

NBER WORKING PAPER SERIES

RECRUITING FOR IDEAS:
HOW FIRMS EXPLOIT THE PRIOR INVENTIONS OF NEW HIRES

Jasjit Singh
Ajay K. Agrawal

Working Paper 15869
<http://www.nber.org/papers/w15869>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2010

The authors contributed equally. We thank INSEAD, the Martin Prosperity Institute, and the Social Sciences and Humanities Research Council of Canada for funding this research. We are grateful to James Costantini, Pushan Dutt, Lee Fleming, Martin Gargiulo, Javier Gimeno, Henrich Greve, Kwanghui Lim, Matt Marx, Illian Mihov, Amine Ouazad, Johannes Pennings, Kamalini Ramdas, Andrew Shipilov, Brain Silverman, Olav Sorenson, Man Zhang, two anonymous referees, and seminar participants at INSEAD and the National University of Singapore for comments. We acknowledge use of the NBER and NUS-MBS patent datasets. Errors remain our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2010 by Jasjit Singh and Ajay K. Agrawal. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires
Jasjit Singh and Ajay K. Agrawal
NBER Working Paper No. 15869
April 2010
JEL No. O31,O32,O33,O34

ABSTRACT

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 the new recruits' prior inventions? Our estimates suggest they do, quite significantly in fact, by approximately 202% 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. 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' pre-move patents. Specifically, we employ a difference-in-differences approach to compare pre-move versus post-move citation rates for the recruits' prior patents and the 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.

Jasjit Singh
INSEAD
3'C{gt'Tclcj 'Cxgpwg
Ukpi cr qtg'35: 898
jasjit.singh@insead.edu

Ajay K. Agrawal
Rotman School of Management
University of Toronto
105 St. George Street
Toronto, Ontario M5S 3E6
CANADA
and NBER
ajay.agrawal@rotman.utoronto.ca

1. INTRODUCTION

The link between recruiting inventors and using their 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 towards exploitation of familiar knowledge with deliberate mechanisms that facilitate exploration of distant knowledge (March 1991), especially since combining ideas drawn from different sources is important for innovation success (Weitzman 1998, Fleming 2001, Chesbrough 2003). Thus, recruiting an individual from outside the organization may enhance a firm's access to external ideas and thus better enable it to complement exploitation of native ideas with exploration of foreign ideas. But how much do firms really increase their use of a new recruit's stock of prior ideas? Furthermore, to the extent they do, how do they do it? To what extent do firms really exhibit "learning-by-hiring"? Moreover, given the temporal nature of diffusion, how does this process evolve over time after the new recruit settles in? We set out to address these questions.

Scholars from a variety of schools of thought have suggested that inter-firm mobility could be a key mechanism driving diffusion of ideas across firms. Noting that knowledge spillovers through mobility often take place even despite the source firm's use of legal measures to prevent them, economist Kenneth Arrow (1962: 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, such as DiMaggio and Powell (1983), suggest that inter-firm 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).¹

At an aggregate level, scholars have also linked regional economic growth to enhanced access to ideas afforded by inter-firm mobility. In her monograph on regional advantage, Saxenian (1994: 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 inter-firm mobility due to non-compete covenants find support for this view (Franco and Mitchell 2008, Marx et al. 2009, Samila and Sorenson 2009). Furthermore, some formal models of economic growth explicitly cite the possibility of knowledge spillovers through inter-firm mobility as one way of justifying a key assumption regarding technological knowledge being at least in part a “non-excludable” good (Grossman and Helpman 1991).

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 intra-regional labor mobility between firms tend to have more localized knowledge flows. In another pioneering study, Song et al. (2003) illustrate that mobile inventors build upon ideas from their previous firms more often than inventors at the hiring firm who have not previously worked at that other firm. In yet another influential article, Rosenkopf and Almeida (2003) examine firm pairs and show that dyads that experience more

¹ Our explicit focus is on an individual moving from one firm to another *existing* firm. It is important to note that an extensive literature also emphasizes knowledge transfer through mobility in the context of new firm formation (Agarwal et al. 2004, Gompers et al. 2005, Klepper and Sleeper 2005).

labor mobility between them also demonstrate greater subsequent knowledge flow. Taken together, these empirical studies have inspired further research employing similar data and methods to sharpen our understanding of various aspects of the mobility-knowledge flow relationship (e.g., Agrawal et al. 2006, Oettl and Agrawal 2008, Agarwal et al. 2009, Corredoira and Rosenkopf 2010).²

The above studies have advanced the field significantly by establishing an empirical link between mobility and knowledge flow and providing insights into particular nuances of this relationship. However, they are subject to two related limitations associated with drawing causal inferences from cross-sectional data on inventor mobility and knowledge use across patents. The first limitation concerns unobserved heterogeneity. The destination firm may be more likely to use an idea that is of inherently higher quality or has greater firm-specific relevance, irrespective of whether the destination firm hires the inventor. However, since the inventor is also likely to be a more attractive target for recruitment, this presents a selection problem wherein the observed cross-sectional correlation between mobility and knowledge use overestimates the true effect of mobility. The second concerns the endogeneity of a firm's decision to hire. 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 of the latter. Not accounting for such endogeneity will once more lead to an upward bias if we interpret the estimates as boosts in the use of ideas caused by an instance of mobility.³

Recognizing these inherent limitations of a cross-sectional research design for making causal inferences, Rosenkopf and Almeida (2003: 764) offer this challenge: "Future research should attempt to

² While we focus on the direct effect of inter-firm mobility on knowledge transfer, it is worth noting that mobility also influences the structure of interpersonal networks within and across firms, which in turn play an important role in further shaping knowledge diffusion patterns (Singh 2005, Fleming et al. 2007, Breschi and Lissoni 2009).

³ 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, as that is surely too coarse to sufficiently capture all relevant characteristics of the underlying knowledge.

utilize fully developed longitudinal databases to explore all possible temporal and causal links.” We take up that challenge here and make significant progress on the first problem (unobserved heterogeneity) and at least some progress on the second (endogeneity).

So, in addition to offering an empirical contribution in terms of estimating the boost in use of a recruit’s prior ideas as well as the fraction of that boost that is due to the recruit herself, our methodology advances existing research on mobility and knowledge flows in three significant ways. First, we base our analysis not on aggregate firm-level citation counts but on disaggregated data that capture changes over time in how individuals in recruiting firms cite a specific patent. Second, we employ a “difference-in-differences” (DD) approach to account for heterogeneity across patents. This avoids making the typical (strong) assumption regarding comparability of the *levels* of citations received by “focal” patents (i.e., those involving an inventor who subsequently moves) and “control” patents (i.e., other similar patents). Instead, our identification strategy relies upon a weaker (and hence more defensible) assumption regarding the legitimacy of comparing *relative changes in* citation rates over time (in particular from the pre-move to the post-move period). Third, to further address concerns regarding comparability of the focal and control patents even despite these weaker assumptions, we replicate our initial analysis (which relies upon a conventional matching approach) using a sample based on a more stringent matching procedure that also takes into account the patent’s pre-move citation patterns and the inventor’s prior patenting history in determining the control sample.⁴

We begin our analysis with a sample we construct using the matching criteria most prior studies employ to identify control patents that are used as a baseline against which to compare focal patents: the three-digit technology classification and the application year.⁵ The typical cross-sectional comparison of the post-move citation levels for focal versus control patents rests on the (often unstated) assumption that

⁴ Admittedly, even our exploitation of longitudinal data, implementation of DD estimation, and use of more stringent matching do not fully establish causality. As discussed in the concluding section, doing so would require a suitable instrument or a natural experiment that randomly assigns the “treatment” (in our case, the “mobility event”).

⁵ 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 any such approach.

the above technology-based match suffices in dealing with the heterogeneity issue mentioned above. However, we find evidence that this assumption is not reasonable; destination firms cite focal patents at a higher rate than matched control patents even *before* the move. In recognition of this systematic difference between focal and control patents, we use the DD approach to essentially “difference out” the pre-move citation premium associated with focal patents. This allows us to distinguish the component of the post-move 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).

To further address concerns regarding the closeness of matching that acts as our baseline, we repeat our DD estimation using a different control sample constructed by employing a more stringent matching procedure. We implement this using “coarsened exact matching” (CEM) (Iacus et al. 2009). Among the additional matching criteria is information related to the patents’ pre-move citations received as well as the inventors’ career histories. As the extensive literature on matching emphasizes, using appropriately stringent matching may reduce endogeneity concerns (though can rarely eliminate them completely) while also reducing the sensitivity of the subsequent regression-based estimation on specific functional form assumptions. In our case, including pre-move citation counts (overall as well as from the destination firm) among these criteria helps to eliminate the pre-move citation rate differences between focal and control patents, improving their comparability in terms of expected future citation rates.

We find that hiring an inventor is associated with a firm subsequently increasing its use of the new recruit’s prior ideas significantly: by approximately 202% on average (even after controlling for changes in the firm’s propensity to cite other similar inventions). In absolute terms, the destination firm increases its annual citation rate from 0.0140 to 0.0396 citations per year for each of the recruit’s prior ideas.⁶ We show that this result is robust to a variety of specifications. In fact, we find the DD estimates using the more stringent matching to be essentially the same qualitatively and very similar quantitatively.

⁶ For the reader who might be wondering why these absolute values look so small, note that these represent citations from a *particular firm* to a *particular patent* in a *particular year*.

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 a part of the firm's overall knowledge base shortly upon the recruit's arrival. However, Simon (1991: 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. Our results support Simon's observation. 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 prior ideas after arriving at her new firm.⁷ Furthermore, when we exclude citations by the recruit's collaborators, we find that other inventors account for only a third of the overall boost realized by the hiring firm. When we further examine the temporal patterns of citations, expecting that the role of the recruit might start to diminish shortly after she arrives as her tacit knowledge diffuses throughout the destination firm, we instead find that her role remains quite persistent over time.

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

⁷Our results are consistent with the findings of Tzabbar et al. (2009), who report that having the recruit on an inventing team in the destination firm significantly increases the likelihood of the team exploiting one of the recruit's prior inventions.

piece of knowledge may confer significant bargaining power to that individual, especially if she also has other outside options (Becker 1962, Lazear 1986, Coff 1997, Moen 2005, Groysberg et al. 2008). The significant role of the recruit in building upon her prior ideas might therefore imply that the recruit rather than the hiring firm captures much of the rents associated with the firm's seemingly better access to her prior knowledge. This echoes Peteraf's (1993: 187) remark: "A 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."

We organize the remainder of the paper as follows. In Section 2, we outline our empirical framework and distinguish it from the traditional cross-sectional approach. In Section 3, we describe the construction of our two datasets, one corresponding to the conventional matching and the other based on more stringent matching criteria. In Section 4, we present our empirical results. Finally, in Section 5, we discuss 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

Micro-level data suitable for examining the link between mobility and knowledge flow is hard to come by. A notable exception is patent data, which researchers have commonly employed for this purpose as they include detailed information on patent characteristics including inventor names, application and grant dates, assignee organization name (if any), technological classification, etc. In particular, patent data enable the researcher to infer inter-firm mobility by chronologically tracing individual inventors as they appear on patents assigned to different firms at different times.

In addition, data on the citations a patent receives allow the researcher to infer when a subsequent invention has used a particular idea.⁸ Admittedly, citations are not a perfect measure of

⁸ Using citations as a *measure* of knowledge flows does not assume that citations are a *mechanism* behind these flows. As an analogy, a PhD student's citations of his advisor's research papers suggests that he has built upon the

knowledge flows. For example, they are often added for reasons such as avoiding litigation or clarifying claims, and many are even added by patent examiners rather than the inventors themselves. Despite this, however, scholars have shown that they correlate quite well with actual knowledge flow, especially when employing large samples (Chapter 12 in Jaffe and Trajtenberg 2002, Duguet and MacGarvie 2005).⁹

2.2. The cross-sectional approach

We begin our empirical analysis by implementing a method of examining the link between mobility and knowledge flow that follows the “best practice” from existing studies. First, we 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 time 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 post-move period) is:

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

Here, $RECRUIT_i$ is an indicator variable that equals one for observations corresponding to a focal patent and zero for the corresponding control patent. Rather than assuming a specific functional form for the temporal patterns of citations, the above model follows a non-parametric approach by more flexibly accounting for patent age (using yearly indicator variables for the gap between patent i 's application year

advisor's ideas, even if he actually has acquired those ideas by working closely with the advisor rather than by reading his papers.

⁹ A specific concern is whether including examiner-added citations biases results (Alcacer and Gittelman 2006). However, to the extent that inventors have strategic motives for omitting citations, including examiner-added citations is actually desirable (Lampe 2008). While 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.

and the citing year t being considered) and citing year (using different indicator variables for each calendar year t).

The baseline expectation is that ψ_F should be positive and significant. In other words, patents by individuals who subsequently move from their original firm (the source firm) to the destination firm will be more highly cited (by inventors at the destination firm after the mobility event) than 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. We describe these variables in Section 3.2 below and list them (except for technology area indicators) for easy reference in Table 1.

Unfortunately, a cross-sectional approach like the one above is not satisfactory for dealing with the 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 sought after as new recruits. 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 that component of the boost in knowledge use that is directly attributable to the move, we next turn to a “difference-in-differences” research design that allows us to distinguish between the component of the boost in post-move citation rate that is more likely due to the “treatment” (attributable to the move) and the component that is instead due to “selection” (attributable to the kind of inventor who is more likely to move).

2.3. Our “difference-in-differences” approach

Our “difference-in-differences” (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 post-move but also in the years preceding the move. While post-move differences confound actual mobility-related knowledge acquisition with differences in quality and/or relevance, we can at least partly identify these

effects by taking into account differences in citation rates that exist even pre-move. In other words, we can use a pre-move difference in citation rates for focal versus control patents as a benchmark against which to compare the post-move difference, helping us better identify the component of the post-move difference that we can attribute to the move itself.¹⁰

We provide Figure 1 to illustrate how the DD approach conceptually differs from the cross-sectional approach described above. For ease of exposition, the figure depicts the effect of mobility as a sudden jump in what would have been the otherwise natural evolution of the citation rate of the focal patent over time (our actual empirical analysis relaxes this assumption).¹¹ The cross-sectional approach above interprets a positive difference in post-mobility citation rates for the recruit’s patent (R_a) versus control patent (C_a) as evidence of knowledge flow attributable to the mobility event. However, that is misleading because had the move not taken place, then the expected citation rate to the recruit’s patent (R_e) would still be greater than that for the control patent (C_e). Comparing citation rates between the recruit and control patents both before and after the move allows us to better estimate how the temporal pattern in citation rate changes from the extrapolated path for the post-move period. In other words, the DD logic suggests that we can better capture the effect attributable to the mobility event not by the extent to which $R_a > C_a$ but by the extent to which $R_a - R_e > C_a - C_e$.

Implementing the above logic needs a patent-year dataset that includes observations from the years not just after but also before the move. Defining $POSTMOVE_{i,t}$ as zero when the year t falls in the pre-move period and one when it falls in the post-move period for a patent i (focal or control), the estimation equation becomes:

$$(2) \quad CITES_{i,t} = f(\psi_R RECRUIT_i + \psi_{RP} RECRUIT_i * POSTMOVE_{i,t} + \psi_P POSTMOVE_{i,t} + \delta_{t=appyear(i)} + \beta_t + \psi_X X_i + \varepsilon_{i,t})$$

¹⁰ For a good overview of the “difference-in-differences” methodology, see Chapter 5 in Angrist and Pischke (2009). See Furman and Stern (2006) and Murray and Stern (2007) for instructive applications in the context of citations.

¹¹ Note that we allow for the possibility of a post-mobility bump even for the control patent. This could happen, for example, if the move coincides with an overall shift in a firm’s technological focus, resulting in a general increase in citation rate for the given technology. In the discussion section, we offer alternate interpretations.

Here, ψ_R captures systematic differences in the focal versus control patents that exist even before the move, and ψ_P captures systematic shifts in how the destination firm increases its use of ideas in this technology area between the pre- and post-move periods, irrespective of recruiting that particular inventor. If mobility really does lead to an increased use in the recruit's prior knowledge, ψ_{RP} should be positive and significant in the estimation.

While the above framework allows for a systematic difference between focal and control patents, we can generalize it further to allow even individual patents to be different in unobserved ways (e.g., due to inventor characteristics we do not observe). In particular, we employ patent fixed effects analysis to detect “abnormal” within-patent changes in the citation rate to a patent after an inventor moves. The estimation relies on a deviation in the patent's post-mobility citation rate from an expected rate, where we derive the expected rate by extrapolating from the pre-move citation rate and assuming a temporal trend analogous to other patents from the same technology-year cohort. We express this model as:

$$(3) \quad CITES_{i,t} = f(\psi_{RP} RECRUIT_i * POSTMOVE_{i,t} + \psi_P POSTMOVE_{i,t} + \delta_{t-appyear(i)} + \beta_t + \gamma_i + \varepsilon_{i,t})$$

Here, γ_i reflects the fixed (time invariant) effect corresponding to patent i . Since the fixed effect absorbs unique patent characteristics, note that 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} as it is based on an interaction effect.¹²

2.4. A more stringent matching approach

Intuitively, it is helpful to think of the above DD research design as replacing the conventional assumption of direct comparability between focal and control patents of post-move citation rate levels with the weaker assumption of comparability in terms of pre- and post-move citation rate trends.

¹² See Murray and Stern (2007) for a similar specification. Since the lag between the patent application and the move varies across focal patents, it is possible to carry out the estimation using just these patents. While we report findings from such an analysis as well (effectively estimating the difference between R_a and R_c in Figure 1), our preferred approach is to use a sample that includes appropriately matched control patents.

However, even this weaker identifying assumption is subject to criticism given that the underlying sample based on the conventional matching criterion (three-digit technology class and application year) might not produce a control sample that is sufficiently similar to the focal patents. In particular, since inter-firm mobility is an endogenous event, it seems particularly important to construct a control sample that matches the focal sample as closely as possible on observed characteristics.

As we describe in detail in Section 3 below, we employ a CEM procedure to construct a control sample based on discrete buckets for pre-move citations received by the patent (both overall and from the destination firm) as well as the patenting history of the inventor (in terms of the number of past patents and age in terms of years since the first patent).¹³ Note that we employ these criteria *in addition to* exact matching (as before) on the conventional criteria of application year and three-digit technology class. We also use the grant delay (i.e., the lag between the application and grant date) as yet another matching criterion for the CEM. This improves comparability of the focal and control patents by allowing for the possibility that the grant delay itself may be a result of important (and possibly unobserved) patent-level characteristics and/or that the start of a patent's "citation clock" is more appropriately modeled using its grant date rather than its application date (Mehta et al. 2010).

2.5. Our preferred specification

Since patent citations involve count data skewed to the right (and over-dispersed relative to Poisson), scholars commonly employ negative binomial models for estimating parameters. For comparability with previous research, we therefore start with negative binomial regressions. Our models employ robust standard errors with clustering on the inventor in order to account for non-independence 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 (OLS) models, which have the benefit of allowing more fine-grained indicator variables for technology (at the three-digit level rather

¹³ For a description of CEM, see Iacus et al. (2009). For a nice application, see Azoulay et al. (2010).

than two-digit level practical for non-linear 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). We first try implementing the above using a conditional fixed effects Poisson framework (Wooldridge 1999).¹⁴ However, in addition to the usual challenges with interpreting DD coefficients in non-linear models (especially when comparing across different models), we run into two issues specific to our setting. First, the model does not converge when we include yearly indicator variables for the patent age and the calendar year of citation to better saturate the models. Second, we must drop a very large fraction of the sample due to no variation in the dependent variable (because of zero citations), and the observations dropped are disproportionately control patents. As a result, our preferred approach is to employ fully saturated linear regression specifications with patent fixed effects.¹⁵

3. DATASET CONSTRUCTION

3.1. Constructing matched samples of focal and control patents

We merge patent data obtained directly from the USPTO with patent data made available by the National Bureau of Economic Research (Chapter 13 in Jaffe and Trajtenberg 2002) and the National University of Singapore-Melbourne Business School patent dataset. We then enhance these along two dimensions. First, for each assigned patent, we determine the assignee organization by

¹⁴ We implement this in Stata using the “xtqml” procedure (written by Tim Simcoe and available for download at <http://people.bu.edu/tsimcoe/code/xtqml.txt>), which corrects the standard errors from a fixed effects Poisson model for over-dispersion. This procedure addresses the concerns regarding interpreting a conditional fixed effects negative binomial model as a true fixed effects estimator (Wooldridge 1999, Allison and Waterman 2002).

¹⁵ We verify that our findings are qualitatively robust to using pooled negative binomial estimation or Poisson conditional fixed effects estimation, even though this typically means that the specification is not saturated as well and/or that we drop a significant fraction of the sample. For further discussion on the tradeoffs involved in employing linear models versus more sophisticated regression techniques, see Angrist and Pischke (2009).

carrying out an assignee name cleanup followed by a parent-subsidary match.¹⁶ Second, we use not just inventor names but also other data fields (i.e., technology classification, inventor address, collaborator names, citation information) to create a unique identifier for each inventor on all patents.¹⁷

In order to obtain a dataset of reasonable size while allowing for a sufficient future time window for observing subsequent inventor mobility and citations, we start with the population of patents with application years 1981-1990 (across all technology classes). In detecting instances of subsequent inter-firm mobility, we follow other researchers and infer mobility through observed changes in the assignee firm on successive patents filed by an inventor (Almeida and Kogut 1999, Song et al. 2003).¹⁸

We restrict our sample (both focal and control) to patents with a single inventor to facilitate straightforward comparability across patents (all have the same number of inventors: one) and an unambiguous conceptualization of inter-firm mobility (or lack thereof) for any given patent. From this subset, we draw our sample of focal patents for which the inventor exhibits inter-firm mobility anytime between (and including) the third and tenth year following the application year (we exclude the first and last two years so that there is at least some pre- and post-move citation data for each patent in order to facilitate meaningful DD calculation even when we employ cited patent fixed effects).

One challenge is that, even when we do 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

¹⁶ We follow the assignee matching procedure used by Singh (2005), who relies upon NBER *Compustat* identifiers, Stopford's *Directory of Multinationals*, *Who Owns Who* directories, and Internet sources. To further reduce any possibility of misclassifying different spellings, name changes, acquisitions, etc. as instances of mobility, we have had a research assistant manually double-check instances of algorithmically detected mobility.

¹⁷ We base our name matching approach on Singh (2008), whose algorithms are similar to procedures implemented by Trajtenberg (2006) and Fleming et al. (2007).

¹⁸ Patent data is only effective for identifying instances of mobility where an inventor successfully files for patents both pre- and post-move. This is sufficient in a setting (like ours) where one does not need to measure the overall extent of mobility. However, even in our setting, systematically ignoring certain types of mobility could produce biases in either direction. One could imagine cases where a recruit does not patent post-move because her role is now just to transfer knowledge to others (e.g., as a technical manager), in which omitting her from the sample leads to a downward bias. On the other hand, if she stops patenting and also stops transferring knowledge (e.g., her knowledge becomes outdated or she moves into a non-technical role), then not including her would produce an upward bias.

or less, as 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 inter-firm mobility events.¹⁹

Lacking precise 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, while the move could have taken place anytime after the start of this window, it would probably have taken place at least a few months before the end of the window due to a lag between an inventor joining a new firm and filing for a patent there. Further, we take the temporal unit of analysis to be the year, rather than the day or month, to avoid any pretence of a precise estimated move date. These two factors lead us to define the beginning of the calendar year of the window midpoint calculated above as our estimate of when a move happened. Given the high uncertainty in the move date, this somewhat ad-hoc assumption is unlikely to do much worse than more sophisticated heuristics. On the other hand, 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 (see Figure 2). We choose the control patent such that it originates in another firm, it also has a single inventor, and its inventor does not exhibit any inter-firm mobility in the 12 years that follow.²⁰ Having a control patent for each mobile inventor patent helps account for general shifts in the technological focus of the destination firm, as those will be reflected in an increased likelihood of citing not just the original 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 patents) of which are associated

¹⁹ 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, since 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 the paper using different window cutoffs. The main results remain qualitatively unchanged.

²⁰ What makes same-firm patents inappropriate as controls is the fact that the mobile inventor is very 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.

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 (see Figure 3). Having 12 observations for each of the above patents results in a patent-year panel of 95,424 observations (we call this the “original sample”).

One might worry that the conventional matching procedure used above might not produce a set of control patents that is sufficiently comparable to the focal patents, despite the weaker comparability requirements inherent in our DD approach (comparison of trends) relative to the commonly employed cross-sectional approach (comparison of levels). A more stringent matching procedure would offer 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, we would reduce concerns about bias due to endogenous treatment (though not eliminate them, as it is impractical to capture all relevant drivers of the move).²¹ Second, more stringent matching would also reduce sensitivity of the findings to specific functional form assumptions. As Moffitt (2004: 1) notes: “compared to least squares estimation of the relationship, with its attendant assumptions of linearity and additivity that can only awkwardly be relaxed in full, matching provides a method by which no functional form restrictions on the relation between [the variables] need be made.”²²

To achieve more stringent matching, we employ the CEM procedure (Iacus et al. 2009) and construct another sample that relies upon not just a match on the three-digit technology class and application year but also on discrete buckets based on five additional criteria: 1) overall pre-move citations received by a patent (four buckets: zero citations, one-two citations, three-seven citations, and eight or more citations), 2) pre-move citations received from the destination firm (four buckets: zero

²¹ The consistency of matching-based estimates relies heavily upon a “selection on observables” assumption. Since the likelihood of treatment might in reality also depend on unobservables or on observables that are impractical to fully match on, any matching-based procedure (in the absence of an exogenous instrument or a natural experiment) can only reduce endogeneity concerns rather than completely eliminate them (Heckman and Navarro-Lozano 2004).

²² The potential benefits from stringent matching must be weighed against the reduced likelihood of finding a match and thus the bias that could result if the focal patents that we match are systematically different from the ones that we drop. In the results section, we present evidence that this is a real concern in our setting. In the concluding section, we return to the implications of and a broader discussion of this point.

citations, one citation, two-three citations, and four or more citations), 3) number of previous patents by this inventor (three buckets: zero or one patents, two-12 patents, and 13 or more patents), 4) number of years since the first patent by this inventor (three buckets: zero or one years, two-nine years, and 10 or more years), and 5) the delay between the application and grant date (three buckets: 490 or less days, 491-832 days, and 833 or more days).²³

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 X 2) cited patents. Again following a procedure similar to the one described above (and illustrated in Figure 3), we construct a patent-year dataset of 61,032 (5,086 X 12) observations, which we call the “CEM sample.” Later in the paper we present statistics to demonstrate that the “CEM sample” is indeed better matched than the “original sample” and therefore use it for carrying out detailed analyses of how firms exploit the prior inventions of new hires.

3.2. Variables

Dependent Variables

Table 1 summarizes our dependent variables as well as other key variables.²⁴ 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. This distinction is important since abundant self-exploitation by a recruit may reflect a setting where the

²³ Choosing the matching criteria involves a tradeoff 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 pre-move 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.

²⁴ To understand why mean values for the dependent variables appear so small, recall that we are dealing with highly disaggregated data: these variables capture citations received by a patent from the firm to which the mobile inventor moves. Naturally, the probability of a *specific firm* citing a *specific patent* in a *specific year* is extremely low. For example, *all cites* has a non-zero value for only around 1.8% of the observations in the “original sample.”

firm is less able to appropriate the benefits from hiring an inventor with a valuable stock of ideas since its ability to exploit her idea may depend on her staying there.

To the extent that the recruit's tacit knowledge does diffuse through the destination firm, we are interested in estimating how widely this occurs. On the one hand, close interpersonal ties such as those formed by collaboration can be very helpful for knowledge diffusion (Hansen 1999, Singh 2005). On the other hand, broader intra-firm interpersonal networks or knowledge diffusion mechanisms might enable and encourage a wider employee base in the destination firm to build upon the recruit's prior ideas. To examine this issue, we construct another dependent variable, *cites excluding inventor self-cites and collaborator cites*, as the count of citations made by those who have also not directly collaborated (as co-authors on previous patents) with the inventor before the citing year under consideration.

Key Explanatory Variables

Our two key explanatory variables are *recruit* and *post-move*. In fact, the estimated coefficient on the interaction of these terms is the focus of our empirical analysis. We use the indicator variable *recruit* to classify 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 as post-move if it occurs after the estimated move date using the indicator variable *post-move*.

Control Variables

For the pooled estimation models, we employ several control variables, though these are not needed in our preferred models that employ patent fixed effects (as they are time-invariant). Specifically, we include controls 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),²⁵ 2) the source firm (overall level of inventive activity), 3) the inventor (number of prior patents, number of years since

²⁵ It is important to control for the destination firm's patent pool since having more patents naturally increases citation likelihood to the focal patent, with this tendency being greatest for firms with more patenting activity in the same technology. However, 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).

first patent, U.S. resident indicator), and 4) the patent itself (number of claims, number of patent references, number of non-patent references, technology area).²⁶ We provide definitions and descriptive statistics for these control variables (except the technology area indicators) in Table 1.

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 estimate of the persistence over time of the recruit as the destination firm’s primary user of her stock of prior ideas.

4.1. Summary Statistics

Before delving into the regression analysis, we present the basic intuition behind our approach using summary statistics for the pre-move and post-move subsamples corresponding to the focal versus control patents. We report these statistics in Table 2. (It is also helpful to refer to Figures 1 and 3 for interpreting these.)

We begin with the means reported for the variable *all cites* for each of the four subsamples. Recall from Table 1 that this variable is a count of *all* citations the focal (or control) patent receives from the destination firm in a given year, including self-citations made by the inventor after she moves. As the column corresponding to the post-move period indicates, 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

²⁶ While the pooled negative binomial models employ technology indicators at the one-digit NBER category level, the pooled linear models employ indicators for the three-digit technology class. Also, we only use *inventor age*, *U.S. inventor*, and *grant delay* in the pooled negative binomial models, as 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 discussion 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. (2010).

between the two. However, it would be misleading to attribute all of this difference to mobility since the annual citation rate is greater even pre-mobility (0.0140 for recruits' patents versus 0.0051 for controls, reflecting a difference of 0.0089). Therefore, rather than attributing the whole post-mobility difference to the move itself, it seems more reasonable to attribute to the move only the difference, 0.0246 (i.e., 0.0335 – 0.0089), between the post-mobility and pre-mobility differences.

We now turn to the analogous calculation for the variable *cites excluding inventor self-cites*, which excludes self-citations made by the mobile inventor to her own patent. The post-move 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 much smaller than the corresponding *all cites* DD statistic of 0.0246. In other words, interpreting overall changes in citation patterns as all being “learning-by-hiring” would be a clear overestimation.

4.2. Regression Analysis: Recruiting and the Citation Premium

Since the above statistics are only meant to illustrate the main intuition behind our DD approach, we now turn to the regression framework, which is more appropriate for a rigorous analysis. Table 3 reports our baseline analysis using the “original sample,” employing *all cites* as the dependent variable. Column (1) reports cross-sectional findings from the post-move subsample, in effect replicating the current “best practice” of employing such comparisons of the post-move citation frequency for focal versus control patents.²⁷ 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.

²⁷ For variables that are highly skewed, the pooled regression analysis employs a logarithmic transformation, first adding one to allow the transformation even in instances where the value could be zero. The results are robust to changing the size of the offset or using the untransformed variables for the analysis.

The findings reported in Column (2) highlight why interpreting the above effect as being driven entirely due to the inventor move itself is inappropriate. We base the results reported in this column on a pooled analysis of the entire “original sample,” with the indicator variable *post-move* used to distinguish the citing years following the move date from those preceding it. The positive and significant coefficient on *recruit*, despite now having a separate term for *recruit X post-move*, demonstrates that focal patents systematically receive more citations even before the move actually takes place. Stated another way, the cross-sectional analysis employed in Column (1) confounds selection effects with treatment effects. Nevertheless, since the estimate of the coefficient on *recruit X post-move* is also positive in Column (2), we still have evidence consistent with an increase in knowledge use associated with mobility even if the effect is smaller than what a cross-sectional analysis would lead us to believe. (Interpreting magnitude of interaction effects in non-linear models, particularly in a way that allows cross-model comparison, is not straightforward. Therefore, we postpone that discussion until we come to our preferred specification below: fully saturated linear models with patent fixed effects.)

We next replicate the pooled analyses from Columns (1) and (2) using linear models before turning to our preferred (fixed effects) specification. As the results in Columns (3) and (4) show, there is again clear evidence that focal patents are systematically more highly cited even pre-move.²⁸ As a benchmark note that the average number of citations received by a control patent in the post-move subsample is 0.0061 (Table 2). Compared with the Column (3) estimate of 0.0356, this represents a 484% ($0.0356/0.0061 - 1$) increase in citations to the focal patents; this estimate 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 estimated

²⁸ 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 the 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 micro-level information we want to exploit. In fact, a low R^2 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).

coefficient for *recruit X post-move* in Column (4) suggests that, once we use *recruit* to account for the systematic difference between focal and control patents that exists even pre-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 2) and controlling for the baseline rise in citations for that technology class as captured by the control patents, this estimate implies a 179% ($0.0251/0.0140$) increase in the citation rate. Comparing the estimates from Columns (3) and (4) indicates that selection accounts for almost half, 42% ($0.0356/0.0251 - 1$), of the effect that cross-sectional models comparable to our Column (3) would (incorrectly) attribute to the inventor move itself.

While cross-sectional analyses typically use post-move citation rates 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 pre-move period as a benchmark. The latter interpretation seems more desirable as it is based directly on the “treatment” of interest (while recognizing that we are still estimating the “treatment on the treated” since mobility is not assigned randomly, a topic we discuss in the conclusion). A comparison of the natural percentage interpretations associated with the estimates from Columns (3) and (4) implies that using a purely cross-sectional comparison of post-move citation rates exaggerates the percentage increase in benefits attributable to mobility by 170% ($484\%/179\% - 1$).

Before turning to our DD regression with patent fixed effects using the complete sample, it is instructive to examine the estimates for a similar model when we use a sample of just focal patents (no controls). This model is still identifiable due to variation in application years and move years across focal patents. We report these results in Column (5). Note that since *recruit* is one for all observations corresponding to the focal patents, we no longer separately identify *recruit X post-move* from *post-move*. Given that the *post-move* effect itself is positive in the Column (4) model employing both focal and control patents, the focal-patent-only sample estimate for *recruit X post-move* in Column (5) is naturally greater. While the difference in magnitudes compared with Column (4) estimates (as well as the estimates in subsequent models) is not trivial, it is also not large enough to change the qualitative conclusions.

The model estimated in Column (6) is our specification of choice; it employs patent fixed effects and includes control patents (although it does not employ our stringent CEM-based matching, an issue we turn to shortly). The DD estimate for the coefficient on *recruit X post-move* is 0.0283, reflecting an increase of 202% ($0.0283/0.0140$) relative to the average number of cites a recruit's invention receives from the destination firm pre-move (after controlling for the baseline increase in citations in that technology class). This is our primary estimate of the "recruitment effect."

Notice that the Column (6) DD estimate is similar in magnitude to the *recruit X post-move* estimate from the pooled analysis reported in Column (4). This offers additional confidence that once we account for observed patent characteristics (including, importantly, the focal patent indicator) and the model has been fully saturated, other (unobserved) sources of heterogeneity do not appear to be too major a concern when we employ the DD approach. Nevertheless, we continue to employ a fixed effects model as the preferred specification going forward.

At this point, it is useful to take stock of the empirical issues we have addressed so far and those that remain outstanding. We have demonstrated that ignoring heterogeneity across patents is problematic when making inferences based on a comparison of post-move citations due to selection bias. We have also shown that applying a DD approach to panel data can mitigate this issue by exploiting cross-patent heterogeneity in pre-move citations to adjust estimates (assuming that focal and control patents are more comparable on trends than levels). However, even our DD estimates do not necessarily imply causality. For example, it is still possible that an unobserved factor (such as a shift in the specific domain focus of the firm that the technology classification match does not capture) drives both the hiring of a specific individual and an increased use of their ideas.

While possibilities such as the one above are difficult to rule out altogether, we can potentially further reduce the latitude for non-causal explanations by examining the temporal patterns of knowledge use. In particular, we look for evidence of the destination firm increasing its use of the inventor's ideas in the years prior to the move actually taking place. Such an uptick would heighten concerns that unexplained time trends not directly related to the move might be driving the DD findings.

We implement the above by including the additional indicator variables *recruit X pre-move period 1* and *pre-move period 1* in the specification reported in Column (7). The estimated coefficient on *recruit X pre-move period 1* suggests an annual citation rate that is clearly much smaller (coefficient 0.0083) in the two years leading up to the move than the average in the years after the move (coefficient 0.0328, an estimate that would be even greater had we separately estimated the effect for just the two-year period post-move). Nevertheless, even this relatively small uptick in citations pre-move causes some concern about a time trend independent of the move, raising issues about the comparability of the focal and control patents even under the (weaker) DD assumptions. We would ideally prefer that the treated and control groups show a similar trend pre-treatment so that we can more confidently assume they would continue to follow parallel paths in the post-treatment period (in the absence of the treatment). This motivates our efforts to implement a more stringent matching procedure, which we turn to next.

4.3. Estimates based on a more stringently matched sample

Table 4a 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 pre-move citations from the destination firm, and 3) are created by more experienced inventors. Clearly there is room for improving the closeness of the match, especially given the relatively large number of patents per three-digit technology class that have the same application year.

As described earlier, we construct an alternate sample (which we call the “CEM sample”) with exactly this in mind. Our CEM procedure relies upon not just a match on the three-digit technology class and application year (already used in constructing the “original sample”) but also on discrete buckets constructed using the first five patent characteristics listed in Table 4a: 1) overall pre-move citations received, 2) pre-move citations received from the destination firm, 3) number of previous patents by the inventor, 4) number of years since the first patent by the inventor, and 5) the delay between the

application and grant dates. A comparison of Table 4b with 4a reveals that the focal and control patent subsamples are indeed much more comparable for the “CEM sample” than for the “original sample.”

Before returning to formal regression analysis for the “CEM sample,” Table 5 reports CEM summary statistics corresponding to those reported in Table 2. The pre-move means for citations received are now very similar for focal versus control patents since cumulative pre-move citation count is one of the matching criteria.²⁹ This makes interpreting the DD statistic calculation easier and the underlying assumptions less stringent, as the estimated DD effect now (roughly) coincides with a direct comparison of the focal and control patent citation rates post-move.

The two main insights from Table 2 persist. 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 2 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 part of the apparent increase associated with the move arises from learning-by-hiring; a significant fraction of the destination firm’s jump in usage of the recruit’s prior ideas is due to self-citations by the recruit herself.

Once more, we turn to a regression approach for a more rigorous analysis and report our findings in Table 6. For ease of readability, we have replicated all the specifications from Table 3 in Table 6. However, one should not compare coefficients across the tables for the first two columns since non-linear regression estimates are not directly comparable across samples. The estimated coefficient for *recruit* in Columns (2) and (4) is now (almost) zero, which is not surprising given that the cumulative pre-move citation count is one of the CEM criteria. Further evidence of the matching working as expected is that, unlike in Table 3, the estimate for *recruit X pre-move period 1* (Column (7)) is now (practically) zero, indicating no pre-move trend.

²⁹ The stringent matching, however, has come at a cost: the focal patents matched during CEM have systematically lower pre-move citations rates (compare Tables 2 and 5). We will return to this point later.

Interestingly, the estimates for the DD coefficient of interest – *recruit X post-move* – are very similar to those reported in Table 3. Comparing our preferred specification, Column (6), across tables, we find that the estimate of interest is 0.0283 earlier and only slightly greater at 0.0296 now. This offers additional confidence in the robustness of the result. We base the remaining analyses reported in the paper on the specification of choice (Column (6)) and the “CEM sample,” although the findings are similar across different specifications and samples.³⁰

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 inventions by the recruit herself versus diffusion to others in the firm. The analysis reported in Table 7 examines this by analyzing the “CEM sample” using fully saturated linear models with patent fixed effects (corresponding to Column (6) from Table 6) and two additional dependent variables (already defined in Table 1): *cites excluding inventor self-cites* and *cites excluding inventor self-cites and collaborators*. For easy comparison, Column (1) reproduces the aggregate findings from Column (6) of Table 6 (using *all cites* as the dependent variable).

Column (2) uses *cites excluding inventor self-cites* as the dependent variable, a measure analogous to *all cites* but excluding inventor self-citations. The estimate for *recruit X post-move* is 0.0161, which represents only 54% (0.0161/0.0296) 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 learning at all. Rather, it is due to the recruit building upon her own prior ideas. While this may be just as useful for overall value creation, it is likely to reduce the destination firm’s relative

³⁰ Comparing the Column (5) findings from Tables 3 and 6, one might wonder why the “CEM sample” demonstrates a greater post-move increase in citation rate than the “original sample” for a focal-patent-only analysis. Recall that the “CEM sample” only includes the subset of original focal patents that get matched in the CEM procedure. Further analysis reveals that the Column (5) difference is driven primarily by the few focal patents that are highly cited pre-move but do not find a CEM match on this dimension. When we repeat the “original sample” analysis using just the same focal patents as the “CEM sample,” then Column (5) estimates are somewhat *greater* than the CEM-based findings. This is consistent with a view that, *for a given sample of focal patents*, less stringent matching leads to a larger (not smaller) estimated effect of mobility. Note that the difference between the two tables is much smaller in our preferred DD specification that includes control patents (Column (6)).

bargaining power in terms of the portion of this value it can appropriate (a point we return to in the conclusion).

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 co-inventors with the recruit on past patents).³¹ The analysis reported in Column (3) uses *cites excluding inventor self-cites or collaborator cites* as the dependent variable. The estimated coefficient on *recruit X post-move* is 0.0095 and represents only 32% (0.0095/0.0296) of the overall increase. This further suggests that learning-by-hiring is actually quite localized; on average, only 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. Do the recruit's ideas diffuse through the hiring firm over time?

Finally, we explore the degree to which our "limited learning" result from the prior section persists over time. In Table 8, we use nine indicator variables (summarized in Table 1) corresponding to five post-move, two-year periods and four pre-move periods, with the two-year period immediately preceding the move (corresponding to *pre-move period 1*) being the reference. For ease of visual interpretation, we plot coefficient estimates for the DD interaction effects in Figure 4.

This figure 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 uptick in the years leading up to it). While the tight timing of the mobility-idea usage relationship does not necessarily imply causality, it is certainly

³¹ In classifying a citation, we use the shortest path of connectedness between the inventors of the cited and citing patents. So, if a cited patent's inventor and some of her former collaborators collectively develop a new patent that cites her previous patent, then 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 citing patent's inventing team includes one or more former collaborators of the cited inventor without involving the cited inventor herself.

consistent with the assertion that recruiting facilitates enhanced access to an individual's ideas. (We discuss caveats for this interpretation in the conclusion.)

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 of the hiring firm's increased use of her ideas right after she arrives.

Third, and perhaps most surprisingly, comparing the findings for *cites excluding inventor self-cites* versus *all cites* over the duration of the post-move period, we see that the recruit continues to account for more than a third of the cites to her prior ideas for at least the first four periods (eight years) after her arrival. 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 only very gradually: during the first eight years there is only a slight increase 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) seems to be relatively persistent over time.

5. DISCUSSION AND CONCLUSION

What are the implications of our findings for managers and policymakers? Let us start with our finding that firms do indeed increase their use of their recruits' prior ideas. From the policymaker's perspective, it is tempting to conclude from this that restrictions on inter-firm mobility, such as non-compete covenants, are detrimental for 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 seems an attractive way to access external ideas. While there may well be merit to all of the above arguments, these conclusions do not unambiguously follow from our results. We describe some of the key nuances 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 worried about inventors leaving and sharing their ideas with competitors.

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 patented while at the source firm may increase the value of those patents, which the source firm could 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 et al. 2006, Oettl and Agrawal 2008, Corredoira and Rosenkopf 2010).

Third, even the implications for the hiring firm are more nuanced than they first appear. Our findings temper the "learning-by-hiring" view prevalent in the literature. Although we find that others at the destination firm do learn about the recruit's ideas, the role the recruit and her immediate collaborators play in realizing the destination firm's use of her prior ideas is particularly prominent, a phenomena that persists over time. This implies that the distribution of bargaining power between the recruit and the destination firm may favor the recruit, particularly when the value of her ideas is significant. The extent to which the firm itself is able to capture rents from better access to the inventor's ideas through recruiting 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 so that more individuals utilize the recruit's ideas 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. Here, it is important to note that 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. 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 under-investing in the sharing of knowledge among 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.

In addition, we must also clarify a general limitation to a causal interpretation of our findings. Although our use of a longitudinal dataset, a DD estimation approach, and more stringent matching move us closer to a causal interpretation than previous related studies afford, we still have a way to go. Due to our research design, we are limited in terms of how far we can extrapolate from our specific context to other settings. 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 whom 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).

Thus, one must be careful about extending our findings from the treated sample to the overall population since the two may have importantly different properties. At a minimum, 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, will be less valuable than those of the average recruited inventor. This is an important caveat to the implication 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.” Hiring an inventor does not cause the firm to increase its use of that individual’s ideas, but it may

facilitate greater use. However, even our (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 the scenario instead is that 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.)

While we have emphasized the benefits of our more stringent CEM-based 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, focal patents that we drop during CEM have higher pre-move citations rates. Thus, the “CEM sample” is comprised of patents that, on average, are of lower quality and/or lower relevance to the destination firm. On the one hand, this raises concerns regarding generalizability of the findings based on this sample. On the other hand, if including the dropped patents is problematic anyway (e.g., if greater pre-move citation is symptomatic of more severe endogeneity issues), excluding them might actually be desirable. We have no direct way of verifying which has more merit. However, it is reassuring to note that our (preferred) DD-based estimates for the “original sample” are not too different from those using the “CEM sample.”

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 non-compete 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 (beyond just the patent-based characteristics already used in our study). For a better understanding of the antecedents as well as consequences of mobility, future research could also benefit from stronger links to the rich literature on job matching (e.g., Jovanovic 1979, Simon and Warner 1992). This involves explicitly incorporating the dynamic process wherein inter-firm mobility is a result of individuals and firms optimizing their match in the labor market.

The focus of our study is on how mobility enhances access to pre-move knowledge associated with the recruit herself. However, we can further (speculatively) interpret our findings as also providing suggestive evidence that hiring facilitates a firm's overall "absorptive capacity" in a knowledge domain (Cohen and Levinthal 1989). Referring back to Column (6) in Table 6, the estimate for *post-move* is significant (even though smaller than the DD estimate corresponding to *recruit X post-move*). A conservative interpretation implicit in our analysis is that this reflects the destination firm's change in technological focus towards 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 may be responsible for improving the firm's absorptive capacity in this knowledge domain, hence driving the post-move boost in citations made even to the control patent. Comparing the coefficient on *post-move* in Column (1) with Column (2) in Table 7, we see that the increase in the firm's use of the control patent is 0.0092 in terms of *all cites* and 0.0054 in terms of *cites excluding the mover*, suggesting that the mover is indeed responsible for a significant portion of the increase in citations to the control patent. This absorptive capacity interpretation is intriguing and worthy of further research

Despite these caveats regarding our research design, our study does offer a methodological contribution for those who study mobility and knowledge flows. Other researchers can apply our idea-level, longitudinal, DD framework for 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 or her original geographic region. 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 or by third parties not directly involved with the inventor.

Extending beyond just usage of knowledge that exists pre-move, researchers could also examine other interesting mobility-related relationships using our approach, such as measuring changes in 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 more accentuated for 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 micro-level 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.

REFERENCES

- Agarwal, R., R. Echambadi, A. Franco, MB Sarkar. 2004. Knowledge transfer through inheritance: spin-out generation, development, and survival. *Academy of Management Journal*. **47**(4) 501–522.
- Agarwal, R., M. Ganco, R. Ziedonis. 2009. Reputations for toughness in patent enforcement: Implications for knowledge spillovers via inventor mobility. *Strategic Management Journal*. **30** 1349–1374
- Agrawal, A., I. Cockburn, J. McHale. 2006. Gone but not forgotten: Labor flows, knowledge spillovers, and enduring social capital. *Journal of Economic Geography*. **6**(5) 571-591.
- Angrist, J.D., A.B. Krueger. 1994. Why do World War II veterans earn more than nonveterans? *Journal of Labor Economics* **12** (1) 74-97
- Angrist, J.D., J. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Alcacer, J., M. Gittelman. 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. *Review of Economics and Statistics*. **88**(4) 774-779.
- Allison, P.D., R.P. Waterman. 2002. *Sociological Methodology*. **32**: 247-265.
- Almeida, P., B. Kogut. 1999. The localization of knowledge and the mobility of engineers in regional networks. *Management Science*. Vol. **45**(7) 905-917.

- Arrow, K. J. 1962. Economic welfare and the allocation of resources for invention. In R. R. Nelson (ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Vol. 13 of NBER Special Conference Series, Princeton University Press, New Jersey, pp. 609-625.
- Azoulay, P., W. Ding, T. Stuart. 2009. The impact of academic patenting on the rate, quality and direction of (public) research output. *Journal of Industrial Economics* **LVII**(4) 637-676.
- Azoulay, P., J. S. Graff Zivin, J. Wang. 2010. Superstar Extinction. *Quarterly Journal of Economics*, Forthcoming.
- Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management*. **17**(1) 99-120.
- Becker, G.S. 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy*. **70**(5) 9-49.
- Breschi, S., F. Lissoni. 2009. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography* **9**(4) 439-468.
- Chesbrough, H. 2003. *Open Innovation*. Boston, Harvard Business School Press.
- Coff, R.W. 1997. Human assets and management dilemmas: Coping with hazards on the road to resource-based theory. *The Academy of Management Review*. **22**(2) 374-402.
- Cohen, W.M., D.A. Levinthal. 1989. Innovation and learning: The two faces of R&D. *Economic Journal*. **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 Journal*. **31** 159-181.
- Dahl, M.S., O. Sorenson. 2010. The migration of technical workers. *Journal of Urban Economics* **67**.
- DiMaggio, P., W.W. Powell. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*. **48**(2) 147-160.
- Dosi, G. 1988. Sources, procedures and microeconomic effects of innovation. *Journal of Economic Literature*. **26**(3) 1120-1171.
- Duguet, E., M. MacGarvie. 2005. How well do patent citations measure knowledge spillovers? Evidence from French innovation surveys. *Economics of Innovation and New Technology*. **14**(5) 375-393.
- Fleming, L. 2001. Recombinant uncertainty in technological search. *Management Science*. **47**(1) 117-132.
- Fleming, L., C. King, A. Juda. 2007. Small worlds and innovation. *Organization Science*. **18**(6) 938-954.
- Franco, A.M., M.F. Mitchell. 2008. Covenants not to compete, labor mobility, and industry dynamics. *Journal of Economics and Management Strategy*. **17** 581-606.
- Furman, J.L., S. Stern. 2006. Climbing atop the shoulders of giants: the impact of institutions on cumulative research. NBER Working Paper 12523.
- Gompers, P., J. Lerner, D. Scharfstein. 2005. Entrepreneurial spawning: public corporations and the genesis of new ventures, 1986 to 1999. *Journal of Finance*. **60**(2) 517-614.
- Grossman, G., E. Helpman. 1991. *Innovation and Growth in the World Economy*. Cambridge, MIT Press.
- Groysberg, B., L.E. Lee, A. Nanda. 2008. Can they take it with them? The portability of star knowledge workers' performance. *Management Science*. 1-18.

- 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. *Economics of Innovation and New Technology*. **16** 159-177.
- Hansen, M.T. 1999. The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*. **44**(1) 82-111.
- Hausman, J., B. Hall, Z. Griliches. 1984. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*. **52**(4) 909-38.
- Heckman, J., S. Navarro-Lozano. 2004. Using matching, instrumental variables, and control functions to estimate economic choice models. *The Review of Economics and Statistics*. **86**(1) 30-67.
- Iacus, S.M., G. King, G. Porro. 2009. Causal inference without balance checking: Coarsened exact matching. Working Paper, Harvard University.
- Imbens, G.W. 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*. **86**(1) 4-29.
- Jaffe, A.B., M. Trajtenberg, 2002. *Patents, Citations & Innovations: A Window on the Knowledge Economy*. MIT Press, Cambridge.
- Jovanovic, B. 1979. Job matching and the theory of turnover. *Journal of Political Economy*. **87**(5)972-90.
- Klepper, S., S. Sleeper. 2005. Entry by spin-offs. *Management Science*. **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. In A.M McGahan & J.A.C. Baum (Ed.), *Advances in Strategic Management*. Vol. 21. Greenwich CT: JAI Press.
- Lampe R. 2008. Strategic citation. SSRN Working Paper #984123.
- Lazear, E.P. 1986. Raids and offer matching. *Research in Labor Economics*. **8** 141-165.
- Levin, S.G., P.E. Stephan. 1991. Research productivity over the life cycle: Evidence for academic scientists. *American Economic Review*. **81** 114-132.
- Llorente, G., R. Michaely, G. Saar, J. Wang. 2002. Dynamic volume-return relation of individual stocks. *The Review of Financial Studies* **15**(4) 1005-1047.
- Malamud, O., C. Pop-Eleches. 2010. General education versus vocational training: Evidence from an economy in transition. *The Review of Economics & Statistics*. **XCII**(1) 43-60.
- March, J.G. 1991. Exploration and exploitation in organizational learning. *Organization Science*. **2**(1) 71-87.
- Marx, M., D. Strumsky, L. Fleming. 2009. Mobility, skills, and the Michigan non-compete experiment. *Management Science*. **55**(6) 875-889.
- Mehta, A., M. Rysman, T. Simcoe. 2010. Identifying the age profile of patent citations: New estimates of knowledge diffusion. *Journal of Applied Econometrics*, Forthcoming.
- Moen, J. 2005. Is mobility of technical personnel a source of R&D spillovers? *Journal of Labor Economics*. **23**(1) 81-114.
- Moffitt, R.A. 2004. Introduction to the symposium on the econometrics of matching. *The Review of Economics and Statistics*. **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. *Journal of Economic Behavior and Organization*. **63**(4) 648-687.
- Nelson, R.R., S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Belknap.
- Oettl, A., A. Agrawal. 2008. International labor mobility and knowledge flow externalities. *Journal of International Business Studies*. **39** 1242-1260.
- Oettl, A. 2009. Productivity and Helpfulness: A New Taxonomy for Star Scientists. Working Paper, Georgia Institute of Technology
- Penrose, E.T. 1959. *The Theory of the Growth of the Firm*. Oxford, England: Blackwell.
- Peteraf, M.A. 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*. **14**(3) 179-191.
- Rosenkopf, L., P. Almeida. 2003. Overcoming local search through alliances and mobility. *Management Science*. **49**(6) 0751-0766.
- Samila, S., O. Sorenson. 2009. Non-compete covenants: Incentives to innovate or impediments to growth. Working paper, Yale School of Management.
- Saxenian, A.L. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.
- Simon, H.A. 1991. Bounded rationality and organizational learning. *Organization Science*. **2**(1) 125-134.
- Simon, C.J., J. Warner. 1992. Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure. *Journal of Labor Economics*. **10**(3) 306-330.
- Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Science*. **51**(5) 756-770.
- Singh, J. 2008. Distributed R&D, cross-regional knowledge integration and quality of innovative output. *Research Policy*. **37**(1) 77-96.
- Song, J., P. Almeida, G. Wu. 2003. Learning by hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Science*. **49**(4) 351-365.
- Thompson, P., M. Fox-Kean. 2005. Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*. **95**(1) 450-460.
- Trajtenberg, M. 2006. The "names game": Harnessing inventors' patent data for economic research. NBER Working Paper 12479.
- Tzabbar, D., B.S. Silverman, B.S. Aharonson. 2009. Learning by hiring or hiring to avoid learning? Organizational exploration via individuals' exploitation. Working Paper.
- Weitzman, M. 1998. Recombinant growth. *The Quarterly Journal of Economics* **CXIII**(2) 331-360.
- Wooldridge, J. M. 1999. Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics*. **90**(1) 77-97.

Table 1. Variable definitions and summary statistics for the “original sample” (N = 95,424)

Dependent variables:		Mean	Std Dev	Min	Max
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 or 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 & 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	1041.9	1433.1	0	8260
destination same-class patents	Number of destination firm patents that belong to the same 3-digit technology class as the focal patent	26.9	75.8	0	1375
original firm patents	Number of patents assigned to the original firm in the five years preceding the focal patent's application year	1120.1	1157.0	0	5504
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
non-patent references	Number of non-patent 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 US 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:					
post-move	Indicator for whether the citing year falls after the estimated move date	0.636	0.481	0	1
pre-move period 1	Indicator for whether the citing year is the 1st or 2nd year before the move	0.167	0.373	0	1
pre-move period 2	Indicator for whether the citing year is the 3rd or 4th year before the move	0.109	0.312	0	1
pre-move period 3	Indicator for whether the citing year is the 5th or 6th year before the move	0.058	0.234	0	1
pre-move period 4	Indicator for whether the citing year is the 7th or 8th year before the move	0.024	0.153	0	1
pre-move period 5	Indicator for whether the citing year is the 9th or 10th year before the move	0.007	0.081	0	1
post-move period 1	Indicator for whether the citing year is the 1st or 2nd year after the move	0.167	0.373	0	1
post-move period 2	Indicator for whether the citing year is the 3rd or 4th year after the move	0.160	0.367	0	1
post-move period 3	Indicator for whether the citing year is the 5th or 6th year after the move	0.143	0.350	0	1
post-move period 4	Indicator for whether the citing year is the 7th or 8th year after the move	0.109	0.311	0	1
post-move period 5	Indicator for whether the citing year is the 9th or 10th year after the move	0.057	0.233	0	1

Notes: We base the summary statistics on the “original sample” constructed as illustrated in Figures 2 and 3 and describe them in the notes corresponding to those figures. The overall sample size of 95,424 arises 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.

Table 2. Annual patent citation frequency for the “original sample”

		Average annual citations received from the destination firm				
		Pre-move		Post-move		
Patents involving move	<i>Subsample mean:</i>			<i>Subsample mean:</i>		<i>First difference (row):</i>
		all cites = 0.0140		all cites = .0396		all cites = 0.0256
	cites excluding inventor self-cites = 0.0140 (N = 17383)		cites excluding inventor self-cites = .0250 (N = 30329)		cites excluding inventor self-cites = 0.0110 (N = 47712)	
Control patents	<i>Subsample mean:</i>			<i>Subsample mean:</i>		<i>First difference (row):</i>
		all cites = 0.0051		all cites = 0.0061		all cites = 0.0010
	cites excluding inventor self-cites = 0.0051 (N = 17383)		cites excluding inventor self-cites = 0.0061 (N = 30329)		cites excluding inventor self-cites = 0.0010 (N = 47712)	
		<i>First difference (column):</i>		<i>First difference (column):</i>		<i>Difference-in-differences:</i>
		all cites = 0.0089		all cites = 0.0335		all cites = 0.0246
		cites excluding inventor self-cites = 0.0089 (N = 34766)		cites excluding inventor self-cites = 0.0189 (N = 60658)		cites excluding inventor self-cites = 0.0100 (N = 95424)

Notes: We base this analysis on the “original sample” constructed as illustrated in Figures 2 and 3 and describe it in the notes corresponding to those figures. Each box in this “2X2” matrix 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 pre-move since 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 pre-move period.

Table 3. Difference-in-differences regression analysis for the “original sample”

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable:	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>
Regression Model:	Negative binomial	Negative binomial	Linear	Linear	Linear (Patent FE)	Linear (Patent FE)	Linear (Patent FE)
Patent-Year Sample:	Original sample (post move only)	Original sample	Original sample (post move only)	Original sample	Original sample (focal patents only)	Original sample	Original sample
<i>recruit</i>	1.896*** (0.137)	1.017*** (0.213)	0.0356*** (0.0034)	0.0101*** (0.0022)			
<i>recruit X post-move</i>		0.907*** (0.230)		0.0251*** (0.0037)	0.0367*** (0.0043)	0.0283*** (0.0027)	0.0328*** (0.0035)
<i>post-move</i>		0.157 (0.207)		0.0033 (0.0025)		0.0045* (0.0027)	0.0091** (0.0036)
<i>recruit X pre-move period 1</i>							0.0083** (0.0042)
<i>pre-move period 1</i>							0.0035 (0.0031)
<i>ln(destination same-class patents)</i>	0.428*** (0.050)	0.497*** (0.048)	0.0085*** (0.0012)	0.0075*** (0.0010)			
<i>ln(destination firm patents)</i>	-0.215*** (0.045)	-0.202*** (0.040)	-0.0039*** (0.0010)	-0.0028*** (0.0007)			
<i>ln(original firm patents)</i>	0.013 (0.040)	0.003 (0.038)	0.0010 (0.0013)	0.0008 (0.0009)			
<i>ln(claims)</i>	0.085 (0.083)	0.096 (0.078)	0.0049** (0.0024)	0.0041** (0.0018)			
<i>ln(patent references)</i>	0.027 (0.083)	0.074 (0.077)	0.0009 (0.0028)	0.0019 (0.0019)			
<i>ln(non-patent references)</i>	0.149 (0.093)	0.073 (0.091)	0.0036 (0.0027)	0.0018 (0.0020)			
<i>ln(inventor patents)</i>	-0.129 (0.092)	-0.112 (0.085)	-0.0013 (0.0018)	-0.0007 (0.0013)			
<i>ln(inventor age)</i>	0.112 (0.105)	0.053 (0.098)					
<i>U.S. inventor</i>	0.662*** (0.156)	0.611*** (0.138)					
<i>ln(grant delay)</i>	0.123 (0.136)	-0.011 (0.131)					
Number of observations	60658	95424	60658	95424	47712	95424	95424
Number of cited patents	7952	7952	7952	7952	3976	7952	7952
Log likelihood	-5238	-6806					
Wald chi2	841.9***	48236***					
R-squared			0.037	0.029	0.006	0.004	0.004
F-statistic			4.47***	5.43***	9.02***	11.76***	11.50***

Notes: We base this regression analysis on the “original sample” already summarized in Tables 1 and 2. All models include yearly indicator variables for the citing year and the lag between the cited and citing year. The pooled linear models (Columns (3) and (4)) also include indicator variables for the inventor age (in years), inventor location (U.S. states or non-U.S. countries), and patent grant delay (in months), though the pooled non-linear models (Columns (1) and (2)) employ *inventor age*, *U.S. inventor*, and *grant delay* variables to ensure convergence. We use robust standard errors clustered on the identity of the inventor in all four pooled models and employ fixed effects for the cited patent in Columns (5) through (7), though results are practically identical if we use fixed effects for the inventor instead. We also obtain qualitatively similar results when we employ a Poisson-based conditional fixed effects regression approach, though we do not choose that as our preferred approach for two reasons: 1) the estimation does not converge if the model is saturated with as many indicators as the linear models, and 2) we drop a large fraction of the sample due to no variation in the dependent variable. Note that the reference period for interpreting the “difference-in-differences” (DD) estimate of interest – *recruit X post-move* – is the entire pre-move period for Columns (2), (4), (5), and (6) but the pre-move period excluding the two years just before the move in Column (7). A comparison of Columns (1) and (2) as well as of Columns (3) and (4) indicates that purely cross-sectional results as reported in Columns (1) and (3) can be misleading since mobile inventor patents receive systematically more citations than control patents even pre-move. Furthermore, the positive and significant estimate for *recruit X pre-move period 1* in Column (7) suggests that the upward trend in citation rate in part starts pre-move, suggesting the need for a more stringent matching for at least partially mitigating concerns regarding endogeneity of the move.

*** p<0.01, **p<0.05, *p<0.1

Table 4a. Mobile inventor patent versus control patent characteristics in the “original sample”

	Patents involving move			Control patents			Difference in means	
	obs	mean	std dev	obs	mean	std dev	magnitude	t-stat
Cumulative overall pre-move citations	3976	2.98	4.50	3976	2.42	3.88	0.57	6.0
Cumulative pre-move citations from destination	3976	0.063	0.461	3976	0.023	0.272	0.040	4.8
Inventor patents	3976	9.13	14.64	3976	5.87	12.27	3.26	10.8
Inventor age	3976	4.98	4.29	3976	3.96	4.32	1.02	10.6
Grant delay	3976	22.78	10.72	3976	22.59	10.91	0.19	0.8
Claims	3976	11.73	9.68	3976	11.32	9.02	0.40	1.9
Patent references	3976	6.92	7.40	3976	6.64	6.58	0.28	0.7
Non-patent references	3976	0.88	2.32	3976	0.84	2.23	0.03	1.8
Originality (NBER measure)	3843	0.34	0.27	3828	0.35	0.28	-0.01	0.9

Notes: The “originality” measure (taken from the NBER database) measures breadth of search across different technology classes and corresponds to one minus the Herfindahl of the backward citations made by a patent to previous patents from different technologies (and is therefore undefined for patents with no backward citations). Note that the means of several key characteristics of the focal and control patents differ significantly, motivating the need for a more stringent matching approach as summarized in Table 4b.

Table 4b. Mobile inventor patent versus control patent characteristics in the “CEM sample”

	Patents involving move			Control patents			Difference in means	
	obs	mean	std dev	obs	mean	std dev	magnitude	t-stat
Cumulative overall pre-move citations	2543	2.55	3.76	2543	2.50	3.83	0.05	0.5
Cumulative pre-move citations from destination	2543	0.010	0.136	2543	0.010	0.136	0.000	0.0
Inventor patents	2543	7.59	11.36	2543	7.40	14.52	0.19	0.5
Inventor age	2543	4.50	4.05	2543	4.61	4.18	-0.12	1.0
Grant delay	2543	23.11	10.62	2543	22.95	10.58	0.16	0.6
Claims	2543	11.39	9.01	2543	11.18	8.67	0.21	0.8
Patent references	2543	6.40	5.97	2543	6.38	5.25	0.02	0.1
Non-patent references	2543	0.85	2.14	2543	0.78	1.98	0.07	1.2
Originality (NBER measure)	2464	0.33	0.28	2457	0.34	0.28	-0.01	1.4

Notes: We construct the “CEM sample” using a CEM procedure that relies upon not just an exact match on the three-digit technology class and application year (used in constructing the “original sample”) but also on an exact match on discrete buckets based on five additional criteria: 1) overall pre-move citations received by a patent (four buckets: zero citations, one-two citations, three-seven citations, and eight or more citations), 2) pre-move citations received from the destination firm (four buckets: zero citations, one citation, two-three citations, and four or more citations), 3) number of previous patents successfully filed for by this inventor (three buckets: zero or one patents, two-12 patents, and 13 or more patents), 4) number of years since the first patent by this inventor (three buckets: zero or one years, two-nine years, and 10 or more years), and 5) the delay between the application date and grant date for the patent (three buckets: 490 or less days, 491-832 days, and 833 or more days). Also note that the overall sample size is now smaller – there are 2,543 focal patents and as many corresponding control patents in contrast to 3,976 focal patents and as many control patents in Table 4a – since we drop the mobile inventor patents for which we find no match in the CEM procedure. A comparison of Tables 4a and 4b reveals that the mobile inventor patents and control patents are much more comparable for the “CEM sample” than for the “original sample” on a variety of key characteristics.

Table 5. Annual patent citation frequency for the “CEM sample”

	Average annual citations received from the destination firm		
	Pre-move	Post-move	
Patents involving move	<i>Subsample mean:</i> all cites = 0.0022 cites excluding inventor self-cites = 0.0022 (N = 10981)	<i>Subsample mean:</i> all cites = .0345 cites excluding inventor self-cites = .0216 (N = 19535)	<i>First difference (row):</i> all cites = 0.0323 cites excluding inventor self-cites = 0.0194 (N = 30516)
Control patents	<i>Subsample mean:</i> all cites = 0.0022 cites excluding inventor self-cites = 0.0022 (N = 10981)	<i>Subsample mean:</i> all cites = .0049 cites excluding inventor self-cites = .0049 (N = 19535)	<i>First difference (row):</i> all cites = 0.0027 cites excluding inventor self-cites = 0.0027 (N = 30516)
	<i>First difference (column):</i> all cites = 0 cites excluding inventor self-cites = 0 (N = 21962)	<i>First difference (column):</i> all cites = 0.0296 cites excluding inventor self-cites = 0.0167 (N = 39070)	<i>Difference-in-differences:</i> all cites = 0.0296 cites excluding inventor self-cites = 0.0167 (N = 61032)

Notes: This analysis replicates the summary analysis from Table 2 (based on the “original sample”) for the “CEM sample.” Note that the pre-move means for citations received are now practically the same for the focal and control patents, since cumulative pre-move citation count is one of the criteria used for the matching done in constructing the “CEM sample.” Also note that the overall sample size is now smaller – 61,032 here versus 95,424 in Table 2 – since we drop the mobile inventor patents for which we do not find the more stringent match when constructing the “CEM sample.”

*** p<0.01, **p<0.05, *p<0.1

Table 6. Difference-in-differences regression analysis for the “CEM sample”

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable:	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>	<i>all cites</i>
Regression Model:	Negative binomial	Negative binomial	Linear	Linear	Linear (Patent FE)	Linear (Patent FE)	Linear (Patent FE)
Patent-Year Sample:	CEM sample (post move only)	CEM sample	CEM sample (post move only)	CEM sample	CEM sample (focal patents only)	CEM sample	CEM sample
<i>recruit</i>	1.927*** (0.185)	0.007 (0.376)	0.0316*** (0.0040)	0.0012 (0.0013)			
<i>recruit X post-move</i>		1.945*** (0.407)		0.0295*** (0.0039)	0.0447*** (0.0048)	0.0296*** (0.0029)	0.0289*** (0.0038)
<i>post-move</i>		0.895*** (0.319)		0.0069** (0.0028)		0.0092*** (0.0029)	0.0115*** (0.0039)
<i>recruit X pre-move period 1</i>							-0.0012 (0.0045)
<i>pre-move period 1</i>							0.0028 (0.0034)
<i>ln(destination same-class patents)</i>	0.450*** (0.067)	0.509*** (0.066)	0.0076*** (0.0013)	0.0056*** (0.0009)			
<i>ln(destination firm patents)</i>	-0.275*** (0.059)	-0.272*** (0.057)	-0.0037*** (0.0011)	-0.0024*** (0.0007)			
<i>ln(original firm patents)</i>	-0.066 (0.051)	-0.063 (0.049)	-0.0016 (0.0018)	-0.0010 (0.0012)			
<i>ln(claims)</i>	-0.014 (0.108)	-0.014 (0.106)	0.0022 (0.0032)	0.0012 (0.0021)			
<i>ln(patent references)</i>	0.207* (0.113)	0.182* (0.107)	0.0050 (0.0035)	0.0033 (0.0023)			
<i>ln(non-patent references)</i>	0.053 (0.125)	0.082 (0.121)	-0.0009 (0.0029)	-0.0000 (0.0019)			
<i>ln(inventor patents)</i>	-0.175 (0.117)	-0.133 (0.114)	-0.0022 (0.0023)	-0.0013 (0.0015)			
<i>ln(inventor age)</i>	0.158 (0.145)	0.092 (0.141)					
<i>U.S. inventor</i>	0.544*** (0.198)	0.411** (0.188)					
<i>ln(grant delay)</i>	-0.103 (0.192)	-0.119 (0.183)					
Number of observations	39070	61032	39070	61032	30516	61032	61032
Number of cited patents	5086	5086	5086	5086	2543	5086	5086
Log likelihood	-4153	-4525					
Wald chi2	988.9***	42198***					
R-squared			0.033	0.025	0.009	0.006	0.006
F-statistic			2.93***	3.48***	13.12***	18.66***	17.64***

Notes: The regression models employed here are identical to the corresponding ones employed in Table 3, including all indicator variables not reported in that table to conserve space, but the sample used is now the “CEM sample.” The sample sizes are now smaller since we drop patents for which the CEM procedure does not find a match when constructing the “CEM sample.” Unlike in Table 3, the estimated coefficient for *recruit* in Columns (2) and (4) are practically zero since the cumulative pre-move citation count is one of the matching criteria used in constructing the “CEM sample.” Further evidence of the matching working well is that the estimate for *recruit X pre-move period 1* (Column (7)) is now practically zero, indicating that there is no pre-move trend in the citation rate.

*** p<0.01, **p<0.05, *p<0.1

Table 7. Learning-by-hiring?

	(1)	(2)	(3)
Dependent Variable:	all cites	cites excluding inventor self-cites	cites excluding inventor self-cites or collaborator cites
Regression Model:	Linear (Patent FE)	Linear (Patent FE)	Linear (Patent FE)
Patent-Year Sample:	CEM sample	CEM sample	CEM sample
<i>recruit X post-move</i>	0.0296*** (0.0029)	0.0161*** (0.0024)	0.0095*** (0.0020)
<i>post-move</i>	0.0092*** (0.0029)	0.0054** (0.0024)	0.0067*** (0.0020)
Number of observations	61032	61032	61032
Number of cited patents	5086	5086	5086
F-statistic	11.31***	6.03***	4.18***

Notes: We base this regression analysis on the “CEM sample” already used in Table 6. We use our preferred specification as in Column (6) of Tables 3 and 6 (including all indicator variables not reported to conserve space). For ease of comparison, Column (1) simply replicates the analysis from Column (6) of Table 6, which uses *all cites* as the dependent variable. In Column (2), we examine the extent to which the estimated boost in citation rates post-mobility diminishes when citation counts exclude inventor self-cites. Similarly, in Column (3), we analyze the extent to which the estimated boost in citation rates post-mobility diminishes when citation counts exclude not just inventor self-cites but also cites made by collaborators in the destination firm. Note that the reference period for interpreting the DD estimate of interest – *recruit X post-move* – is the entire pre-move period.

*** p<0.01, **p<0.05, *p<0.1

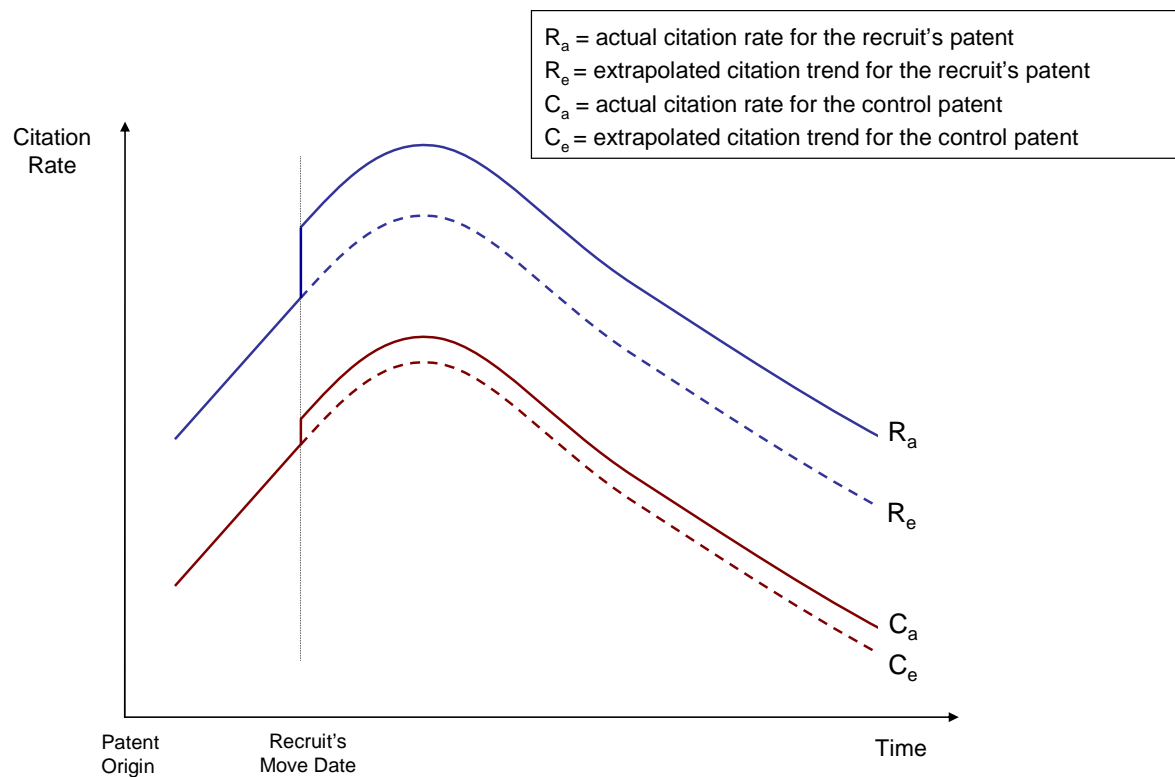
Table 8. Does the importance of self-exploitation by the recruit persist over time?

	(1)	(2)	(3)
Dependent Variable:	all cites	cites excluding inventor self-cites	cites excluding inventor self-cites or collaborator cites
Regression Model:	Linear (Patent FE)	Linear (Patent FE)	Linear (Patent FE)
Patent-Year Sample:	CEM sample	CEM sample	CEM sample
<i>recruit X post-move period 1</i>	0.0378*** (0.0045)	0.0197*** (0.0037)	0.0165*** (0.0032)
<i>recruit X post-move period 2</i>	0.0348*** (0.0045)	0.0196*** (0.0037)	0.0118*** (0.0032)
<i>recruit X post-move period 3</i>	0.0326*** (0.0047)	0.0186*** (0.0038)	0.0064* (0.0033)
<i>recruit X post-move period 4</i>	0.0125** (0.0051)	0.0078* (0.0042)	0.0054 (0.0036)
<i>recruit X post-move period 5</i>	0.0157** (0.0065)	0.0131** (0.0053)	0.0086* (0.0045)
<i>recruit X pre-move period 2</i>	0.0023 (0.0052)	0.0023 (0.0042)	0.0032 (0.0036)
<i>recruit X pre-move period 3</i>	0.0046 (0.0066)	0.0034 (0.0054)	0.0034 (0.0046)
<i>recruit X pre-move period 4</i>	0.0047 (0.0098)	0.0049 (0.0080)	0.0043 (0.0069)
<i>recruit X pre-move period 5</i>	0.0017 (0.0184)	0.0002 (0.0151)	-0.0007 (0.0130)
<i>post-move period 1</i>	0.0053 (0.0042)	0.0032 (0.0034)	0.0031 (0.0030)
<i>post-move period 2</i>	0.0090 (0.0063)	0.0043 (0.0051)	0.0038 (0.0044)
<i>post-move period 3</i>	0.0122 (0.0087)	0.0049 (0.0071)	0.0041 (0.0061)
<i>post-move period 4</i>	0.0160 (0.0112)	0.0066 (0.0092)	0.0052 (0.0079)
<i>post-move period 5</i>	0.0207 (0.0139)	0.0080 (0.0114)	0.0041 (0.0098)
<i>pre-move period 2</i>	-0.0041 (0.0045)	-0.0017 (0.0036)	-0.0012 (0.0031)
<i>pre-move period 3</i>	-0.0042 (0.0070)	0.0009 (0.0057)	0.0021 (0.0049)
<i>pre-move period 4</i>	-0.0108 (0.0104)	-0.0032 (0.0085)	-0.0012 (0.0073)
<i>pre-move period 5</i>	-0.0189 (0.0167)	-0.0087 (0.0136)	-0.0058 (0.0117)
Number of observations	61032	61032	61032
Number of cited patents	5086	5086	5086
F-statistic	8.46***	4.39***	3.21***

Notes: Again using the “CEM sample” and the same underlying regression model (including all unreported indicator variables), this analysis extends from Table 7 by employing fine-grained indicator variables for a series of two-year time windows pre-move as well as post-move in order to examine the temporal patterns of citation rate changes associated with the move. Note that, unlike in Table 7, the reference period for interpreting the regression estimates is now *pre-move period 1* (the omitted timing category). Figure 4 graphically illustrates the DD results (i.e., the interaction term coefficients) from this table.

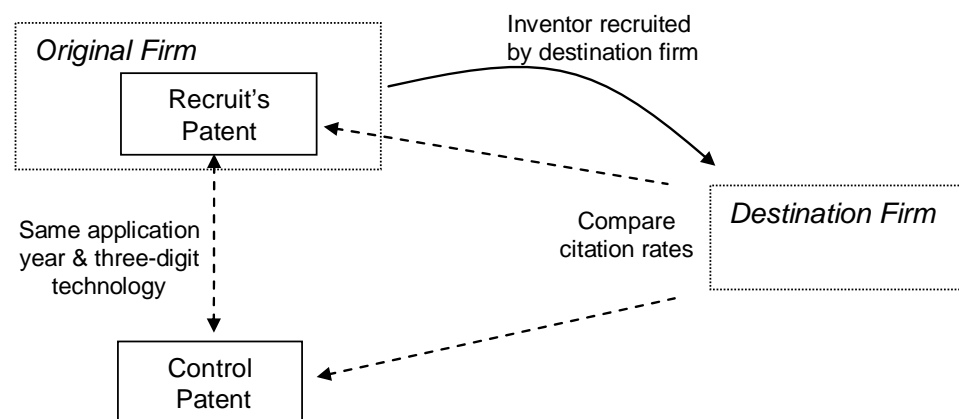
*** p<0.01, **p<0.05, *p<0.1

Figure 1. Motivating a “difference-in-differences” approach for estimating the effect of mobility



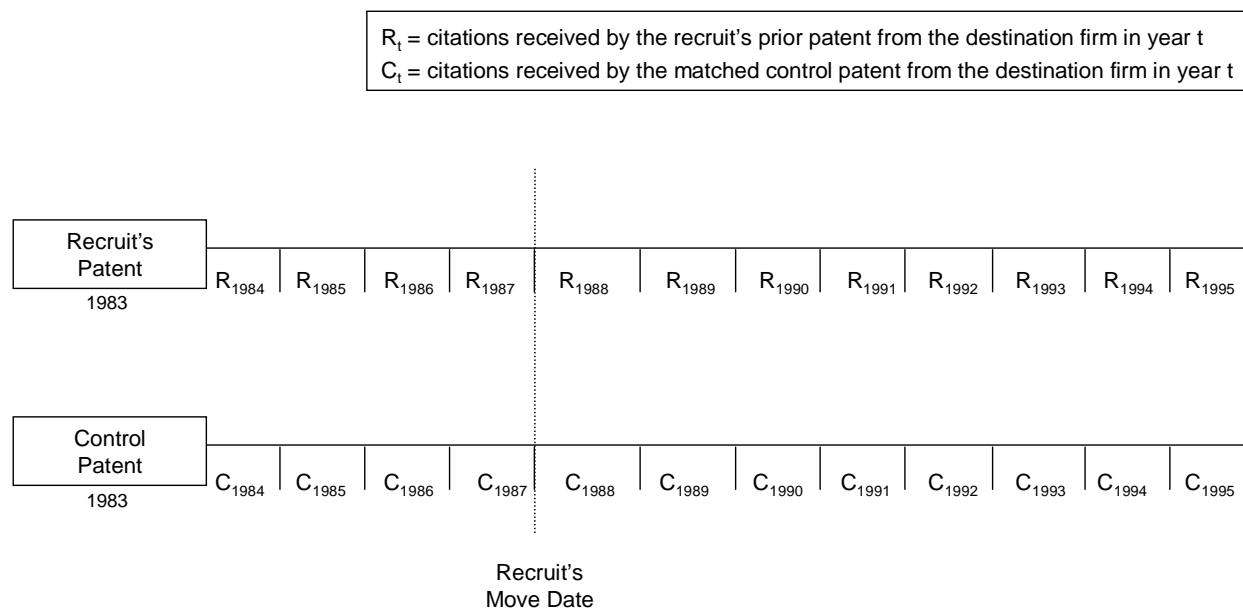
Notes: This figure motivates our DD approach for examining how firms exploit the prior inventions of new recruits. The key point is that a post-move comparison of the citation rates for focal versus control patents ignores the possibility that these citation rates may differ not just due to the recruit’s move *per se* but also because focal versus control patents might be inherently different to start with. Our DD approach in controlling for the latter involves extrapolating from information on citations received pre-move to estimate the likely citation trends if the move does not take place. Crudely speaking, we can better estimate the effect of mobility as a DD statistic $(R_a - R_e) - (C_a - C_e)$ instead of calculating it as just a post-move difference $(R_a - C_a)$. We can further reduce potential biases in the estimates that might still arise from the endogeneity of the move by matching the focal and control patents more stringently using pre-move citation trends and inventor-related information. We do this in the construction of our “CEM sample.”

Figure 2. Constructing a matched sample of potentially cited patents



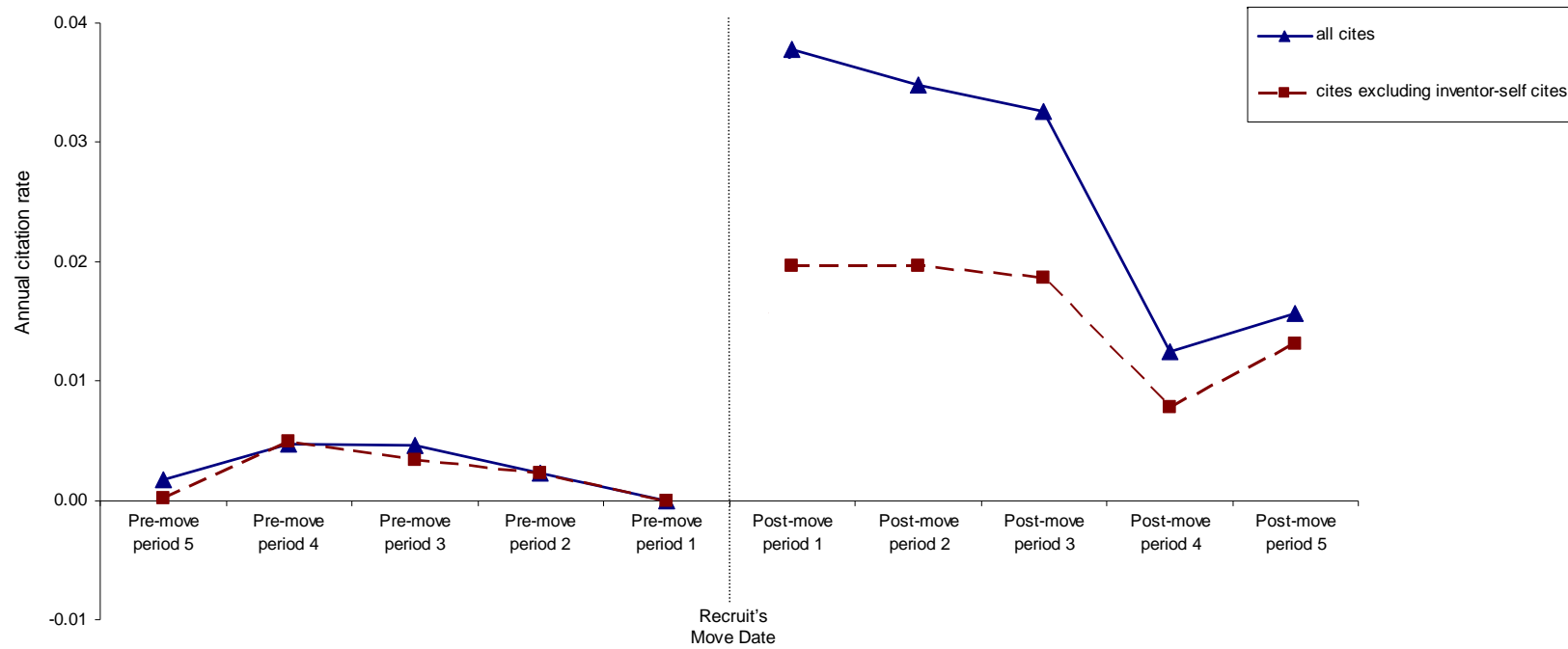
Notes: We begin the sample construction process by identifying a set of focal patents, each created by a single inventor who subsequently moves to another firm between the third and 10th year after the original patent's application date. (Excluding the first two years ensures a pre-move observation window, 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 each control patent has the same application year and three-digit primary technology class as the corresponding focal patent. Construction of the "CEM sample" follows a similar procedure except that we use additional observables (including pre-move citation and inventor information) to carry out a more stringent match between focal and control patents using the CEM technique.

Figure 3. Constructing a longitudinal dataset of citations received from the destination firm



Notes: This figure illustrates how we use either of the matched samples of patents as described in the notes to Figure 3 to construct a longitudinal patent-year dataset of citations to employ for our DD analysis. The resulting dataset has 12 yearly observations for each (focal or control) cited patent, corresponding to the 12 years directly following the patent's application year. In the above example, the focal patent (involving a future move) and the corresponding control patent originate in year 1983, so each of them gives rise to one observation per year for the period 1984-1995. We show the estimated move date to be January 1, 1988, so we classify the four annual observations corresponding to the period 1984-1987 as being pre-move (indicator *post-move* = 0) and the eight annual observations corresponding to the 1988-1995 as post-move (indicator *post-move* = 1). Following a similar procedure, we construct the "original sample" to have 95,424 observations (3,976 X 12 X 2) corresponding to 3,976 focal patents and as many corresponding control patents. In contrast, the "CEM sample" has 61,032 observations (2,543 X 12 X 2) corresponding to 2,543 focal patents and as many corresponding control patents, since we drop the focal patents that cannot find a match when employing the stringent CEM procedure.

Figure 4. Estimated temporal trends in citations received from the destination firm



Notes: This figure plots the DD estimates from regression analysis of temporal patterns of citation rate changes associated with the move, as reported in Table 8. Specifically, it plots the interaction coefficients of *recruit* with different *pre-move* and *post-move* periods as listed in the table. Since the reference period for the regression analysis is *pre-move period 1*, the DD coefficient for that period is set to zero in plotting the regression coefficients above. We base these results on analysis of the “CEM sample.”