Journal of International Business Studies (2008) 39, 1242–1260

www.jibs.net

International labor mobility and knowledge flow externalities

Alexander Oettl[1,2] and
Ajay Agrawal[1,3]

[1]Rotman School of Management, University of Toronto, Toronto, Canada; [2]Max Planck Institute of Economics, Jena, Germany; [3]National Bureau of Economic Research, Cambridge, MA, USA

Correspondence:
A Oettl, Rotman School of Management, University of Toronto, 105 St George Street, Toronto, ON, Canada M5S 3E6.
Tel: +1 416 978 7019;
Fax: +1 416 978 5433;
E-mail: aoettl@rotman.utoronto.ca

Abstract
Although knowledge flows create value, the market often does not price them accordingly. We examine "unintended" knowledge flows that result from the cross-border movement of inventors (i.e., flows that result from the move, but do not go to the hiring firm). We find that the inventor's new country gains from her arrival above and beyond the knowledge flow benefits enjoyed by the firm that recruited her (National Learning by Immigration). Furthermore, the firm that lost the inventor also gains by receiving increased knowledge flows from that individual's new country and firm (Firm Learning from the Diaspora). Surprisingly, the latter effect is only twice as strong when the mover moves within the same multinational firm, suggesting that knowledge flows between inventors do not necessarily follow organizational boundaries, thus creating opportunities for public policy and firm strategy.
Journal of International Business Studies (2008) **39,** 1242–1260.
doi:10.1057/palgrave.jibs.8400358

Keywords: labor mobility; knowledge flows; immigration; diaspora; inventors

INTRODUCTION

Knowledge flows are economically important because they increase the efficiency of the innovation process. The recombination of knowledge drives innovation; thus wider access to knowledge facilitates more efficient innovation by reducing the need to re-create what already exists elsewhere. In fact, contemporary economic theory focuses on knowledge spillovers – knowledge flows that occur outside market mechanisms – as the central determinant of economic growth (Romer, 1986, 1990).

Surprisingly, given the acknowledged importance of knowledge flows, we know very little about how they move through the economy and the mechanisms that influence flow patterns. However, we are reasonably certain about one feature of knowledge flow patterns: prior research shows, with reasonably conclusive empirical evidence, that such flows stay geographically localized (Agrawal & Cockburn, 2003; Almeida & Kogut, 1999; Audretsch & Feldman, 1996; Jaffe, Trajtenberg, & Henderson, 1993; Thompson & Fox-Kean, 2005).

The localization finding is important for a number of reasons. First, it provides insight into general flow patterns; knowledge does not flow uniformly across geographic space. Second, the finding implies that knowledge does not flow freely across the marketplace; public policy and firm strategy may influence flow patterns in self-serving ways. Finally, this finding offers insight into the

Received: 6 September 2005
Revised: 30 May 2007
Accepted: 13 June 2007
Online publication date: 7 February 2008

mechanisms that cause knowledge to flow the way it does; despite its apparent costless-to-disseminate properties, knowledge flows faster locally. Some may not find this intuitive, since researchers often place knowledge in the public domain by way of widely accessible monographs, journals, patents, etc. Moreover, some types of inventor – particularly scientists – have strong incentives to disseminate their findings as fast as possible (Dasgupta & David, 1987, 1994). Why does knowledge flow this way?

Knowledge includes both codified and non-codified components. Even inventors able to codify knowledge frequently do not do so, because of a lack of incentives (Agrawal, 2006). In order for inventors to apply knowledge, they often need access to both the codified and the non-codified components. The non-codified components of knowledge are likely to contribute to geographic stickiness, as non-codified knowledge often requires direct interaction with the inventor for effective transfer.

This need to interact with the inventor may explain why knowledge frequently flows locally. Prior research shows that scientists and engineers impart knowledge among their peers, or "invisible college", particularly if they share a personal social relationship (Crane, 1969), and one can assume that co-located inventors more often maintain such social relationships. Indeed, recent empirical evidence suggests that social relationships at least partly mediate localized knowledge flows (Agrawal, Kapur, & McHale, 2006b; Almeida & Kogut, 1999; Singh, 2005; Zucker, Darby, & Brewer, 1998).[1]

If social relationships among co-located individuals contribute to the geographic localization of knowledge flows, what happens to knowledge flows when an individual moves? In this study, we examine knowledge-flow patterns that occur when inventors move across borders. We base our hypotheses on three conjectures:

(1) Social relationships facilitate knowledge flows.
(2) Inventors are more likely to establish social relationships with colleagues (same firm) and other co-located individuals (same country[2]) than with random individuals from the overall population, conditional on working in related fields.
(3) Inventors' relationships with colleagues and other co-located individuals may persist after they move.

Based on these conjectures, we address three specific questions in this paper, each with respect to an inventor moving from a source firm in a source country to a receiving firm in a receiving country. Since we examine multiple scenarios with labor and flows moving in different directions, we provide an example with illustrations to clarify. Imagine an inventor who moves from a research lab at Siemens in Germany to a lab at IBM in Canada (Figure 1). In this case, Siemens represents the source firm, Germany the source country, IBM the receiving firm, and Canada the receiving country.

Figure 1 Cross-border labor mobility.

First, we focus on the flows from Siemens in Germany to inventors in Canada, above and beyond the increase in flows from Siemens to IBM specifically, that result from the move (Figure 2). Since we conjecture that social relationships facilitate knowledge flows, and that the mover will at least partially maintain relationships with the colleagues she just left and will create new relationships with others located in her new country, including some who do not work for her new employer, we hypothesize that knowledge flows will increase from Siemens to Canada beyond the growth in flows from Siemens to IBM.[3] Thus, more generally, we provide our first hypothesis:

> **Hypothesis 1**: We expect cross-border labor movement to increase knowledge flows from the source firm to the receiving *country*, above and beyond any additional flows to the receiving firm. We refer to this externality as *National Learning by Immigration*.

Second, we focus on the flows from Canada, not including those from IBM specifically, back to Siemens in Germany that result from the move (Figure 3). Since we conjecture that social relationships facilitate knowledge flows and that the mover will at least partially maintain relationships with the colleagues she just left and will create new relationships with others located in her new country, including some who do not work for her new employer, we hypothesize that knowledge

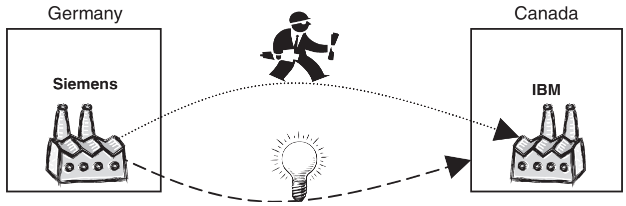

Figure 2 National learning by immigration.

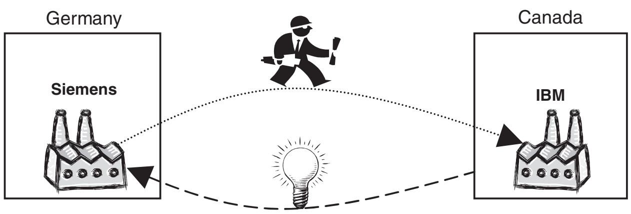

Figure 3 Learning from the diaspora: receiving country.

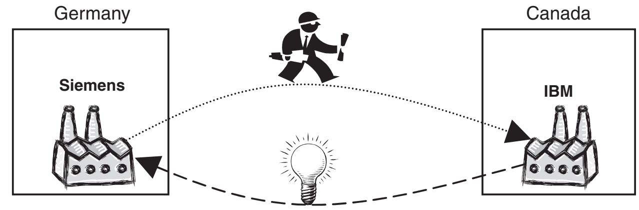

Figure 4 Learning from the diaspora: receiving firm.

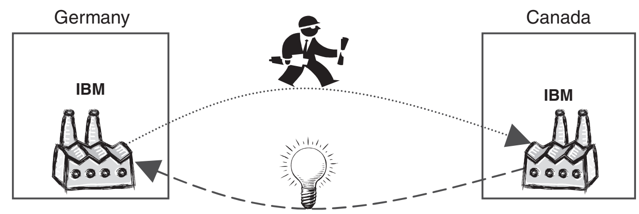

Figure 5 Learning from the diaspora: multinational.

flows will increase from Canada to Siemens. Thus, more generally, we posit our second hypothesis:

> **Hypothesis 2a**: We expect cross-border labor movement to increase knowledge flows from the receiving *country* back to the source firm, above and beyond any addition in flows from the receiving firm.

Next, as an extension of this hypothesis, we focus on the increased flows from IBM in Canada back to Siemens in Germany that result from the move (Figure 4). Since we conjecture that social relationships facilitate knowledge flows, and that the mover will at least partially maintain relationships with the colleagues she just left and will create new relationships with colleagues at her new firm, we hypothesize that knowledge flows will increase from IBM to Siemens. Thus, more generally, we extend our second hypothesis:

> **Hypothesis 2b**: We expect cross-border labor movement to increase knowledge flows from the receiving firm back to the source firm.

We refer to Hypotheses 2a and 2b as *Firm Learning from the Diaspora*.[4]

Finally, as a further extension of the second hypothesis, we consider backward knowledge flows associated with cross-border movement within the same firm. In other words, instead of moving from Siemens, we examine the case where an inventor moves from IBM Germany to IBM Canada; we focus on how that move affects knowledge flows from IBM Canada back to IBM Germany (Figure 5). Since we conjecture that social relationships facilitate knowledge flows, and that the mover will at least partially maintain relationships with the colleagues she just left and also will create new relationships with her new colleagues, we hypothesize that knowledge flows will increase from IBM Canada to IBM Germany. Note that in this case knowledge flows between the two locations may also be mandated and managed by the firm. Thus, more generally, we put forward the final extension to our second hypothesis:

> **Hypothesis 2c**: We expect within-firm, cross-border labor movement to increase knowledge flows from the receiving location back to the source location.

With respect to prior literature, only a few empirical studies, all of which use patent citation data, focus on estimating the relationship between labor mobility and knowledge flows.[5] Our work builds primarily on these papers. In particular, Almeida and Kogut (1999) present results suggesting that regions such as Silicon Valley that experience higher-than-average levels of inter-firm mobility tend also to experience a greater degree of knowledge localization, implying a direct relationship between labor mobility and knowledge flows. Song, Almeida, and Wu (2003) find evidence of firm ''learning by hiring'', and show the significance of this phenomenon by examining a firm's ability to access technologically distant knowledge from other firms through the recruitment of engineers.

Rosenkopf and Almeida (2003) also present evidence of learning by hiring. They explore two mechanisms that firms can utilize to expand their search for new knowledge: alliances and recruiting. While they find inconclusive evidence with respect to alliances, they discover clear outcomes with respect to mobility: recruiting firms benefit not only from the knowledge of the person they hire, but also from increased access to knowledge from the mobile inventor's prior firm. They also show that knowledge flows are greatest in technologies outside the recruiting firm's area of expertise.

The above papers all focus on knowledge flows the market might price. In other words, Song et al. (2003) and Rosenkopf and Almeida (2003) examine flows to the hiring firm. Hiring firms could include the expected value of increased knowledge flows from the mover's prior firm in the price they pay for the employee (through salary or other forms of compensation). Similarly, recruiting firms might price the knowledge flows associated with Almeida and Kogut's (1999) high inter-firm mobility regions. However, one study focuses on labor mobility and knowledge flows that firms certainly do not price. Agrawal, Cockburn, & McHale (2006a) report findings that suggest knowledge generated by an inventor who moves locations is more likely to flow back to that inventor's prior city than if the inventor had never lived there ("gone but not forgotten"). Clearly, the market does not price flows from an inventor back to her prior city: thus this represents an externality.

Our study builds on this set of prior papers. Like the papers on learning by hiring, our first hypothesis examines knowledge that flows in the same direction that the mover moves. However, unlike the learning by hiring papers, we focus on knowledge flow externalities. Whereas those studies measure the increase in flows from the source firm to the receiving firm that results from a move, we measure the increase in flows from the source firm to the receiving *country*, above and beyond those that go to the receiving firm. We discuss the policy implications of this externality in the final section of the paper.

In addition, like the "gone but not forgotten" hypothesis, our Hypotheses 2a, 2b and 2c examine knowledge that flows in the opposite direction to that of the mover. However, while that previous study focuses on flows back to the mover's prior city, we concentrate on flows that go back to the mover's prior *firm*. Since these flows go to the firm that the mover left, the market does not price them. Thus our paper extends the current literature on labor mobility and knowledge flows by examining the international *externalities* caused by the cross-border movement of inventors.

We organize the remainder of the paper as follows. In the next section, we describe our empirical methodology, including a discussion of our econometric approach and variables. In the Data and Variables section, we detail how we construct our data set, mostly from US patent data. In the Results section, we discuss descriptive statistics associated with our key variables, namely those that measure labor mobility and knowledge flows; we also interpret the regression results associated with tests for each of our hypotheses. Finally, we conclude by discussing the policy and strategy implications of these findings.

METHODOLOGY

We aim to deepen our understanding of the relationship between cross-border labor mobility and knowledge flows:

$$\text{Knowledge Flows} = f(\text{Labor Flows}) \qquad (1)$$

For measurement purposes, we use patent citation counts as a proxy for knowledge flows, and counts of "patenting inventors" who cross national borders as a proxy for labor flows. We describe the construction of these and other measures in the Data and Variables section. Here, we describe our method for empirically exploring the relationship between these two variables.

Specifically, we seek to address two questions. First, to what degree does immigration influence national knowledge inflows? And, second, to what degree do the migration patterns of a company's diaspora influence knowledge flows into the firm? In order to address these questions, we designed our study around the cross-border movement of inventors.

Our unit of analysis is the firm–country dyad. In other words, we determine the unit of analysis by the specific firm from which the mover moved (source firm) and the specific country to which the mover moved (receiving country). Therefore we distinguish between a mover who leaves IBM Germany for Canada and one who leaves IBM USA for Canada.[6] We use the same unit of analysis to address our two main research questions.

We explore these questions empirically using a 21-year panel data set (1980–2000) and the following base specification:

$$E[K|M, X, C] = \exp[\alpha(M) + \beta(X) + C + \varepsilon] \quad (2)$$

where the expected knowledge flow (K) is a function of the number of movers (M) from a *previous* time period, a vector of control variables (X), a full set of dyad fixed effects (C), and an error term ε.

We use a count model for our empirical specification, since the dependent variable is a count of patent citations between the source firm and receiving country. Therefore we assume that patent citations occur by means of a Poisson process. We model the expected level of knowledge flow as an exponential function of the number of movers to ensure non-negativity of knowledge flow, in line with similar studies of this nature (Henderson & Cockburn, 1996; Wooldridge, 2002). Because of this log-linear relationship, we interpret the estimated coefficient of movers (α) as the percentage increase in knowledge flow due to one more mover.

Overall, we aim to estimate the degree to which labor flows influence knowledge flows. In order to estimate this relationship properly, we must isolate any idiosyncratic heterogeneity that may exist between dyad members. If we believe this heterogeneity exists largely as a time-invariant effect, then a fixed-effects model, which estimates coefficients using within-dyad variation, will yield consistent estimates (Wooldridge, 2002).[7] Consequently, we drop dyads with no variation in knowledge flows across our sample time period.

While fixed-effects estimation captures time-invariant heterogeneity, we must also control for time-varying factors. We include patent flow measures, which capture a firm's and a country's patenting output, since we believe higher patenting entities are more likely to receive or provide knowledge spillovers. Additionally, we use patent stock variables for both firms and countries to capture the greater tendency of firms or countries with larger stocks to provide knowledge spillovers. Furthermore, we include measures that control for the degree to which movers themselves generate knowledge flows (since we seek to estimate the indirect effect of movers).[8] Lastly, we construct a technological similarity index to control for time-varying characteristics between the two members of the dyad that may influence their propensity to receive or provide knowledge flows to one another. We describe the construction of these variables in the following section.

Returning to our specific research questions, we first examine our National Learning from Immigration question (i.e., estimating the degree to which the arrival of a mover influences the receiving country's knowledge flows). For the dependent variable, we use a count of the number of citations per year made by the receiving country to the immigrant's source firm over the 21-year period under investigation (1980–2000, inclusive). We also examine our Firm Learning from the Diaspora question (i.e., estimating the degree to which the location decision of the mover who left influences the source firm's knowledge flows). To examine this phenomenon, we use as our dependent variable a count of the number of citations per year made by the mover's source firm to the mover's receiving firm (and country).

We focus our attention on the statistical significance and economic importance of the coefficient α. We interpret the value of this coefficient as indicating the degree to which movers influence knowledge flows. We remain concerned, however, by the potentially endogenous relationship between labor mobility and knowledge flows. While we do employ a lagged data structure in which we compare labor mobility in one period with knowledge flows in the following period, we interpret our results cautiously and discuss this issue further in the Results section.

DATA AND VARIABLES

Data Source

Our examination of knowledge flows begins with data commonly used in this setting: the patent data set compiled by the United States Patent and Trademark Office (USPTO) and refined by Hall and others associated with the National Bureau of Economic Research (Jaffe & Trajtenberg, 2002). Specifically, we aim to study the relationship between international labor mobility and knowledge flows. To accomplish this we construct a data set conditioned on labor movement so that we can estimate the effect of that movement on knowledge flows.

Unit of Analysis

Our unit of analysis is the source firm/country to receiving country dyad year. For example, Siemens Germany to Canada, 1990 represents an observation. Thus, in the context of our first hypothesis, National Learning by Immigration, we examine the degree to which a mover from Siemens Germany to a receiving firm in Canada, say IBM, in 1990 will increase the knowledge flows from Siemens Germany to firms in Canada in 1991, above and beyond the increase in flows that go directly to IBM that year. The potential number of dyads is extremely large. However, we condition our sample on dyads that actually experience a move during our 21-year sample period.

Sample Construction

We begin with the full set of patents issued by the USPTO between the years 1976 and 2004, inclusive. Approximately 3.2 million such patents exist. From these data, we identify "movers" by examining the inventor names on all patents and finding matches. Specifically, we search for identical inventor name pairs that have addresses in different countries (i.e., the inventor name character strings match exactly on first name, last name, and middle initial, if included). For example, if a patent lists someone as an inventor located in Germany in 1991 and exactly the same name appears on a patent as an inventor located in Canada in 1993, we flag this inventor as a potential mover.

Matching inventors purely on their names introduces the risk of type I errors (inventors may use multiple spelling permutations of their name: therefore we may miss actual movers) and type II errors (different inventors may have the same name: therefore we may erroneously flag someone as a mover). We do not address type I errors, and thus our sample serves as a conservative estimate of the overall levels of inventor migration. However, since we do not expect the likelihood of recording different name spellings across multiple patents to be associated with citation propensities, we do not expect this measurement error to bias our main result.

To minimize type II errors, we add the sampling restriction that the inventor's multiple patents must fall in similar fields as defined by:

(1) a match at the international classification subclass level (normally only one international classification per patent exists);
(2) a match at the US classification primary three-digit level; or
(3) a match between one of the patent's primary three-digit classifications and one of the secondary classifications of the other patent.[9]

Based on these criteria (name matching and field matching), we identify 37,200 moves,[10] some across the same dyad. For example, an individual moves from Siemens Germany to Canada in 1990 and another individual makes the same move two years later. Since we base dyads on moves in the sample, we clarify that the 37,200 moves represent 12,943 unique dyads. In other words, 24,257 moves occur across dyads already traversed by a prior mover.

Since we aim to identify the effect of movers on knowledge flows, we limit the inclusion of dyads to only those where no previous inventors moved during the period 1975–1979. We do this to minimize the effect of individuals who moved just prior to our period of analysis, which begins in 1980. Furthermore, we require that the source firm patented at least once between 1975 and 1979 as well as at least once between 2001 and 2004. This ensures the firm's existence by 1980 and continued existence in 2000: thus the firm was capable of receiving or providing citations throughout our period of analysis.

After imposing the data restrictions described above, our sample contains 2143 unique dyads.[11] Since we collect 21 years of data (1980–2000, inclusive) for each dyad, our panel data set consists of $21 \times 2143 = 45{,}003$ observations. However, owing to the lagged data structure employed in the regression analysis, we use only 20 years of data (since we measure knowledge flows at time t and labor flows at time $t-1$). Therefore we conduct regression analysis on $20 \times 2143 = 42{,}860$ observations.[12]

Dependent Variables

Our dependent variables measure knowledge flows. Specifically, following in the tradition of previous studies, we use a count of patent citations as a proxy for knowledge flows (Almeida & Kogut, 1999; Jaffe & Trajtenberg, 2002; Jaffe et al., 1993; Rosenkopf & Almeida, 2003; Song et al., 2003). The paper that pioneered the empirical measurement of knowledge flows using patent citations describes the methodology: "Thus, in principle, a citation of Patent X by Patent Y means that X represents a piece of previously existing knowledge upon which Y builds" (Jaffe et al., 1993: 580).[13] While patent citations do not perfectly quantify knowledge flows (in fact, they are rather noisy), we still regard them as useful measures of real knowledge flows (Jaffe, Trajtenberg, & Fogarty, 2002), and employ them for systematically gauging flows over large samples.

Since we explore two basic research questions (the second derives the third), we employ two different, but related, dependent variables. First, we use the variable Citations_{CjFit} to measure knowledge flows to the receiving country from the source firm (National Learning by Immigration). This variable counts the number of times inventors in the receiving country j cite the source firm i in year t. Specifically, we identify all issued US patents applied for in year t where at least one inventor lives in country j. Then, we count the number of citations made by this set of patents to prior patents assigned to firm i.[14]

Similarly, we use the variable Citations_{FiCjt} to measure the reverse knowledge flows from the receiving country to the source firm (Firm Learning from the Diaspora). This variable counts the number of times source firm i cites receiving country j in year t. Specifically, we identify all issued US patents applied for in year t where firm i is the patent assignee. Then, we count the number of citations made by this set of patents to prior patents where at least one of the inventors resides in country j. We remove inventor self-cites from the data and do not include them in these citation counts.[15]

Key Explanatory Variable

We use Movers_{FiCj} as our key explanatory variable for all research questions; it counts the number of inventors who move from source firm i to receiving country j in year $t-1$. Thus, as a flow variable, this measure reflects the number of movers in a given year, as opposed to the stock or cumulative number of movers who have migrated over time. Since we condition our sample on mobility, we describe the identification of movers in the sample construction subsection above. The construction of this variable simply counts those movers by dyad by year.

Control Variables

We estimate the relationship between annual flows of labor and annual flows of knowledge, paying particular attention to macro-level shifts in the economy that lead to changes in mobility and knowledge flows over time. We include a year trend in all specifications, including base specifications, to control for this possibility. This control operates as a counter from 0 to 20, where 0 corresponds to the year 1980 and 20 corresponds to the year 2000.[16]

In all of our fully specified models, we also include a measure of technological similarity. We do this to control for shifts in technology focus by the source firm, the receiving country, or both. If, for example, a US firm changed technology focus and suddenly increased its innovative output in the area of wireless voice communications, we might observe an increase in both knowledge flows and labor flows to and from Finland (home of Nokia and related companies). We include a control for technological similarity so that any change in knowledge flows, due to a change in technology focus, remains separate from the effect of labor mobility.

TechOverlap_{ijt} is a five-year moving average measure of the technological similarity between source firm i and receiving country j between the years $t-2$ and $t+2$:

$$\text{TechOverlap}_{ijt} = \frac{\sum_s p_{ist}p_{jst}}{\sqrt{\sum_s (p_{ist})^2 \sum_s (p_{jst})^2}} \qquad (3)$$

where p_{ist} reflects the share of patents issued to source firm i between years $t-2$ and $t+2$ that belong to NBER subclassification s.[17,18] A value of 1 denotes a perfect technological overlap between source firm i and receiving country j, and a value of 0 denotes no overlap.

Since we examine the effect of labor mobility on knowledge flows, and we hypothesized that knowledge flows may go in both the same and opposite directions as the labor movement, we need to control for any labor movement in the opposite direction to that on which we focus. In other words, in each of our research questions, we examine a particular type of knowledge flow associated with movement from a source firm to a receiving country. However, mobility may occur from the receiving country to the source firm. Therefore we include a control that measures the number of movers *to* the source firm in year $t-1$. We construct this reverse-mover variable in the same way we did our primary mover variable except that we count moves in the opposite direction.

In addition to the time trend, technology overlap, and reverse-mover control variables, which we include in our full estimation models for all research questions, we apply additional controls for each specific hypothesis. For the National Learning by Immigration question, we employ three additional controls. First, we control for the overall stock of innovation by the source firm, since firms with a greater stock of innovation are more likely to send and receive knowledge flows and also more likely to send and receive labor. Although we measure only within-dyad variation, stocks may change over time. So we construct PatentStock_{Fit} as a control, which counts the cumulative number of patents granted to the source firm between 1975 and year *t*, inclusive.

Next, since we focus the National Learning by Immigration hypothesis on knowledge flow externalities as opposed to knowledge flows that go to the receiving firm (as measured in Song et al., 2003), we control for flows that go to: (1) receiving firms; and (2) the mobile inventor herself. The former, $\text{Citations}_{\text{F}j\text{F}it}$, counts the number of citations the receiving companies' patents, applied for in year *t*, make to the stock of patents assigned to the mover's source firm. Similarly, $\text{MoverCitations}_{\text{F}j\text{F}it}$ counts the number of citations the receiving companies' patents, applied for in year *t*, on which the mover is an inventor, make to the stock of patents assigned to the source firm *i*. Thus, employing these two controls, we can interpret the coefficient on $\text{Movers}_{\text{F1C2}}$ as the effect of labor mobility on knowledge flows to the receiving country, above and beyond the flows that go directly to the receiving firm.

For the Learning from the Diaspora question, we add five additional control variables similar in spirit to the three additional controls employed in the previous model. Source firms that perform more innovation are more likely to draw upon knowledge from their diaspora and their diaspora's new network. We therefore include $\text{PatentFlow}_{\text{F}it}$, which counts the number of patents source firm *i* applies for in year *t* to control for the annual level of innovation at the source firm and thus the firm's propensity to receive knowledge flows.

As a related issue, the source firm is more likely to obtain knowledge flows from the receiving country or receiving firm when the receiving side generates more knowledge. We therefore include $\text{PatentStock}_{\text{C}jt}$, which counts the number of patents receiving country *j* applies for from 1975 up to and including year *t*, to control for the accumulated amount of innovation by the receiving country and thus the country's propensity to generate knowledge flows. Similarly, we include $\text{PatentStock}_{\text{F}jt}$, which we construct in an identical manner but at the receiving firm level rather than that of the receiving country.

Finally, we focus on measuring knowledge flows from the mover's new network back to the source firm, as opposed to flows from the receiving firm specifically or from the mover herself. This allows us to control for flows generated by: (1) the receiving firm; and (2) the mover. We construct $\text{Citations}_{\text{F}i\text{F}jt}$ and $\text{MoverCitations}_{\text{F}i\text{F}jt}$ as controls using counts of citations made by source firm patents, applied for in year *t*, that cite patents issued to the receiving firm. The latter counts those specific patents that list the mover as an inventor.

RESULTS

We begin by providing descriptive statistics of our key measures: labor flows and knowledge flows. Next, we offer summary statistics of all variables used in the multivariate analyses. Finally, we report regression results and discuss our interpretation of the coefficient estimates.

Descriptive Statistics

The data presented in Tables 1 and 2 allow us to examine worldwide trends in labor mobility and knowledge flows. Beginning with the aggregate international labor mobility data presented in Table 1, we see that countries vary substantially in both their inflow and outflow of inventors. Perhaps unsurprisingly, the United States received more than three times the number of inventors as Japan, the country with the second largest volume of inventor inflow. In fact, the four countries with the largest inflow of inventors – the US, Japan, Germany, and Great Britain – account for more than half (57%) of the inflows to the 26 countries present in our sample.[19] In other words, we find highly skewed inventor inflows across nations. We also find highly skewed inventor outflows. The same four countries account for 82% of the outflows, meaning that 82% of the movers in our sample moved out of one of those four countries.

Additionally, we find skewed knowledge flows, as seen in the data presented in Table 2. However, in this case, the US received only a little more than twice the level of knowledge flows as Japan, the second largest recipient. Yet the four countries with the largest knowledge inflows – the US, Japan, Germany, and France – account for more than three-quarters (78%) of the inflows to the 26 countries present in our sample. Moreover, the same four countries account for 81% of the knowledge outflows.

Interestingly, although the US exports more inventors overall than it imports (import:export ratio 0.72), it is a net importer of knowledge (ratio 1.25). Japan and Germany are net exporters of both inventors and knowledge. Great Britain has a balanced level of importing and exporting inventors, but exports more knowledge. France imports more inventors, but exports more knowledge. These data help in orienting the reader to the

Table 1 Worldwide movers

To

FROM		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	Total Outflow
1	Austria	x	8	0	0	2	0	20	0	0	0	0	1	0	0	0	0	0	0	0	0	5	0	0	0	0	0	36
2	Australia	8	x	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	15	0	24
3	Belgium	0	0	x	0	1	0	22	1	0	0	8	0	0	0	0	0	0	1	0	0	9	0	0	0	15	0	57
4	Canada	0	1	1	x	1	0	3	0	0	0	5	3	0	0	0	1	0	2	2	0	0	1	0	0	113	0	133
5	Switzerland	2	1	2	3	x	1	73	0	0	0	6	7	0	0	0	0	0	10	1	0	14	9	0	0	23	0	152
6	China	0	0	0	0	0	x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	Germany	53	5	58	17	119	1	x	9	13	1	50	43	5	2	5	0	6	21	79	8	20	1	7	4	199	0	726
8	Denmark	0	0	1	0	0	0	3	x	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	3	0	9
9	Spain	0	0	0	0	0	0	0	0	x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	8
10	Finland	0	0	0	0	0	0	0	3	0	x	1	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	8
11	France	0	2	8	6	8	0	20	1	4	0	x	13	0	0	0	2	0	2	16	0	13	0	0	0	103	0	198
12	Great Britain	1	11	8	14	6	0	10	0	0	0	39	x	1	0	10	1	0	3	3	1	9	0	2	1	272	0	392
13	Hong Kong	0	0	0	0	0	0	0	0	0	0	0	0	x	0	0	1	0	0	0	0	0	0	0	0	1	0	2
14	Hungary	0	0	0	0	0	0	1	0	0	0	3	3	0	x	0	1	0	1	0	0	0	0	0	0	11	0	20
15	Ireland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x	0	0	0	0	0	0	0	0	0	0	0	0
16	Israel	0	0	0	0	2	0	2	0	0	0	2	0	0	0	2	x	0	0	0	0	0	0	0	0	46	0	54
17	India	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x	0	0	0	0	0	0	0	0	0	0
18	Italy	0	0	0	0	9	0	7	0	0	0	2	0	1	0	0	0	0	x	2	0	0	0	0	0	11	0	32
19	Japan	2	7	16	19	6	2	48	0	0	0	19	67	8	0	5	1	1	4	x	13	17	1	6	7	714	0	963
20	Rep. of Korea	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x	0	0	0	0	0	0	0
21	Netherlands	10	2	48	3	0	5	3	0	2	0	18	21	4	0	1	0	0	1	4	0	x	0	3	2	9	0	136
22	Sweden	0	0	0	3	3	0	3	0	2	0	4	1	0	0	0	0	0	1	4	0	2	x	0	0	25	0	48
23	Singapore	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x	0	1	0	1
24	Taiwan	0	0	0	0	0	5	0	0	0	0	0	1	0	0	0	0	0	0	5	0	0	0	0	x	12	0	23
25	United States	7	54	58	204	81	18	253	6	29	2	125	228	19	5	14	145	50	60	392	85	106	15	53	180	x	1	2190
26	Yugoslavia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x	0
Total Inventor Inflow		83	91	200	269	238	32	468	20	50	3	282	391	38	7	37	152	57	106	508	107	195	31	71	194	1581	1	5212
Total Inventor Outflow		36	24	57	133	152	0	726	9	8	8	198	392	2	20	0	54	0	32	963	0	136	48	1	23	2190	0	5212
Ratio		2.31	3.79	3.51	2.02	1.57	0.00	0.64	2.22	6.25	0.38	1.42	1.00	19.00	0.35	0.00	2.81	0.00	3.31	0.53	0.00	1.43	0.65	71.00	8.43	0.72	0.00	

Note: Countries with ratios greater than 1 are net importers of inventors.

Table 2 Worldwide citations

To

FROM		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	Total Inflow
1	Austria	x	162	43	280	561	4	2720	49	22	133	712	481	10	13	9	36	1	389	3151	54	132	295	2	47	8530	4	17840
2	Australia	204	x	63	1174	487	2	3040	88	30	237	1142	1125	55	34	21	111	5	391	5126	109	274	506	17	117	27346	3	41707
3	Belgium	41	61	x	307	388	8	1818	40	17	45	743	602	12	14	17	73	3	258	5452	48	198	165	4	41	12494	0	22849
4	Canada	362	939	471	x	2216	45	9861	377	104	1082	5151	5306	197	104	114	730	8	1906	26658	844	1178	2966	31	563	134912	14	196139
5	Switzerland	449	362	370	1356	x	15	11557	460	128	361	3823	2792	243	92	76	230	17	1855	18426	248	986	1991	21	286	65929	6	112079
6	China	6	13	3	99	89	x	329	10	8	29	271	126	281	4	7	19	2	83	1405	130	63	37	5	199	4168	1	7387
7	Germany	2020	1549	1579	6136	11241	117	x	1213	419	2859	19594	15328	350	344	194	1082	29	6959	125162	2067	3505	5898	126	1766	297089	28	506654
8	Denmark	64	73	35	297	382	3	1673	x	16	103	586	561	25	26	7	75	21	170	2282	55	273	275	3	33	11534	0	18572
9	Spain	27	56	24	150	163	0	850	35	x	29	361	251	18	11	6	46	0	182	1195	34	74	94	1	30	4723	7	8367
10	Finland	110	160	88	1357	696	41	3780	133	20	x	1189	1750	133	35	6	84	2	323	8440	673	354	2488	10	555	28869	3	51299
11	France	590	728	633	3064	3644	52	18868	397	266	725	x	6125	185	173	121	470	20	2999	36615	1109	1586	1971	42	585	131148	27	212143
12	Great Britain	344	806	539	2995	2637	25	15636	458	136	457	6729	x	214	125	119	476	11	1713	28587	492	1197	1882	28	301	123632	10	189549
13	Hong Kong	20	57	12	372	229	72	694	21	48	49	457	334	x	9	1	52	1	1533	3509	158	88	89	12	399	10475	1	18692
14	Hungary	7	10	19	40	78	0	274	6	8	11	129	97	0	x	0	15	8	45	361	3	25	31	0	0	1534	7	2708
15	Ireland	10	39	16	146	125	2	414	50	9	9	145	195	6	8	x	69	1	95	956	20	35	97	1	26	5284	0	7758
16	Israel	44	176	117	803	544	13	2218	60	11	164	1095	1000	48	39	33	x	1	343	7670	268	237	494	21	435	31579	4	47417
17	India	29	16	12	52	96	10	223	21	15	15	153	177	1	12	17	21	x	71	650	39	46	26	0	14	2566	1	4283
18	Italy	397	204	279	1089	1961	28	8410	178	133	186	3302	2046	100	99	33	184	8	x	15729	537	637	718	32	585	45080	4	81959
19	Japan	2463	2829	4258	16658	15824	430	115125	1667	619	3762	38277	31255	1782	614	375	2855	71	13680	x	25364	7013	10491	752	13216	991699	61	1301140
20	Rep. of Korea	90	182	190	1741	553	76	4417	110	53	819	2212	1730	148	25	30	226	12	1155	69632	x	517	1021	296	4724	81098	2	171059
21	Netherlands	155	251	267	1090	1175	20	5409	218	43	368	2106	1722	107	51	21	197	13	830	14768	666	x	760	48	414	50495	2	81196
22	Sweden	196	377	243	2346	1879	18	7377	234	41	1629	2607	2439	63	82	41	456	5	780	15813	429	849	x	18	226	59110	8	97266
23	Singapore	3	17	7	95	73	4	272	6	9	29	163	176	23	0	5	30	0	57	2340	379	47	39	x	1107	7320	0	12201
24	Taiwan	66	172	100	1274	565	187	3608	53	55	535	1621	972	396	15	37	215	4	987	33379	5983	288	455	787	x	74673	8	126435
25	United States	10513	25463	15218	129698	82111	1580	419954	14225	3860	21412	186790	178235	7887	3866	3886	21301	578	57190	1381372	42863	42274	63877	3911	43238	x	261	2761563
26	Yugoslavia	1	3	0	5	10	0	14	0	1	1	14	5	1	1	0	3	0	7	36	0	1	2	0	0	108	x	213
Total Knowledge Outflow		18211	34705	24586	172624	127727	2752	638541	20109	6071	35049	279372	254830	12285	5796	5176	29056	821	94001	1808714	82572	61877	96668	6168	68907	2211395	462	6098475
Total Knowlde Inflow		17840	41707	22849	196139	112079	7387	506654	18572	8367	51299	212143	189549	18692	2708	7758	47417	4283	81959	1301140	171059	81196	97266	12201	126435	2761563	213	6098475
Ratio		0.98	1.20	0.93	1.14	0.88	2.68	0.79	0.92	1.38	1.46	0.76	0.74	1.52	0.47	1.50	1.63	5.22	0.87	0.72	2.07	1.31	1.01	1.98	1.83	1.25	0.46	

Note: Countries with ratios greater than 1 are net importers of knowledge.

relative magnitudes of labor and knowledge flows between countries in our sample. However, our analysis focuses on flows that originate at the firm level. We therefore turn next to descriptive statistics of the variables used in the empirical analysis.

In Table 3, we see that on average the mover's source firm produces approximately 100 patents per year. In addition, the source firm cites the mover's new country (receiving country) approximately 19 times per year, of which a little more than one citation (1.33) refers to the mover's receiving firm and an even smaller fraction of a citation (0.03) refers to the mover. In terms of the reverse, we see that on average the mover's receiving country cites her source firm approximately 24 times per year. Also, we see that a move occurs between the source firm and the receiving country on average 0.121 times a year. We must keep these mean values in mind when assessing the marginal impact of movers on knowledge flows estimated below.

National Learning by Immigration

Table 4 presents panel regression results to address the National Learning by Immigration question in which our dependant variable represents the flow of knowledge from source firm to receiving country as measured by patent citations. We employ maximum likelihood negative binomial regression analysis, a common estimation technique with these types of count data.[20] We include the patent stock of the source firm, defined as the cumulative count of all patents issued to the firm up until time *t*, in all regressions to control for the influence of a firm's patenting behavior on its likelihood of being cited. In addition, we account for source firm to receiving country dyad-specific heterogeneity

Table 3 Descriptive statistics

	Obs	*Mean*	*Std. dev.*	*Min*	*Max*
Firm-level data					
Patenting flow of source firm (1000s)	45,003	0.103	0.252	0	3.455
Patenting flow of receiving firm (1000s)	45,003	0.022	0.121	0	3.455
Patent stock of source firm (1000s)	45,003	1.192	2.662	0.001	28.844
Patent stock of receiving firm (1000s)	45,003	0.265	1.387	0	28.844
Patent stock of receiving firms (1000s)[a]	45,003	0.592	2.529	0	90.958
Country-level data					
Patenting flow of source country (1000s)	45,003	28.630	25.762	0	86.084
Patenting flow of receiving country (1000s)	45,003	12.815	21.489	0	86.084
Patent stock of source country (1000s)	45,003	337.064	337.646	0.015	1229.6
Patent stock of receiving country (1000s)	45,003	150.906	270.981	0	1229.6
Dyad-level data					
Citations by source firm to receiving country	45,003	19.493	121.495	0	5442
Citations by receiving country to source firm	45,003	23.601	126.140	0	4808
Citations by source firm to receiving firms[a]	45,003	1.329	23.471	0	2055
Citations by source firms to source firm[a]	45,003	1.455	20.576	0	1487
Citations by source firm to mover	45,003	0.032	0.540	0	42
Citations by mover to source firm	45,003	0.107	2.929	0	569
Citations by source firm to receiving firm[b]	45,003	0.510	4.430	0	371
Citations by receiving firm to source firm[b]	45,003	0.529	4.423	0	371
Citations by source firm to mover at receiving firm	45,003	0.023	0.439	0	42
Citations by mover at receiving firm to source firm	45,003	0.077	2.626	0	514
Movers from source firm to receiving country	45,003	0.121	0.447	0	17
Movers from receiving country to source firm	45,003	0.079	0.453	0	18
Movers from source firm to receiving firm	45,003	0.052	0.278	0	12
Movers from receiving firm to source firm	45,003	0.038	0.264	0	12
Technology overlap between source firm and receiving country	43,578	0.445	0.215	0	0.988
Technology overlap between source firm and receiving firm	21,186	0.578	0.297	0	1

[a]Often dyads contain more than one mover. Consequently, these movers may move to multiple firms. Receiving firms refers to this set of firms.
[b]This refers to the case where source firm and receiving Firm reflect different offices within the same multinational firm.

Table 4 National learning by immigration (H1): Conditional fixed effects negative binomial regressions, 1980–2000

Dependent variable	*Knowledge flows from source firm to receiving country*	
	(1)	*(2)*
Movers to receiving country	0.053 (0.006)***	0.044 (0.006)***
Movers to source firm		0.025 (0.006)***
Patent stock source firm	0.032 (0.002)***	0.023 (0.002)***
Cites by mover to source firm		0.001 (0.0003)***
Cites by receiving firms[a] to source firm		0.001 (0.0001)***
Technology overlap		0.922 (0.038)***
Year	0.122 (0.001)***	0.127 (0.001)***
Constant	−242.776 (1.697)***	−251.475 (1.754)***
Observations	39,160	38,232
Number of groups	1,958	1,953
Log likelihood	−82,500.67	−81,222.00
Chi squared	31,189.84	32,228.46
Prob > chi squared	0.0000	0.0000

[a]Often dyads contain more than one mover. Consequently, these movers may move to multiple firms. Receiving firms refers to this set of firms.
Note: All specifications contain dyad fixed effects. Standard errors are in parentheses.
***Significant at 1% level.

by using a fixed-effects estimation, utilizing the variation in knowledge flow within the dyad group across time to estimate the coefficients. Consequently, we drop dyads that exhibit no variation in knowledge flows across time.

Columns 1 and 2 present both base and full-model specifications. As evidenced in both columns, the number of movers produces a positive and statistically significant effect on the level of knowledge flows from the source firm to the receiving country. Since we construct all independent variables as levels, we interpret coefficients as the percentage change in the dependent variable, given a one-unit increase in the independent variable. The estimates presented in column 1 indicate that the arrival of a single mover results in an approximate 5% increase in knowledge flows from the source firm to the receiving country. While the magnitude of this coefficient may seem small, one must recall that we are measuring the effect of a single person on changes in knowledge flows at the country level.

In column 2 we present the full-model specification, wherein we add four additional controls. First, we include a measure that captures the degree to which movers from the receiving country to the source firm (i.e., ''reverse movers'') influence knowledge flows. Second, we include a control for knowledge flows from the source firm that the mover generates herself. Third, we control for the level to which the mover's receiving firm generates knowledge flows from the source firm to the receiving country. Lastly, we include a control for the technological overlap between the source firm and the receiving country in order to capture time-varying characteristics not captured by the dyad fixed effects.

The total number of observations drops, since 928 dyad years have no technological overlap index. The technological overlap measure depends on patenting activity: thus if either a country or a firm does not patent within a five-year window, then we cannot construct the index – in which case we drop the observation from our analysis. Overall, the results largely hold wherein the arrival of a single mover at time $t-1$ increases knowledge flows from the source firm to the receiving country at time t by more than 4%.

In addition to these two specifications, we run four sets of robustness checks.[21] First, we model the specification using a zero-inflated negative binomial approach (ZINB).[22] ZINB produces slightly stronger results than those presented in Table 4. Second, we loosen our restrictions on the lag structure of movers from the source firm to the receiving country. That is, we examine the extent to which movers who arrive at times $t-5$ through $t-1$, influence knowledge flow activity. The arrival of a single mover from the source firm to the receiving country at time $t-1$ still produces an approximate 4% effect on knowledge flows from the source firm to the receiving country.[23]

Third, we include a one-year lag of knowledge flows from the source firm to the receiving firm (our dependent variable) as an independent variable. The coefficient of movers to receiving country decreases from 0.044 to 0.038, but remains significant at the 1% level. Lastly, in an attempt to further capture possible time-variant dyad-level relationships not picked up by the dyad fixed effects, we include a control for mutual trade-bloc membership.[24] We set a dummy variable to 1 if the source firm country and the receiving country both

Table 5 Learning from the diaspora (H2a and H2b): Conditional fixed effects negative binomial regressions, 1980–2000

Dependent variable	*Knowledge flows from receiving country to source firm*		*Knowledge flows from receiving firms to source firm*[a]	
	(1)	*(2)*	*(3)*	*(4)*
Movers to receiving country	0.033 (0.008)***	0.030 (0.007)***	0.038 (0.014)***	0.037 (0.013)***
Movers to source firm		0.018 (0.007)**		0.037 (0.012)***
Patenting flow source firm	0.939 (0.015)***	0.962 (0.016)***	0.797 (0.032)***	0.764 (0.032)***
Patent stock receiving country	−0.0004 (0.0000)***	−0.0003 (0.0000)***		
Patent stock receiving firms[a]			0.005 (0.002)**	0.001 (0.003)
Cites by source firm to mover		0.037 (0.004)***		0.048 (0.005)***
Cites by source firm to receiving firms[a]		−0.001 (0.0000)***		
Technology overlap		1.397 (0.043)***		1.054 (0.109)***
Year	0.080 (0.001)***	0.080 (0.001)***	0.101 (0.003)***	0.102 (0.003)***
Constant	−159.340 (2.204)***	−160.700 (2.188)***	−201.970 (5.333)***	−203.286 (5.356)***
Observations	37,600	36,816	15,220	14,965
Number of groups	1880	1880	761	759
Log likelihood	−80,451.22	−79,239.98	−14,323.98	−14,172.54
Chi squared	14,743.46	15,807.44	3,503.20	3,751.89
Prob > chi squared	0.0000	0.0000	0.0000	0.0000

[a]Often dyads contain more than one mover. Consequently, these movers may move to multiple firms. Receiving firms refers to this set of firms.
Note: All specifications contain dyad fixed effects. Standard errors are in parentheses.
***, **Significant at 1 and 5% levels, respectively.

belong, during time *t*, to one of the following organizations: WTO, NAFTA, EU, or USIS. Not surprisingly, mutual membership in a trade bloc significantly affects knowledge flows, but the effect of a single mover from the source firm to the receiving country on knowledge flows persists, with only a slightly smaller magnitude of 3.8%.[25]

Firm Learning from the Diaspora

To investigate the Learning from the Diaspora question, we turn our attention to Table 5. The first two columns estimate the influence of the movement of an inventor on knowledge flows from the receiving country to the source firm (Hypothesis 2a). In columns 3 and 4, we separately test the extent to which knowledge flows from the mover's receiving firm to her source firm increase with inventor mobility (Hypothesis 2b).

Column 1 presents the basic Learning from the Diaspora specification estimates, where knowledge flows from the receiving country to the source firm at time *t* are a function of the number of movers at time $t-1$, the yearly patenting activity of the source firm, the patent stock of the receiving country, and a year trend. We control for time-invariant dyad heterogeneity using fixed effects. The estimated coefficient on movers indicates that a mover increases knowledge flows from the receiving country back to the source firm by approximately 3%.

Column 2 augments this specification by including the four additional controls. First, we include the number of movers from the receiving county to the source firm. Second, cites by the source firm to the mover controls for the possibility that the source firm does not learn from the mover's new environment, but just from the mover herself. Third, the source firm's citations to the receiving firm controls for flows from the mover's receiving firm; we examine the degree to which movers themselves increase flows from their receiving country, not just from their receiving firm. Finally, technological overlap controls for time-variant dyadic heterogeneity in technology profile. The main result continues to hold after adding these additional controls.

The last two columns shift from the firm–country to the firm–firm level of analysis, allowing us to examine the effect of movers on flows from their receiving firms back to their source firms. Column 3 establishes the base case, controlling for the yearly patenting activity of the source firm, a yearly time trend, and the patent stock of the mover's receiving firm.[26] Column 4 extends the base specification to include three additional controls:

(1) the reverse movers (movers from the receiving country to the source firm);
(2) the cites by the source firm to the mover; and
(3) the technological overlap control.

These estimations suggest that movers cause an approximate 4% increase in knowledge flows from the receiving firm back to the source firm.[27]

Table 6 Intra-firm learning from the diaspora (H2c): Conditional fixed effects negative binomial regressions, 1980–2000

Dependent variable	*Knowledge flows from receiving firm to source firm*	
	(1)	*(2)*
Movers to receiving firm	0.079	0.059
	(0.021)***	(0.021)***
Movers to source firm		0.094
		(0.020)***
Patenting flow source firm	0.907	0.821
	(0.041)***	(0.043)***
Patent stock receiving firm	0.029	0.026
	(0.007)***	(0.008)***
Cites by source firm to mover		0.049
		(0.006)***
Technology overlap		0.953
		(0.095)***
Year	0.103	0.093
	(0.003)***	(0.003)***
Constant	−206.685	−186.537
	(6.397)***	(6.856)***
Observations	13,200	10,370
Number of groups	660	622
Log likelihood	−10,774.58	−10,002.29
Chi squared	2,505.67	2,278.36
Prob > chi squared	0.0000	0.0000

Note: All specifications contain dyad fixed effects. Standard errors are in parentheses.
***Significant at 1% level.

Firm Learning from the Diaspora: Examining the Effect of Within-Firm Movers

We further explore Learning from the Diaspora in Table 6. In this case, we focus on movers who move across borders, but remain employed by the same firm (Hypothesis 2c). In other words, they locate to a new geographic site within the same multinational company. We use the same estimation technique and specifications as in Table 5. However, the coefficient on movers is approximately two times greater than in Table 5. We maintain a mixed reaction to the magnitude of this difference. On the one hand, it indicates that, as one might expect, firms manage knowledge flows more effectively within their boundaries than outside them. On the other hand, within the firm, labor movement *only* results in twice the increase in knowledge flows (as opposed to, say, ten times), suggesting perhaps that knowledge flows are difficult to control and manage. From a different perspective, if we consider the within-firm flow premium high, we could interpret this as suggesting that multinational firms do poorly at managing knowledge flows, since firm knowledge should flow between locations regardless of labor mobility, such that a within-firm move should not increase flows. Alternatively, if we perceive the premium as low, perhaps firms do a good job of managing cross-location knowledge flows. Clearly, we need to study further the implications of the magnitude of this difference.

Causality

We have interpreted the statistical correlation between labor flows and knowledge flows as the result of a causal relationship. That is, we assume labor flows cause knowledge flows. Of course, we must address issues of potential endogeneity and omitted variable bias when making such claims. We discuss these issues here.

We acknowledge potential endogeneity concerns, particularly with respect to National Learning by Immigration. While we assume that the movement of inventors from the source firm to the receiving country causes an increase in knowledge flows from the source firm to the receiving country, perhaps an increase in knowledge flows causes mobility. For example, because more inventors in the receiving country build on the ideas of the source firm, and possibly even on the potential mover specifically, firms within the receiving country have a higher propensity to recruit the mover. Moreover, the mover has a higher likelihood of being attracted to the receiving country.

We are also concerned about the potential for omitted variable bias. For example, knowledge flows and labor flows might both be correlated with national technology policy: a country that initiates a policy to foster a semiconductor industry by way of research grants, subsidies, and tax incentives might stimulate activity in this technology area that increases both the absorptive capacity of the nation (such that its inventors cite foreign inventors more frequently) and the propensity for local firms to hire related inventors from abroad. [28]

Ideally, to address these concerns, we would have an instrument correlated with movers, but not citations. In the absence of such an instrument, we rely on a lagged data structure, dyad-level fixed effects, and a control for technology overlap. We describe these approaches next.

First, we employ a lagged measure of labor mobility in our regression models. In other words, we examine the effect of labor flows at time $t-1$ on knowledge flows at time t. We employ this lagged time structure to reflect the causal relationship that we believe exists between labor and knowledge flows. However, our measures are messy, particularly for movers. Recall that we do not actually know the precise year when movers moved. We only know the last year they applied for a patent in their source country and the first year they applied for a patent in their receiving country. In other words, we may be late in estimating when they actually moved, but never early. However, this noise in the data will bias our results downwards.

Second, we employ firm–country fixed effects that control for time-invariant characteristics of the dyad. This modeling technique addresses some potential sources of omitted variable bias, such as distance between source firm and receiving country. Third, we include a control for technology overlap. In other words, if either a firm or a country changes (in response to a new policy, for example) such that the composition of its technological activity alters and becomes more similar to that of the other side in the dyad, our technology overlap measure will capture this.

CONCLUSIONS

We have found preliminary evidence of knowledge flows caused by labor mobility that occur outside intended market mechanisms. While the receiving *firm* may price the additional knowledge flows that it expects to receive as a result of the mover (reflected in the mover's salary, for example), the market does not price the flows we have identified that go to the receiving *country* above and beyond those that go to the receiving firm (Hypothesis 1; National Learning by Immigration); these flows represent an externality. We speculate that social relationships formed between individuals due to co-location that persist after separation are at least partly responsible for the knowledge flow patterns identified here. These externalities underscore the inability of firms to fully control or appropriate knowledge flows between inventors, even though they may try to impose restrictions on information dissemination. Although knowledge flows – a critical input for economic growth – are difficult to control, we see a clear role for policy in terms of influencing flow patterns to optimize social welfare.

Since a firm will invest in recruiting inventors only up to the point where the marginal benefit equals the marginal cost, and consider only the marginal benefit to itself and not to the nation in which it is situated, it is likely to under-invest in recruiting foreign talent from a welfare perspective. Thus, to the extent to which access to knowledge flows facilitated by movers creates significant externalities, welfare-enhancing policies might create incentives for companies to recruit more inventors than they otherwise would (up to the point where the marginal social benefit from an additional mover equals the marginal cost).

Countries with national strategies that involve focusing the allocation of resources on particular technological areas might find such policies particularly effective. For example, Canada has declared biotechnology, alternative energy, and wireless communications to be areas of strategic importance. Through federally funded centers of excellence, and other programs, the government encourages the formation and growth of firms in these areas. Lowering the cost for firms to hire international talent in these specific areas (through tax breaks, subsidies, etc.) may particularly benefit Canada, since the country has many companies working in these areas, often clustered geographically, and thus as a nation is likely to encompass sufficient absorptive capacity to exploit spillover knowledge flows accessible through the social networks of new immigrant inventors.

In addition, some countries, such as Canada, employ a points-based immigration system;

governments evaluate applicants based on a variety of criteria such as the value of their skills and education. To the extent that knowledge fuels economic growth and access to new knowledge encourages competitiveness, such evaluation systems might benefit from also considering the value of an applicant's network to their domestic economy. In other words, while two applicants might have identical skills and education, one might have access to a more valuable social network by virtue of working at a university or company with greater knowledge production capabilities; such an applicant might therefore facilitate more valuable knowledge flows to her receiving firm and receiving country. Thus the evaluation system could benefit from taking such networks into consideration.

In addition to the flows to the receiving country discussed above, the market does not price knowledge flows that go back to the source firm (that has lost the mover) from both the receiving firm and the receiving country. We refer to such flows as Firm Learning from the Diaspora (Hypotheses 2a, 2b and 2c). Although the market does not price these flows, representing externalities, they are directed to a specific firm – the source firm. Therefore firm strategy – not public policy – should be concerned about investments required to optimize these flows.

What types of strategy could enhance knowledge flows from the new networks of former employees who immigrate to other countries? Firms may consider investing in updating their relationships with their diaspora in order to increase the half-life of their relationships and thus the access to their network of knowledge flows. For example, consulting firms such as McKinsey and Company and the Boston Consulting Group offer "alumni" events designed to strengthen ties between current and former employees. Technology-oriented firms could make similar investments. While movers may no longer have an interest in their prior employer, they might maintain significant interest in their former colleagues and be quite willing to share knowledge that ultimately benefits their former employer.

More generally, as the locus of innovation continues to spread beyond the boundaries of countries such as the United States, Japan, Germany, and the United Kingdom to nations such as India, China, Israel, and Ireland, immigration patterns and the resultant knowledge flows will become an increasingly important feature of national innovation systems (Nelson & Rosenberg, 1993). For example, India will no longer play a role in the US innovation system as simply a source of educated and motivated students who emigrate to attend American universities and then stay on to work for American firms; instead, the US will value Indian immigrants for the access they provide to networks of knowledge creators located in India at organizations such as Wipro, Tata, Infosys, Ranbaxy, and the renowned Indian Institutes of Technology.

In other words, as the tight oligopoly of first-world innovation weakens, and the sources of knowledge creation become more geographically diverse, the management of knowledge flows will become increasingly complex and important. Nations and firms better able to harness these flows will enjoy a competitive advantage. In this paper we shed some light on the relationship between labor mobility and knowledge-flow patterns. However, this literature remains in its infancy. We use only a crude empirical estimation of the relationship, and have only a rudimentary understanding of the mechanisms that actually facilitate flows. Given the importance of knowledge flows to competitiveness and growth, much work remains.

ACKNOWLEDGEMENTS

We thank the *JIBS* departmental editor, Tom Murtha, and three anonymous reviewers for helpful suggestions and comments. The Social Sciences and Humanities Research Council of Canada (Grant No. 410-2004-1770), Human Resources and Skills Development Canada, and Industry Canada (Grant No. 537-2004-1001) funded this research. We gratefully acknowledge their support. Errors and omissions are our own.

NOTES

[1]Research on social capital provides a useful framework for understanding knowledge-sharing networks more generally. This research has been impressively multidisciplinary, with important contributions by sociologists (Burt, 1992; Coleman, 1988; Granovetter, 1973), political scientists (Putnam, 2000), and economists (Glaeser, Laibson, & Sacerdote, 2002; Knack & Keefer, 1997). In particular, the concepts of structural holes (Burt, 1992) and weak ties (Granovetter, 1973), which highlight the special role of individuals who provide access to non-redundant networks, offer a useful conceptual framework for explaining why the international movers studied here impact on knowledge flows so significantly: they provide access to knowledge networks that neither the receiving firm

and country nor the source firm and country might otherwise have.

[2]Jaffe et al. (1993) find that knowledge flows disproportionately within the city, state, and even the country of the inventor.

[3]No gender assumptions should be inferred from our hypothetical inventor being female; references to the feminine should be understood to include the masculine and vice versa.

[4]We use the term ''diaspora'' in this paper to describe groups of individuals who share a common history in terms of the firm by which they used to be employed and by the country in which they used to live. So, for example, the IBM Canada diaspora refers to the former employees of IBM Canada who now work for other firms and perhaps in other countries. Other scholars have used the term ''diaspora'' in a similar context, such as Kapur and McHale (2005).

[5]These papers focus directly on the relationship between labor mobility and knowledge flows. However, other empirical papers address the related link between social relationships and knowledge flows, such as those by Zucker et al. (1998) and Singh (2005). These papers also relate closely to our topic of interest, since we conjecture that labor mobility matters because of the effects of residual social relationships that persist after separation.

[6]We treat member nations of the European Union as distinct countries.

[7]By allowing each dyad member to have its own intercept in the regression specification, we control for omitted time-invariant variables, such as geographic distance and cultural characteristics.

[8]These measures include the number of citations made by the source firm to the mover herself, as well as the number of citations made by the mover to the source firm.

[9]Although we make considerable efforts to minimize measurement errors with respect to identifying movers, our process is by no means perfect. We offer three points regarding the nature and implications of this measurement error. First, we intend for the technology field matching process to remove from the sample individuals who share the same name, but who do not work in the same technology area. By employing this process, we do not, for example, falsely identify Michelle Scott as a mover if there is actually one Michelle Scott who works in textiles in Canada and another who works in electrical connectors in the United States. However, if both Michelle Scotts work in electrical connectors, we still will wrongly identify her as a mover.

Second, measurement errors, such as the Michelle Scott example, will bias our main result downwards. In other words, if we erroneously assume an individual is a mover when two different people actually exist, then we will increase the mover variable, but we cannot reasonably expect a related increase in knowledge flows. This will weaken the estimated coefficient on movers. Also, errors in the other direction (we miss actual movers if they spell their names differently before *vs* after the moves, for example) will add noise to the data but will not systematically bias the results in favor of our hypotheses.

Third, to offer the reader some sense of the potential magnitude of the measurement error, we have calculated the fraction of ''suspicious'' instances where the same name from our sample patented during the same year from two different organizations (1.32%). We also have calculated the fraction of ''suspicious'' mover instances where an inventor moves from firm A to firm B and later moves back to firm A (2.32%). While the small fractions calculated here do not prove a small error (in fact, the measured phenomenon might indicate moves that actually occurred), they offer comfort that the potential error is not obviously large.

[10]If the same inventor moves from country A to country B and then to country C, we observe two direct moves (from A to B and B to C) and one indirect move (from A to C). We do not distinguish between direct and indirect moves for the purposes of this analysis.

[11]As described in this section, we condition our sample on firms that existed throughout the study period (1980–2000). This results in dropping a significant fraction of observations. As a result, we bias our sample towards movers who leave larger, older firms. This will not obviously bias our estimated relationship between knowledge flows and labor mobility in either direction, but we note this potential concern and offer the caveat that the generalization of our results to firms of all sizes be considered with caution.

[12]Note that if no variation in the dependent variable exists (i.e., the number of citations remains constant over the 20-year period; this usually occurs when a dyad receives zero citations over the temporal period of our sample), then we drop the observation. This explains why the number of observations falls below 42,860 in certain cases. For example, 1958 groups exist in the base specification in Table 4 instead of 2143. Furthermore, when we employ the full specification, we drop additional observations because the technology overlap index cannot always be computed since we base this measure on a five-year moving average, and sometimes either end of the dyad does not produce any patents during that period.

[13]Although similar in spirit to academic article citations, patent citations serve as a more strict measure of knowledge exchange. For academic article citations, including an additional citation costs close to zero. However, for a patent the cost may be higher, since an additional citation may further reduce the scope of the claims over which it grants the inventor protection, thus reducing its value. We therefore expect fewer spurious citations in patents than in academic article citations.

[14]We note that this measure counts citations, not patents. In other words, we count the number of citations to firm *i*, conditional on the patent having application year *t* and at least one of the inventors listed as residing in country *j*. A patent with such characteristics may not cite firm *I*, or may cite more than one piece of prior art belonging to firm *i*, and thus such a patent can increment the citation count by an integer value of 0, 1, or more than 1.

[15]Also, if a cited patent lists multiple inventors located in multiple countries, we count each country.

[16]We also use year dummies instead of a year trend variable. The results remain largely unchanged, but we achieve maximum likelihood estimation convergence more consistently using a time trend.

[17]We adopt the NBER patent classification schema (Jaffe & Trajtenberg, 2002), which aggregates the approximately 420 three-digit USPTO Utility Classes into 36 classes. Whereas the USPTO schema is intended to aid patent examiners with prior art research, the NBER schema aims to reflect basic technology application categories. For example, the NBER classification code of 46 corresponds to Semiconductor Devices, which consists of four USPTO Utility classes: Active Solid-State Devices (257); Electronic Digital Logic Circuitry (326); Semiconductor Device Manufacturing: Process (438); and Superconductor Technology: Apparatus, Material, Process (505).

[18]Jaffe (1986) created this index, and referred to it as an ''uncentered correlation coefficient''. Whereas we use the index to measure the technological distance between the source firm and receiving country, Jaffe uses it to measure the technological distance between a focal firm and another firm in its industry. Jaffe employs this to develop a measure of the potential spillover pool available to a firm by multiplying the technological distance measure by each dyad member's R&D spending: the closer a focal firm exists to another firm in technology space, the more it will benefit from the other firm's R&D spending. We follow the more recent literature that has built upon this measure to estimate technological positions between two patenting entities (Acs, Audretsch, & Feldman, 1994; Branstetter, 2001; Peri, 2005; Wu, Levitas, & Priem, 2005).

[19]The 26 countries listed in Table 1 represent all nations that movers in our sample moved *from*.

[20]We find the Poisson assumption of first and second moment equality too strong for these data. While we still obtain consistent parameter estimates through a Poisson regression model, we greatly underestimate the standard errors, making hypothesis testing difficult. Instead, we adopt the negative binomial regression model, which allows the expected mean of knowledge flows to be proportional to the expected variance (Hausman, Hall, & Griliches, 1984).

[21]We do not present the robustness check tables here, but will provide them upon request.

[22]ZINB, as developed by Greene (1994), assumes that the dependent variable consists of two states unknown to the researcher. In the first regime the likelihood of a variable taking on a value above zero is low, while in the second regime the variable follows a Poisson distribution, where the variable can take on values of both zero and greater. As a result, ZINB estimation involves two distinct parts. The first part distinguishes which regime the observation falls into, in turn ''inflating'' the zero. We follow tradition and estimate this process using a logit regression. We then use a negative binomial regression to provide coefficient estimates.

[23]Movers from the source firm to the receiving country at times $t-5$, $t-4$, $t-3$ and $t-2$ are all insignificant.

[24]Andrew Rose kindly provides these data on his website: http://faculty.haas.berkeley.edu/arose/

[25]We perform these robustness checks on the specifications presented in Tables 5 and 6 as well. Results hold throughout.

[26]Recall that the source firm may have multiple movers to Country 2. In this case, ''patent stock'' is the sum of the patent stocks of each recipient firm in Country 2.

[27]As we examine the effect of movement on knowledge flows in an aggregate sense, we remain agnostic as to the motivation for the move. However, different reasons for moving certainly may result in different flow patterns. For example, if an individual leaves a firm on unfriendly terms because of a falling out, she may sever ties with former colleagues and thus be much less likely to facilitate knowledge flows back to the source firm than another inventor who leaves due to a spouse relocating or for other such reasons. Thus motivation for moving may play an important part in terms of predicting the resultant knowledge flow patterns.

[28]Another example is developing nation ''catch-up'' policies. It is possible that such policies influence the

behavior of both knowledge flows and labor flows in our data set, since we study a reasonably long period (1980–2000) during which many countries made explicit efforts to increase their participation in the innovation-oriented economy. However, since we use dyad fixed effects in our estimations, we take a conservative approach and consider only within-dyad variation. So, empirically, we have no reason to discount or control for public policies that ''artificially'' increase labor flows to a particular country, which in turn cause an increase in knowledge flows to that country. We want to capture that and attribute the increase in knowledge flows to the increase in immigration. That will not introduce bias into our measure. However, if the policy directly increases both immigration and knowledge flows (in other words, some policy mechanism separate from immigration increases knowledge flows), this presents a problem. For such an effect to bias our measure, the policy would have to increase knowledge flows the year after it increases immigration, and some mechanism other than immigration would have to influence those knowledge flows. We note this possibility but do not consider it a likely occurrence in our data.

REFERENCES

Acs, Z. J., Audretsch, D. B., & Feldman, M. P. 1994. R&D spillovers and recipient firm size. *Review of Economics and Statistics*, 76(2): 336–340.

Agrawal, A. 2006. Engaging the inventor: Exploring licensing strategies for university inventions and the role of latent knowledge. *Strategic Management Journal*, 27(1): 63–79.

Agrawal, A., & Cockburn, I. 2003. The anchor tenant hypothesis: Exploring the role of large, local, R&D intensive firms in regional innovation systems. *International Journal of Industrial Organization*, 21(9): 1227–1253.

Agrawal, A., Cockburn, I., & McHale, J. 2006a. Gone but not forgotten: Knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, 6(5): 593–617.

Agrawal, A., Kapur, D., & McHale, J. 2006b. *Birds of a feather – Better together? How co-location and co-ethnicity influence knowledge flow patterns*. Mimeo, University of Toronto, Toronto, ON.

Almeida, P., & Kogut, B. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45(7): 905–917.

Audretsch, D. B., & Feldman, M. P. 1996. R&D spillovers and the geography of innovation and production. *American Economic Review*, 86(3): 630–640.

Branstetter, L. G. 2001. Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the US and Japan. *Journal of International Economics*, 53(1): 53–79.

Burt, R. S. 1992. *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Coleman, J. S. 1988. Social capital in the creation of human capital. *American Journal of Sociology*, 94(Supplement): S95–S120.

Crane, D. 1969. Social structure in a group of scientists: A test of the ''invisible college'' hypothesis. *American Sociological Review*, 34(3): 335–352.

Dasgupta, P., & David, P. 1987. Information disclosure and the economics of science and technology. In G. Feiwel (Ed.), *Arrow and the ascent of modern economic theory*: 519–542. New York: New York University Press.

Dasgupta, P., & David, P. 1994. Towards a new economics of science. *Research Policy*, 23(5): 487–521.

Glaeser, E. L., Laibson, D., & Sacerdote, B. I. 2002. An economic approach to social capital. *Economic Journal*, 112(483): F437–F458.

Granovetter, M. S. 1973. The strength of weak ties. *American Journal of Sociology*, 78(6): 1360–1380.

Greene, W. 1994. *Accounting for excess zeros and sample selection in Poisson and negative binomial regression models*. NYU Department of Economics Working Paper Number EC-94-10, New York University, New York.

Hausman, J., Hall, B. H., & Griliches, Z. 1984. Econometric models for count data with an application to the patents–R&D relationship. *Econometrica*, 52(4): 909–938.

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

Jaffe, A. B. 1986. Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review*, 76(5): 984–1001.

Jaffe, A. B., & Trajtenberg, M. 2002. *Patents, citations, and innovations: A window on the knowledge economy*. Cambridge, MA: MIT Press.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, 108(3): 577–598.

Jaffe, A. B., Trajtenberg, M., & Fogarty, M. S. 2002. The meaning of patent citations: Report on the NBER/Case-Western Reserve survey of patentees. In A. B. Jaffe & M. Trajtenberg (Eds), *Patents, citations, and innovations*: 379–402. Cambridge, MA: MIT Press.

Kapur, D., & McHale, J. 2005. *Give us your best and brightest: The global hunt for talent and its impact on the developing world*. Washington, DC: Brookings Institution Press.

Knack, S., & Keefer, P. 1997. Does social capital have an economic payoff? *The Quarterly Journal of Economics*, 112(4): 1251–1288.

Nelson, R., & Rosenberg, N. 1993. Technical innovation and national systems. In R. Nelson (Ed.), *National innovation systems*: 3–22. Oxford: Oxford University Press.

Peri, G. 2005. Determinants of knowledge flows and their effect on innovation. *Review of Economics and Statistics*, 87(2): 308–322.

Putnam, R. D. 2000. *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

Romer, P. M. 1986. Increasing returns and long run growth. *Journal of Political Economy*, 94(5): 1002–1037.

Romer, P. M. 1990. Endogenous technological change. *Journal of Political Economy*, 98(5 Part 2): S71–S102.

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

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

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

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

Wooldridge, J. M. 2002. *Econometrics analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Wu, S., Levitas, E., & Priem, R. L. 2005. CEO tenure and company invention under differing levels of technological dynamism. *Academy of Management Journal*, 48(5): 859–873.
Zucker, L., Darby, M., & Brewer, M. 1998. Intellectual capital and the birth of US biotechnology enterprises. *American Economic Review*, 88(1): 290–306.

ABOUT THE AUTHORS

Alexander Oettl (MSc, Queen's University, Canada) is a PhD candidate in Strategic Management at the Rotman School of Management, University of Toronto. His research interests include the economics of innovation, knowledge spillovers, the economics of talent, and economic geography. Born in Canada, he holds both Canadian and German citizenships. He can be reached at aoettl@rotman.utoronto.ca.

Ajay Agrawal (PhD, University of British Columbia, Canada) is the Peter Munk Professor of Entrepreneurship at the Rotman School of Management, University of Toronto. Professor Agrawal works on topics that include the economics of innovation, the economics of creative industries, the market for ideas, and entrepreneurship. He is a Canadian citizen and can be reached at Ajay.Agrawal@rotman.utoronto.ca.

Accepted by Thomas Murtha, Departmental Editor, 13 June 2007. This paper has been with the authors for three revisions.