge use does not reflect the true effect of mobility. The second concerns the endogeneity of a firm's decision to hire. For example, a firm may be more inclined to recruit an individual who works in a domain that the firm intends to focus on in the future. This could coincide with the firm also employing other mechanisms to improve its access to external knowledge in that domain (including the recruit's prior stock of ideas). This would again produce a mobility–knowledge flow correlation without the former being a (fully) causal precursor to the latter. Not accounting for such endogeneity will lead to an upward bias if we interpret the estimates as boosts in the use of ideas caused by an instance of mobility.[2]

Recognizing these inherent limitations of a cross-sectional research design for making causal inferences, Rosenkopf and Almeida (2003, p. 764) offer this challenge: "Future research should attempt to utilize fully developed longitudinal databases to explore all possible temporal and causal links." We take up that challenge here. Therefore, in addition to offering a conceptual contribution, this study also advances the methodology employed in prior empirical research on mobility and knowledge flow in three significant ways. First, rather than basing our analysis on aggregate citation counts, we demonstrate how the use of disaggregated longitudinal citation data generates sharper insights into how different individuals in a recruiting firm use a specific piece of knowledge associated with a mobile inventor. Second, we employ a "difference-in-differences" (DD) approach to account for heterogeneity across patents. This avoids making the typical (strong) assumption regarding cross-sectional comparability of the *levels* of citations received by "focal" patents (i.e., those involving an inventor who subsequently moves) and

[2] To be clear, prior literature does attempt to deal with the aforementioned issues by using technologically matched control patents as a benchmark for cross-sectional comparison. However, given necessarily imperfect matching, such challenges are unavoidable in any cross-sectional research design. The concerns become particularly salient in the typical mobility-related study that employs the relatively aggregate three-digit technology match, because that is likely too coarse to sufficiently capture all relevant characteristics of the underlying knowledge.

"control" patents (i.e., other similar patents). Instead, our identification strategy only relies upon comparing *changes* in citation rates over time. Third, to further address concerns regarding comparability of the focal and control patents, we replicate our initial analysis, which is based on a conventional matching approach, with a more stringent matching procedure.

We begin our analysis with a sample based on the matching criteria most prior studies employ: the three-digit technology classification and the application year.[3] This method, predicated on the comparison of postmove citation levels for focal versus control patents to estimate the effect of mobility, rests on the assumption that a three-digit technology match suffices for addressing heterogeneity across patents. However, we find evidence that this assumption is not reasonable; destination firms cite focal patents at a higher rate than control patents even *before* the move. In recognition of this systematic difference between focal and control patents, we use the DD approach to "difference out" the premove citation trend associated with focal patents. This allows us to distinguish the component of the postmove citation rate that is more likely due to the "treatment" (attributable to the move) from the component that is due to "selection" (attributable to the kind of inventor who is more likely to move). Our DD analysis reveals that hiring an inventor is associated with a firm increasing its use of the new recruit's prior ideas significantly: by approximately 219% on average.

We repeat our DD estimation using an alternative sample based on a more stringent matching procedure employing "coarsened exact matching" (CEM) (Iacus et al. 2009). The additional matching criteria we employ are based on information related to the patent's premove citations, the inventor's career history, and the lag between the patent's application and grant dates. As the extensive literature on matching emphasizes, using appropriately stringent matching reduces endogeneity concerns as well as the sensitivity of the subsequent regression-based estimation on specific functional form assumptions. In our case, eliminating the premove citation rate differences between focal and control patents particularly improves their comparability in terms of expected future citation rates. We find the DD estimates based on this more stringent matching to be very similar to those found using the previous analysis.

Having estimated the extent to which a destination firm uses its new recruit's prior ideas, we next turn our attention to examining the mechanism through which the firm accomplishes this. The extant literature emphasizes the idea of "learning by hiring" (Song et al. 2003, Rosenkopf and Almeida 2003), which effectively assumes that the recruit's tacit knowledge diffuses internally and becomes part of the firm's overall knowledge base shortly after the recruit's arrival. However, Simon (1991, p. 126) cautions against such assumptions: "We must be careful about reifying the organization and talking about it as 'knowing' something or 'learning' something. It is usually important to specify where in the organization particular knowledge is stored, or who has learned it.... Since what has been learned is stored in individual heads, its transience or permanence depends on what people leave behind them when they depart from an organization or move from one position to another. Has what they have learned been transmitted to others or stored in ways that will permit it to be recovered when relevant?"

The above remark motivates our next set of analyses, where we distinguish between self-exploitation of the recruit's prior ideas and increased usage by others in the firm. We discover that almost half of the boost in the use of the recruit's prior ideas is due to the recruit herself building upon her own prior ideas after arriving at her new firm.[4] To the extent that the recruit's tacit knowledge does diffuse through the destination firm, we are also interested in estimating how widely this occurs. On the one hand, knowledge may diffuse narrowly, through close interpersonal ties such as those formed by collaboration (Hansen 1999, Singh 2005, Singh et al. 2010). On the other hand, the presence of alternate intrafirm diffusion mechanisms might reduce the reliance on such ties, enabling a wider employee base to build upon the recruit's prior ideas. When we exclude idea use by the recruit as well as by the collaborative network she forms in her new firm, we find that other inventors account for only about a third of the overall boost realized by the hiring firm. Next, when we focus specifically on the temporal pattern of citations, expecting that the role of the recruit might diminish shortly after she arrives as her tacit knowledge diffuses throughout the destination firm, we instead find that her role persists over time to a surprising degree.

These findings suggest the need to temper the prevalent learning-by-hiring view with recognition that, at least in our sample, we must attribute almost half of the knowledge boost to what can better be described as "exploiting by hiring." This has important implications for how we should interpret the mobility–knowledge flow link, given that reliance on

[3] Jaffe et al. (1993) pioneer this matching procedure. Thompson and Fox-Kean (2005) propose refinements to this technology-based matching while also discussing the inherent challenges of using such a matching approach.

[4] Our results are consistent with the findings of Tzabbar et al. (2009), who report that having the recruit on an inventing team significantly increases the likelihood of the team exploiting one of the recruit's prior inventions.

a single employee for an important piece of knowledge may confer significant bargaining power to that individual (Becker 1962, Lazear 1986, Coff 1997, Moen 2005, Groysberg et al. 2008). This echoes Peteraf's (1993, p. 187) remark: "For example, a brilliant, Nobel Prize–winning scientist may be a unique resource, but unless he has firm-specific ties, his perfect mobility makes him an unlikely source of sustainable advantage." In addition, to the extent that the individual also chooses her collaborators (who seem to enjoy preferential access to her ideas), she might further influence the knowledge diffusion process in a way that is suboptimal for the firm.

We organize the remainder of the paper as follows. In §2, we outline our empirical framework and distinguish it from the traditional cross-sectional approach. In §3, we describe the construction of our two data sets, one corresponding to the conventional matching procedure and the other based on more stringent matching criteria. In §4, we present our empirical results. Finally, in §5, we discuss the implications of our findings, limitations of our study, and potential directions for future research.

2. Empirical Framework

2.1. Patent Data, Mobility, and the Use of Ideas

Microlevel data suitable for examining the link between mobility and knowledge flows are hard to come by. A notable exception is patent data, which researchers have commonly employed for this purpose. These data include detailed information on each patent, including inventor names, application and grant dates, assignee organization (if any), technological classification, etc. This information is particularly useful on two dimensions. First, it enables the researcher to infer interfirm mobility by chronologically tracing individual inventors as they appear on patents assigned to different firms at different times. Second, tracking citations that a patent receives allows inference regarding subsequent use of a particular idea. Admittedly, both of these patent-based measures are far from perfect, a point we elaborate on in the discussion section.

2.2. A Cross-Sectional Comparison of Citation Rates

We begin by implementing a method of examining the link between mobility and knowledge flows that follows "best practice" from existing studies. To do so, we first create a sample of patents representing ideas developed by inventors who subsequently change firms. We refer to these as "focal patents." Second, we match each focal patent with a comparable "control patent" with the same three-digit technology class and application year. Finally, we examine cross-sectional differences between the focal and control patents in terms of the number of citations they receive from the destination firm (the firm that recruits the focal inventor) in the period following the move.

Formally, we define $CITES_{i,t}$ as the number of citations a patent i (focal or control) receives from the destination firm in year t (any year following the move). The cross-sectional estimation equation we employ (using observations just from the postmove period) is

$$CITES_{i,t} = f\big(\psi_R RECRUIT_i + \psi_X X_i + \delta_{t-app\,year(i)} + \beta_t + \varepsilon_{i,t}\big). \quad (1)$$

Here, $RECRUIT_i$ is an indicator variable that equals 1 for a focal patent and 0 for the corresponding control patent. Rather than assuming a specific functional form for the temporal pattern of citations, the above model follows a nonparametric approach in accounting for patent age using yearly indicator variables for the gap between patent i's application year and the citing year t being considered, and citing year using different indicator variables for each calendar year t.[5]

The baseline expectation is that ψ_R should be positive and significant. In other words, the destination firm should cite prior patents of recruits more than the corresponding control patents. The vector X_i represents the variables we use to control for variation in key observables associated with the destination firm, the source firm, the inventor, and the patent.

Although commonly employed, such a cross-sectional approach leaves open a concern regarding unobserved heterogeneity across inventors and their patents. For example, better inventors may generate higher-quality ideas (that naturally receive more citations from others, including from the destination firm) and might also be more aggressively recruited. A focal patent is also more likely to come from a specific knowledge domain of greater relevance for the destination firm, despite the corresponding control being drawn from the same broad three-digit technology class. In either case, the recruit's patent would receive more citations from the destination firm, irrespective of the move. To better identify the component of the boost in knowledge use that is directly attributable to the move, we next turn to a difference-in-differences

[5] Our goal is to appropriately control for patent age and citing year effects without necessarily identifying them separately. Given that perfect collinearity would result if patent age and citing year effects are included as full sets of indicator variables, we omit one of the indicator variables. All our findings are robust to instead following the Rysman and Simcoe (2008) and Mehta et al. (2009) approach of using indicator variables only for the citing year and accounting for patent age by including the squared, cubic, and fourth terms of patent age.

research design that allows us to distinguish between the component of the boost in postmove citation rate that is more likely due to "treatment" (i.e., the move) versus "selection" (i.e., differences in the kind of inventors who move versus those who do not).

2.3. A Difference-in-Differences Approach for Examining the Effect of Mobility

A DD approach for examining the mobility–knowledge flow link exploits the fact that we observe citations received by the focal and control patents not just postmove but also in the years preceding the move. Although postmove differences confound mobility-related knowledge acquisition with differences in quality and/or relevance, we can disentangle these effects by taking into account differences in citation rates that exist before the move. In effect, the premove difference in citation rates for focal versus control patents can serve as a benchmark against which to examine the postmove difference to identify the component of the latter difference that is attributable to the move itself.[6]

Implementing this logic needs a patent-year data set that includes observations from the years not just after but also before the move. Defining a new indicator variable, $POSTMOVE_{i,t}$, as equal to 1 for observations that involve a patent i from an inventor who actually moves and are from a citing year t that falls in the postmove period, the estimation equation is

$$CITES_{i,t} = f\big(\psi_R\, RECRUIT_i + \psi_{RP}\, POSTMOVE_{i,t} + \psi_P\, PAIR\_POSTMOVE_{i,t} + \psi_X X_i + \delta_{t-app\,year(i)} + \beta_t + \varepsilon_{i,t}\big). \quad (2)$$

Here, ψ_R captures systematic differences in focal versus control patents that exist even before the move. The additional indicator variable, *PAIR_POSTMOVE*, is defined as equal to 1 for focal and also control patent observations that occur in the (focal) postmove period for a matched pair. As a result, ψ_P estimates the "counterfactual" change in the citation rate at the time of the move had the move not actually taken place. Although the above equation can be estimated using a pooled model, the more natural approach is to use matched-pair fixed effects. The coefficient of real interest is ψ_{RP}, which will be positive and significant if mobility really does lead to the destination firm's increased use of the recruit's prior ideas.

The above estimation relies upon an assumption of strict one-to-one comparability of focal and control patents. Such an assumption is necessary when carrying out univariate DD analysis (of the kind reported later in Tables 4 and 5). However, because control patents do not actually involve a move, matching does not ensure perfect comparability. Therefore, our preferred approach is not to include a separate *PAIR_POSTMOVE* variable, but rather we use the control patents only to help account for effects associated with a patent's application year and age (captured using a series of indicator variables).

Also, although the above framework allows for a systematic difference between focal and control patents using the *RECRUIT* variable, we can generalize it to allow individual patents to be different in unobserved ways (e.g., because inventor characteristics we do not observe). In particular, we can employ patent fixed effects analysis to detect "abnormal" within-patent changes in the citation rate to a patent after an inventor moves. The estimation would therefore rely only on a deviation in the patent's post-mobility citation rate from its own estimated expected rate, where the expected rate is derived by extrapolating from the premove citation rate and assuming a temporal trend analogous to other patents with similar technological and temporal characteristics. We express the resulting model as

$$CITES_{i,t} = f\big(\psi_{RP}\, POSTMOVE_{i,t} + \gamma_i + \delta_{t-app\,year(i)} + \beta_t + \varepsilon_{i,t}\big). \quad (3)$$

Here, γ_i reflects the fixed (time invariant) effect corresponding to patent i. Because the fixed effect absorbs unique patent characteristics, the model no longer includes direct effects of time-invariant, patent-level variables ($RECRUIT_i$ and X_i in the previous models). However, we can still estimate the DD coefficient ψ_{RP}.[7]

2.4. A More Stringent Matching Procedure for Sample Construction

The above DD research design essentially replaces the conventional (strong) assumption of comparability of postmove citation levels of focal versus control patents with a weaker assumption regarding comparability in terms of changes in the citation rate over time. However, one might still worry whether the conventional matching criteria (three-digit technology class and application year) really produce a

[6] In the online appendix provided in the e-companion to this paper (which is available as part of the online version that can be found at http://mansci.journal.informs.org/), Figure A1 and related notes provide further intuition of how the DD approach conceptually differs from the cross-sectional approach described above. For a more general overview of the DD methodology, see Angrist and Pischke (2009, Chap. 5). See Murray and Stern (2007) and Furman and Stern (2011) for instructive applications in the context of citations.

[7] See Murray and Stern (2007) for a similar specification. Because the lag between the application and the move varies even across focal patents, we can estimate ψ_{RP} using just the focal patent subsample as well. Although we do report findings from such analysis, particularly to ensure that our findings are not too sensitive to the specific control patents included, our preferred approach is to include appropriately matched control patents in the sample.

control sample comparable to the focal patents even under these weaker assumptions. As described in detail in §3.2, we construct an alternate set of control patents based on a more stringent match and use that sample in the subsequent empirical analysis.

A more stringent matching procedure offers two potential benefits (Imbens 2004). First, to the extent that the likelihood of "treatment" (in our case, the "mobility event") correlates with the additional matching criteria employed, it reduces concerns about endogeneity-related biases. Second, more stringent matching reduces sensitivity of the findings to specific functional form assumptions (Moffitt 2004).

Given that a key concern above is comparability of patents in terms of citation rates, our more stringent matching procedure explicitly includes among the additional matching criteria premove citations received by focal and control patents, both overall and from the destination firm in particular. We also match on the patenting history of the inventor in terms of the number of past patents and years since the first patent. Finally, we also use the grant delay (i.e., the lag between the application and grant date) as a matching criterion to allow for the possibility that the grant delay may itself be a result of important patent characteristics or that the start of a patent's "citation clock" is more appropriately modeled using its grant date rather than application date (Mehta et al. 2009).

The consistency of matching-based estimates normally relies upon a "selection on observables" assumption. Because the likelihood of "treatment" might also depend on unobservables or on observables that are impractical to fully match on, even seemingly stringent matching only reduces endogeneity concerns, rather than eliminating them (Heckman and Navarro-Lozano 2004). Therefore, given the weaker underlying assumptions discussed above, we continue to employ a DD approach even when using the more stringently matched sample.

2.5. Our Preferred Regression Model

Because patent citations involve count data skewed to the right (and overdispersed relative to Poisson), scholars commonly employ negative binomial models for estimating parameters. For comparability with previous research, we therefore start with negative binomial regressions that illustrate some of our key points using models that are commonly employed in related literature. We employ robust standard errors, with clustering on the inventor to account for nonindependence of observations pertaining to the same cited patent and to different patents involving the same inventor. We then implement analogous regressions using the corresponding linear (ordinary least squares) models, which have the benefit of allowing more fine-grained indicator variables for technology (at the three-digit level rather than two-digit level practical for nonlinear models), inventor location (U.S. state or non-U.S. country), inventor age (in years), and grant delay (in months).

For our preferred implementation of DD, however, we employ patent fixed effects to better address concerns regarding unobserved heterogeneity across patents (e.g., differences in the intrinsic quality or relevance to the destination firm that the pooled models do not capture). Our preferred approach is to employ linear regression specifications with patent fixed effects, though the results reported in this paper are robust to employing a conditional fixed effects Poisson framework instead (Wooldridge 1999).[8] The reason we prefer linear models here is that, in addition to the usual challenges related to including too many fine-grained indicator variables or interpreting DD coefficients (especially when comparing across different models), such nonlinear conditional fixed effects models also drop a large fraction of the sample that has no within-patent variation in the dependent variable (because of zero citations being so prevalent, especially among observations on control patents).[9]

3. Data Set Construction

3.1. Constructing the "Original Sample" of Focal and Control Patents

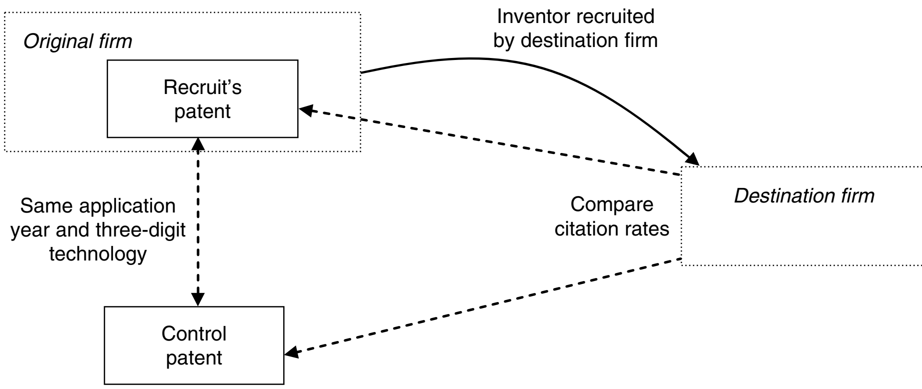
We merge patent data obtained directly from the United States Patent and Trademark Office with the National Bureau of Economic Research (NBER) patent data set (Jaffe and Trajtenberg 2002) and the National University of Singapore/Melbourne Business School patent, data set. We enhance these along two dimensions. First, for each assigned patent, we determine the assignee organization by carrying out an assignee name cleanup followed by a parent-subsidiary match.[10] Second, we use not just inventor names but also other data fields (i.e., technology classification, inventor address, collaborator names,

[8] We implement this in Stata using the "xtqmlp" procedure (written by Tim Simcoe and available for download at http://people.bu.edu/tsimcoe/code/xtpqml.txt), which corrects the standard errors from a fixed effects Poisson model for overdispersion (Rysman and Simcoe 2008). This addresses concerns regarding interpreting a conditional fixed effects negative binomial model as a true fixed effects estimator (Wooldridge 1999, Allison and Waterman 2002).

[9] For an excellent discussion on the trade-offs involved in choosing between linear and nonlinear models in such settings, see Angrist and Pischke (2009).

[10] We build upon the assignee matching procedure used by Singh (2005, 2007), who relies upon NBER Compustat identifiers, different corporate ownership directories, and Internet sources. To further reduce any possibility of misclassifying spelling differences, name changes, or acquisitions as instances of mobility, we manually double-checked instances of algorithmically detected mobility.

Figure 1 Sample Construction



Notes. We begin by identifying a set of focal patents ("Recruit's patents"), each created by a single inventor who subsequently moves to another firm between the 3rd and 10th year after the focal patent's application date. (Excluding the first two years ensures a premove observation window, data upon which the DD approach relies.) In constructing the original sample, we match each focal patent with a "control" patent (from a different firm and by a single inventor who shows no evidence of moving in the first 12 years) such that the control patent has the same application year and three-digit primary technology class. We follow a similar procedure in constructing the CEM sample except that we use additional observables (including premove citation and inventor information) to carry out a more stringent match.

citation information) to create a unique identifier for each inventor on all patents.[11]

To obtain a data set of reasonable size while allowing a sufficient time window for observing subsequent inventor mobility and citations, we start with patents with application years 1981–1990 (across all technology classes). In detecting instances of subsequent interfirm mobility, we follow the prior literature and infer mobility through observed changes in the assignee firm on successive patents filed by individual inventors (Almeida and Kogut 1999, Song et al. 2003).[12]

We restrict our sample (both focal and control) to patents with a single inventor to facilitate unambiguous conceptualization of interfirm mobility (or lack thereof) for any given patent. From this subset, we draw our sample of focal patents for which the inventor exhibits interfirm mobility any time between (and including) the 3rd and 10th years following the application year (we exclude the first and last two years so that there is at least some pre- and postmove citation data for each patent to facilitate meaningful DD estimation when we employ cited patent fixed effects).

One challenge is that even when we observe two successive patents from the same inventor but at different firms, we cannot pinpoint the inventor's exact move date within this window. We thus base all analyses reported in this paper on mobility events for which this window is four years or less, because a detailed temporal examination would not be as useful for cases where the move date is too uncertain. Therefore, we drop about 30% of all observed interfirm mobility events.[13]

Because we do not have exact information about the move date, we start by calculating the halfway point between the last observed date at the source firm and the first observed date at the destination firm. However, although the move could have taken place any time after the start of this window, it would probably have taken place at least a few months before the end of the window because of a lag between an inventor joining a new firm and filing a patent there. Furthermore, we take the temporal unit of analysis to be the year, rather than the day or month, to avoid any pretence of a precisely estimated move date. These two factors lead us to define the beginning of the calendar year of the window midpoint calculated above as our "move date" estimate. Given the uncertainty, this somewhat ad hoc estimate is unlikely to do much worse than more sophisticated heuristics. Moreover, it facilitates analysis by allowing classification of each calendar year as being either completely before or after a given move, allowing us to work with a patent-year panel.

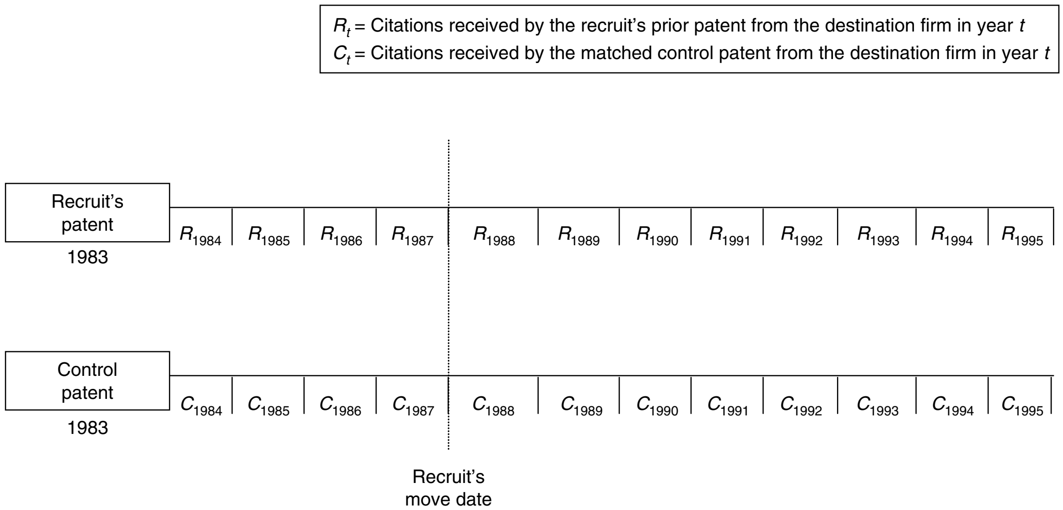
As the next step, we follow previous research in matching each focal patent with a corresponding control patent (Figure 1). We choose the control patent

[11] We base our name matching approach on Singh (2008), whose algorithms are similar to procedures implemented by Trajtenberg (2006) and Fleming et al. (2007).

[12] Patent data are only effective for identifying instances of mobility where an inventor successfully files for patents both before and after a move. In the discussion section, we elaborate on the potential concerns arising from this.

[13] Of the remaining, the uncertainty is zero to two years for 65% of the cases and three to four years for the rest. Thus, an estimate of the move year based on the midpoint is off by not more than one year in two-thirds of the cases and not more than two years for the rest. One might worry about representativeness of the final sample. For example, because longer time windows imply fewer patents per year, the dropped observations could pertain to less productive inventors. To rule out the possibility of any resulting biases, we redo the analyses reported in this paper using different window cutoffs. The main results remain qualitatively unchanged.

Figure 2 A Longitudinal Data Set of Citations Received from the Destination Firm



Notes. We use each of the two matched samples described in the notes to Figure 1 to construct a longitudinal patent-year data set of citations. Each of the two resulting data sets has 12 yearly observations for each cited patent, corresponding to the 12 years directly following the patent's application year. In the above example, the focal and corresponding control patent originate in year 1983, so each of them gives rise to one observation per year for the period 1984–1995. Taking the move date as January 1, 1988, we classify the observations corresponding to 1984–1987 as premove and the observations corresponding to 1988–1995 as postmove. Following this procedure, we construct the original sample with 95,424 observations (3,976 × 12 × 2) corresponding to the 3,976 focal patents and as many control patents. In contrast, the CEM sample has 61,032 observations (2,543 × 12 × 2), because we drop the focal patents that cannot find a more stringent match.

such that it originates in another firm, it also has a single inventor, and its inventor does not exhibit any interfirm mobility in the 12 years that follow.[14] We use control patents to account for general shifts in the technological focus of the destination firm because such shifts will be reflected in an increased likelihood of citing not just the focal patent but also the control patent. In the relatively infrequent cases where we cannot match the focal patent with a control patent on the above criteria, we drop the patent from the sample. These steps lead to a final sample of 7,852 patents, exactly half (3,976) of which are associated with an inventor who subsequently moves (focal patents), and the other half are the corresponding control patents. For each of these, we count the annual number of citations made by the destination firm in the 12 years following the application year (Figure 2). Having 12 observations for each patent results in a patent-year panel of 95,424 observations (the "original" sample).

Table 1 summarizes the key variables.[15] Our first dependent variable, *all cites*, includes *all* citations the recruit's patent receives from the hiring firm in the focal year. We define the next dependent variable, *cites excluding inventor self-cites*, analogously to *all cites* but exclude self-citations made by inventors to their own patents. By eliminating an inventor's citations, this variable measures the use of an idea by individuals at the destination firm other than the inventor herself, hence excluding self-exploitation by the recruit. To examine the breadth of intrafirm knowledge diffusion, we construct another dependent variable, *cites excluding inventor self-cites and collaborator cites*, as the count of citations made by those in the destination firm that have not directly collaborated (as coauthors on previous patents) with the inventor before the citing year.[16]

Our two key explanatory variables are *recruit* and *postmove*. We use the indicator *recruit* to identify inventions made by individuals who are subsequently recruited and move to a new firm, the so-called destination firm. In other words, we use recruit to distinguish between our "treated" versus "untreated," or control, patents. We classify a given patent-year observation for a focal patent as occurring after the estimated move date using the indicator variable *postmove*. In initial models, we also employ a separate indicator variable *pair postmove* that is one

[14] Same-firm patents are inappropriate as controls because the mobile inventor is likely to carry knowledge about not just her own patent but also about other same-technology patents from the same firm. So using those as controls would systematically underestimate the benefits from mobility.

[15] Note that the average citation rates reported are small because the probability of a *specific firm* citing a *specific patent* in a*specific year* is low. For example, *all cites* has a nonzero value for only around 1.8% of the observations in the original sample.

[16] If a cited patent's inventor and some of her former collaborators develop a new patent that cites her previous patent, we classify that citation as an inventor self-cite rather than a collaborator cite. We classify a citation as a collaborator cite only when the cited inventor herself is not involved.

Table 1 Variable Definitions and Summary Statistics

Variables		Mean	Std. dev.	Min	Max
Dependent variables					
All cites	All citations to the focal patent from the destination firm (in the given year)	0.018	0.200	0	11
Cites excluding inventor self-cites	Destination firm citations coming from individuals other than the patent's inventor herself	0.013	0.164	0	11
Cites excluding inventor self-cites and collaborator cites	Destination firm citations coming from individuals other than the patent's inventor herself or any of her prior collaborators	0.011	0.147	0	11
Explanatory and control variables					
Recruit	Indicator for whether the given patent is a focal patent involving a move or a control patent	0.50	0.50	0	1
Destination firm patents	Number of patents assigned to the destination firm in the five years preceding the inventor's move year	1,041.9	1,433.1	0	8,260
Destination same-class patents	Number of destination firm patents that belong to the same three-digit technology class as the focal patent	26.9	75.8	0	1,375
Original firm patents	Number of patents assigned to the original firm in the five years preceding the focal patent's application year	1,120.1	1,157.0	0	5,504
Claims	Number of claims made by the focal patent	11.53	9.36	1	155
Patent references	Number of backward citations that the focal patent makes to other patents	6.78	7.01	0	171
Nonpatent references	Number of nonpatent references made by the focal patent	0.86	2.28	0	37
Inventor patents	Number of previous patents successfully applied for by the same inventor	7.76	13.07	0	202
Inventor age	Number of years since the inventor applied for her first successful patent	4.47	4.34	0	26
U.S. inventor	Indicator for whether the focal patent's inventor has a U.S. address or not	0.54	0.50	0	1
Grant delay	The delay between application date and grant date for the focal patent (in months)	22.69	10.82	1	136
Timing-related variables					
Postmove	For a focal patent, this is an indicator for whether the citing year falls after the move. For a control patent, this is always defined as 0.	0.318	0.466	0	1
Pair postmove	For both the focal and the control patents in a matched pair, this is an indicator for whether the citing year falls after the move for the focal patent in that pair.	0.636	0.481	0	1
Premove period 1	For a focal patent, this is an indicator for whether the citing year is the 1st or 2nd year before the move. For a control patent, this is always 0.	0.083	0.276	0	1
Premove period 2	For a focal patent, this is an indicator for whether the citing year is the 3rd or 4th year before the move. For a control patent, this is always 0.	0.055	0.227	0	1
Premove period 3	For a focal patent, this is an indicator for whether the citing year is the 5th or 6th year before the move. For a control patent, this is always 0.	0.029	0.168	0	1
Premove period 4	For a focal patent, this is an indicator for whether the citing year is the 1st or 2nd year before the move. For a control patent, this is always 0.	0.012	0.109	0	1
Premove period 5	For a focal patent, this is an indicator for whether the citing year is the 7th or 8th year before the move. For a control patent, this is always 0.	0.003	0.057	0	1
Postmove period 1	For a focal patent, this is an indicator for whether the citing year is the 1st or 2nd year after the move. For a control patent, this is always 0.	0.083	0.276	0	1
Postmove period 2	For a focal patent, this is an indicator for whether the citing year is the 3rd or 4th year after the move. For a control patent, this is always 0.	0.080	0.271	0	1
Postmove period 3	For a focal patent, this is an indicator for whether the citing year is the 5th or 6th year after the move. For a control patent, this is always 0.	0.071	0.257	0	1
Postmove period 4	For a focal patent, this is an indicator for whether the citing year is the 7th or 8th year after the move. For a control patent, this is always 0.	0.054	0.227	0	1
Postmove period 5	For a focal patent, this is an indicator for whether the citing year is the 9th or 10th year after the move. For a control patent, this is always 0.	0.029	0.167	0	1

Notes. We report summary statistics based on the original sample constructed as illustrated in Figures 1 and 2. This sample has 95,424 observations, arising from 12 yearly observations for each of the 3,976 focal cited patents and another 12 yearly observations for each of the 3,976 control patents.

for not just the focal but also the control patents in each matched pair for observations corresponding to the focal patent's postmove period. However, as already discussed in the previous section, our preferred models (that use patent fixed effects) do not include this variable, thus avoiding the implied assumption of strict one-to-one comparability of focal and control patents. Note that this issue is irrelevant for models estimated using the focal patent subsample, where a separate *pair postmove* variable cannot be estimated in any case.

For the pooled estimation models, we employ several control variables (that are time invariant within a patent, and hence dropped in the preferred models that employ patent fixed effects). Specifically, we control for key observed characteristics of (1) the destination firm (overall level of inventive activity, level of inventive activity in the same technology class as the recruit's patent),[17] (2) the source firm (overall level of inventive activity), (3) the inventor (number of prior patents, number of years since first patent, U.S. resident indicator), and (4) the patent itself (number of claims, number of patent references, number of nonpatent references, technology area).[18]

3.2. Constructing the "CEM Sample" Using More Stringent Matching

Table 2 summarizes several characteristics of focal versus control patents for the original sample and shows that the two subsamples differ significantly on several key dimensions. In particular, focal patents systematically (1) receive more overall citations, (2) receive more premove citations from the destination firm, and (3) are created by more experienced inventors.

To achieve more stringent matching, we employ the CEM approach (Iacus et al. 2009).[19] Specifically, we construct a sample that matches patents not just on the three-digit technology class and application year but also on discrete buckets based on five additional criteria: (1) number of premove citations received overall (four buckets: 0, 1–2, 3–7, and 8 or more); (2) number of premove citations received from the destination firm specifically (four buckets: 0, 1, 2–3, and 4 or more); (3) number of previous patents by this inventor (three buckets: 0 or 1, 2–12, and 13 or more); (4) number of years since the first patent by this inventor (three buckets: 0 or 1, 2–9, and 10 or more); and (5) number of days delay between the application and grant date (three buckets: 490 or less, 491–832, and 833 or more).[20]

Table 2 Focal vs. Control Patents in the "Original Sample" (3,976 Patent Pairs)

	Focal patents		Control patents	
	Mean	Std. dev.	Mean	Std. dev.
Cumulative overall premove citations	2.98	4.50	2.42	3.88
Cumulative premove citations from destination	0.063	0.461	0.023	0.272
Inventor patents	9.13	14.64	5.87	12.27
Inventor age	4.98	4.29	3.96	4.32
Grant delay	22.78	10.72	22.59	10.91
Claims	11.73	9.68	11.32	9.02
Patent references	6.92	7.40	6.64	6.58
Nonpatent references	0.88	2.32	0.84	2.23
Originality	0.34	0.27	0.35	0.28

Notes. The "originality" measure (taken from the NBER database) computes the breadth of search as one minus the Herfindahl index of the backward citations made by a patent to different technology classes, and is undefined for the small fraction of patents (less than 4% of the sample) that have no backward citations. The means of a majority of the reported characteristics (all except *grant delay*, *nonpatent references*, and *originality*) differ between the focal and control subsamples at the 5% significance level, motivating the need for a more stringent matching approach.

With this set of criteria and corresponding buckets, we are able to find matches for 2,543 of the 3,976 mobile inventor patents from the original sample, resulting in a final sample of 5,086 ($2{,}543 \times 2$) cited patents. Again following a procedure similar to the one illustrated in Figure 2, we construct a patent-year data set of 61,032 ($5{,}086 \times 12$) observations, which we call the "CEM sample." As the statistics reported in Table 3 demonstrate, the focal and control patents in this sample are indeed better matched than in the original sample summarized in Table 2.

4. Results

We now report our findings on four topics: (1) a comparison of our estimates for the mobility–knowledge flow relationship using the traditional cross-sectional method versus our DD approach, (2) an estimate of the sensitivity of our findings to the more stringently matched CEM sample, (3) an estimate of the fraction of the boost in use of the recruit's prior ideas that is due not to "learning" but rather to exploitation of those ideas by the recruit herself, and (4) an examination of temporal trends, including the persistence over time of the recruit as the destination firm's primary user of her stock of prior ideas.

4.1. Summary Statistics for Citation Rates

Before delving into the regression analysis, we present the basic intuition behind a DD approach using summary statistics for the premove and postmove subsamples corresponding to the focal as well as control

[17] To prevent circularity, wherein the move itself affects the size of the patent pool, we calculate the patent pool as of the move year itself (rather than as of a citing year subsequent to the move).

[18] Whereas the pooled negative binomial models employ technology indicators at the two-digit NBER subcategory level, the pooled linear models employ indicators at the three-digit technology class level. Also, we only use *inventor age*, *U.S. inventor*, and *grant delay* in the pooled negative binomial models, but the pooled linear estimation allows us to use a full set of indicator variables for inventor age (in years), inventor location (U.S. state or non-U.S. country), and grant delay (in months). For more on the importance of saturating such models, particularly on the temporal dimension, see Levin and Stephan (1991), Hall et al. (2007), and Mehta et al. (2009).

[19] For a recent application of this technique, see Azoulay et al. (2010).

[20] Choosing the matching criteria involves a trade-off between the stringency of the match and the fraction of the sample for which a match can be found. We choose to be most stringent on criteria (1) and (2) to ensure that premove citation patterns are practically identical for the recruit versus control patents. For criteria (3), (4), and (5), we form boundaries for the three buckets using the 25th and 75th percentiles as cutoffs.

Table 3 **Focal vs. Control Patents in the "CEM Sample" (2,543 Patent Pairs)**

	Focal patents		Control patents	
	Mean	Std. dev.	Mean	Std. dev.
Cumulative overall premove citations	2.55	3.76	2.50	3.83
Cumulative premove citations from destination	0.010	0.136	0.010	0.136
Inventor patents	7.59	11.36	7.40	14.52
Inventor age	4.50	4.05	4.61	4.18
Grant delay	23.11	10.62	22.95	10.58
Claims	11.39	9.01	11.18	8.67
Patent references	6.40	5.97	6.38	5.25
Nonpatent references	0.85	2.14	0.78	1.98
Originality	0.33	0.28	0.34	0.28

Notes. The CEM sample has fewer patent pairs than the original sample (2,543 in contrast to 3,976 in Table 2) because we drop focal patents for which we find no match. The CEM procedure requires a match not just on the three-digit technology class and application year but also on discrete buckets based on five additional criteria: (1) number of premove citations received overall (four buckets: 0, 1–2, 3–7, and 8 or more); (2) number of premove citations received from the destination firm specifically (four buckets: 0, 1, 2–3, and 4 or more); (3) number of previous patents by this inventor (three buckets: 0, 1, 2–12, and 13 or more); (4) number of years since the first patent by this inventor (three buckets: 0 or 1, 2–9, and 10 or more); and (5) number of days delay between the application and grant date (three buckets: 490 or less, 491–832, and 833 or more). None of the means differ between the focal and control subsamples at the 5% significance level, indicating a closer match than that in Table 2.

patents.[21] Table 4 reports the means for *all cites* for each of these four subsamples from the original sample. Recall from Table 1 that this variable is a count of *all* citations a patent receives from the destination firm in a given year, including self-citations made by the recruit. For the postmove period, focal patents have a greater average annual citation rate (0.0396) than control patents (0.0061), reflecting a difference of 0.0335 citations per year. However, it is worth noting that the annual citation rate is greater for the focal patents even in the premobility period (0.0140 versus 0.0051, reflecting a difference of 0.0089). The DD intuition is that, rather than attributing the whole postmobility difference to the move itself, it is more reasonable to attribute to the move only the difference, 0.0246 (i.e., 0.0335 − 0.0089), between the postmobility and premobility differences.

[21] By construction, inventors of control patents do not move between firms. So, in reporting summary statistics and initial regression models that are directly motivated by the summary statistics, we apply the move date of each focal patent to its corresponding matched control patent as a "counterfactual" move date. However, recognizing that control patents do not actually involve a move, all our preferred regression models (on which we base our key conclusions) drop this approach and instead "throw the controls back in the pot."

We now turn to the analogous calculation for the variable *cites excluding inventor self-cites*, which excludes self-citations made by the mobile inventor. Note that *all cites* and *cites excluding inventor self-cites* are identical premove because the inventor is not yet an employee of the destination firm; therefore, by construction, there are no inventor self-cites from the destination firm that involve this inventor in the premove period. The postmove mean for *cites excluding inventor self-cites* for recruits' patents is 0.0250, which is significantly smaller than the corresponding *all cites* mean of 0.0396, suggesting that roughly 37% (i.e., 1 − (0.0250/0.0396)) of the overall citations are self-cites by the mover. Repeating the earlier calculation for *cites excluding inventor self-cites* leads to a DD summary statistic of 0.0100, which is less than half the corresponding *all cites* DD statistic of 0.0246. In other words, inventor self-citations appear to play a prominent role in the overall citation patterns seemingly associated with mobility.

Table 5 reports the corresponding summary statistics for the CEM sample. The premove means for citations received are now very similar for focal versus control patents because the cumulative premove citation count is one of the matching criteria for CEM.[22] Thus, the estimated DD effect now (roughly) coincides with a direct comparison of the focal and control patent citation rates after the move.

The two main insights from Table 4 persist in Table 5. First, there is a positive DD effect for *all cites*, suggesting an increase in citation rate attributable to mobility (with the magnitude 0.0296 being somewhat greater but not too dissimilar from the value of 0.0246 reported in Table 4 for the original sample). Second, comparing the DD statistic for *cites excluding inventor self-cites* with that for *all cites* (0.0167 versus 0.0296, respectively) again suggests that only a fraction of the apparent increase associated with the move arises from "learning by hiring"; a significant component of the destination firm's jump in usage of the recruit's prior ideas is due to the recruit herself.

4.2. Regression Analysis Using the "Original Sample"

We now turn to the regression framework for a more rigorous analysis of the link between mobility and knowledge flow. Table 6 reports our baseline analysis using the original sample, employing *all cites* as the dependent variable. Column (1) reports cross-sectional findings from the postmove subsample, in effect replicating the current "best practice" of

[22] Note that the focal patents successfully matched during CEM have systematically lower premove citation rates (compare Tables 4 and 5). We offer possible implications of this in the discussion section.

Table 4 Annual Patent Citation Frequency for the "Original Sample"

	Average annual citations received from the destination firm: Premove	Average annual citations received from the destination firm: Postmove	
Focal patents	Subsample mean: All cites = 0.0140 Cites excluding inventor self-cites = 0.0140 ($N = 17{,}383$)	Subsample mean: All cites = 0.0396 Cites excluding inventor self-cites = 0.0250 ($N = 30{,}329$)	First difference (row): All cites = 0.0256 Cites excluding inventor self-cites = 0.0110 ($N = 47{,}712$)
Control patents	Subsample mean: All cites = 0.0051 Cites excluding inventor self-cites = 0.0051 ($N = 17{,}383$)	Subsample mean: All cites = 0.0061 Cites excluding inventor self-cites = 0.0061 ($N = 30{,}329$)	First difference (row): All cites = 0.0010 Cites excluding inventor self-cites = 0.0010 ($N = 47{,}712$)
	First difference (column): All cites = 0.0089 Cites excluding inventor self-cites = 0.0089 ($N = 34{,}766$)	First difference (column): All cites = 0.0335 Cites excluding inventor self-cites = 0.0189 ($N = 60{,}658$)	Difference in differences: All cites = 0.0246 Cites excluding inventor self-cites = 0.0100 ($N = 95{,}424$)

Notes. This analysis is based on the original sample described in Table 2. Each cell summarizes the average value of the *all cites* and *cites excluding inventor self-cites* variables in the corresponding subsample (see Table 1 for variables definitions). Note that *all cites* and *cites excluding inventor self-cites* are identical premove because the inventor is not yet an employee of the destination firm; therefore, by construction, there are no inventor self-cites from the destination firm in the premove period.

employing such comparisons of the postmove citation frequency for focal versus control patents.[23] The estimates appear to be both statistically and economically quite significant: the citation rate implied by the negative binomial regression coefficients is 566% ($e^{1.896} - 1$) greater in the focal sample than in the control sample.

However, the DD analysis reported in column (2) highlights why interpreting the above effect as being entirely due to the inventor move itself is inappropriate. The positive and significant coefficient on *recruit*, despite having a separate term for *postmove*, demonstrates that focal patents systematically receive more citations even before the "mobility event" takes place. Stated another way, the cross-sectional analysis employed in column (1) confounds selection effects with treatment effects. Nevertheless, because the estimate of the coefficient on *postmove* is positive and significant in column (2), we have evidence consistent with an increase in knowledge use associated with mobility (even though the effect is smaller than a cross-sectional analysis would lead us to believe). Because interpreting the magnitude of effects in nonlinear models (particularly in cross-model comparison) is not straightforward, we postpone that discussion until we come to our preferred specification below: linear models with patent fixed effects.

We next replicate the pooled analyses from columns (1) and (2) using linear models before turning to fixed effects specifications. As the results in column (4) show, there is again clear evidence that focal patents are systematically more highly cited even premove.[24] As a benchmark, note that the average number of citations received by a control patent in the postmove subsample is 0.0061 (Table 4). Compared with the column (3) estimate of 0.0356, this represents a 584% (0.0356/0.0061) increase in citations to the focal patents; this is similar in magnitude to the column (1) estimate discussed earlier.

The specification employed in column (4) again takes into account the concern that a cross-sectional association between mobility and citations confounds selection with treatment. The estimates suggest that once we use *recruit* to account for the systematic difference between focal and control patents that exists even before the move, a move is associated with a smaller but still significant increase of 0.0251 citations per year. Compared to the average number of citations received by the focal patents prior to the mobility event (0.0140 from Table 4), this estimate implies a 179% (0.0251/0.0140) increase in the citation rate. The results in column (4) also suggest that selection accounts for almost one-third, 29% (0.0101/(0.0101 + 0.0251)), of the effect that a cross-sectional approach like column (3) would (incorrectly) attribute to the inventor move.

[23] For variables that are highly skewed, the pooled regression analysis employs a logarithmic transformation, first adding one to allow for transformation even in instances where the value could be zero. The results are robust to changing the size of the offset or using untransformed variables for the analysis.

[24] Given that we are working with highly disaggregated data (citations to a specific patent from a specific firm in a specific year), the low R^2 values are not a surprise. The more important statistics here are the F-statistic for the model as a whole and the t-statistics for key variables, which are statistically significant. If increasing R^2 were an end in itself, we would use a more aggregate unit of analysis (e.g., firm pairs) to remove unsystematic individual-level noise. However, doing so would not utilize the longitudinal microlevel information we want to exploit. In fact, a low R^2 value is common when using such disaggregated data; prominent examples include labor market outcomes (e.g., Angrist and Krueger 1994, Malamud and Pop-Eleches 2010) and short-term stock returns (e.g., Llorente et al. 2002).

Table 5 **Annual Patent Citation Frequency for the "CEM Sample"**

	Average annual citations received from the destination firm		
	Premove	Postmove	
Focal patents	Subsample mean: All cites = 0.0022 Cites excluding inventor self-cites = 0.0022 (N = 10,981)	Subsample mean: All cites = 0.0345 Cites excluding inventor self-cites = 0.0216 (N = 19,535)	First difference (row): All cites = 0.0323 Cites excluding inventor self-cites = 0.0194 (N = 30,516)
Control patents	Subsample mean: All cites = 0.0022 Cites excluding inventor self-cites = 0.0022 (N = 10,981)	Subsample mean: All cites = 0.0049 Cites excluding inventor self-cites = 0.0049 (N = 19,535)	First difference (row): All cites = 0.0027 Cites excluding inventor self-cites = 0.0027 (N = 30,516)
	First difference (column): All cites = 0 Cites excluding inventor self-cites = 0 (N = 21,962)	First difference (column): All cites = 0.0296 Cites excluding inventor self-cites = 0.0167 (N = 39,070)	Difference-in-differences: All cites = 0.0296 Cites excluding inventor self-cites = 0.0167 (N = 61,032)

Notes. This analysis replicates the summary analysis from Table 4, but for the CEM sample. Note that the premove means for citations received are practically the same for the focal and control patents, because cumulative premove citation count is one of our CEM matching criteria. Also note that the overall sample size is smaller—61,032 here versus 95,424 in Table 4—because we drop the focal patents for which we do not find a match when constructing the CEM sample.

The model estimated in column (5) employs matched-pair fixed effects to carry out the above identification using just within-pair variation for the matched pairs; the results remain almost identical to those in column (4). In fact, the estimate of *postmove* turns out to be exactly as before (0.0251), so the implied mobility-related increase in citation rate (179%) is also the same. The estimate for *recruit* is also almost the same as before (0.0097 instead of 0.0101), once more highlighting selection concerns.

Note that, whereas a cross-sectional analysis uses the average postmove citation rate for control patents as a benchmark, our longitudinal design enables the use of citation rates for the focal patents themselves (adjusted for time trends) during the premove period as a benchmark. The latter interpretation seems more desirable as it is based directly on the "treatment" of interest (i.e., the mobility event). A comparison of the natural percentage interpretations associated with the estimates from column (3) versus from column (4) or (5) implies that using a purely cross-sectional comparison of postmove citation rates exaggerates the percentage increase in benefits attributable to mobility by 226% (584%/179% − 1).

Note that the models in columns (4) and (5) include, in addition to the variable of interest (*postmove*), a separate variable *pair postmove* that is equal to 1 for both the focal and control patents in a matched pair for the period after the move date for the focal patent's inventor. The model estimated in column (6), which is our specification of choice, effectively "throws these control patents back into the pot" by dropping the *pair postmove* variable. Instead of pair fixed effects, it therefore relies upon (single) patent fixed effects, using within-patent variation to estimate *postmove*, while also employing the control sample to appropriately account for fixed effects of the different application years and lags (through indicator variables not reported in the tables). The DD estimate for the coefficient on *postmove* goes up slightly to 0.0306, reflecting an increase of 219% (0.0306/0.0140) relative to the average number of cites a recruit's invention receives from the destination firm before the move. This is our primary estimate of the "recruitment effect" when employing the original sample.

Before turning to our regression analysis using a more stringently matched sample, it is useful to check the robustness of our main result to using a fixed effects model based on the focal patent only subsample, which does not include controls. (The model is still identifiable because of variation in application and move years across focal patents.) This analysis is reported in column (7). Given that the control patents that at least partly help account for endogeneity of the move are no longer included, the focal-patent-only sample estimate for *postmove* in column (7) is naturally greater. Although the difference in magnitudes between columns (6) and (7) is not trivial, it is not large enough to change the qualitative conclusions. (As discussed later, these estimates are also not too different from those obtained using the matched sample based on more stringent matching.)

It is useful to summarize our main points from Table 6. First, making inferences based on a cross-sectional comparison of postmove citations is problematic, because ignoring preexisting heterogeneity across patents leads to a selection bias. Second, a DD approach, which relies on a weaker assumption of comparability on trends rather than levels of focal and control citation rates, mitigates this issue by exploiting information on premove citation rates to adjust the estimates.

Table 6 Regression Analysis ("Original Sample")

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	*All cites*	*All cites*	*All cites*	*All cites*	*All cites*	*All cites*	*All cites*
Regression model:	Negative binomial	Negative binomial	Linear	Linear	Linear (pair FE)	Linear (patent FE)	Linear (Patent FE)
Patent-year sample:	Original sample (postmove only)	Original sample	Original sample (postmove only)	Original sample	Original sample	Original sample	Original sample (focal only)
Recruit	1.896***	1.017***	0.0356***	0.0101***	0.0097***		
	(0.137)	(0.213)	(0.0034)	(0.0022)	(0.0021)		
Postmove		0.907***		0.0251***	0.0251***	0.0306***	0.0367***
		(0.230)		(0.0037)	(0.0028)	(0.0023)	(0.0043)
Pair postmove		0.157		0.0033	0.0061***		
		(0.207)		(0.0025)	(0.0023)		
ln(*destination same-class patents*)	0.428***	0.497***	0.0085***	0.0075***			
	(0.050)	(0.048)	(0.0012)	(0.0010)			
ln(*destination firm patents*)	−0.215***	−0.202***	−0.0039***	−0.0028***			
	(0.045)	(0.040)	(0.0010)	(0.0007)			
ln(*original firm patents*)	0.013	0.003	0.0010	0.0008	0.0011		
	(0.040)	(0.038)	(0.0013)	(0.0009)	(0.0009)		
ln(*claims*)	0.085	0.096	0.0049**	0.0041**	0.0046**		
	(0.083)	(0.078)	(0.0024)	(0.0018)	(0.0021)		
ln(*patent references*)	0.027	0.074	0.0009	0.0019	0.0018		
	(0.083)	(0.077)	(0.0028)	(0.0019)	(0.0024)		
ln(*nonpatent references*)	0.149	0.073	0.0036	0.0018	−0.0002		
	(0.093)	(0.091)	(0.0027)	(0.0020)	(0.0029)		
ln(*inventor patents*)	−0.129	−0.112	−0.0013	−0.0007	0.0006		
	(0.092)	(0.085)	(0.0018)	(0.0013)	(0.0022)		
ln(*inventor age*)	0.112	0.053					
	(0.105)	(0.098)					
U.S. inventor	0.662***	0.611***					
	(0.156)	(0.138)					
ln(*grant delay*)	0.123	−0.011					
	(0.136)	(0.131)					
Number of observations	60,658	95,424	60,658	95,424	95,424	95,424	47,712
Number of cited patents	7,952	7,952	7,952	7,952	7,952	7,952	3,976
Log likelihood	−5,238	−6,806					
Wald chi^2	841.9***	48,236***					
R^2			0.037	0.029	0.014	0.004	0.006
F-statistic			4.47***	5.43***	5.76***	12.06***	9.02***

Notes. All models include yearly indicator variables for the citing year and patent age. The pooled linear models in columns (3) and (4) do not include variables *inventor age*, *U.S. inventor*, and *grant delay* used in the columns (1) and (2) because more fine-grained indicator variables for the inventor age (in years), inventor location (U.S. states or non-U.S. countries), and patent grant delay (in months) have been used instead but not shown. We use robust standard errors clustered on the identity of the inventor in all four of these pooled models. We employ matched pair fixed effects (FE) (with standard errors clustered for each pair) in column (5), and employ patent fixed effects in columns (6) and (7). Whereas column (6) reports our preferred estimation approach that employs a matched sample, the column (7) analysis is a robustness check using just the focal patent subsample. Detailed timing analysis corresponding to column (6) appears in Figure 3, revealing an upward trend in focal patent citations even before the move year, motivating the need for more stringent matching.

** $p < 0.05$; *** $p < 0.01$.

However, even a DD approach leaves important endogeneity-related concerns unaddressed. For example, an unobserved factor (such as a shift in the specific domain focus of the firm that the technology match does not capture) might drive both the hiring of a specific individual and an increased use of their ideas. One way to investigate this further is to examine the temporal pattern of citations in more detail. Such a detailed examination (described later in the context of Figure 3) reveals evidence that the destination firm does indeed increase its use of the inventor's ideas even in the years leading up to the move. This heightens concerns that unexplained time trends not directly related to the move might in part be driving the DD findings.

4.3. Estimates Based on More Stringent Matching ("CEM Sample")

It is desirable that focal and control patents show a similar citation trend before a move so that one can more reasonably assume they would have followed parallel paths in the postmove period had the move

Table 7 Regression Analysis ("CEM Sample")

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	*All cites*	*All cites*	*All cites*	*All cites*	*All cites*	*All cites*	*All cites*
Regression model:	Negative binomial	Negative binomial	Linear	Linear	Linear (Pair FE)	Linear (Patent FE)	Linear (Patent FE)
Patent-year sample:	CEM sample (postmove only)	CEM sample	CEM sample (postmove only)	CEM sample	CEM sample	CEM sample	CEM sample (focal only)
Recruit	1.927***	0.007	0.0316***	0.0012	0.0008		
	(0.185)	(0.376)	(0.0040)	(0.0013)	(0.0010)		
Postmove		1.945***		0.0295***	0.0290***	0.0341***	0.0447***
		(0.407)		(0.0039)	(0.0029)	(0.0025)	(0.0048)
Pair postmove		0.895***		0.0069**	0.0095***		
		(0.319)		(0.0028)	(0.0026)		
ln(*destination same-class patents*)	0.450***	0.509***	0.0076***	0.0056***			
	(0.067)	(0.066)	(0.0013)	(0.0009)			
ln(*destination firm patents*)	−0.275***	−0.272***	−0.0037***	−0.0024***			
	(0.059)	(0.057)	(0.0011)	(0.0007)			
ln(*original firm patents*)	−0.066	−0.063	−0.0016	−0.0010	0.0010		
	(0.051)	(0.049)	(0.0018)	(0.0012)	(0.0012)		
ln(*claims*)	−0.014	−0.014	0.0022	0.0012	0.0029		
	(0.108)	(0.106)	(0.0032)	(0.0021)	(0.0026)		
ln(*patent references*)	0.207*	0.182*	0.0050	0.0033	0.0029		
	(0.113)	(0.107)	(0.0035)	(0.0023)	(0.0022)		
ln(*nonpatent references*)	0.053	0.082	−0.0009	−0.0000	0.0027		
	(0.125)	(0.121)	(0.0029)	(0.0019)	(0.0025)		
ln(*inventor patents*)	−0.175	−0.133	−0.0022	−0.0013	0.0007		
	(0.117)	(0.114)	(0.0023)	(0.0015)	(0.0023)		
ln(*inventor age*)	0.158	0.092					
	(0.145)	(0.141)					
U.S. inventor	0.544***	0.411**					
	(0.198)	(0.188)					
ln(*grant delay*)	−0.103	−0.119					
	(0.192)	(0.183)					
Number of observations	39,070	61,032	39,070	61,032	61,032	61,032	30,516
Number of cited patents	5,086	5,086	5,086	5,086	5,086	5,086	2,543
Log likelihood	−4,153	−4,525					
Wald chi^2	988.9***	42,198***					
R^2			0.033	0.025	0.021	0.006	0.009
F-statistic			2.93***	3.48***	6.15***	11.35***	7.94***

Notes. The regression models employed here are similar to the ones employed in the corresponding columns in Table 6, but using the CEM sample. The sample sizes are smaller since patents for which the CEM procedure does not find a match are not included. Unlike in Table 6, the estimates for *recruit* in columns (2), (4), and (5) are practically zero because the cumulative premove citation count is now one of the matching criteria. Further evidence of CEM-based matching working well is that Figure 4, which depicts more detailed timing analysis corresponding to column (6), shows no premove trend in the citation rate. FE, fixed effects.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

not taken place. We address this using the CEM sample, which largely eliminates premove citation rate differences by construction.

We replicate all the specifications from Table 6 in Table 7 by employing the CEM sample instead of the original sample. The estimated coefficient for *recruit* in the DD models shown in columns (2), (4), and (5) are now close to zero, which is not surprising given that premove citations are among the matching criteria for our CEM procedure. Furthermore, as an analysis of the temporal pattern shows (illustrated in Figure 4 and described in §4.5), the CEM sample demonstrates no evidence of the destination firm increasing its use of the focal patents more than the control patents in the years leading up to the move. This reduces significantly the above endogeneity-related concerns.

Interestingly, the linear estimates for the DD coefficient of interest—*postmove*—are very similar in the CEM sample to those reported in the corresponding models for the original sample in Table 6. (Note that the nonlinear coefficient estimates in the first two columns are not directly comparable across tables.) For example, comparing our preferred specification—column (6)—across tables, we find that the estimate of interest (*postmove*) is 0.0306 in Table 6 and only slightly greater, 0.0341, in Table 7. The fact that the

Table 8 Learning by Hiring?

	(1)	(2)	(3)
Dependent Variable:	*All cites*	*Cites excluding inventor self-cites*	*Cites excluding inventor self-cites and collaborator cites*
Regression model:	Linear (patent FE)	Linear (patent FE)	Linear (patent FE)
Patent-year sample:	CEM sample	CEM sample	CEM sample
Postmove	0.0341***	0.0188***	0.0128***
	(0.0025)	(0.0021)	(0.0018)
Number of observations	61,032	61,032	61,032
Number of cited patents	5,086	5,086	5,086
F-statistic	11.35***	6.06***	3.97***

Notes. This analysis is based on the CEM sample. (We report the analogous analysis for the original sample in Table A1 of the online appendix.) In column (1), we replicate the results from column (6) of Table 7 using *all cites* as the dependent variable. (We have again not reported the estimates for the indicator variables to conserve space.) In column (2), we examine the extent to which the estimated boost in citation rates after mobility diminishes when citation counts exclude inventor self-cites. In column (3), we analyze the extent to which the estimated boost in citation rates after mobility diminishes when citation counts also exclude cites made by the recruit's collaborators in the destination firm. FE, fixed effects.

***$p < 0.01$.

estimates do not seem sensitive to the control sample employed offers additional confidence in using the DD approach. We base the remaining analyses reported in this paper on specifications akin to those in column (6) in Table 7, i.e., using linear models with patent fixed effects and employing the CEM sample, although the key findings are robust to alternative specifications and samples.[25]

4.4. Learning by Hiring?

Next, we explore the extent to which the increase in use of the recruit's prior ideas is due to subsequent self-exploitation by the recruit herself versus broader learning by hiring (i.e., diffusion to others in the destination firm). This analysis, using the CEM sample, is reported in Table 8.

For ease of comparison, we reproduce the aggregate findings (using *all cites* as the dependent variable) from column (6) of Table 7 in column (1) of Table 8. The analysis reported in column (2) of Table 8 uses *cites excluding inventor self-cites* as the dependent variable instead. The estimate for *postmove* is now 0.0188, which represents only 55% (0.0188/0.0341) of the destination firm's overall increase in the use of the recruit's prior ideas. In other words, almost half of the boost in use of the recruit's prior ideas has nothing to do with firm learning at all; rather, it is due to the recruit building upon her own prior ideas.[26]

We further explore the limits to widespread diffusion of the tacit knowledge associated with the recruit's prior ideas by excluding not only the inventor's self-citations but also citations made by the inventor's collaborators (individuals who appear as coinventors with the recruit on past patents). With this goal, the analysis reported in column (3) uses cites excluding inventor self-cites and collaborator cites as the dependent variable. The estimated coefficient on *postmove* is 0.0128, and represents only 38% (0.0128/0.0341) of the overall *postmove* effect from column (1). This further suggests that learning by hiring is actually quite localized; on average, only just over a third of the hiring firm's boost in the use of a recruit's stock of prior ideas is due to individuals other than the recruit and her collaborators.

4.5. A Closer Look at Temporal Trends in Knowledge Use

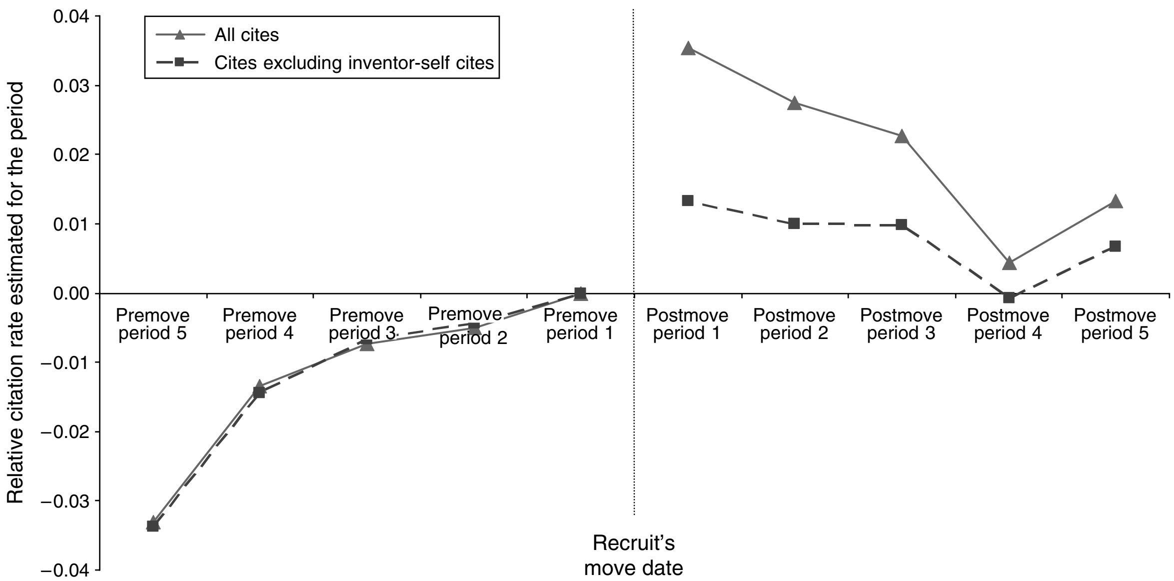
To examine the temporal patterns of citations received by a patent from a destination firm, we use indicator variables (summarized in Table 1) corresponding to five postmove two-year periods (*postmove period* 1 through *postmove period* 5) and four premove two-year periods (*premove period* 2 through *premove period* 5) in the fixed effects regression analysis, with the two-year period immediately preceding the move (corresponding to *premove period* 1) being the reference. For ease of visual interpretation, we plot the coefficient estimates for the original and CEM samples in Figures 3 and 4, respectively, with the corresponding regression estimates reported in the online appendix as Tables A2 and A3, respectively.

The analysis based on the original sample (Figure 3) reveals evidence of the destination firm increasing its use of the inventor's ideas to some extent even in the years leading up to the move. In contrast, the corresponding analysis for the CEM sample (Figure 4) demonstrates no evidence of the destination firm increasing its use of the focal patents more than the control patents in the years leading up to the move. This is consistent with the CEM procedure reducing endogeneity-related concerns through use of stringent matching. Therefore, we base the rest of the discussion on Figure 4 (though the temporal effects

[25] Comparing Tables 6 and 7, one might wonder why the CEM sample demonstrates a (slightly) *greater* postmove increase in citation rate than the original sample. This difference is driven primarily by a few focal patents that are highly cited before a move but do not find a match using the stringent CEM criteria. When we repeat the Table 6 analysis using the same focal patents (but original controls) as in the CEM sample, the estimates (not reported) are actually somewhat *greater* than the corresponding CEM-based findings in Table 7. This is in line with the intuition that, *for a given sample of focal patents*, less stringent matching leads to a larger estimated effect of mobility.

[26] As Table A1 in the online appendix shows, the results remain qualitatively similar if we employ the original sample instead, with the relative importance of self-exploitation (versus learning by hiring) now estimated to be even greater.

Figure 3 Estimated Temporal Trends in Destination Firm Citations for the "Original Sample"



Notes. This figure plots the detailed temporal pattern behind the DD finding from column (6) of Table 6. The coefficient estimates plotted here are reported in Table A2 in the online appendix. Each of the periods plotted on the horizontal axis is two years long (see Table 1). Because the reference period is *premove period* 1, the DD coefficient for that period is zero by definition. The upward trend in citation rate before the move year suggests the need for more stringent matching.

associated with mobility seem qualitatively similar for Figure 3).

Figure 4 facilitates three observations. First, with respect to both of our dependent variables, *all cites* and *cites excluding inventor self-cites*, there is a sharp discontinuity in the annual citation rate at the time of the mobility event (without any discernable increase in the years leading up to it). Although the tight timing of the mobility–idea usage relationship does not necessarily imply causality, it is certainly consistent with the assertion that recruiting facilitates enhanced access to an individual's ideas.

Second, a comparison of the estimated citation rate for *cites excluding inventor self-cites* versus *all cites* in the period immediately following the mobility event illustrates that the recruit accounts for approximately half (45%) of the hiring firm's increased use of her ideas right after she arrives, i.e., during the first two years.

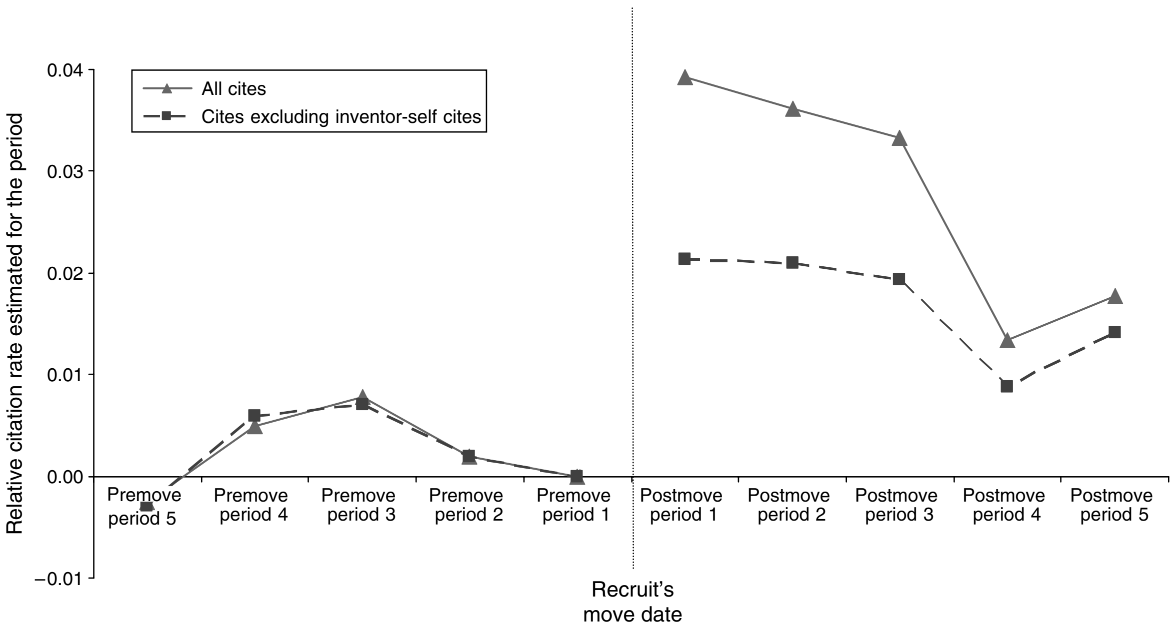
Third, and perhaps most surprisingly, comparing the findings for *cites excluding inventor self-cites* versus *all cites* over the duration of the postmove period, we see that the recruit continues to account for more than a third of the citations (between 34% and 42%) to her prior ideas for the next three periods (six years). Although one might expect the recruit's relative role to diminish rapidly as the knowledge associated with her ideas diffuses across her new firm, this actually occurs quite gradually: during the first eight years after a move there is only a gradual increase (from 55% up to 66%) in the fraction of cites that are not inventor self-cites. In other words, the importance of the recruit to the hiring firm (in terms of her direct involvement in the firm's use of her prior ideas) is surprisingly persistent over time.

5. Discussion and Conclusion

What are the implications of our findings for managers and policymakers? Let us start with our main finding that firms do indeed increase their use of their recruits' prior ideas. From the policy maker's perspective, it is tempting to conclude from this that restrictions on interfirm mobility, such as noncompete covenants, are detrimental to the circulation of ideas and knowledge spillovers that would otherwise enhance regional growth. From the point of view of the source firm, one might infer that mobility of employees leads to "leakage" of ideas to competitors, an outcome that must be guarded against. Finally, for the hiring firm, poaching employees from others might seem an attractive way to access external ideas. Although there may be merit to all of the above arguments, one should be cautious in drawing strong normative prescriptions directly from our results. We describe some of the key nuances in our interpretation below.

First, the policy implications are not obvious. On the one hand, when mobility is restricted, ideas are less likely to circulate among different firms in the form of knowledge spillovers, which may indeed inhibit growth. On the other hand, firms may be more willing to invest in innovation and human capital development for their employees when they are less

Figure 4 Estimated Temporal Trends in Destination Firm Citations for the "CEM Sample"



Notes. Following a procedure analogous to Figure 3, this figure plots the detailed temporal patterns behind the DD findings from column (6) of Table 7. The estimates are reported in Table A3 in the online appendix. There is no evidence of an upward premove citation trend as seen in Figure 3, indicating that the more stringent matching is working as expected.

worried about inventors leaving and sharing their ideas with competitors. Our study in itself is insufficient to determine which effect dominates.

Second, the implications for the source firm are also nuanced. Although the source firm has lost an employee, that individual's use and diffusion of the ideas that she generated while at the source firm may increase the value of those ideas, some of which the source firm might realize through licensing, sale of the intellectual property, or some type of strategic partnership. In addition, if she maintains ties with individuals in the source firm, then these ties might become a conduit enabling the source firm to better access external ideas (Agrawal and Henderson 2002, Singh 2005, Agrawal and Goldfarb 2008, Agrawal et al. 2008, Corredoira and Rosenkopf 2010).

Third, even the implications for the hiring firm are not straightforward. Our findings temper the "learning-by-hiring" view prevalent in the literature. Although others at the destination firm do learn about the recruit's ideas, we find that the role the recruit and her immediate collaborators play in realizing the destination firm's use of her prior ideas is particularly prominent. This implies that the distribution of bargaining power between the recruit and the destination firm may favor the recruit. The extent to which the firm itself is able to capture rents from better access to the recruit's ideas is therefore not obvious (Becker 1962, Lazear 1986, Coff 1997).

Managers may be tempted to conclude from the above that active steps are needed to ensure that the ideas of new recruits are more widely dispersed and that a smart strategy is therefore to increase investments in systems for sharing knowledge and consequently reduce the firm's dependence on the recruit herself. However, before adopting such a view we must consider why we observe the recruit playing such a prominent role. Although sharing knowledge creates benefits, it also incurs costs. Such costs may arise from an increased need for coordination and the opportunity cost of inventors' time required by the knowledge transfer process. Individuals engage in knowledge sharing up to the point where their private marginal benefit from doing so equals their private marginal cost. Therefore, the absence of knowledge sharing among particular individuals in a firm reflects that the benefits from doing so do not outweigh the costs from the perspective of at least one of the individuals involved. Firms should only intervene in the knowledge-sharing process if they have reason to believe that employees are underinvesting in the sharing of knowledge with each other relative to what is best for the firm. Employee incentives or investments in systems to promote knowledge sharing may be very effective, but managers should keep in mind that pursuing such a strategy relies on the existence of a divergence between the objective of the firm and that of its employees.

More broadly, we must clarify a general limitation to a causal interpretation of our findings. Recall that we construct our sample by identifying patents associated with inventors who subsequently move to another firm, interpret this "mobility event" as the "treatment" and examine its affect on patterns of knowledge flow. But mobility is endogenous, not random; firms make deliberate choices about who to recruit for a reason. So what we estimate is not the "treatment effect" per se (the effect of recruitment on knowledge flow under random assignment of mobility), but rather the "treatment effect on the treated" (the effect of recruitment conditional on a sample of individuals that are actually recruited). Although our use of a longitudinal data set, a DD estimation approach, and more stringent matching moves us closer to a causal interpretation than previous related studies, we have not resolved the causality issue completely.

One must be careful about extending our findings from the treated sample to the overall population because the two may have importantly different properties. Even if the qualitative results reported in the paper hold for the overall population, the magnitude of the effect could be quite different. For example, if firms recruit the inventors whose ideas they value most, then the ideas of the marginal inventor, the "last" inventor to be recruited, may be less valuable than those of the average recruited inventor. This is an important caveat to the simplistic interpretation that firms can increase their access to particular ideas by recruiting the inventor.

In terms of our own interpretation, we prefer using the term "facilitates" rather than "causes." Even if hiring an inventor does not cause the firm to increase its use of that individual's ideas, it may facilitate greater use thereof. However, even this (weaker) interpretation that mobility facilitates knowledge use has limitations. For example, consider a scenario where a firm increases its use of a particular inventor's ideas (without help from the inventor) at exactly the same time that it hires the inventor (for other reasons). Although the increase in idea use would take place even in the absence of recruitment, our DD model is not able to distinguish this scenario from an observationally similar one where the recruitment does facilitate a subsequent rise in idea use. (However, if instead the firm begins to use the ideas *before* the move takes place, then our methodology *will* correctly spot that the move does not drive the boost in the use of the idea.)

Although we have emphasized the benefits of more stringent matching in making progress on some of the above issues, the fact that an appropriate match cannot be found for a significant fraction of the focal patents also raises questions. As already noted, the CEM sample is comprised of patents that, on average, are of lower quality and/or relevance to the destination firm than those that comprise the original sample. On the one hand, this raises concerns regarding generalizability of the findings. On the other hand, if greater premove citation is symptomatic of more severe endogeneity issues, excluding them might be desirable. We have no way of determining which sample is most suitable. However, it is reassuring to note that our (preferred) DD-based estimates for the original sample are not too different from those produced using the CEM sample.

The focus of our study is on how mobility enhances access to premove knowledge associated with the recruit herself. However, while acknowledging previous caveats regarding including a *pair postmove* variable in some of our models, one can (speculatively) interpret the related findings as providing suggestive evidence that hiring facilitates a firm's overall "absorptive capacity" in a knowledge domain (Cohen and Levinthal 1989). Referring back to column (5) in Table 7, note that even the estimate for *pair postmove* is significant (even though smaller than the main DD estimate *postmove*). One possible (conservative) interpretation is that this reflects the destination firm's change in focus toward knowledge domains related to the recruit's expertise (from which we also draw the control patent). However, we can also give this estimated coefficient a (more speculative) interpretation that the recruit is responsible for improving the firm's absorptive capacity in this domain, hence driving the postmove boost in citations made even to the control patent. Indeed, if we carry out regression analyses analogous to Table 8 using matched-pair fixed effects (instead of patent fixed effects) to separately identify a *pair postmove* effect, we would find that the increase in the firm's use of the control patent goes from 0.0095 in terms of *all cites* to 0.0052 in terms of *cites excluding inventor self-cites*, suggesting that the mover is responsible for a significant portion (almost half) of the increase in citations to the control patent as well. This absorptive capacity interpretation is intriguing and worthy of further research.

One promising way forward on the causality issue is to find sources of exogenous variation in mobility, such as closure of establishments (Dahl and Sorenson 2010), changes in noncompete laws (Marx et al. 2009), or death of close colleagues (Azoulay et al. 2010, Oettl 2009). Even in the absence of such natural experiments or instruments, however, progress can be made by more explicitly modeling the likelihood of being exposed to the (endogenous) treatment. For example, Azoulay et al. (2009) apply the "inverse probability of treatment weighted" estimation approach to model a life scientist's selection into patenting behavior by employing rich data on time-varying characteristics of individuals. Researchers could use similar methodologies for modeling mobility events based on observables, although doing so convincingly is likely to involve extensive compilation of individual-level data. Future research on mobility could also benefit from stronger links to the rich literature on job matching (e.g., Jovanovic 1979, Simon and Warner 1992), explicitly incorporating the dynamic process wherein interfirm mobility is a result of individuals and firms optimizing their match in the labor market.

In addition to the econometric issues raised above, we also acknowledge measurement issues that the literature using patent-based measures is still grappling with, and that are beyond the scope of the current study to resolve. First, patent citations are far from perfect as a measure of knowledge flows. One justification for their use is that they have been found to correlate well with actual knowledge flows (Jaffe and Trajtenberg 2002, Chap. 12; Duguet and MacGarvie 2005). Note that using citations as an indicator of knowledge flows entails a weaker assumption than claiming that they are the *mechanism* behind these flows.[27] Nevertheless, it is worth recognizing that citations may be added for reasons such as avoiding litigation or clarifying claims, with a large fraction added by lawyers or patent examiners rather than the inventors themselves (Alcacer and Gittelman 2006). It is still not clear how to interpret such citations. For example, to the extent that inventors have strategic motives for omitting citations, including examiner-added citations might actually be desirable (Lampe 2011). Although we would have liked to do a robustness analysis dropping examiner-added citations, these data are not available in machine-readable form for our study's time period.

Using patent data to detect instances of mobility is not without problems. These data are only effective for identifying instances of mobility where an inventor successfully files patents both before and after a move.[28] Even in a setting (like ours) where one does not need to measure the overall extent of mobility, systematically ignoring certain types of mobility could produce biases—which could go in either direction. For example, omitting individuals that do not patent after a move because their role is just to transfer knowledge to others (e.g., as a technical manager) would lead to a downward bias. On the other hand, omitting individuals who stop patenting and also stop transferring knowledge (e.g., their knowledge becomes outdated or they move into a nontechnical role) would instead produce an upward bias.

Despite the above caveats, our paper does offer a methodological contribution to the study of the mobility–knowledge flow link that is more generally applicable than just the specific question examined here. Other researchers can apply our idea-level, longitudinal, DD framework in examining a broad range of mobility-related questions. For example, one could examine the destination firm's use of not just the recruit's own prior inventions (as in this study), but of *any* prior knowledge that the mobile inventor could serve as a carrier for. This might include access to other knowledge originating from the mover's original firm, network, or geographic context. In addition, one might examine how mobility affects not just the destination firm's use of the recruit's stock of ideas, but also use by the source firm and third parties. Extending beyond just usage of knowledge that existed before a move, researchers could also examine other interesting mobility-related relationships using our approach, such as measuring changes in the productivity of the mover herself or of other employees who might benefit or lose out because of the move (Groysberg et al. 2008). Identification issues are likely to be accentuated when examining such questions (Lacetera et al. 2004), although our framework, perhaps with adjustments, might still be useful.

More generally, any line of inquiry related to the antecedents and consequences of mobility will greatly benefit from investigating longitudinal data and emphasizing temporal patterns that bring us closer to uncovering underlying relationships and microlevel mechanisms. In this paper, we have offered one approach for moving this topic forward. The relationship between inventor mobility and the use of ideas is economically important for both firm strategy and public policy, and we still have much to learn.

6. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

Acknowledgments

The authors contributed equally. They thank INSEAD, the Martin Prosperity Institute, and the Social Sciences and Humanities Research Council of Canada for funding this research. They are grateful to departmental editor Kamalini Ramdas and two anonymous reviewers for

[27] As an analogy, a PhD student's citation of his advisor's research papers may suggest that he built upon the advisor's ideas, even if he actually acquired those ideas by working closely with the advisor rather than by reading his publications.

[28] Because patents only capture a fraction of mobility instances, we do not use these data to estimate the overall impact of mobility. For example, it is inappropriate to use patent data to compare the relative importance of mobility versus alternative channels of knowledge diffusion. In our sample, the fraction of citations received from the destination firm compared to all citations to a recruit's prior patent is 2.1% before a move versus 4.7% after a move (consistent with our main finding that there is a mobility-related jump in citations). However, given the inherent undermeasurement of recruiting activity when employing patent data, such statistics cannot be used to infer the relative importance of mobility versus other channels of knowledge diffusion. Furthermore, we know that mobility shapes the structure of interpersonal networks, which in turn affects knowledge flows (Singh 2005, Fleming et al. 2007, Breschi and Lissoni 2009), yet our method does not capture these indirect effects of mobility.

thoughtful and constructive feedback. They also thank Rajshree Agarwal, James Costantini, Iain Cockburn, Pushan Dutt, Lee Fleming, Jeff Furman, Martin Gargiulo, Javier Gimeno, Avi Goldfarb, Henrich Greve, Bill Kerr, Kwanghui Lim, Matt Marx, Ilian Mihov, Amine Ouazad, Johannes Pennings, Andrew Shipilov, Brian Silverman, Olav Sorenson, Scott Stern, Man Zhang, and seminar participants at Duke University, Georgia Tech, Harvard Business School, INSEAD, National University of Singapore, the NBER Productivity Lunch, the University of North Carolina, and the Wharton Technology Conference for comments. They acknowledge use of the NBER and National University of Singapore–Melbourne Business School patent data sets. Errors remain their own.

References

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

Agarwal, R., R. Echambadi, A. Franco, M. B. Sarkar. 2004. Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Acad. Management J.* **47**(4) 501–522.

Agrawal, A., A. Goldfarb. 2008. Restructuring research: Communication costs and the democratization of university innovation. *Amer. Econom. Rev.* **98**(4) 1578–1590.

Agrawal, A., R. Henderson. 2002. Putting patents in context: Exploring knowledge transfer from MIT. *Management Sci.* **48**(1) 44–60.

Agrawal, A., I. Cockburn, J. McHale. 2006. Gone but not forgotten: Labor flows, knowledge spillovers, and enduring social capital. *J. Econom. Geography* **6**(5) 571–591.

Agrawal, A., D. Kapur, J. McHale. 2008. How do spatial and social proximity influence knowledge flows? Evidence from patent data. *J. Urban Econom.* **64**(2) 258–269.

Alcacer, J., M. Gittelman. 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. *Rev. Econom. Statist.* **88**(4) 774–779.

Allison, P. D., R. P. Waterman. 2002. *Sociol. Methodology* **32**(1) 247–265.

Almeida, P., B. Kogut. 1999. The localization of knowledge and the mobility of engineers in regional networks. *Management Sci.* **45**(7) 905–917.

Angrist, J. D., A. B. Krueger. 1994. Why do World War II veterans earn more than nonveterans? *J. Labor Econom.* **12**(1) 74–97.

Angrist, J. D., J. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

Arrow, K. J. 1962. Economic welfare and the allocation of resources for invention. R. R. Nelson, ed. *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Vol. 13. NBER Special Conference Series, Princeton University Press, Princeton, NJ, 609–625.

Azoulay, P., W. Ding, T. Stuart. 2009. The impact of academic patenting on the rate, quality and direction of (public) research output. *J. Indust. Econom.* **57**(4) 637–676.

Azoulay, P., J. S. Graff Zivin, J. Wang. 2010. Superstar extinction. *Quart. J. Econom.* **125**(2) 549–589.

Barney, J. 1991. Firm resources and sustained competitive advantage. *J. Management* **17**(1) 99–120.

Becker, G. S. 1962. Investment in human capital: A theoretical analysis. *J. Political Econom.* **70**(5) 9–49.

Breschi, S., F. Lissoni. 2009. Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *J. Econom. Geography* **9**(4) 439–468.

Chesbrough, H. 2003. *Open Innovation*. Harvard Business School Press, Boston.

Coff, R. W. 1997. Human assets and management dilemmas: Coping with hazards on the road to resource-based theory. *Acad. Management Rev.* **22**(2) 374–402.

Cohen, W. M., D. A. Levinthal. 1989. Innovation and learning: The two faces of R&D. *Econom. J.* **99**(397) 569–596.

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

Dahl, M. S., O. Sorenson. 2010. The migration of technical workers. *J. Urban Econom.* **67**(1) 33–45.

DiMaggio, P., W. W. Powell. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *Amer. Sociol. Rev.* **48**(2) 147–160.

Dosi, G. 1988. Sources, procedures and microeconomic effects of innovation. *J. Econom. Literature* **26**(3) 1120–1171.

Duguet, E., M. MacGarvie. 2005. How well do patent citations measure knowledge spillovers? Evidence from French innovation surveys. *Econom. Innovation New Tech.* **14**(5) 375–393.

Fleming, L. 2001. Recombinant uncertainty in technological search. *Management Sci.* **47**(1) 117–132.

Fleming, L., C. King, A. Juda. 2007. Small worlds and innovation. *Organ. Sci.* **18**(6) 938–954.

Franco, A. M., M. F. Mitchell. 2008. Covenants not to compete, labor mobility, and industry dynamics. *J. Econom. Management Strategy* **17**(3) 581–606.

Furman, J. L., S. Stern. 2011. Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *Amer. Econom. Rev.* Forthcoming.

Gompers, P., J. Lerner, D. Scharfstein. 2005. Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *J. Finance* **60**(2) 517–614.

Groysberg, B., L.-E. Lee, A. Nanda. 2008. Can they take it with them? The portability of star knowledge workers' performance. *Management Sci.* **54**(7) 1213–1230.

Hall, B. H., J. Mairesse, L. Turner. 2007. Identifying age, cohort and period effects in scientific research productivity: Discussion and illustration using simulated and actual data on French physicists. *Econom. Innovation New Tech.* **16**(2) 159–177.

Hansen, M. T. 1999. The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Admin. Sci. Quart.* **44**(1) 82–111.

Heckman, J., S. Navarro-Lozano. 2004. Using matching, instrumental variables, and control functions to estimate economic choice models. *Rev. Econom. Statist.* **86**(1) 30–67.

Hoisl, K. 2007. Tracing mobile inventors—The causality between inventor mobility and inventor productivity. *Res. Policy* **36**(5) 619–636.

Iacus, S. M., G. King, G. Porro. 2009. Causal inference without balance checking: Coarsened exact matching. Working paper, Harvard University, Cambridge, MA.

Imbens, G. W. 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *Rev. Econom. Statist.* **86**(1) 4–29.

Jaffe, A. B., M. Trajtenberg. 2002. *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. MIT Press, Cambridge, MA.

Jaffe, A. B., M. Trajtenberg, R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quart. J. Econom.* **108**(3) 577–598.

Jovanovic, B. 1979. Job matching and the theory of turnover. *J. Political Econom.* **87**(5) 972–990.

Klepper, S., S. Sleeper. 2005. Entry by spin-offs. *Management Sci.* **51**(8) 1291–1306.

Lacetera, N., I. M. Cockburn, R. M. Henderson. 2004. Do firms change capabilities by hiring new people? A study of the adoption of science-based drug discovery. A. M. McGahan, J. A. C. Baum, eds. *Advances in Strategic Management*, Vol. 21. JAI Press, Greenwich, CT, 133–159.

Lampe, R. 2011. Strategic citation. *Rev. Econom. Statist.* Forthcoming.

Lazear, E. P. 1986. Raids and offer matching. *Res. Labor Econom.* **8** 141–165.

Levin, S. G., P. E. Stephan. 1991. Research productivity over the life cycle: Evidence for academic scientists. *Amer. Econom. Rev.* **81**(1) 114–132.

Llorente, G., R. Michaely, G. Saar, J. Wang. 2002. Dynamic volume-return relation of individual stocks. *Rev. Financial Stud.* **15**(4) 1005–1047.

Malamud, O., C. Pop-Eleches. 2010. General education versus vocational training: Evidence from an economy in transition. *Rev. Econom. Statist.* **92**(1) 43–60.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organ. Sci.* **2**(1) 71–87.

Marx, M., D. Strumsky, L. Fleming. 2009. Mobility, skills, and the Michigan non-compete experiment. *Management Sci.* **55**(6) 875–889.

Mehta, A., M. Rysman, T. Simcoe. 2009. Identifying the age profile of patent citations: New estimates of knowledge diffusion. *J. Appl. Econometrics*, ePub ahead of print September 1, http://onlinelibrary.wiley.com/doi/10.1002/jae.1086/abstract.

Moen, J. 2005. Is mobility of technical personnel a source of R&D spillovers? *J. Labor Econom.* **23**(1) 81–114.

Moffitt, R. A. 2004. Introduction to the symposium on the econometrics of matching. *Rev. Econom. Statist.* **86**(1) 1–3.

Murray, F., S. Stern. 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *J. Econom. Behav. Organ.* **63**(4) 648–687.

Nelson, R. R., S. G. Winter. 1982. An evolutionary theory of economic change. Belknap, Cambridge, MA.

Oettl, A. 2009. Productivity and helpfulness: A new taxonomy for star scientists. Working paper, Georgia Institute of Technology, Atlanta.

Oettl, A., A. Agrawal. 2008. International labor mobility and knowledge flow externalities. *J. Internat. Bus. Stud.* **39** 1242–1260.

Penrose, E. T. 1959. *The Theory of the Growth of the Firm*. Blackwell, Oxford, UK.

Peteraf, M. A. 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management J.* **14**(3) 179–191.

Rosenkopf, L., P. Almeida. 2003. Overcoming local search through alliances and mobility. *Management Sci.* **49**(6) 0751–0766.

Rysman, M., T. Simcoe. 2008. Patents and the performance of voluntary standard-setting organizations. *Management Sci.* **54**(11) 1920–1934.

Samila, S., O. Sorenson. 2011. Non-compete covenants: Incentives to innovate or impediments to growth. *Management Sci.* Forthcoming.

Saxenian, A. L. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128.* Harvard University Press, Cambridge, MA.

Simon, C. J., J. Warner. 1992. Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure. *J. Labor Econom.* **10**(3) 306–330.

Simon, H. A. 1991. Bounded rationality and organizational learning. *Organ. Sci.* **2**(1) 125–134.

Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Sci.* **51**(5) 756–770.

Singh, J. 2007. Asymmetry of knowledge spillovers between MNCs and host country firms. *J. Internat. Bus. Stud.* **38**(5) 764–786.

Singh, J. 2008. Distributed R&D, cross-regional knowledge integration and quality of innovative output. *Res. Policy* **37**(1) 77–96.

Singh, J., L. Fleming. 2010. Lone inventors as sources of breakthroughs: Myth or reality? *Management Sci.* **56**(1) 41–56.

Singh, J., M. T. Hansen, J. M. Podolny. 2010. The world is not small for everyone: Inequity in searching for knowledge in organizations. *Management Sci.* **56**(9) 1415–1438.

Song, J., P. Almeida, G. Wu. 2003. Learning by hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Sci.* **49**(4) 351–365.

Thompson, P., M. Fox-Kean. 2005. Patent citations and the geography of knowledge spillovers: A reassessment. *Amer. Econom. Rev.* **95**(1) 450–460.

Trajtenberg, M. 2006. The "names game": Harnessing inventors' patent data for economic research. NBER Working Paper 12479, National Bureau of Economic Research, Cambridge, MA.

Tzabbar, D., B. S. Silverman, B. S. Aharonson. 2009. Learning by hiring or hiring to avoid learning? Organizational exploration via individuals' exploitation. Working paper, University of Central Florida, Orlando.

Weitzman, M. 1998. Recombinant growth. *Quart. J. Econom.* **113**(2) 331–360.

Wooldridge, J. M. 1999. Distribution-free estimation of some nonlinear panel data models. *J. Econometrics* **90**(1) 77–97.