The Effect of International Migration on Productivity: Evidence from Randomized Allocation of U.S. Visas to Software Workers at an Indian Firm

Michael A. Clemens[*]

Center for Global Development, Email: mclemens@cgdev.org

December 3, 2012

Abstract: Why do Indian software workers employed by U.S. firms earn more in the U.S. than in India? There are several possibilities, among which is the pure effect of location on workers' economic product. This study seeks to isolate the location effect from other effects in a single setting via a natural experiment: a randomized allocation of temporary U.S. visas among one group of Indian software programmers. In this setting, outputs are close to perfectly tradable, workers in the U.S. and India are observably and unobservably identical in expectation, effects like Baumol's cost disease and cost-of-living compensating differentials are less relevant, and some of the plausible effects of place on productivity (such as access to technology) are identical for both groups. The large majority of the earnings gap remains, suggesting that these workers are several times more economically productive solely due to working in the U.S. rather than in India. This effect is measured for a single firm and external validity is circumscribed. Further study of the effect of location on economic product has implications for the economic gains to migration, trade, and technology transfer.

JEL Classification Numbers: O1, O4, F22, J61.

1 Introduction

Why would an Indian worker in the United States earn more than an Indian worker in India? Economists have proposed numerous explanations.

The two workers might use different technology (e.g. Hall and Jones 1999). They might work in industries whose output is imperfectly tradable, and demand for that output could be higher in the U.S. The worker in the U.S. might have better reserve options in other sectors—as in Baumol's cost disease (Baumol and Bowen 1966). Nominal wage differences could simply reflect differences in the cost of living (such as due to the Balassa-Samuelson effect). The two workers could differ inherently, such as in observed or unobserved skill, due to self-selection and sorting of Indian migrants (e.g. Borjas 1987; Grogger and Hanson 2011). Finally, the workers' locations themselves could affect the economic value of their marginal output—such as through differences in a variety of nontradable inputs to production such as local public goods of institutions, geography, and agglomeration economies (e.g. Krugman 1991; Kremer 1993; Ellison et al. 2010; Kerr 2010).

This paper exploits a unique natural experiment to test a simple model of the effects of country-of-work on the earnings and economic productivity of one group of software workers from India. Indian software workers in general have much higher wages in the United States than in India. A quirk in the administrative process for granting temporary skilled-worker ("H-1B") visas to the United States in 2007 and 2008 caused the U.S. government to randomize which visa applications it processed. This created two groups of Indian workers identical in every way, in expectation, except the country where each would work after being hired. I observe these workers in the personnel records of a single major India-based multinational software consulting firm supplying workers to U.S. clients. This allows a test of the relative importance of country-of-work, in isolation, to the economic productivity of these workers.

This paper does *not* seek to precisely decompose each of the various possible causes of

U.S.-India wage differences in the software industry. It has the more modest goal to rule out, in a single setting, some of the above channels as the principal cause of those wage differences. The external validity of its results is tightly circumscribed to the setting examined.

In this unusual setting, international differences in technology are negligible: the Indians in the U.S. are employees of the same software consulting firm, staffed to the same project, as the Indians in India. Their output is almost perfectly tradable. The U.S. clients who purchase the services of these workers need not compete with reserve options in other U.S. sectors, because U.S. client firms hire the workers out of India as consultants on work visas tied to their specialized knowledge of software. Higher prices in the U.S. have limited relevance to a U.S. employer's decision to bring a worker from India to the U.S., because the employer has the reserve option to hire workers to complete the software project *in India*. And most uniquely, the Indian workers in this setting are in expectation identical, both in observed and unobserved traits, because those sent to the U.S. are randomly chosen.

I find that for this group of Indian software workers at one firm, country-of-work alone accounts for roughly three quarters of the U.S.-India wage gap. This suggests large differences in workers' economic productivity caused exclusively by local, nontraded inputs accessible in different countries.

Intuitively, this conclusion begins from the observation that U.S. clients are willing to pay the Indian consulting firm about US$75,000 more, per worker per year, to have an Indian worker physically present in the U.S. rather than in India. It would be rational for U.S. clients to do this if the workers who come to the U.S. are more skilled or use better technology, but in the present setting workers are identical in expectation and use identical technology. It would *not* be rational for them to do this if the US$75,000 substantially reflected something other than increased marginal revenue product for the client, such as a compensating differential to keep the Indian worker from leaving for other employment opportunities or pay a higher cost of living. If that were the case, the client would get less

additional revenue than its additional cost—and would lose money for each Indian worker it brought to the United States. The perfect tradability of the output means that clients always have the reserve option of having the same software written in India.

Together, these imply that the workers in the U.S. have a much higher marginal revenue product for the client, for reasons other than their individual characteristics or the technology they use. Their economic product is higher mostly because of where they are. Kerr and Lincoln (2010) and Hunt (2011) offer evidence that an important channel of H-1B workers' economic contribution to the U.S. client firm arises through stimulating innovation. Paserman (2011) offers suggestive evidence that different high-skill workers, in Israel, raise productivity at the destination firm.

Methodologically, this study lies in a recent literature using natural randomization of visas (Gibson et al. 2009; Clingingsmith et al. 2009; McKenzie et al. 2010) and housing vouchers (Kling et al. 2007) to isolate the effects of location alone on workers and families. Its finding relates to different strands of literature. Work in growth and development economics has long asked how much of international gaps in earnings results from differences in human capital and technology, and how much it depends on nontraded complementary inputs (Clark 1987; Lucas 1990; Young 1995; Easterly and Levine 2001; Easterly 2004). Work in urban and labor economics has investigated what portion of spatial differences in worker productivity arises purely from location rather than from human capital (Rauch 1993; Glaeser and Maré 2001; Freeman 2006; Clemens et al. 2008; Ortega and Peri 2009; Ashenfelter 2012), and has suggested that spatial agglomeration accounts for large portions of domestic differences in labor productivity (Ciccone and Hall 1996).

I begin with a very simple model of international wage gaps in the software industry. I proceed to describe the natural experiment and the firm that provided the data. I present the results on wage effects of country-of-work, followed by arguments for why they should be interpreted primarily as differences in the marginal revenue product of U.S. client firms.

3

I conclude by discussing two important challenges to the internal validity of the results—attrition and contamination biases—and by delimiting the external validity of the findings.

2 The U.S.-India wage gap for software workers

Commander et al. (2008) report that developers and programmers in the U.S. software industry earn about US$71,000 more (10.9 times) what their counterparts in India earn in similar jobs. Adjusted for differences in purchasing power, workers in the U.S. earn $62,240 (4.4 times) what those in India earn.[1] Differences this large demand explanation.

2.1 Three classes of theory

Consider three broad classes of explanations for spatial differences in earnings, reviewed in a general context by Rosenthal and Strange (2004).

First, differences in wages between countries may exist without differences in the productivity of equivalent workers, for several reasons. Labor markets or output markets can be less competitive within some countries than others. Workers may produce goods that are nontradable, or simply nontraded, between countries. There may be barriers to the flow of technology from one country to another. Workers in different countries can have different levels of portable individual ability, either as formal training or tacit skill, acquired before

---

[1]Commander et al. (2008, Table 5) report the earnings of software developers, defined thus: "Developers work on specifying, designing and constructing information technology artefacts. They tend to be responsible for working out the programming inputs of a project and managing the allocation of tasks." They report average annual US earnings of $58,395 in 2002 dollars, equal to $78,402 in 2009 dollars adjusted by the US Consumer Price Index (CPI). This agrees well with the May 2009 mean US wage for "computer programmers" of $74,690 and "computer systems analysts" of $80,430 reported by the U.S. Bureau of Labor Statistics (2010). Commander et al. also report the earnings of comparable workers in India at US$5,360 in 2002 dollars (at market exchange rates), equal to US$7,196 in 2009 dollars inflated by the US CPI. This agrees well with data gathered Kelly Services (2009) on earnings in India's information technology sector, such as "analyst programmers" and "systems engineers" with a tertiary degree and about three years of experience (that is, roughly 25 years old) in Table 8. Adjusted for price differences between the U.S. and India, the same workers earn $80,754 in the U.S. and $18,515 in India (in Purchasing Power Parity dollars inflated to 2009 dollars with the US CPI). The difference betweeen Commander et al.'s U.S. PPP figure and U.S. exchange-rate dollar figure arises from the fact that they use wage data from 2002 but a PPP conversion from 2000.

entering the labor force.

Second, the productivity of initially equivalent workers could indeed differ across countries, because nontraded inputs to production in some places cause workers to acquire more human capital on the job. This could include knowledge spillovers from working with knowledgeable colleagues, or changes in workers' expectations of themselves or others. Glaeser and Maré (2001) call this a "growth effect" of location on human capital.

Third, again the productivity of initially equivalent workers could differ across countries, but for a different reason: that local nontraded inputs augment a worker's existing human capital rather than form new, portable human capital. Glaeser and Maré call this a "level effect" of place on productivity. These could include local externalities purely from working with high human capital workers, lower transactions costs, better matching of workers with complementary skills, local rules and institutions, and advantages of geography and infrastructure.

One way to distinguish empirically among these theories is to investigate an episode of exogenous movement of workers from one country to another. Suppose this occurs in a setting where barriers to trade in outputs and barriers to technology flow are very low, and labor markets are reasonably competitive in both countries.

In such a setting, 1) if wage differences between countries do not arise from differences in productivity caused by location—arising instead from differences in human capital—then an exogenous change in country-of-work will not substantially affect wages. 2) If instead differences in wages arise from differences in productivity, and these are caused by the effect of local nontraded inputs on portable human capital, then an exogenous change in country-of-work will affect wages. Furthermore, this effect will persist if workers later move back to the country of origin. 3) If differences in earnings arise from differences in productivity, and these are caused by strictly local forces, then an exogenous change in country-of-work will affect wages, and this effect will not persist if they later move back to the country of origin.

In subsequent sections I attempt to disentangle these models by observing changes in wages during and after an exogenous change in country-of-work.

2.2   A simple model

This test can be formalized as simply as possible with a spatial equilibrium model of labor demand and supply, following Rosen (1979), Roback (1982), Glaeser and Maré (2001), and related to the offshoring and outsourcing models of Antràs and Helpman (2004) and Grossman and Rossi-Hansberg (2008). Suppose that a firm in the U.S. chooses between two options: 1) produce a good in the U.S. with Indian workers brought to the U.S., or 2) produce the good in India with Indian workers, and bring the good to the U.S.

The firm's production function takes a standard Cobb-Douglas form and it maximizes profits $\pi = (1 - \tau) K^{\frac{\sigma-1}{\sigma}} (\theta h L)^{\frac{1}{\sigma}} - wL - rK$, where K is capital, L is labor, $\sigma > 1$ is a production parameter, and τ reflects the cost of transporting the good to the U.S. The parameter $\theta > 0$ describes the effect of location on a worker's productivity; the parameter $h > 0$ describes the workers' human capital, which influences productivity regardless of location. The wage is w and the cost of capital is r.

The first order condition for profit maximization fixes the capital-labor ratio, and the assumption of free entry (zero-profit condition) fixes the wage in terms of the capital labor ratio. Combined, these conditions determine wages

$$w = \varphi \left(1 - \tau\right)^{\sigma} r^{1-\sigma} h\theta, \tag{1}$$

where $\varphi \equiv (\sigma - 1)^{\sigma-1} \sigma^{-\sigma}$. The simplicity of (1) requires substantial competition in the output market and the labor market, both of which do characterize the software industry. In both India and the U.S. it is a vibrant and dynamic industry in which thousands of firms of all sizes compete, entry and exit is frequent, and workers often change firms, leave the

sector, and enter the sector (Arora et al. 2001; Arora and Athreye 2002; Commander et al. 2008).

Other features of the experimental setting in this study allow even greater simplification. Let subscript u represent Indian workers in the U.S. and subscript i represent Indian workers in India. Any difference in wages between the two workers may be expressed as

$$w_u - w_i = A_u \theta_u - A_i \theta_i, \tag{2}$$

where $A_k \equiv \varphi_k (1 - \tau_k)^{\sigma_k} r_k^{1-\sigma_k} h_k$ for $k \in \{u, i\}$. But in the case at hand, the two workers are employed by the same company in the same industry, using the same technology within the same work-team, so that $\varphi_u \approx \varphi_i$. The output, software, is instantly and almost costlessly transmissible to any point on earth, so that $\tau_u \approx \tau_i$. The firm's cost of borrowing does not depend on which worker it hires, so that $r_u \approx r_i$. Finally, the exogenous sorting of workers across countries afforded by randomized visa processing implies that, in expectation, $h_u \approx h_i$. Together, these imply that $A_u = A_i \equiv A$.

Without loss of generality, choose units such that $A = 1$. In this unusual setting, spatial equilibrium in labor demand then requires

$$w_u - w_i = \theta_u - \theta_i. \tag{3}$$

That is, under the strong assumptions above, differences in exchange-rate dollar earnings between these workers in different countries reflect differences in the effect of country-of-work on the productivity of otherwise equivalent workers.

The corresponding spatial equilibrium condition for labor supply requires the Indian workers to be indifferent at the margin to going to work in the United States. This will be

true if workers move to maximize real wages and

$$w_u - w_i = (p_u - p_i) + c, \tag{4}$$

where c is the price equivalent of any barriers to labor mobility, whether natural or policy-induced.

By (3) and (4), equilibrium in this labor market implies $\theta_u - \theta_i = (p_u - p_i) + c$. This is shown graphically in Figure 1. There, a migration quota $\bar{L}$ limits the movement of Indian labor to the U.S. The curve $D_{\text{all workers in US}}$ shows the firm's demand for any worker, while the curve $D_{\text{identical Indians in the US}}$ represents demand for the exact same workers as those in India, if they were in the U.S., described by $D_{\text{Indians in India}}$. The infinitely elastic supply curve $S_{\text{Indians in US}}$ reflects the fact that Indians will supply their labor in the U.S. if paid enough to offset the price differential. The value c is the price equivalent of the migration quota $\bar{L}$.

This suggests a simple way to test among the models of wage differences in subsection 2.1. The model predicts that in this unique setting, if wage differences do not arise from productivity differences, random selection of workers into the U.S. will yield wage gaps of zero on the left-hand side of (3). If wage differences arise from forces specific to place, the left-hand side of (3) will exceed zero. If those place-specific forces act by augmenting existing human capital rather than through the formation of new and portable human capital, wage gaps between workers in India who have worked in the U.S. and those who have not will exceed zero.

3 Research design: Randomized processing of high-skill U.S. work visas

The visa considered in this study is the H-1B visa to work in the United States, the largest single channel for admission of skilled foreign workers. This is a temporary or "non-

immigrant" work visa, created in 1990, that grants entry to foreign nationals in a "specialty occupation". Such an occupation is defined by U.S. Citizenship and Immigration Services (USCIS) as one that

> "requires theoretical and practical application of a body of specialized knowledge along with at least a bachelor's degree or its equivalent. For example, architecture, engineering, mathematics, physical sciences, social sciences, medicine and health, education, business specialties, accounting, law, theology, and the arts are specialty occupations."

With immaterial exceptions, an individual worker can only hold H-1B status for a period of six continuous years, after which the worker must remain outside the U.S. for a period of at least one year.

Starting with visas awarded in 2003, the U.S. Congress set a cap of 65,000 H-1B visas for new employment that could be awarded annually to a particular group of foreign nationals: those working in the for-profit sector who did not hold a master's degree from a U.S. institution. After a small number of these were set aside for Chileans and Singaporeans according to bilateral treaty obligations, the remainder was issued each year on a first-come, first-served basis following the opening of the application process on a set date in the spring. In 2003 through 2006, the visa quota was exhausted between two and six months after applications were opened.

A crucial change in policy occurred for the H-1B application process that began in the spring of 2007 (for visas awarded in Fiscal Year 2008). April 2, 2007 was the first business day on which USCIS considered H-1B applications. The selection process was to be identical to that of previous years, with an added stipulation. Anticipating the possibility of a large number of applications, USCIS stated in a press release before the opening of the application process,

> "If the final receipt date [the date the cap is exceeded] is the same as the first date that petitions may be filed, USCIS will randomly apply all of the numbers among the petitions filed on the final receipt date and the following day. This

means that, should the cap be reached on April 2, the first day filings can be received, USCIS will perform a random selection of petitions filed on April 2 and April 3. . . ”

In the event, USCIS was inundated with applications and the cap was reached on opening day, April 2. On April 12, USCIS conducted a random lottery among 123,480 applications received on April 2 and 3 to determine which would be processed.[2] No applications received on or after April 4 were entered in the lottery.

A similar lottery occurred in the spring of 2008 (for visas awarded in Fiscal Year 2009). This time, USCIS determined in advance that if the cap were met in the first five business days after it began accepting applications, the agency would conduct a random lottery among those visa applications received on any of those five days to determine which would be processed. The cap was indeed met within those five days, April 1-4 and 7, 2008. The lottery was conducted on April 14, and unselected applications were returned unprocessed. No such lottery was conducted in 2009, when the global economic slowdown reduced demand for H-1B visas.

4   Data

The data for this study are taken from the confidential, internal personnel records of a major international India-based information technology/software firm that I will call Anaamika Systems.[3] Anaamika Systems is a major multinational information technology services firm based in India. Its large staff performs technology-related services for other firms in India and all over the world, including designing and producing business software and hardware, managing databases, managing service-based business processes such as accounting, and providing technical and management consulting under contract.

Indian technology firms have recently experienced rapid and globally extensive growth.

---

[2]This and all subsequent discussion of the H-1B applications refers only to workers without a postgraduate degree from a U.S. educational institution. Workers with a U.S. postgraduate degree faced different rules.

[3]A pseudonym used at the firm's request. “Anaamika” is a Hindi word meaning “nameless”.

Indian software firms' revenues grew annually at 56% in the 1990s, accounting for 7% of Gross National Income growth in India (Arora and Athreye 2002), and capturing a quarter of the world market for customized software (Arora et al. 2001). This growth has been driven by international linkages: Exports account for two thirds of this revenue growth (Arora and Athreye 2002), and information technology became the largest sector of India's outward Foreign Direct Investment, whose overall stock has grown by two orders of magnitude since 1990 (Nayyar 2008).

Arora et al. (2001) estimate that 57% of the Indian software industry's service exports involve the temporary physical presence of Indian workers at an overseas worksite ("onsite work"), rather than products created in India and sent overseas ("offshore work"). The extent of offshore work is constrained by—among other reasons—limits to clear and nuanced long-distance communication with client teams (Arora and Asundi 1999), the difficulty of timely back-and-forth interaction when the team is split between different time zones, concerns about data security and intellectual property (Athreye 2005; Majumdar et al. 2007), and limits to the contractability of long and inherently complex software production processes (Banerjee and Duflo 2000).

This means that access to the labor market of the service-importing countries lies at the heart of the industry's success, and the H-1B visa has been the portal of entry for the largest share of these workers into the United States, the number one importer of Indian high-tech services. While the H-1B lasts three years and may be renewed for an additional three years, few workers sponsored by Anaamika Systems stay in the United States for even the initial three years. The typical onsite project lasts between six and fifteen months, after which the large majority of Anaamika employees return to India rather than being assigned to a different onsite project.

Anaamika Systems provided data on the full universe of Indian citizens that it sponsored in 2007 and 2008 for an H-1B visa subject to the general-category visa cap—that is, those

11

without a postgraduate degree earned in the United States. These records describe 16,182 visa applications, representing 14,799 unique individuals (1,383 of the applications in 2008 are repeat applications for losers of the 2007 lottery). In essentially all cases, lottery winners and losers were assigned ex ante to the same client work team, an assignment that was almost never affected by the lottery outcome. Winners on the team were sent to do on-site portions of the project (such as high-level design) while losers were assigned to off-site portions of the project (such as writing computer code).

For each individual, the data give 1) age, highest degree completed, and job title of the worker at the time the visa application was filed in 2007 or 2008, 2) the result of the visa lottery, and 3) employment status, country of location, job title, and earnings in June of 2009. All were located in India at the time the application was filed. No other traits of any of the workers are observed. No post-lottery information is available for 13.7% of these workers because they left Anaamika Systems before June of 2009.

The first step in analysis is to check that the information available in the personnel records is consistent with genuine randomization of visa processing. If USCIS conducted a truly random lottery to determine which visas should be processed, it would be unlikely for us to see substantial differences in pre-treatment characteristics between workers whose application was accepted for processing and workers whose application was rejected. Table 1 assesses these differences for all three worker traits on which the database contains pre-treatment information: age, highest degree completed, and level of job title. The null hypothesis of equality of means is not rejected in any case except one: In 2007 only, there is a statistically significant difference between level of job title between the treatment and control groups. This difference is extremely small, amounting to one twentieth of a point on a five-point seniority scale. This evidence is consistent with truly random selection of visa applications by USCIS. This accords with reports from Anaamika Systems that applications rejected in the lottery were returned unopened and unexamined.

5 Results

The analysis proceeds in three steps. First, it assesses the effect of the lottery on workers' location. Second, it uses this exogenous change in location to estimate the effect of location on earnings. Finally, it considers several empirical approaches to assessing threats to the internal validity of the results.

5.1 The effect of visa processing on location

Table 2 shows the effect of each of the two lotteries on workers' locations in June of 2009. The upper half of the table shows entrants to the 2008 lottery, observed one year and two months later. The lower half shows entrants to the 2007 lottery, observed two years and two months later. In both tables, the location "Other" represents any country in the world besides India and the United States. The most common of these are the United Kingdom, Switzerland, Australia, Canada, Germany, Singapore, the Netherlands, Japan, mainland China, Belgium, France, and Hong Kong (China).

Among workers whose 2009 location is known, workers in the group denied an H-1B visa are 22.2 percentage points less likely to be in the United States a year after the lottery, and 30.5 percentage points less likely to be in the United States two years after the lottery. A year after the 2008 lottery, 1.3% of those who lost the lottery are in the United States anyway, on L-1 visas.[4] Two years after the 2007 lottery, 14.3% of those denied an H-1B visa in 2007 are in the United States anyway. For about half of these (7.6 percentage points) this is because they re-entered the lottery in 2008 and won it the second time; the rest (6.6 percentage points) are on L-1 visas. There is no evidence of crossover, in the sense that no worker who lost the lottery somehow obtained an H-1B visa.

---

[4]The L-1 visa can substitute for the H-1B only in a small number of cases. It is designed for intracompany transfers of corporate managers and workers with highly specialized skills, but requires that the primary worksite of the individual be the sponsoring company, and cannot be used for contract work at a different worksite. This latter condition precludes its use for most onsite work projects by companies like Anaamika Systems. The L-1 allows entry for seven years and there is no annual limit on the number these visas.

One noteworthy feature of Table 2 is that the fraction of lottery winners present in the U.S. rises from 23.5% a year after the lottery to 44.8% two years thereafter. The rise is to be expected, since many workers granted a visa do not immediately begin on-site work at the client. The fact that 100% of lottery winners are not present in the U.S. two years after the lottery is also to be expected. The typical contract length is 6 to 15 months, so many of these have gone to the U.S. and returned by the time they are observed. The firm estimates that roughly half of the 2008 lottery winners who were in India in June of 2009 had previously worked abroad on the visa, and greater than 90% of the 2007 lottery winners who were in India in June of 2009 had previously worked abroad on the visa. The rest of winners located in India have not yet left for abroad—either because a suitable project for them has not been found, because the original onsite project has been delayed or (rarely) canceled, or because (very rarely) their petition for an H-1B visa was considered and denied.

Another feature of Table 2 is that a large fraction of workers denied entry to the United States with an H-1B visa do not remain in India. Workers denied an H-1B visa in 2007 are 30.5 percentage points less likely to be in the United States in 2009, but are only 12.5 percentage points more likely to be in India working for Anaamika Systems in 2009. The rest of the drop in the number of workers who would have been expected to have entered the United States if they had been treated with a visa is accounted for by an 8.9 percentage point rise in the number who migrate from India to work for Anaamika in another, non-U.S. country, and a 9.1 percentage point rise in the number who leave Anaamika Systems prior to June 2009. In other words, Table 2 suggests that among workers whose location is known with certainty, for every ten workers whose visa rejection caused them not to be in the United States two years after the lottery, three ended up in other countries outside India.

5.2   The effect of location on earnings

Randomized acceptance of visa applications fits a standard randomized encouragement experimental design. The simple difference in mean income between the group whose ap-

14

plication was accepted and the group whose application was rejected is the effect of visa processing on earnings. This number, divided by the change in the probability of working abroad, is the effect of working outside India on workers' average earnings—what Angrist et al. (1996) call the Local Average Treatment Effect. This is equivalent to the coefficient on a dummy variable signifying location abroad in a two-stage least squares regression with earnings as the dependent variable, where location is instrumented by acceptance of the worker's visa application for processing.

Table 3 assesses the local average effect of location outside India on individual workers' annual earnings as of June 2009. Data are presented in detail for exchange-rate dollars, and in summary form for PPP dollars. The table shows that processing of a worker's visa application causes that workers' 2009 earnings to rise by US$10,675 one year after the 2008 lottery, and by US$12,641 two years after the 2007 lottery. This effect is on the order of a 100% increase in earnings for a worker in India.

The effect of location outside India on 2009 earnings is US$54,949 after one year and US$58,203 after two years, roughly a sixfold increase in earnings, measured at market exchange rates. Measured in PPP dollars, the same treatment effects are PPP$38,958 after one year, and PPP$38,674 after two years, representing an increase in real wages by a factor of between 2 and 2.5. The accompanying standard errors show that all of these estimates are statistically precise well below the 5% level.

The table also juxtaposes these estimates with the simple differences in earnings for comparable software workers between the U.S. and India reported by Commander et al. (2008). The effect of location outside India on exchange-rate earnings is over three quarters of the simple difference in earnings. This suggests that a large majority of the unconditional difference in earnings is due to the place that those workers are located, and not to any other observable or unobservable difference between those workers.

15

5.3 The effect of overseas experience on earnings

How does foreign experience—working temporarily in the high-productivity environment of a developed country—affect the earnings of a worker in India? Table 4 explores the close similarity between the earnings, in India, of winners and losers in the lottery. One year after the lottery, winners make just 1.8% more, and two years after, 3.2% more.

In the personnel records as provided by the firm, workers in India who have previously been abroad cannot be distinguished between workers in India who have not previously been abroad. But as previously noted, the firm estimates that over half of the 2008 lottery winners who were in India in June of 2009 had previously worked on a contract abroad, and that the same fraction exceeds 90% for the 2007 lottery winners observed to be in India in June 2009.

This allows bounding of the effect of short-term experience abroad on the earnings of workers in India. Again following the randomized encouragement design, if greater than 50% of these workers have been abroad, the effect of prior experience abroad on earnings in India must be less than $360 per year (3.5% of earnings) for the 2008 lottery cohort, and less than $792 per year (6.3% of earnings) for the 2007 lottery cohort. This suggests that very little of the vast increase in earnings caused by working abroad accompanies the worker to India upon his or her return.

6 Interpretation: From wages to productivity

The preceding results suggest a very large effect of location on wages, but need not suggest similar effects on productivity. A workers' wage can differ from his or her marginal revenue product. In this case, however, there are many reasons to believe that the estimated effect of location on wages reflects a comparable or greater effect on productivity. This interpretation requires a series of assumptions. These are the assumptions underlying equation (3).

In this limited setting, many of those assumptions are *prima facie* reasonable. Large differences in technology are unlikely given that the winners and losers of the lottery are in the same sector, employed by the same firm, working on the same software project to serve the same client within the same workteam. Large differences in the tradability of the output are implausible; all of these workers are producing software. It is hard to see how a single firm's cost of capital could greatly vary depending on the location of its workers. And random sorting of workers by the visa lottery makes it difficult to posit substantial differences between workers' inherent traits, observable or not, in different locations.

Less clear is the approximate truth of other assumptions. The derivation of equation (3) relies on assuming a high degree of competition in three markets: the software output market, the labor market for Indian software professionals in the client firms' country (outside India), and the labor market for software professionals inside India. Each of these assumptions deserves scrutiny.

First, it is possible that the client firms paying for these workers hold some degree of power in the output market, violating the zero profit condition. A long literature has suggested that oligopoly in output markets exerts upward pressure on wages (e.g. Long and Link 1983; Heywood 1986; Belman and Heywood 1990). It is unclear, however, why upward pressure on w_u and w_i would systematically tend to raise or lower the gap $w_u - w_i$, given that each type of worker is producing an output for the same client firm. It is less clear how any such effect could produce a sixfold difference between w_u and w_i.

Second, there may be distortions in the labor market for Indian software professionals in the client firms' country, such as the United States. For example, firms employing H-1B workers in the U.S. are constrained in wage-setting by a legal requirement that they pay those workers the "prevailing wage" for their occupation and industry. It is reasonable to assume, however, that client firms do not pay those workers more than their marginal revenue product, for this would reduce profits. So while various distortions might shape

17

the divergence between w_u and the marginal revenue product of Indian workers in the client firms' country, any such distortions would only change the degree to which the wage difference $w_u - w_i$ understates the productivity difference $\theta_u - \theta_i$.

Third, there may be distortions in the Indian labor market for software professionals. It could be, for example, that Anaamika Systems holds substantial power in hiring. This could mean that differences in the clients' willingness to pay for workers in different locations are not comparable to differences in those workers' wages. The in-India wage w_i could then be substantially less than the marginal revenue product of that worker to Anaamika Systems, reducing the productivity gap relative to the wage gap.

But this phenomenon cannot be generating the preceding results. Indirect evidence of this is the fact that there is little sign that Anaamika Systems holds large market power in hiring. Five large firms in India, and numerous smaller firms, compete to hire workers for similar projects. Quitting one firm to work for another, or to found another, is frequent in the industry. More direct evidence is that the difference in client firms' willingness to pay for on-site workers and willingness to pay for workers in India is at least as large as the wage differences estimated here. The amount that clients pay for each worker covers not only that worker's wages but also overhead, transportation costs, and profits for the shareholders of Anaamika Systems. Arora and Asundi (1999) find that across this industry, foreign client firms are typically willing to pay \$60,000-\$65,000 more, per worker per year, to have work performed on-site rather than in India. This exceeds the wage differences in Table 3. For Anaamika Systems in particular this difference in clients' willingness to pay is roughly \$75,000—that is, roughly 1.3 times the wage difference.

Another distortion in the Indian labor market for software professionals might be that client firms abroad hold oligopsony power when contracting workers in India, power that they do not hold when contracting workers in their own countries. Suppose for example that subsidies to high-tech education in India produce an artificial glut of software workers there,

workers who are prevented from international movement by migration barriers. This would bid down the wage of software workers in India relative to their marginal revenue product for foreign clients when they serve those clients from within India.

But this phenomenon, also, could not generate the wage gaps in Table 3 in the absence of comparable productivity gaps. The reason is that essentially all of the work teams in Anaamika Systems that contain workers outside India serving clients on-site also contain workers inside India serving the same client off-site. Certain tasks, such as the writing of extensive portions of computer code, are almost never performed at the client site. If the client's marginal revenue product of having members of that team on-site were less than the wage difference required to shift a member of the team from India to on-site, the client's profits would rise by shifting a worker within the team from on-site to off-site. Thus the marginal revenue product to the client of having a worker outside India must exceed the wage differences in Table 3.

The assumptions underlying equation (3) are strong. But many of them are reasonable in the experimental setting at hand. Others are more likely to be violated in some measure, but if they are, this would only suggest that the effect of location on productivity exceeds the effect of location on wages in Table 3.

7 Internal validity

Two important challenges to establishing the internal validity of these results are attrition and the possibility of contamination. First, 13.7% of those who entered the 2007 and 2008 lotteries were no longer employed at Anaamika Systems in June of 2009, and their location and earnings cannot be established with certainty. Second, since Anaamika Systems exerts influence over both the treatment group and control group in the lottery, it is possible that lottery losers were treated differently by the firm than lottery winners, which would change the interpretation of the treatment effect.

Table 5 shows suggestive correlations between attrition and observable traits. It presents logit regressions in which the dependent variable equals 1 if the individual left the firm before post-test observation, and the independent variables are all observable pre-test traits. Older workers are slightly more likely to have left the firm one year after the 2008 lottery; but two years after the 2007 lottery, attrition by older workers cannot be statistically distinguished from attrition by younger workers.

The one clear result of the table is that workers with a master's degree are substantially less likely to leave the firm. This is consistent with anecdotal evidence that many workers leave the firm in order to pursue a master's degree in India. Because workers with a master's degree in a field related to information technology might be less willing to switch sectors, this is also consistent with anecdotal evidence that many workers leave the firm in order to join financial services firms in India.

This evidence is only suggestive, however. It leaves open the possibility that a numerous unobserved traits that might influence workers' location choice after leaving the firm are correlated with attrition. It also leaves open the possibility that some of those leaving the firm to pursue a master's degree do so abroad. Table 5 compels further investigation of the possibility of attrition bias to the estimates presented above.

7.1 Lee bounds on attrition bias

This study uses four methods to investigate the degree of attrition bias.[5] First, Lee (2009) presents a nonparametric method for bounding treatment effects in the presence of sample attrition, and Lee (2002) adapts the method to binary discrete outcomes. Here I apply the method to multiple discrete outcomes. The core assumption of Lee's method—that the effect of treatment on attrition is monotonic—means in the present case that denial of a U.S. visa only increases the probability of attrition and never decreases it.

Table 2 demonstrates that attrition is indeed much higher in the group whose visa applications are returned, and it is plausible that this effect is close to monotonic. While a few workers who are granted visas and go to the U.S. do quit while abroad, the firm estimates that roughly 90% of attrition happens in India. Commander et al. (2008) report that only about 5% of Indian information technology workers who leave their firms do so to take up employment abroad. This evidence suggests that the negative effect of winning the lottery on the probability of attrition dominates any positive effect.

Applying Lee's method to the case of multiple discrete outcomes, I first note that 15.8% of those who were granted a visa in the 2007 lottery quit the firm by 2009, while 24.9% of those denied a visa did so. Under monotonicity, visa denial only increased attrition, and thus caused the difference—9.1 percentage points—to quit the firm. One way to bound the influence of this attrition on the treatment effect is by allocating all of those who quit to

---

[5]The most straightforward method would be to conduct a representative survey of those who left the company prior to June 2009, but the firm does not maintain contact information for ex-employees. Another approach is to use one of a set of standard parametric methods to correct for attrition bias, but these have very strong data requirements. The method of Inverse Probability Weighting (e.g. Rubin and Little 1987) attaches greater weight to individuals whose location is less likely to be observed, but requires the assumption that after controlling for observed individual traits, the probability of attrition is independent of location. This requires extremely rich data on individual traits, absent from the present circumstance. Alternatively, the selection correction method of Hausman and Wise (1979) requires that at least one observed trait of each worker be correlated with attrition propensity but independent of location choice, as an exogenous source of variance in attrition propensity. The only individual-level observed traits of each worker in the present data base are age, education level, job title, and earnings; theory can suggest plausible reasons why each could be related to both attrition and location choice, and empirical evidence presented below confirms this possibility. There is no clear observed exogenous determinant of attrition.

each location outcome, one at a time.

Table 6 conducts this exercise for both lottery cohorts, assuming in each case that unobserved workers allocated to a particular location have the same average earnings as observed workers in that location. The estimated effects of location on earnings are not statistically significantly different from those in Table 3.

This analysis is not sensitive to the assumed earnings for unobserved workers. Those who quit while abroad can reasonably be assumed to be earning no less than their wage at Anaamika Systems, meaning that the above assumption about their earnings could only bias downward the estimated effect of location on earnings.

What if unobserved workers in India earn more than observed workers in India? Anecdotal reports from Anaamika Systems suggest that indeed most workers who leave the firm remain in India, either working either in the high-technology field, in banking and finance, or in other private sector positions, or attending graduate school. Table 8 reports typical annual base salaries, by years of experience, for Indian workers near the beginning of their careers with an undergraduate qualification in the Information Technology sectors and competing sectors. Even allowing for generous bonuses beyond these base salaries in the Banking and Finance sector or the Sales and Marketing sector, it appears unlikely that workers near the beginning of their careers who leave Anaamika Systems for other jobs in India earn more than their earnings at Anaamika.[6]

7.2   Other approaches to attrition bias

Table 6 suggests that, under the assumption of monotonicity in the effect of visa processing on attrition, the estimated effects of location on earnings do not materially depend on whether workers who quit the firm stay in India or go abroad. That said, further evidence

---

[6]If the Lee bounds exercise in Table 6 is repeated assuming that workers who leave Anaamika but remain in India earn 150% of their earnings at Anaamika, the estimated effect of location outside India on exchange-rate dollar earnings is $53,933 one year after the lottery, and $54,461 two years after the lottery. Neither of these is statistically significantly different from the estimates in Table 3

suggests that the large majority of workers who quit the firm remain in India. This can be seen in three independent ways.

The first alternative method to assess attrition bias relies on a simple assumption about worker behavior: Suppose that some workers desire to work outside of India in a rich country, and important reason that those workers join and remain at Anaamika Systems is because Anaamika offers opportunities for such abroad. If this is the case, it could be that workers who quit Anaamika because they lost the U.S. visa lottery did so in order to find other ways to work abroad. This would mean that the tendency to locate outside India in the unobserved workers denied a visa could be greater than the tendency for unobserved workers granted a visa to locate outside India, biasing estimates of the effect of visa processing on location and of location on earnings.

But this assumption has a testable implication: If an important reason that workers join and remain at Anaamika is the opportunity to work abroad, then a large exogenous change in the opportunity to work abroad should change the attrition rate. Part a of Table 7 allows us to test this condition for the 2007 lottery, comparing $Pr\{$Quit$|$Win, Non-US$\}$ to $Pr\{$Quit$|$Lose, Non-US$\}$. The first of these gives the probability that a worker leaves Anaamika after winning the 2007 visa lottery, conditional on not working for Anaamika in the U.S. in 2009. The second gives the same conditional probability for *losers* of the lottery. In other words, a worker's probability of quitting if his 2007 visa application to work in the United States was processed but was not there in 2009[7] is not substantially different from his probability of quitting if he never had the opportunity in the first place. The fact that the opportunity to work in the United States barely changes the attrition rate suggests that the opportunity to work abroad per se is not an important reason that workers remain at Anaamika.

To make the same argument more formal, let Q represent a worker who quits, W a worker

---

[7]This could be either because he went and returned by 2009 or had not yet gone, or in a few cases because his application was processed and later rejected.

who wins the lottery but is not in the U.S. in 2009, and L a worker who loses the lottery and is not in the U.S. in 2009. Suppose there are two types of workers, a type for whom the ability to work abroad affects job choice (τ_1) and a type for whom it does not (τ_0). For brevity, let $p_1 = Pr\{\tau_1\}$, $p^z = Pr\{Q \mid z\} \ \forall \ z \in \{W, L\}$, and $p_i^z = Pr\{Q \mid z \cap \tau_i\}$, $\forall \ i \in \{0, 1\}$, $z \in \{W, L\}$. We then have a simple identity, derived in the appendix, for the proportion of workers for whom the ability to work abroad is an important determinant of job choice:

$$p_1 = \frac{\left(p^W - p^L\right) - \left(p_0^W - p_0^L\right)}{\left(p_1^W - p_1^L\right) - \left(p_0^W - p_0^L\right)} \tag{5}$$

By the definition of τ_0 we can restrict $p_0^W - p_0^L = 0$, because the only difference between lottery winners and losers in expectation is that winners have more opportunities to work abroad, which does not affect job choices by workers of type τ_0. Furthermore, by (5), $p^W - p^L \approx 0$. Thus $p_1 \approx 0$.

This in turn suggests that the propensity to seek work abroad differs little among lottery winners who quit the firm and lottery losers who quit. And the first of these must be very low: It is unlikely that a lottery winner would quit the firm to pursue work outside India. If she wishes to work abroad and had an alternative job that gave her a higher probability of being able to work abroad than Anaamika, then it would have made little sense to join Anaamika; and if she wishes to work abroad and Anaamika gives her the highest probability of doing so, then it would make little sense to leave Anaamika. Thus the propensity for both lottery winners and losers who leave the firm to be abroad must be low. That is, the large majority of those who leave the firm by 2009 must still be in India in 2009.

More formally, let A represent a worker who locates abroad (outside India), so that the proportion of workers who go abroad conditional on having left the firm is $Pr\{A \mid Q\}$. For brevity let $a = Pr\{A \mid Q\}$, $a_i = Pr\{A \mid Q \cap \tau_i\} \ \forall \ i \in \{0, 1\}$, and $p_1^Q = Pr\{\tau_1 \mid Q\}$. This gives an identity, derived in the appendix, for the fraction of workers leaving the firm who

go abroad by 2009:

$$a = p_1^Q a_1 + \left(1 - p_1^Q\right) a_0 \tag{6}$$

But the above result $p_1 \approx 0 \implies p_1^Q \approx 0$. And theory suggests that $a_0 \approx 0$ as well: If work abroad is not an important determinant of a worker's job choice, it is unclear why the worker quitting Anaamika would go abroad. If the worker had an alternative choice abroad that offered better conditions than Anaamika unrelated to its location abroad, it is unclear why such a worker would have joined Anaamika to begin with. In these conditions, $a \approx 0$. Few workers who leave the firm leave India.

A second alternative approach to addressing attrition bias in the effect of visas on location choice relies on the assumption that any workers who left the firm in order to go abroad foresaw this fact at the time they left the firm. Under this assumption, the evidence on changes in attrition over time is incompatible with a large number of departures from India by those who left the firm.

The reason for this is that international demand for Indian high-tech workers changed substantially between the period April 2007 to April 2008 and the period April 2008 to June 2009 due to the global financial crisis, but attrition changed little.[8] If employees were leaving Anaamika Systems in order to work abroad, they would have had a much more difficult time finding employment abroad between April 2008 and June 2009 than they would have between April 2007 and April 2008. Coupled with the assumption that employees who left the firm to go abroad knew that this was their plan at the time they left, this means that markedly fewer employees would have left the firm during the 2008-2009 period than prior to it.

But this is not the case. Part *b* of Table 7 shows that 14.3% of those caused not to be in the United States by losing the 2008 lottery left the firm. If the attrition rate between April of 2007 and April of 2008 were much higher because lottery losers were leaving the firm to

---

[8]One sign of the change in international demand for Indian labor is the fact that, while in April 2007 and April 2008 the United States received over 120,000 employer-sponsored H-1B visa petitions in the first few days the visas were available, in 2009 applications opened in April but the quota of 65,000 was not filled until December.

take then-plentiful jobs outside of India, the attrition rate for the period April 2007 to June 2009 should be substantially higher than 28.2%—which would have been the rate over that whole period if it proceeded at the same rate per month as during April 2008 to June 2009. But the true attrition rate since 2007, among those caused not to be in the United States by losing the lottery, is 29.8%. This is barely higher than 28.2% and the difference is not statistically significant. In other words, international demand for these workers' labor fell substantially between 2007-8 and 2008-9, but the attrition rate of the lottery losers changed little between those periods. This is incompatible with a model in which large numbers of workers who leave the firm do so in order to work abroad, and suggests that the vast majority of workers who left the firm remained in India.

A third alternative method to address attrition bias takes advantage of information contained in the behavior of workers who lost the 2007 lottery and re-entered the lottery in 2008. Of the 3,068 workers who lost the 2007 lottery, 1,383 (45.1%) stayed with the firm another year and entered the 2008 lottery. Of these second-time entrants, 45.6% were denied a visa again, and some of these left the firm by 2009.

It is striking that their attrition rate, as part c of Table 7 shows, is not statistically distinguishable from that of their colleagues who had lost the lottery after having entered it just once. Likewise, the attrition rate among lottery winners in 2008 who had been forced to wait a year for the opportunity to go abroad is not distinguishable from the attrition rate among 2008 winners who had not been waiting. This suggests that very few of either the first-time lottery losers or the second-time lottery losers were leaving the firm in order to seek work outside India. A worker's expected probability of employment at Anaamika leading to work opportunities abroad would be lower if she had been denied the opportunity two years running that that of an equivalent worker denied just once, which would necessarily affect the propensity to leave Anaamika if leaving were for the purpose of seeking an employer that would give the opportunity to work abroad.

All three of these methods give suggestive evidence that the large majority of unobserved workers are in India. According to Table 6, this indicates that attrition makes the estimates in Table 3 tend to slightly overstate the effect of location on earnings. Any corrected estimate, however, would lie well within one standard error from the original estimate.

7.3   Contamination

Both the treatment group and control group in this study are employees of one company and therefore under the influence of a single decision maker. It is conceivable that the company treats lottery losers differently than lottery winners, for example by taking actions that compensate losers for the loss. To the extent that this occurs, it complicates the interpretation of the control group as a true counterfactual for the treatment group.

There is little evidence that lottery losers experience faster wage growth in India than their colleagues located in India who won the lottery. We have already seen in Table 4 that earnings among lottery losers in India are lower than earnings among lottery winners. To the extent that Anaamika raises lottery losers' wages to compensate them for the loss, this would little alter the estimates of treatment effect on wages. Even under the extreme assumption that winners' wages in India would grow at 0% per year, and that all wage growth for losers represents compensation for the loss, the estimated treatment effects would barely change.

The data also permit an assessment of the effect of U.S. visa processing on promotion from one job level to the next. Anaamika's personnel records classify workers in 53 different "roles" or job titles, each of which is given an integer "level" score between 1 and 5, in order of increasing seniority. A level of 1 is the least senior and includes the role of "junior test engineer". A level of 2 is the most common and includes the role of "Software Engineer". A level of 3 includes the role of "Programmer Analyst" and "Associate Consultant". Level 4 includes "Project Manager" and "Senior Associate", and level 5 includes "Business Manager" and "Principal Architect". Each person's role at the time he or she entered the visa lottery is

given, as the role of each worker in June 2009 who remains employed at Anaamika Systems.

Table 9 shows the effect of visa provision on the change in job level, on the five-point scale, between lottery entry and June 2009. One year after the 2008 lottery, there is no evidence to reject the null hypothesis of no effect of visas on promotion. Two years after the 2007 lottery, there appears to be a tiny but statistically significant *negative* effect of visa processing on promotion. The mechanism for this is unclear but might reflect any number of forces. It might be that workers who spend more time in India have more "face time" with the managers who promote them. It might also be that managers selectively give promotions to those who have not had the opportunity to work in the United States—evidence of contamination. At any rate the magnitude of this phenomenon is extremely small: 0.04 points on a five-point scale.

8 External Validity

There are obvious, strict limits to external validity for a study of applicants to one visa from one firm in one country. Further research is needed before the quantitative or even qualitative conclusions reached here could extend to other firms in the software industry, to other industries, or to other countries.

That said, the study does consider a globally-important industry and labor market: The country that issued the visa (the United States) is the top destination country for all international migrants as well as for all tertiary-educated migrants; the country of migrant origin (India) is the number one origin country for tertiary educated foreign-born residents of the United States;[9] the visa considered is the principal pathway of admission for temporary entry of skilled foreign workers to the United States;[10] and the firm providing personnel

---

[9]11% of all foreign-born people in the US with a bachelor's degree or higher are from India—more than from any other single country. Calculated using the US Bureau of the Census American Community Survey 2006 Public-Use Microdata Sample.

[10]According to the US Department of Homeland Security's Yearbook of Immigration Statistics 2008, in 2008 there were 409,619 admissions with H-1B, 382,776 admissions with L-1, and much smaller numbers of

records is one of a small group that have been the principal users of this visa.[11]

The study's use of two lotteries at two different moments in time is also attractive for external validity: King and Behrman (2009) highlight the importance of measuring treatment effects at different timepoints when theory gives little guidance on precisely when those effects might be realized. Another point in favor of external validity is that it was not clear to anyone in advance, including the U.S. government, that the 2007 lottery would occur. Visa applicants thus could not be sure that they were entering a lottery. This helps alleviate what Heckman (1992) calls "randomization bias".

9 Conclusion

This paper discusses the results of an unusual natural experiment in randomized processing of visas for work in the United States by highly skilled information technology professionals from India. It uses the effects of visa processing on location to estimate the effect of location on earnings.

This effect is very large, and it constitutes at least three quarters of the simple difference between U.S. earnings and India earnings exhibited by observably and unobservably different workers in the software industry. This suggests that only a small portion of this earnings gap can be attributed to worker traits of any kind, notably human capital, or to differences in the technology of production. The workers analyzed here are employees of the same firm using identical technology.

This evidence is difficult to reconcile with a model in which these software workers' marginal economic product does not depend heavily on location. The workers under con-

---

entries with O-1, O-2, E-1, and E-3. Indian nationals account for 38% of entries to the United States in 2008 with an H-1B visa—the principal mode of temporary entry for skilled workers (Calculated from the database underlying Docquier et al. (2009)).

[11]In FY 2007, eight of the top ten firms by number of H-1B visas sponsored were India-based high-technology firms. Kirkegaard (2007, 50) reports that the top 25 employers sponsoring H-1B visas in 2006—of which 13 where Indian high-tech firms—accounted for about half of the visas issued to the top 200 employers.

sideration are producing an almost perfectly tradable good: software. This means that differences in the cost of living between the U.S. and India can safely be ignored when calculating the economic productivity of these workers; U.S. clients who purchase software from them care only about the number of dollars a given software output cost, not about the cost of living of those producing it. There is evidence, reviewed above, that the labor market for hiring these workers is roughly competitive both in the U.S. and India. All of this suggests that large differences in workers' earnings, in exchange-rate dollars, imply comparably large differences in those workers' productivity.

If indeed these conditions hold, and earnings differences caused by location broadly reflect productivity differences caused by location (3), this finding has implications for the urban and labor literature on spatial differences in productivity. It suggests that the "level effect" of location on productivity (Glaeser and Maré 2001) can be extremely large between countries, and dominates the "growth effect" of location on portable human capital. The finding also has implications for the growth and development literature: Very little of the large income difference between a developed country and a developing country considered here can be attributed to any difference in individual human capital or technology. Working directly with high-productivity workers in a center of productivity appears to explain the majority of the gap, as has been observed at the domestic level (Ciccone and Hall 1996) and theorized at the international level (Kremer 1993).

Finally, it provides one data point relevant to the international literature on the price equivalent of quotas (Edwards 1999; Anderson and van Wincoop 2004): the PPP-dollar estimates of the effect of location on earnings reported at the bottom of Table 3 constitute a rare estimate of the price equivalent of all migration barriers facing the workers under study: the visa quota, combined with natural barriers. In the present case, most of this figure is likely the effect of the quota, as these workers are not subject to many of the natural barriers that normally impede labor mobility in the absence of policy barriers—such as language (all have excellent English), job search (jobs are lined up for them), risk aversion (unemployment

risk is essentially nil), and travel costs (paid by their employer). This in turn suggests large efficiency losses occasioned by the quota.

Appendix

Derivation of equation (5)

The effect of winning the lottery on attrition, conditional on not working for Anaamika in the United States as of 2009, is

$$
Pr\{Q|W\} = Pr\{Q \cap \tau_1|W\} + Pr\{Q \cap \tau_0|W\} - \Big(Pr\{Q \cap \tau_1|L\} + Pr\{Q \cap \tau_0|L\}\Big)
$$

$$
= Pr\{\tau_1\} Pr\{Q|W \cap \tau_1\} + Pr\{\tau_0\} Pr\{Q|W \cap \tau_0\}
$$

$$
- \Big(Pr\{\tau_1\} Pr\{Q|L \cap \tau_1\} + Pr\{\tau_0\} Pr\{L|W \cap \tau_0\}\Big)
$$

$$
= Pr\{\tau_1\} \Big(Pr\{Q|W \cap \tau_1\} - Pr\{Q|L \cap \tau_1\}\Big)
$$

$$
+ \Big(1 - Pr\{\tau_1\}\Big)\Big(Pr\{Q|W \cap \tau_0\} - Pr\{L|W \cap \tau_0\}\Big)
$$

where $Pr\{\tau_1 \mid W\} = Pr\{\tau_1\}$ because treatment is randomized. Letting $p_1 = Pr\{\tau_1\}$, $p^z = Pr\{Q \mid z\} \ \forall \ z \in \{W, L\}$, and $p_i^z = Pr\{Q \mid z \cap \tau_i\} \ \forall \ i \in \{0, 1\}, \ z \in \{W, L\}$ allows rewriting the above as:

$$
p^W - p^L = p_1 \left(p_1^W - p_1^L\right) - (1 - p_1)\left(p_0^W - p_0^L\right)
$$

Solving for p_1 gives

$$
p_1 = \frac{\left(p^W - p^L\right) - \left(p_0^W - p_0^L\right)}{\left(p_1^W - p_1^L\right) - \left(p_0^W - p_0^L\right)}.
$$

Derivation of equation (6)

The probability of being located abroad conditional on leaving the firm is

$$
Pr\{A|Q\} = Pr\{A \cap \tau_1|Q\} + Pr\{A \cap \tau_0|Q\}
$$

$$
= Pr\{\tau_1|Q\} Pr\{A|Q \cap \tau_1\} + Pr\{\tau_0|Q\} Pr\{A|Q \cap \tau_0\}
$$

Letting $p_i^Q = Pr\{\tau_i \mid Q\} \ \forall \ i \in \{0, 1\}$, $a = Pr\{A \mid Q\}$, and $a_i = Pr\{A \mid Q \cap \tau_i\} \ \forall \ i \in \{0, 1\}$, and noting that $p_0^Q = \left(1 - p_1^Q\right)$, we have

$$
a = p_1^Q a_1 + \left(1 - p_1^Q\right) a_0.
$$

References

Anderson, James E. and Eric van Wincoop, "Trade Costs," *Journal of Economic Literature*, September 2004, *42* (3), 691–751.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, "Identification of causal effects using instrumental variables," *Journal of the American Statistical Association*, 1996, *91* (434), 444–455.

Antràs, Pol and Elhanan Helpman, "Global Sourcing," *Journal of Political Economy*, June 2004, *112* (3), 552–580.

Arora, Ashish and Jai Asundi, "Quality Certification and the Economics of Contract Software Development A Study of the Indian Software Industry," NBER Working Papers 7260, National Bureau of Economic Research, Inc July 1999.

_ **and Suma S. Athreye**, "The software industry and India's economic development," *Information Economics and Policy*, 2002, *14* (2), 253–273.

_ **, V. S. Arunachalam, Jai Asundi, and Ronald Fernandes**, "The Indian software services industry," *Research Policy*, 2001, *30* (8), 1267–1287.

Ashenfelter, O., "Comparing Real Wages," NBER Working Paper 18006, National Bureau of Economic Research, Cambridge, MA 2012.

Athreye, Suma S., "The Indian software industry and its evolving service capability," *Industrial and Corporate Change*, 2005, *14* (3), 393–418.

Banerjee, Abhijit V. and Esther Duflo, "Reputation Effects And The Limits Of Contracting: A Study Of The Indian Software Industry," *The Quarterly Journal of Economics*, August 2000, *115* (3), 989–1017.

Baumol, W.J. and W.G. Bowen, *Performing Arts: The Economic Dilemma; A study of problems common to theater, opera, music, and dance*, New York: Twentieth Century Fund, 1966.

Belman, Dale and John S Heywood, "The Concentration-Earnings Hypothesis: Reconciling Individual and Industry Data in U.S. Studies," *Oxford Bulletin of Economics and Statistics*, August 1990, *52* (3), 293–302.

Borjas, George J., "Self-selection and the earnings of immigrants," *The American Economic Review*, 1987, *77* (4), 531–553.

Ciccone, Antonio and Robert E. Hall, "Productivity and the Density of Economic Activity," *American Economic Review*, March 1996, *86* (1), 54–70.

Clark, Gregory, "Why isn't the whole world developed? Lessons from the cotton mills," *Journal of Economic History*, 1987, *47* (1), 141–173.

Clemens, Michael, Claudio Montenegro, and Lant Pritchett, "The Place Premium: Wage Differences for Identical Workers across the US Border," Working Paper Series 148, Center for Global Development 2008.

Clingingsmith, David, Asim I. Khwaja, and Michael Kremer, "Estimating the Impact of the Hajj: Religion and Tolerance in Islam's Global Gathering," *Quarterly Journal of Economics*, 2009, *124* (3), 1133–1170.

Commander, Simon, Rupa Chanda, Mari Kangasniemi, and L. Alan Winters, "The Consequences of Globalisation: India's Software Industry and Cross-border Labour Mobility," *The World Economy*, 2008, *31* (2), 187–211.

Docquier, Frédéric, B. Lindsey Lowell, and Abdeslam Marfouk, "A Gendered Assesment of Highly Skilled Emigration," *Population and Development Review*, 2009, *35* (2), 297–321.

Easterly, William, "Channels from Globalization to Inequality: Productivity World versus Factor World," in "Brookings Trade Forum" JSTOR 2004, pp. 39–81.

___ **and Ross Levine**, "What have we learned from a decade of empirical research on growth? It's Not Factor Accumulation: Stylized Facts and Growth Models," *The World Bank Economic Review*, 2001, *15* (2), 177.

Edwards, Sebastian, "How Effective Are Capital Controls?," *Journal of Economic Perspectives*, Fall 1999, *13* (4), 65–84.

Ellison, G., E.L. Glaeser, and W.R. Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 2010, *100* (3), 1195–1213.

Freeman, R.B., "People Flows in Globalization," *Journal of Economic Perspectives*, 2006, *20* (2), 145–170.

Gibson, John, David McKenzie, and Steven Stillman, "The Importance of Selectivity and Duration-Dependent Heterogeneity When Estimating the Impact of Emigration on Incomes and Poverty in Sending Areas: Evidence from the Samoan Quota Migration Lottery," Paper presented at the Second International Conference on Migration and Development, Washington DC, Sept. 10-11, World Bank 2009.

Glaeser, Edward L. and David C. Maré, "Cities and Skills," *Journal of Labor Economics*, April 2001, *19* (2), 316–42.

Grogger, J. and G.H. Hanson, "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, 2011, *95* (1), 42–57.

Grossman, Gene M. and Esteban Rossi-Hansberg, "Trading tasks: a simple theory of offshoring," *American Economic Review*, 2008, *98* (5), 1978–1997.

Hall, Robert E. and Charles I. Jones, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *The Quarterly Journal of Economics*, February 1999, *114* (1), 83–116.

Hausman, Jerry A. and David A. Wise, "Attrition bias in experimental and panel data: the Gary income maintenance experiment," *Econometrica: Journal of the Econometric Society*, 1979, *47* (2), 455–473.

Heckman, James J., "Randomization and social policy evaluation," in Charles F. Manski and Irwin Garfinkel, eds., *Evaluating Welfare and Training Programs*, Cambridge, MA: Harvard University Press, 1992.

Heywood, John S., "Labor Quality and the Concentration-Earnings Hypothesis," *The Review of Economics and Statistics*, May 1986, *68* (2), 342–46.

Hunt, J., "Which Immigrants Are Most Innovative and Entrepreneurial? Distinctions by Entry Visa," *Journal of Labor Economics*, 2011, *29* (3), 417–457.

Kelly Services, *Indian Salary Handbook 2008/09: A practitioners guide to salaries across industries*, Haryana: Kelly Services India Corporate Office, 2009.

Kerr, W.R., "Breakthrough inventions and migrating clusters of innovation," *Journal of Urban Economics*, 2010, *67* (1), 46–60.

_ **and W.F. Lincoln**, "The supply side of innovation: H-1B visa reforms and US ethnic invention," *Journal of Labor Economics*, 2010, (3), 473–508.

King, Elizabeth M. and Jere R. Behrman, "Timing and duration of exposure in evaluations of social programs," *The World Bank Research Observer*, 2009, *24* (1), 55–82.

Kirkegaard, Jacob Funk, *The accelerating decline in America's high-skilled workforce: implications for immigration policy*, Peterson Institute, 2007.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Kremer, Michael, "The O-ring theory of economic development," *The Quarterly Journal of Economics*, 1993, *108* (3), 551–575.

Krugman, Paul, "Increasing Returns and Economic Geography," *Journal of Political Economy*, June 1991, *99* (3), 483–99.

Lee, David S., "Trimming for Bounds on Treatment Effects with Missing Outcomes," Center for Labor Economics Working Paper 51, University of California Berkeley 2002.

_ , "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Long, James E. and A. N. Link, "The impact of market structure on wages, fringe benefits, and turnover," *Industrial and Labor Relations Review*, January 1983, *36* (2), 239–250.

Lucas, Robert E., "Why Doesn't Capital Flow from Rich to Poor Countries?," *American Economic Review*, May 1990, *80* (2), 92–96.

Majumdar, Sumit K., Ashok Nag, and Kenneth L. Simons, "Bodyshopping versus Offshoring: Determinants of Outsourcing Behavior by Indian Software and Information Technology Firms," Dept. of Economics Working Paper, Rensselaer Polytechnic Institute 2007.

McKenzie, D., S. Stillman, and J. Gibson, "How Important is Selection? Experimental versus Non-Experimental Measures of the Income Gains from Migration," *Journal of the European Economic Association*, 2010, *8* (4), 913–945.

Nayyar, Deepak, "The internationalization of firms from India: investment, mergers and acquisitions," *Oxford Development Studies*, 2008, *36* (1), 111–131.

Ortega, Francesc and Giovanni Peri, "The Causes and Effects of International Labor Mobility: Evidence from OECD Countries 1980-2005," MPRA Paper 19183, University Library of Munich, Germany March 2009.

Paserman, M.D., "Do High-Skill Immigrants Raise Productivity? Evidence From Israeli Manufacturing Firms, 1990-1999," Working Paper, Department of Economics, Boston University 2011.

Rauch, James E., "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities," *Journal of Urban Economics*, November 1993, *34* (3), 380–400.

Roback, Jennifer, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, December 1982, *90* (6), 1257–1278.

Rosen, Sherwin, "Wage-based indexes of urban quality of life," in Peter Miezkowski and Mahlon R. Straszheim, eds., *Current Issues in Urban Economics*, Baltimore: Johns Hopkins University Press, 1979.

Rosenthal, Stuart S. and William C. Strange, "Evidence on the nature and sources of agglomeration economies," in J. Vernon Henderson and Jacques-Franois Thisse, eds., *Handbook of Regional and Urban Economics*, Amsterdam: North-Holland, 2004.

Rubin, Donald B. and Roderick J. A. Little, *Statistical analysis with missing data*, New York: John Wiley and Sons, 1987.

U.S. Bureau of Labor Statistics, "May 2009 National Occupational Employment and Wage Estimates," May 2010. http://www.bls.gov/oes/current/oes_nat.htm.

Young, Alwyn, "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *The Quarterly Journal of Economics*, August 1995, *110* (3), 641–80.

Figure 1: **Spatial equilibrium conditions for labor demand and labor supply**
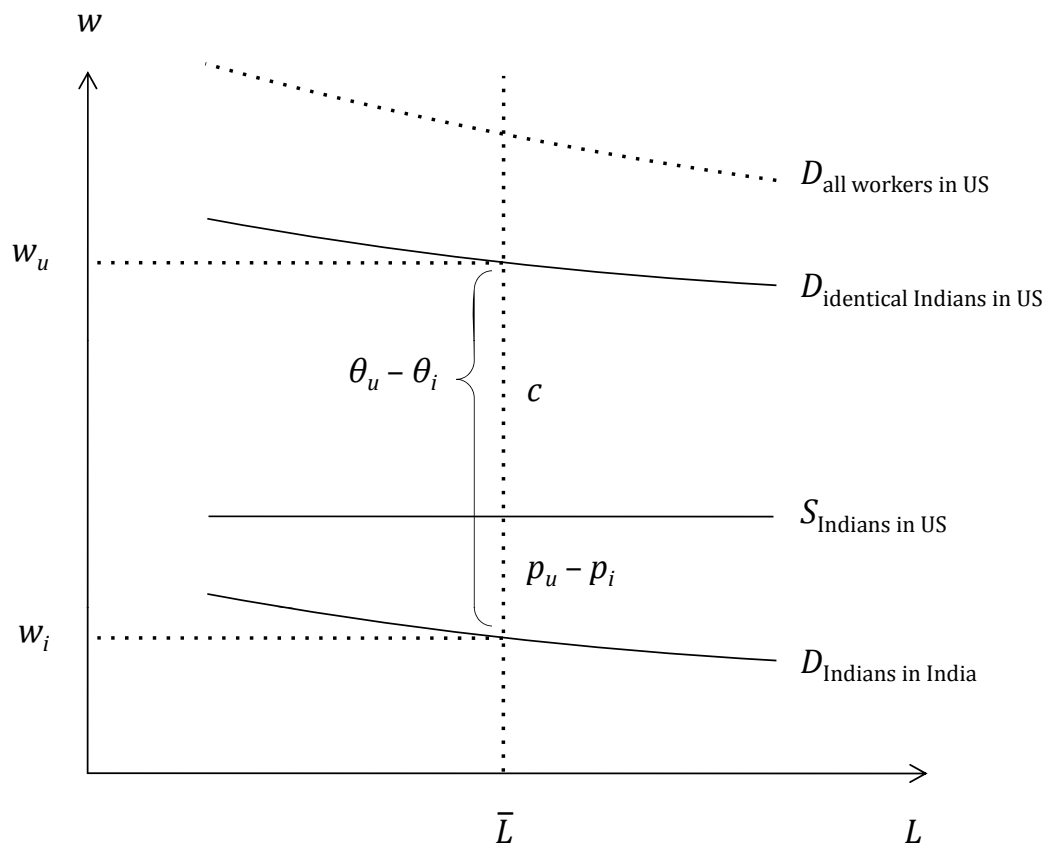
Table 1: **Balance in pre-lottery observables**

	Accepted		Rejected			
	Mean	N	Mean	N	Difference	S.E.
2008 lottery						
Age	25.06	4,750	25.03	4,183	0.0254	0.0544
Master's degree	0.162	4,750	0.158	4,183	0.004	0.007
Position level (1-5)	2.30	4,750	2.29	4,183	−0.011	0.012
2007 lottery						
Age	25.14	4,181	25.19	3,068	−0.0538	0.0667
Master's degree	0.203	4,181	0.195	3,068	0.008	0.009
Position level (1-5)	2.32	4,181	2.34	3,068	−0.021	0.014

"Master's degree" is 1 if the worker has a master's degree, zero otherwise. Position level is a 1-5 score given to the worker's job title by the firm, with a higher number indicating greater seniority.

Table 2: **The effect of visa processing on location**

Location 2009	Accepted		Rejected				If removed from U.S.
	Percent	N	Percent	N	Diff.	S.E.	
2008 lottery							
USA	23.5%	1,118	1.3%	55	−22.2%	(0.7%)	
India	62.5%	2,967	77.7%	3,250	15.2%	(1.0%)	68.5%
Other	6.6%	314	10.4%	436	3.8%	(0.6%)	17.1%
Quit	7.4%	351	10.6%	442	3.2%	(0.6%)	14.4%
Total	100.0%	4,750	100.0%	4,183			100.0%
2007 lottery							
USA	44.8%	1,874	14.3%	438	−30.5%	(1.1%)	
India	35.1%	1,469	47.7%	1,462	12.5%	(1.2%)	41.0%
Other	4.3%	179	13.2%	405	8.9%	(0.6%)	29.2%
Quit	15.8%	659	24.9%	763	9.1%	(0.9%)	29.8%
Total	100.0%	4,181	100.0%	3,068			100.0%

The location "Other" comprises, in descending order of frequency: United Kingdom, Switzerland, Australia, Canada, Germany, Singapore, Netherlands, Japan, China, Belgium, France, Hong Kong (China), United Arab Emirates, Malaysia, Norway, Finland, Italy, Mexico, Oman, Croatia, Ireland, and Sweden.

Table 3: **The effect of location on earnings, 2009 US$/year**

	One year after 2008 lottery			Two years after 2007 lottery		
	Mean earnings	S.E.	N	Mean earnings	S.E.	N
Accepted	27,394	(377)	4,399	43,834	(456)	3,522
Rejected	16,718	(287)	3,741	31,194	(533)	2,305
In India	10,346	(71)	6,217	12,726	(123)	2,931
Accepted, in India	10,440	(104)	2,967	12,924	(175)	1,469
Rejected, in India	10,260	(96)	3,250	12,528	(173)	1,462
Outside India	61,740	(131)	1,923	65,257	(137)	2,896
Accepted, outside India	62,520	(142)	1,432	65,952	(154)	2,053
Rejected, outside India	59,466	(281)	491	63,565	(276)	843
% of accepted outside India			32.6%			58.3%
% of rejected outside India			13.1%			36.6%
Change			19.4%			21.7%
Exchange-rate dollars						
Effect, visa processing	10,675	(474)		12,641	(701)	
Effect, location outside India	54,949	(2,439)		58,203	(3,230)	
$(w_{India} + \text{loc. effect}) / w_{India}$	6.31	(0.058)		5.57	(0.067)	
Location effect, % observed gap ($71,206)	77%			82%		
PPP dollars						
Effect, visa processing	7,569	(328)		8,399	(443)	
Effect, location outside India	38,958	(1,686)		38,674	(2,042)	
Location effect / India wage	2.42	(0.004)		2.14	(0.004)	
Location effect, % observed gap (PPP$62,240)	63%			62%		

The PPP-to-market exchange rate ratio used is 0.376, for a basket of 122 goods typically consumed by urban white-collar professionals in New York vs. Delhi, reported by UBS/Warburg, *Prices and Earnings 2009*, p. 14. It is similar to the factor of 0.3328 reported for the whole of the U.S. versus the whole of India reported by the World Bank in *World Development Indicators 2008* (for 2006).

Table 4: **The effect of experience abroad on earnings in India**

	Mean wage if in India 2009	S.E.	N	Assumption
Accepted 2008	10,440	(104)	2,967	
Rejected 2008	10,260	(96)	3,250	
Effect of 2008 acceptance	180	(141)		
Predicted effect of experience abroad	<360	(282)		>50% of winners have been abroad
Accepted 2007	12,924	(175)	1,469	
Rejected 2007	12,528	(173)	1,462	
Effect of 2007 acceptance	396	(246)		
Predicted effect of experience abroad	<792	(492)		>50% of winners have been abroad

Table 5: **Comparison of attrition by observable pre-lottery traits**

Logit regression *Dependent variable:* *Quit by June 2009*	Elapsed time	
	One year (2008 lottery)	Two years (2009 lottery)
Age	0.056	0.013
	(0.026)	(0.019)
Master's degree	−0.423	−0.223
	(0.126)	(0.084)
Job level (1-5)	−0.146	−0.094
	(0.106)	(0.083)
Constant	−3.335	−1.479
	(0.473)	(0.349)
No. of observations	8,933	7,249
No. with firm in 2009	8,140	5,827
No. left firm by 2009	793	1,422
Pseudo R^2	0.003	0.002

All independent variables measured at the time the lottery was held. Robust standard errors in parentheses. "Master's degree" is 1 if the worker has a master's degree, zero otherwise. Position level is a 1-5 score given to the worker's job title by the firm, with a higher number indicating greater seniority.

Table 6: **Lee nonparametric bounds on attrition bias to earnings effects**

	One year after 2008 lottery				Two years after 2007 lottery			
	Base	All to India	All to USA	All to Other	Base	All to India	All to USA	All to Other
Exchange-rate dollars								
Effect, visa processing	10,675	10,897	9,114	9,221	12,641	14,659	8,961	9,335
Effect, location outside India	54,949	54,818	55,412	56,062	58,203	57,100	60,304	62,815
Location effect / India wage	6.3	6.3	6.4	6.4	5.6	5.5	5.7	5.9
PPP dollars								
Effect, visa processing	7,569	7,714	6,515	6,622	8,399	9,595	6,146	6,519
Effect, location outside India	38,958	38,805	39,610	40,260	38,674	37,376	41,358	43,869
Location effect / India wage	2.4	2.4	2.4	2.5	2.1	2.1	2.2	2.3

Table 7: **Patterns of attrition informative on the location of unobserved workers**

Group	Period	Attrition	S.E.	Conditional on	Basis
a) *Attrition by treatment conditional on no U.S. placement*					
Won lottery 2007	2007-9	28.6%	(0.9%)	No U.S. placement	Observed
Lost lottery 2007	2007-9	29.0%	(0.8%)	No U.S. placement	Observed
	Diff.	0.4%	(1.2%)		
b) *Attrition and the global crisis*					
Lost lottery 2008	2008-9	14.3%	(2.7%)	Removed from U.S.	Observed
Lost lottery 2007	*2007-9*	*28.2%*		*Removed from U.S.*	*Predicted**
Lost lottery 2007	2007-9	29.8%	(3.1%)	Removed from U.S.	Observed
c) *Repeat lottery entrants*					
Won 2008, 1st try	2008-9	7.3%	(0.4%)	—	Observed
Won 2008, 2nd try	2008-9	8.0%	(1.0%)	—	Observed
	Diff.	−0.7%	(1.1%)		
Lost 2008, 1st try	2008-9	10.6%	(0.5%)	—	Observed
Lost 2008, 2nd try	2008-9	10.5%	(1.2%)	—	Observed
	Diff.	0.1%	(1.3%)		

*Assumes that the annualized attrition rate from April 2008 to June 2009 applies for the period April 2007 to June 2009.

Table 8: **Base annual salaries in IT and competing sectors, 2008 US$**

Years of experience	Information Technology		Banking and Finance		Sales and Marketing	
	Min.	Max.	Min.	Max.	Min.	Max.
0-1			2,206	3,310		
1-2			2,620	5,344		
1-3	3,723	10,067	2,413	5,378	1,379	5,516
2-3			3,172	7,952	2,206	5,240
2-4	5,516	13,100	2,758	7,033	2,942	10,113
2-5	5,516	17,927	2,206	5,516		
3-5	6,501	15,681	3,723	11,721		
4-5			8,274	16,548		
4-6			7,584	19,995	9,653	22,064
5-7	11,032	28,315			9,653	26,890
5-9			8,274	33,096	9,653	34,475

Source: Kelly Services (2009). Rupee figures have been converted at 43.51 Rs/US$, the average exchange rate for 2008 reported in Economist Intelligence Unit, *Country Report: India*, October 2009, p. 17.

Table 9: **The effect of visa processing on promotion within the firm**

	Accepted		Rejected			
	Mean	N	Mean	N	Diff.	S.E.
One year after 2008 lottery	0.133	4,398	0.127	3,738	0.006	(0.007)
Two years after 2007 lottery	0.435	3,522	0.473	2,305	-0.038	(0.013)

Promotion is defined as a positive change in position level between the time of the lottery and June 2009. Position level is a 1-5 score given to the worker's job title by the firm, with a higher number indicating greater seniority.