Firm-Induced Migration Paths and Strategic Human-Capital Outcomes

Prithwiraj (Raj) Choudhury, Tarun Khanna and Victoria Sevcenko[1]

Abstract

Firm-induced migration typically entails firms relocating workers to fill value-creating positions at destination locations. But such relocated workers are often exposed to external employment opportunities at their destinations, possibly triggering turnover. We conceptualize the *firm-induced migration path*, consisting of the relocated workers' place of origin and destination, as relevant in determining worker performance and turnover post-relocation. Using a unique dataset from a large Indian technology firm that hires talent from both large cities and smaller towns, we document robust econometric patterns by exploiting the firm's randomized assignment of workers to production centers across the country. These production centers are located in the largest technology cluster in India (Bangalore), smaller technology clusters, and non-cluster locations. We find that the firm-induced migration path shapes both worker performance and turnover. Compared to workers from large cities, workers from smaller towns achieve higher performance when relocated to Bangalore than to other production centers, but are also more likely to join competing firms. Fine-grained data on employment and human-capital-augmentation opportunities at workers' destination locations, and on socioeconomic conditions in workers' places of origin, help us rule in an abductive explanation: across firm-induced migration paths, differences in external labor-market opportunities between workers' places of origin and their destinations, as well as intrafirm skill-development opportunities at the destination, are related to heterogeneous human-capital outcomes.

[1] Choudhury Prithwiraj: Harvard Business School, (e-mail: pchoudhury@hbs.edu). Khanna Tarun, Harvard Business School. Sevcenko Victoria, INSEAD. The authors would like to thank Lauren Cohen, Lawrence Katz, Bill Kerr, Dylan Minor, Pian Shu, Claudia Steinwender; seminar participants at Duke, Harvard University, the International Economic Association (IEA) meetings in Beijing, INSEAD, London School of Economics, the Seventh International Conference on Migration and Development, University of Oxford, NYU Stern School of Business, Washington University St. Louis, Wharton, and the World Bank DECTI seminar for their comments on a previous draft.

1. Introduction

Firms play an important role in moving workers between geographic regions, and firm-induced migration has been shown to positively affect both workers' and firms' productivity (Wang 2015; Choudhury 2016; Hernandez and Kulchina 2020). Prior work suggests that accessing high-quality talent in distant locations, and matching relocated workers to intrafirm opportunities and resources, creates value for firms (Rosenkopf and Almeida 2003; Kerr et al. 2015). At the same time, turnover is often attributable to workers learning about the quality of their labor-market matches via discovery of new outside options (Jovanovic 1979; Lee and Mitchell 1994). As firms search for talent more widely, and relocate workers to regions with diverse opportunities, more workers may discover superior external matches and depart the focal firm. This raises the question explored in this paper: *What patterns of firm-induced migration increase, or decrease, worker performance and turnover to competitors?*

Two related literatures provide starting points for answering this question. The literature in economics and strategy suggests that the characteristics of the *destination* locations play an important role in shaping worker performance and turnover. Firms that relocate workers to clusters (regions where numerous firms compete for human capital) often face a trade-off: workers in clusters are incentivized to expend effort to develop value-creating skills, but are also prone to turnover to competing firms (Ciccone and Hall 1996; Kerr and Robert-Nicoud 2020). Relatedly, workers relocated to company headquarters often enjoy disproportionate access to intrafirm resources and opportunities, which can positively affect their performance (Karim and Williams 2012; Choudhury 2017). Meanwhile, however, little is known about how relocated workers' *origins*[2] impact strategic human-capital outcomes, such as on-the-job performance and turnover.

A large and vibrant literature in sociology and economics explores the role of migrants' *place of origin* in both their migration choices and their subsequent outcomes. These studies do not focus on firm-induced migration, and do not shed light on on-the-job performance and turnover (see Kerr et al. 2015). However, they do articulate the concept of a *migration path*, and document that individuals typically move, via intermediate steps, from rural locales to large cities that offer vaster arrays of employment and human-

[2] Traditionally, firms have hired talented workers from large cities (Jensen 2012) and moved them to clusters (Saxenian 2000; Bresnahan and Gambardella 2004). Increasingly, however, firms are aggressively hiring workers from smaller towns (Singh 2018; Zhao et al 2018) and relocating them to non-cluster locations (Choudhury 2020). These shifts in firms' hiring patterns suggest that labor-market opportunities for relocated workers are likely to continue to change in diverse ways.

capital-augmentation opportunities (Ravenstein 1885; Conway 1980; Paul 2011).[3] Along the migration path, the characteristics of workers' origin and destination locations affect outcomes such as wages (Borjas 1994, building on Roy 1951). From the perspective of firms, however, we know little about how moving workers along diverse firm-induced migration paths affects on-the-job performance and turnover.

This paper combines insights from both literatures. We extend the concept of the migration path to firm-induced migration, and posit that human-capital outcomes relevant to the firm—on-the-job performance and turnover of relocated workers—are likely to depend on the *joint effect* of workers' origin and destination locations, that is, on their *firm-induced migration path*. Specifically, we examine whether workers exhibit heterogeneous performance and turnover outcomes when relocated to production centers located in a large cluster, in smaller clusters, and in non-cluster locations. We also examine whether such outcomes vary depending on relocated workers' geographic origins—that is, whether a worker moved from a smaller town or from a large city. Finally, we examine whether these outcomes vary with the availability of skill-development resources at the destination production center. Specifically, workers moved to the firm headquarters might enjoy greater opportunities to work on research and development (R&D) projects that can enhance knowledge and skills that firms value (Collis et al. 2007).

Empirically, it is often difficult to observe differences in performance and turnover that are causally related to patterns of firm-induced migration, because assignment to production centers is correlated with observable and unobservable characteristics of the individual and the production center. We address such endogeneity by leveraging a natural experiment at a large Indian technology firm (hereafter INDTECH). This firm employed over 120,000 people worldwide in 2007 (the first year of our study period) and hires widely across cities and towns in India. Crucially for our empirical design, INDTECH has a unique policy of randomly assigning workers to its ten production centers in India to ensure that its end customers—mostly U.S.-based firms—will be indifferent to which center executes their projects and to prevent sociolinguistic cliques at its production centers. This randomized protocol allows us to circumvent econometric concerns and to compare workers' performance and turnover at different production centers.

[3] Subsequent literature in strategy and economics documents that differences in wages between the home and host regions (e.g., Borjas 1994) and preferences for living nearer to family and friends (Dahl and Sorenson 2012) are important considerations for workers charting migration paths.

We construct three sets of measures. First, we determine whether a worker was randomly assigned to the Indian IT industry's largest geographic cluster, Bangalore, or to one of the other production centers in smaller clusters and non-cluster locations. Second, we measure variation in workers' geographic origins by determining whether a worker received primary, secondary, and tertiary education in a smaller town or in a large city. Third, to capture variation in the characteristics of destination production centers, we code Bangalore as INDTECH's headquarters.[4]

We report three sets of results. First, we show that the performance effects of a Bangalore placement do indeed vary by place of origin: relative to placement elsewhere, placement in Bangalore has a positive and statistically significant effect on the performance of smaller-town workers, but a negative and statistically significant impact on workers from large cities.

Second, we report results on turnover to competitors. We find that workers assigned to Bangalore are significantly more likely to move to competitors than counterparts relocated elsewhere, and that Bangalore's higher rate of turnover is driven mostly by workers from smaller towns. Workers relocated from large cities to Bangalore are about 3.3 percentage points more likely to depart for a competing firm than counterparts at other production centers; those hired from smaller towns are 7 percentage points more likely to do so. This pattern is largely unique to Bangalore: compared to workers from larger cities, workers from smaller towns exhibit similar or lower turnover rates in the smaller clusters of Chennai, Hyderabad, and Pune and at non-cluster production centers.

Third, guided by prior theory, and in the spirit of "red-state papers" (Mitchell and Tsui 2012) and recent calls to apply abductive explanations to empirical patterns (King et al. 2019), we aim to offer plausible explanations of our results.[5] In short, across firm-induced migration paths, differences in external labor-market opportunities between the origin and the destination, and intrafirm skill-development opportunities at the destination, may be related to heterogeneous human-capital outcomes. We build on Jovanovic (1979), Lee and Mitchell (1994), and Carnahan et al. (2017), who argue that turnover occurs when "mismatches

[4] Bangalore is both the largest Indian technology cluster and the firm's headquarters. Thus we cannot cleanly disaggregate the effect of being assigned to the largest cluster from that of being assigned to the headquarters. In robustness tests, we attempt to mitigate this concern by distinguishing the outcomes of assignment to non-headquarters cluster locations (Chennai, Hyderabad, and Pune), all of which are smaller technology clusters than Bangalore. A later section discusses this limitation in detail.

[5] To quote King et al. (2019, p. 24), "research that uses abduction to develop plausible explanations is well suited to the management research setting. . . . Pre-specification is impractical for most research conducted on archival datasets." Mitchell and Tsui (2012, p. 2) define red-state research thus: "The focus of red research is on the phenomenon while existing theory or theories provide a means by which to focus on and obtain an understanding of the phenomenon."

become apparent to the workers and/or the firms" (Carnahan et al. 2017, p. 6), a learning process driven largely by new outside options. This scenario is especially applicable to workers from smaller towns, who are apt to have encountered fewer potential employers prior to joining INDTECH than workers from large cities. This circumstance may result in lower-quality initial matches for workers from smaller towns (Dauth et al. 2018), and in a higher likelihood of subsequent turnover—particularly among entry-level workers who are still exploring their preferences (Andersson and Thulin 2013; Bleakley and Lin 2012; Wheeler 2006).

To examine whether this abductive explanation is consistent with our findings, we leverage additional data sources and study secondary empirical questions. Figure 1 presents a schema of our research question, the main pattern of evidence, and the pattern of evidence pertinent to the abductive explanation. First, we examine whether the smaller-town construct is a proxy for limited economic opportunities and poor living conditions, which might motivate workers to seek better career opportunities once they arrive at a cluster. Second, to shed light on whether smaller-town workers encountered higher-quality employment options when they moved to Bangalore, we assess its relative concentration of domestic and multinational technology firms during the study period. Finally, we study whether workers from smaller towns disproportionately avail themselves of skill-development opportunities in Bangalore.

[Insert Figure 1 Here]

Our study provides causal evidence of how firm-induced migration pathways affect workers' on-the-job performance and turnover. It also suggests that a traditional firm-induced migration path—one that transfers workers from regions with fewer opportunities to those with greater opportunities—might enhance performance and skill development but also promote turnover to competitors. Our study is of interest not only to scholars of strategic human capital but also to managers: our findings can inform firms' decisions about where to move workers from different places of origin and about firms' future geographic footprints. Our findings also have implications for the literatures on firm-induced migration, value creation and capture from human capital, and turnover.

2. Empirical Context and Data

We obtain our data from INDTECH, one of India's largest IT firms, whose customers span the globe. Every year, INDTECH hires about 10,000 new college graduates with no prior full-time employment experience. We collect data on one such entry-level cohort, employees hired in 2007, and follow their production-center assignments and performance as well as turnover outcomes for a period of three years.

Upon entry, the new recruits are subject to two random assignments. First, they are assigned to one of three technological areas—.NET, Java, and Mainframe—that constitute INDTECH's core businesses; they then receive four months of induction training.[6] The company's training center in the southern city of Mysore has a 337-acre campus, 400 instructors, and 200 classrooms. According to our field interviews, INDTECH spent around $3,500 during the period of study to train each new hire for four months on such topics as relational databases, client-server concepts, and programming languages. Though assignment to one of the three technological areas is uncorrelated with observable individual characteristics, trends in demand for and supply of each technology may affect worker performance. We therefore restrict our data-collection exercise to the 1,665 entry-level workers[7] trained in a single area, .NET, which accounts for about 17 percent of the 10,000 new hires in 2007.

Following training, new recruits undergo a second random assignment to one of INDTECH's ten production centers,[8] located in Bangalore, Bhubaneshwar, Chandigarh, Chennai, Hyderabad, Jaipur, Mangalore, Mysore, Pune, and Trivandrum. Each production center executes projects in all three technological areas, and entry-level workers can be assigned to any of them. During the study period, however, no new recruits trained in .NET were assigned to Jaipur. Our final sample thus consists of 1,665 workers assigned to one of the nine other production centers.

3. Measures

3.1. Dependent Variables

We create two dependent variables to capture workers' (i) performance and (ii) voluntary turnover to competitors. We measure performance using annual performance ratings. All workers who worked on a coding/testing project for at least nine months of the preceding year are rated on a 1–5 scale; 5 is the highest rating. Field interviews with the head of talent development, a senior HR manager, and several

---

[6] .NET (pronounced *dot net*) is a software framework developed by Microsoft that runs primarily on Microsoft Windows. It provides interoperability across several programming languages—that is, each language can use code written in other languages—and includes a large library (http://en.wikipedia.org/wiki/.NET_Framework).

[7] In 2007, 1,696 of the graduates hired by INDTECH were assigned to .NET. However, 34 dropped out during the initial training and were never assigned to a production center. Thus, we drop these workers from the sample.

[8] Though assignment to production centers is random, the characteristics of workers hired by INDTECH from smaller towns and large cities might raise selection concerns. We thus perform a series of matched-sample analyses, matching workers from different places of origin on their observable pre-assignment characteristics, and then examine their differential responses to exogenous production-center assignments. These results appear in Tables A4–A6, A10–A12, and A17 in the Appendix; they replicate our main findings. The Results section discusses these analyses in more detail.

workers indicate that performance ratings for entry-level employees are based on objective measures, including the quality of coding and/or testing (measured using mistakes recorded by automated software), timeliness and completeness of coding/testing, and documentation (also measured by automated software). Each worker's manager provides an initial performance rating based on the objective criteria; HR managers check the rating against the underlying scores (i.e., scores of error rates and coding completeness). In the words of a senior HR manager, *"For workers in the study, performance evaluation is strictly based on objective metrics."*

We follow workers throughout their first three years at INDTECH. Thus our variable of interest, *Average Performance,* is the average of three annual performance ratings (unless the worker was dismissed or quit sooner). We standardize this variable to have a mean of 0 and a standard deviation of 1, yielding *Average Performance (z-score).* Table A3 in the Appendix replicates our results by disaggregating average three-year performance into separate performance ratings for 2008, 2009, and 2010; our results continue to hold.

To measure voluntary turnover to competitors, we leverage data provided by INDTECH about the recorded reason for each departure. We use these data to measure *Quit to a Competitor,* which takes a value of 1 if the worker quit to move to a competing firm and 0 otherwise.[9] Tables A7 and A8 report on additional reasons for departures, namely all types of voluntary departure, *Quit by Choice*, and departure for further study, *Quit for Further Study.*

3.2. Independent Variables

For each worker, we construct two independent variables. For the first, whether a worker's place of origin can be classified as a smaller town (*From Smaller Town*), we first obtain detailed résumés, which include the locations of a worker's primary school, high school, and college; this data is available for 93 percent of 2007 hires. We then classify cities and towns using a system promulgated by the government of India.[10] The system assigns each of India's cities and towns to one of three categories: the six largest metropolitan areas (Bangalore, Chennai, Delhi, Hyderabad, Kolkata, and Mumbai) (Tier 1); the next-largest cities (Tier 2); and smaller towns (Tier 3).[11] We code *From Smaller Town* as 1 if a worker attended (i) primary school, (ii) high school, and (iii) college in a Tier 3 location and 0 otherwise. Assuming that being from a small town

[9] A total of 207 workers departed INDTECH (either by choice or due to dismissal) before receiving their first performance rating. Thus, the number of observations in our regressions on performance and turnover differ.

[10] The Indian government formalized this system in 2008. State-owned entities and government departments use it to determine employees' cost of living.

[11] This categorization system is fully described at http://www.referencer.in/PayCommission/Reports/OM_Allowances.pdf.

correlates with *ex-post* superior performance (an assumption confirmed in Table 2, discussed below), this is the most conservative way of coding the variable *From Smaller Town*. Workers who attended primary school and high school in a Tier 3 town but attended college in a Tier 1 or Tier 2 city (or vice-versa) are coded as part of the control group, because they are likely to have been exposed to a different set of opportunities than workers whose schooling took place entirely in smaller towns.[12]

We adhere to the long tradition in urban economics of conceptualizing city size as relative to a country's population (Gabaix 1999; Eeckhout 2004). Though many Tier 3 towns are home to hundreds of thousands of people, they are considered small relative both to India's total population of more than 1.25 billion as of 2011 (the census nearest in time to our study period), and to its large metropolitan areas. Table 1, Panel A, reports the average populations of towns/cities in our sample by tier. The population of a Tier 3 town averages about 353,000 inhabitants (about 0.03 percent of the nation's population in 2011), relative to nearly 8 million on average in Tier 1 cities (about 0.66 percent of the 2011 population) and 1.4 million on average in Tier 2 cities (about 0.12 percent of the 2011 population). Panel B shows the breakdown of production-center locations by tier, the percentage of workers assigned to each production center, and each city's 2011 population and distance from Bangalore. Note that none of the production centers is in a Tier 3 town. Figure 2 shows the locations of production centers and other cities and towns in our sample.

[Insert Table 1 and Figure 2 Here]

In addition to our discrete measure of the size of workers' places of origin, we replicate our results using a continuous population measure (see the Robustness section). These results appear in Table A2. Because Tier 2 includes large metropolitan areas as well as smaller cities, we confirm that our results are robust to treating workers who attended school or college in Tier 2 areas as being from smaller towns. (These results are available on request.)

Our second independent variable is whether or not a worker was assigned to Bangalore. The resulting variable, *Placed in Bangalore*, takes a value of 1 if the worker was assigned to Bangalore and 0 if assigned elsewhere.[13] We also explore the effects of relocation to smaller technology clusters: Chennai, Hyderabad, and Pune. These results appear in the Robustness section.

[12] The quality of a college is unrelated to its location. For more information on colleges' rankings and locations, see Table A1 in the Appendix.
[13] As noted, no new hires trained in .NET were assigned to Jaipur during our study. Thus our final sample consists of nine production centers.

3.3. Control Variables

To rule out alternative explanations, we include a battery of controls. First, we add controls for workers' pre-entry performance on INDTECH's standardized recruitment test. The test has two components, verbal and logical; we recorded the scores separately as *Verbal Score* and *Logical Score*. We also control for cumulative grade-point average, *CGPA Training*, which captures performance during induction training; these scores are expected to correlate positively with on-the-job performance. To account for the possibility that large-city origin is associated with placement in one's hometown, we also create a dummy variable *Hometown,* which takes a value of 1 if a worker was assigned to his/her school or university location and 0 otherwise. Finally, we control for gender (*Male*) and production center (dummies) to which workers were assigned.

4. Identification Strategy

Our identification strategy exploits INDTECH's random computer-generated talent-allocation protocol. Appendix C outlines the steps in the assignment algorithm. This policy ensures that assignment of a worker to a particular location does not correlate with measures of observed ability, such as test scores or place of origin. Table 4 verifies this assumption empirically.

Using this identification strategy, we run the following specification:

(1) $$Performance_i \mid Turnover_i = \beta_0 + \beta_1 From\ Smaller\ Town_i + \beta_2 Placed\ in\ Bangalore_i + \beta_3 From\ Smaller\ Town_i * Placed\ in\ Bangalore_i + \sum_{j=1}^{J} \beta_{j+3}\, K_{ji} + \epsilon_i$$

where i refers to each individual and K_{ji} is a vector of J control variables. We also run separate subsample analyses of the effect of a Bangalore placement for workers from smaller towns and large cities; and, for both of our dependent variables, we run a series of matched-sample models for workers matched on their pre-assignment observable characteristics. For all our main analyses, we rely on OLS estimation[14] with and without production-center fixed effects and errors clustered at the production-center level.[15] The tables provide an alternative estimation of p-values in brackets below each of our main variables, based on a procedure developed by Roodman et al. (2019) that addresses the small number of clusters in our data.

5. Main Results: Workers' Place of Origin, Placement in Bangalore, Performance, and Turnover

[14] All of our results are further robust to nonlinear specifications for our binary and categorical variables (available in Table A3 and Figures A1–A4 in the Appendix).

[15] Table A9 further replicates our results using an alternative error-clustering method that also helps to account for the small number of clusters; our results continue to hold.

Table 2 presents descriptive statistics for the entire sample and for two subsamples, by place of origin and production-center placement. Columns 1–3 compare personnel and performance data for workers from smaller towns and large cities. The results reported in Panel A, Column 3, indicate that gender, hometown placement, and Bangalore placement are not significant predictors of smaller-town origin. Those reported in Panel B, Column 3, indicate that workers from smaller towns earned higher scores on the logical portion of the recruitment test than those from large cities (difference = 0.543, $p < 0.01$). Panel C indicates that smaller-town workers achieved higher performance ratings than those from large cities in 2008 and 2009, and on average in their first three years at INDTECH (differences = 0.153, $p < 0.001$; 0.146, $p < 0.01$; and 0.156, $p < 0.001$ respectively), and were not significantly more likely to move to competing firms.

Columns 4–6 compare personnel and performance data of workers placed in Bangalore to that of those placed elsewhere. The descriptive statistics reported in Panels A and B, Column 6, show that the pre-entry characteristics of workers placed in Bangalore do not differ significantly from those of workers placed elsewhere, further validating INDTECH's random-assignment protocol. Panel C indicates that workers placed in Bangalore neither perform significantly differently post-entry nor are more likely to move to competitors than counterparts placed elsewhere. Though suggestive, these results are based on pairwise comparisons, omitting a battery of controls for pre-entry characteristics and location fixed effects.

[Insert Table 2 Here]

Next we present robust econometric results. We begin by graphically exploring raw tabulations of worker performance and turnover. Figure 3 plots the average 2008–2010 performance of workers placed in Bangalore and elsewhere by place of origin. Workers from large cities placed in Bangalore perform worse than counterparts placed elsewhere; workers from smaller towns perform better in Bangalore.[16] T-tests of group differences in means indicate that performance differences between Bangalore and other production centers are marginally significant (p-value = 0.078) for workers from large cities but not for those from smaller towns.

[Insert Figure 3 Here]

[16] The confidence intervals displayed are those for the means in each group, not for the difference in their means; thus, the significance of the differences between groups cannot be inferred from them.

Figure 4 depicts overall rates of voluntary turnover to competitors by workers' places of origin and by placement location. Workers placed in Bangalore exhibit higher rates of turnover regardless of their places of origin. However, workers from smaller towns exhibit larger increases in propensity to leave for competing firms when placed in Bangalore than when placed elsewhere. This difference is marginally significant (p-value = 0.075). For workers from large cities, it is insignificant.

[Insert Figure 4 Here]

Table 3 reports the effects of placement location and place of origin on performance and on propensity to move to competing firms, using specification (1) above. Columns 1 and 2 examine the differential effects of placement in Bangalore on the performance of workers from smaller towns and large cities. As Column 1 shows, workers from smaller towns show similar average performance regardless of placement location. But when we interact placement location with city of origin in Column 2, we see patterns similar to those in Figure 3; workers from smaller towns exhibit larger increases in performance in Bangalore (relative to elsewhere) than their large-city counterparts. This difference is statistically significant at conventional values, even when we include production-center fixed effects in Column 3. The economic significance of these effects is also meaningful: the difference in performance across production centers between workers from smaller towns and large cities is about 0.3 standard deviations.

The models in Columns 1–3 do not account fully for possible heterogeneous effects of our control variables by place of origin. (That is, the effects of CGPA and Logical Score on Average Performance may vary in important ways by place of origin). Therefore, Columns 4 and 5 replicate Column 2 using subsample analyses, splitting workers into the two mutually exclusive and exhaustive groups, *Large-City Sample* and *Smaller-Town Sample*. The results in Column 3 are driven both by the inferior performance of large-city workers placed in Bangalore (by about 0.158 standard deviations) and by the superior performance of workers from smaller towns (by about 0.156 standard deviations). We verify that these estimates differ significantly (p-value = 0.007) using a cross-model Wald test that accounts for covariance between models (that is, using seemingly unrelated estimation).

Columns 6–10 report results for turnover to competing firms. Column 6 shows that workers from smaller towns are just as likely to quit for competing firms as are workers from large cities. Moreover, placement in Bangalore increases the likelihood of such turnover for all workers, regardless of origin. When we interact place of origin and placement location in Column 7, however, we see that higher overall turnover in Bangalore is driven at least in part by workers from smaller towns, whose departure rates are significantly higher in Bangalore than elsewhere. In fact, workers from smaller towns appear less likely to depart INDTECH if placed anywhere but Bangalore, as shown in Columns 7 and 8. We also perform a split-sample analysis to examine placement in Bangalore separately for workers from large cities and smaller towns. As Columns 9 and 10 show, both groups experience an increased propensity in Bangalore for turnover to competing firms, but workers from smaller towns exhibit a much higher likelihood of such turnover than those from large cities (7 percentage points relative to 3.3 percentage points). These estimates differ significantly (p-value = 0.018) using a cross-model Wald test that accounts for covariance between models (that is, using seemingly unrelated estimation).

[Insert Table 3 Here]

6. Robustness Checks

6.1. Validation of Random Assignment to Production Centers

An important robustness check is validation of INDTECH's talent-allocation protocol (i.e., that assignment to a particular production center is not correlated with observable individual characteristics, including performance during recruitment and training). As Table 4 shows, assignment to Bangalore after induction training is *not* correlated with observable characteristics, such as being *From a Smaller Town*. Likewise, assignment to Bangalore is correlated neither with observable measures of performance (such as CGPA at the end of training or standardized test scores) nor with gender or hometown location. These findings validate the random talent-allocation policy underlying our study.

[Insert Table 4 Here]

6.2. Matched-Sample Analyses

A second concern about the estimations in Table 3 is that, in all our analyses, only workers' production-center placements are randomized, not their places of origin. We thus replicate our results with a series of matched-sample models—matching workers from smaller towns and large cities on their pre-entry and pre-placement characteristics, using coarsened exact matching with different breakpoints—and repeat our analyses. As Tables A4–A6 in the Appendix show, the conclusions in Table 3 continue to hold.

6.3. Standard-Error Estimation

A third possible issue with the estimations in Table 3 is the small number of production centers on which we cluster our standard errors, which can lead to challenges in accurately estimating standard errors. To address this concern, we first follow Roodman et al. (2019) and recalculate p-values for our main variables using a wild bootstrap-t procedure that accounts for the small number of clusters. These results appear in brackets below the main variables in Table 3; our conclusions continue to hold with these re-estimated p-values. Table A9 in the Appendix uses a second method, paired bootstrap-t clustered errors, to account for the small number of clusters in our data; our results continue to hold.

6.4. Non-Linear Estimation

A fourth concern has to do with the distribution of dependent variables. Our measure of performance is a categorical variable that takes values from 1 to 5; our measure of turnover is binary. To ensure that our results are not driven by the linear specification used in Table 3, and to provide further estimates of the effect sizes of our independent variables, Table A3 in the Appendix replicates our results using non-linear specifications. We perform these analyses separately for each year for which we have performance data; the general patterns persist (though the results are stronger for 2008 and 2010 than for 2009). Our turnover results continue to hold with non-linear estimation. Figures A1–A4 in the Appendix also depict the marginal effects on our interaction terms, following advice in Hoetker (2007) and Gomila (2020); Tables A7 and A8 examine alternative measures of performance and other possible reasons for turnover.

6.5. Cluster or Headquarters? Replication of Results for Chennai, Hyderabad, and Pune

Finally, we attempt to disentangle whether the patterns of worker performance, turnover, and human-capital augmentation that we observe in Bangalore are driven by its status as the largest cluster or as INDTECH's headquarters. We begin by studying whether Bangalore's effects might be replicated at production centers located in smaller clusters. INDTECH has production centers in three smaller clusters; if the results in Tables 3 are driven mainly by Bangalore's status as a cluster, we should be able to replicate the patterns reported above at each of those production centers.

Table 5 replicates the Bangalore results reported in Table 3 for Chennai, Hyderabad, and Pune, using subsample analyses for easier interpretation. As shown, the Bangalore effect appears to be replicated only for performance in Hyderabad but not for turnover there, or for performance or turnover in Chennai and Pune. Neither performance nor turnover increases significantly for smaller-town workers in Chennai or Pune, as opposed to other production centers. These patterns persist when we exclude Bangalore from the sample and compare Chennai, Hyderabad, and Pune only to the remaining five production centers. Jointly, these results suggest that opportunities for human-capital augmentation that may enhance workers' performance, and thus promote turnover, are likely to be more prevalent and, in the case of R&D, more available in Bangalore than in the three next-largest cluster locations; this pattern may help explain the smaller differences in turnover between locations other than Bangalore.

[Insert Table 5 Here]

7. Examining Evidence Supportive of Abductive Explanations of the Main Results

We turn examine evidence consistent with an abductive explanation, guided by King et al. (2019, pp. 8–9): "An explanation is a conjecture about an observed pattern of evidence. . . . Abduction allows only a basis for making a 'promising explanatory conjecture' which then must be 'subject to further test.'" We attempt to rule in a likely abductive explanation of our results: that across firm-induced migration paths, differences in external labor-market opportunities between workers' places of origin and destinations, and differences in intrafirm opportunities at the destinations, are related to heterogeneous human-capital outcomes. Specifically, we document four pieces of evidence consistent with this explanation. Section 7.1 documents that, during hiring, smaller-town workers exhibit observable human capital superior to that of their counterparts from larger cities (higher scores on standardized tests); this might explain differences in performance post-relocation. Section 7.2 then reports that workers from smaller towns enjoy fewer socioeconomic opportunities in their hometowns than counterparts from larger cities; this might explain why INDTECH is able to hire more talented workers from smaller towns. Section 7.3 then shows that workers encounter increased employment opportunities when relocated to Bangalore; this might explain heterogenous turnover outcomes post-relocation. Finally, Section 7.4 documents that smaller-town workers disproportionately avail themselves of human-capital-augmentation opportunities in Bangalore; this might explain both heterogenous performance and turnover outcomes post-relocation.

7.1. Pre-Hire Differences in Observable Human Capital: Smaller-Town and Large-City Workers

We begin by documenting that, at the time of hiring, smaller-town workers earn superior standardized test scores. As Table 2, Panel B, Column 3 shows, their scores on a standardized test of logical ability administered during recruitment are about one-third of a standard deviation higher than those of large-city counterparts. Moreover, as discussed in Section 6.2: Matched Sample Analyses, smaller-town workers also perform better post-hire even after we match workers from smaller towns and large cities on their pre-entry and pre-placement characteristics, using coarsened exact matching with different breakpoints. (These results appear in Tables A4–A6 in the Appendix.)

7.2. Socioeconomic Opportunities at Places of Origin

Next, we examine the role of socioeconomic conditions at workers' places of origin in driving the empirical patterns we found in performance and turnover. Prior literature suggests that the smaller-town construct might be a proxy for, among other things, scarce economic opportunities and poor living conditions. This might explain both why firms are able to hire more talented workers from smaller towns and why workers from smaller towns are more motivated to seek better career opportunities once they arrive in a city or cluster.[17] This section examines whether this is the case.

Smaller towns are often characterized by relatively poor social infrastructure, such as limited access to high-quality education (Banerjee et al. 2007; Chaudhury et al. 2006; Muralidharan and Sundararaman 2015), high crime (Chattopadhyay and Choudhury 2017), and a poor quality of life (Khan 2014). A 2018 government report documents that smaller towns lack "critical size for attracting investments, economic activity, large infrastructure spending" (Ministry of Housing and Urban Affairs 2018, p. 37).

We validate these observations by consulting two additional datasets. The first is a report by the Indian Ministry of Housing and Urban Affairs on living conditions in 111 cities and towns in 2018.[18] The data consists of indices assigned to each city to measure economic, social, institutional, and physical living conditions. These indices are aggregated into an overall "ease-of-living" index that consists of a weighted average measure across the four metrics. The index assigns absolute values from 0 to 100 to each city, as measured in 2017–2018. Our second source of data is aggregate crime rate data for 35 cities[19] collected by the National Crime Records Bureau, Ministry of Home Affairs, in 2008,[20] with a specific focus on

---

[17] In seminal work on moving toward opportunity, Katz et al. (2001) document that fear of crime and a desire for better housing and schooling can act as powerful motivators for geographic mobility. Florida (2005) documents that clusters' superior amenities and quality of life motivate migrants to move to such areas and to strive to enter the creative class. The literature on migration (e.g., Chiswick 1978) has documented that the career returns from migration are higher for more motivated migrants.

[18] The 111 cities were selected as representative of living conditions across India; they include a broad cross-section in terms of population size (from cities with fewer than 500,000 inhabitants to those with more than 5 million inhabitants). For more information on the construction of the index and its scope, see https://www.ipsos.com/en-in/ease-living-index-2018.

[19] Data on locales other than the 35 largest cities and towns was not available for our study period; thus this sample is considerably smaller.

[20] Data are available at https://ncrb.gov.in/en/crime-in-india-table-addtional-table-and-chapter-contents?page=1.

"cognizable" crimes that require immediate police investigation and potential arrest without warrant.[21] Crime rates are adjusted for population size and expressed as the number of incidents per 100,000 inhabitants. Thus, we can directly compare crime rates across cities. A list of the cities and their associated crime rates appears in Table A13 in the Appendix.

We construct four measures of socioeconomic well-being for workers' hometowns, defined as their school location[22]: *Overall Score (reversed)* takes values from 1 to 100, with 1 representing the highest quality of life and 100 the lowest; *Overall Score Bottom 20%* takes a value of 1 for hometowns that fall in the bottom 20 percent for overall ease of living and 0 otherwise. *Crime Rate* is the number of reported crimes per 100,000 inhabitants; *Crime Rate Top 20%* takes a value of 1 for hometowns among the 35 cities whose crime statistics we obtained whose crime rates fall in the top 20 percent and 0 otherwise.

As Table 6 reports, workers from smaller towns are more likely to be from places that score poorly on the overall ease-of-living index and suffer high crime rates.

[Insert Table 6 Here]

Next, we replace our measure of smaller-town origin with each of the continuous metrics that capture living conditions in the worker's city of origin, and rerun our results for turnover to competing firms and average performance. Though the *Overall Score* and *Crime Rate* measures both appear to be correlated with *From Smaller Town*, results in Table 7 show weak correlation with average performance and stronger correlation with turnover only for *Overall Score*. However, when we distinguish between locations in the bottom and top fifth of the distribution for overall ease of living and crime rates, a strong pattern replicates our results for the smaller-town construct. Table 8 reports these results. Like workers from smaller towns, those from cities/towns in the bottom 20 percent for overall ease of living are more likely to exhibit better performance in their first three years on the job and to move to competing firms when placed in Bangalore than counterparts placed elsewhere. Further, workers from locations with crime rates in the top 20 percent perform better and are more likely to move to competitors, on average, when placed

[21] The National Crime Records Bureau defines a cognizable crime as "one which an officer in-charge of a police station may investigate without the order of a magistrate and effect arrest without warrant. . . . Cognizable crimes are broadly categorised as those falling either under the 'Indian Penal Code (IPC)' or under the 'Special and Local Laws (SLL)'" (National Crime Records Bureau, 2008, p. 15). Cognizable crimes include bodily harm (attempted murder or injury and sexual offenses), property theft, and forgery.

[22] We use workers' school locations to construct this measure because they are likely to represent the conditions that workers experienced throughout childhood and adolescence; they are also apt to be where their families still reside and where they are apt to return if they lose their jobs at INDTECH and cannot find a replacement. In unreported results, we replicate these analyses separately for university locations; the results are similar in the direction of their effects but considerably weaker in magnitude and statistical significance.

in Bangalore than counterparts placed elsewhere or workers from cities with lower crime rates who are placed in Bangalore.

To ensure that these results are not driven by our arbitrary cut-off of the towns/cities distribution at 20 percent, we replicate the analyses in Tables A14 and A15 with cut-offs at 30 percent and 10 percent. The results are weaker for performance, but hold for turnover to competitors. Table A16 disaggregates the ease-of-living index into its four components and shows that our results are largely driven by the social component, which captures the quality and availability of education, the quality and availability of health care, safety, and security, and the fraction of the municipal budget allocated to cultural heritage. These results, along with our findings for crime rates, suggest that the smaller town construct is a proxy for differences between smaller towns and large cities on quality-of-life indicators.

[Insert Tables 7 and 8 Here]

Appendix B validates and extends our insights using qualitative data on employment opportunities for college graduates in smaller towns. Our survey of engineering colleges finds evidence, reported in Table B1, that—with the exception of INDTECH and a few other firms—technology firms in India, especially multinationals, do not hire from smaller towns.[23] This finding is consistent with prior research (e.g., Jensen 2012). Our survey also finds that mean salaries for 2011 and 2012 are significantly higher for large-city college graduates than for smaller-town graduates. We validate that this difference is statistically significant in a t-test comparison of means. Given the higher test scores of smaller-town workers reported in Section 6.1, smaller-town workers might have experienced an initial mismatch at INDTECH.

7.3. Employment Opportunities at Destinations

Next, we examine whether, during the study period, Bangalore was indeed a hub with a disproportionate number of technology firms and MNCs relative to other production centers, which might explain heterogenous turnover outcomes post-relocation. We collected data on member companies of the National Association of Software and Service Companies (NASSCOM),[24] a trade association of information-technology and business-process outsourcing companies established in 1988. We hand-collected two pieces

[23] We randomly selected 10 large-city colleges and 10 smaller-town colleges from the list of colleges where INDTECH hires, and contacted representatives at each via phone and/or e-mail. Eleven of the 20 agreed to telephone interviews lasting about 30 minutes. We interviewed either the head of the college or the head of the group responsible for organizing recruitment. We asked about total graduating-class size, starting salaries for undergraduate engineers in 2011 and 2012, which technology firms (Indian and foreign multinational) hired from the college, and how many students were hired by each firm.

[24] For more information on NASSCOM, see https://nasscom.in/.

of information for each member firm: its founding date and its headquarters location. We wanted to determine whether by 2011 Bangalore was home to more technology firms and more MNCs than the three next-largest clusters where INDTECH had production centers (Chennai, Hyderabad, and Pune). Figures 5 and 6 present our results. Figure 5 shows that, even by 2011, Bangalore hosted 32 percent more technology firms (296) than the runner-up city, Chennai (224), and significantly more than the next two, Pune (181) and Hyderabad (175). In Figure 6, which looks only at MNC IT firms (IT firms with foreign headquarters), the gap between Bangalore and the other cities is even larger: By 2011, Bangalore was home to 147 MNCs, Hyderabad to 88, and Chennai and Pune to 68 each.

[Insert Figures 5 and 6 Here]

To examine whether Bangalore's greater concentration of technology firms and MNCs is associated with INDTECH workers' turnover there, we collect data on former employees from LinkedIn. Specifically, we manually collect data on all employees between 2008 and 2011 who had moved to another employer by 2012. From an initial sample of 2,182 employees, we exclude 1,158 whose profiles did not specify their next employer's location or characteristics—industry, founding date, public/private status, and headquarters location. We also restrict our sample to employees who left to join private or public corporations, excluding educational entities, government, and charities (147 profiles report pursuing further education or joining governmental entities or non-profits). We then construct two samples. The first consists of 877 active profiles, which we use in Table 7, Panel B, to analyze the characteristics of former INDTECH workers' subsequent jobs (if such jobs were in India). The second sample is a subset of 282 workers from among these 877 active profiles whose INDTECH placement location is one of the nine production centers in our sample. (We exclude workers with no INDTECH placement location or a location other than the nine production centers in our sample.)

We then create nine variables. *Next Job in Tech* takes a value of 1 if the focal worker's next job was in a technology industry, defined as one of several categories identified by LinkedIn—information technology and services, internet, computer software, computer networking, and computer and network security—and 0 otherwise. *Next Job in a Startup* takes a value of 1 if the worker's next employer was founded during or after 2008 and 0 otherwise. *Next Job in a Public Firm* takes a value of 1 if the next employer is listed as a public firm on LinkedIn and 0 otherwise. *Next Job in Bangalore/India/Abroad* take a value of 1 if the next employer is located in Bangalore, India excluding Bangalore, and outside of India, respectively, and 0 otherwise; *HQ in Bangalore/India/Abroad* takes a value of 1 if the next employer's headquarters is located in Bangalore, India excluding Bangalore, and outside of India, respectively, and 0 otherwise.

Using the two samples, we examine two related questions. First, how do employment opportunities vary with placement locations? Table 9, Panel A, presents the results for the subsample of 282 workers whose placement location we were able to determine. We find that workers placed in Bangalore had a 6.2 percentage-point-higher likelihood of moving on to a startup, a 42.2 percentage-point-higher likelihood of moving on to a different firm in Bangalore, and a 12.1 percentage-point-higher likelihood of moving on to an MNC, than counterparts at other INDTECH production centers. Second, do the employment opportunities workers pursue in Bangalore differ systematically from the employment opportunities they pursue outside of Bangalore? Table 9, Panel B, presents corresponding results for the full sample of 877 workers. On average, workers whose next jobs were located in Bangalore were more apt to move on to a technology firm (5.7 percentage points more likely), a startup (6.9 percentage points), or a firm headquartered in Bangalore (12.2 percentage points) or abroad—that is, an MNC—(10 percentage points) than workers whose next jobs were located elsewhere.

These additional analyses further support the abductive explanation that departing INDTECH workers in Bangalore enjoyed better options, such as joining MNCs and startups, than departing INDTECH workers elsewhere. In unreported results, we confirm that these results are specific to Bangalore; they do not hold for Chennai, Hyderabad, or Pune, the three next-largest technology hubs.

[Insert Table 9 Here]

7.4. Human-Capital-Augmentation Opportunities at Destinations

Finally, we examine whether workers from smaller towns exhibit a greater propensity to augment their human capital after relocation to Bangalore, relative to other locations, than counterparts from large cities, which might explain heterogenous performance and turnover outcomes post relocation. We leverage two sources of data: (i) microdata on workers' time allocation across activities and (ii) data on workers' voluntary coursework. INDTECH's microdata on workers' time allocation captures how much time per week each worker dedicates to project work, R&D, training, and other activities during the first three years on the job. We use these data to compare time allocation at different production centers. As Figure 7 shows, workers in Bangalore do not spend disproportionately more time on active client projects (production) and do not receive disproportionately more training; they are, however, more likely to spend time on R&D activities. Workers in Bangalore spend about 4.5 percent of their time on R&D activities, relative to 3.5 percent in Trivandrum and less than 2 percent at all other locations. The difference between time spent on R&D in Bangalore and elsewhere is statistically significant (p-value = 0.000).

[Insert Figure 7 Here]

To examine whether these differences are consequential, and how they differ between workers from smaller towns and large cities, we implement specification (1) and regress worker-level time use on places of origin, Bangalore location, the interaction of the two, and the usual battery of control variables. Table 10, Columns 1 and 2, confirm that workers from both large cities and smaller towns spend more time on R&D in Bangalore than elsewhere (by about 1.4 and 5.5 percentage points respectively), but that the increase in time spent on R&D is larger for workers from smaller towns (p-value = 0.000) based on a cross-model Wald-test comparison using seemingly unrelated estimation. Columns 3 and 4 confirm that these differences are significant—that workers from smaller towns are more likely to engage in R&D than their large-city counterparts when assigned to Bangalore, by about 3.5 percentage points relative to an unconditional mean time spent on R&D of 4.5 percentage points. These results persist even with the addition of production-center fixed effects and the recalculation of standard errors to account for the small number of clusters (recalculated p-values are reported in brackets below our key variables). Table A17 replicates our results with coarsened exact matching to account for INDTECH's non-random selection of workers from smaller towns and large cities; our results continue to hold.

[Insert Table 10 Here]

To understand how engagement with voluntary training varies with workers' places of origin and placement locations, we examine online coursework provided to workers free of charge. We collect data on the number of courses that workers enrolled in (average is about 3.3 courses each) and the percentage they passed (average is about 81%), as well as the number of Business English courses they enrolled in and the difficulty level of those they completed. On average, about 1 in 7 workers enrolled in at least one English course, at the second level of difficulty, during their first three years at INDTECH.

Table 11 presents our results. Columns 1 and 2 examine the numbers of courses that workers from smaller towns enroll in when placed in Bangalore versus other production centers. Such workers enroll in more courses overall, but their propensity to do so does not vary between Bangalore and elsewhere. Columns 3 and 4 compare percentages of courses passed by workers' places of origin and placement locations, and reveal a similar pattern: workers from smaller towns pass more courses, but their pass rate does not appear to vary between Bangalore and elsewhere.

When we focus on the propensity of workers from smaller towns to take courses that enhance general human capital and are thus easily transferable to other firms—namely, courses in Business English—we find that variation between Bangalore and other placement locations. (English-language skills are highly sought after, especially in the IT sector, which is geared toward serving international customers, and especially at MNCs, which require workers to collaborate with colleagues overseas.) By contrast, according to HR managers and workers at INDTECH, the content of functional courses, such as Banking or Telecommunications, is apt to be less transferable to competitors.) Columns 5 and 6 show that workers from smaller towns do not take more English courses on average than workers from large cities (Column 5), but do take more English courses when posted to Bangalore. Columns 7 and 8 show that, conditional on taking at least one English course, workers from smaller towns enroll in higher-level courses on average than those from large cities, and that this difference is driven largely by workers from smaller towns placed in Bangalore. Results in Columns 1–8 are robust to recalculation of standard errors, which accounts for the small number of clusters in our data (p-values appear in brackets below key variables). In unreported results, the results for the number of English courses taken (Columns 5 and 6) and their difficulty (Columns 7 and 8) are robust to re-estimation with Poisson and ordered logit estimators respectively.

[Insert Table 11 Here]

Finally, to tease out whether the Bangalore effect is driven by its status as the largest technology cluster or as INDTECH's headquarters, we replicate the results in Tables 10 and 11 for each production center in smaller clusters that workers could be assigned to: Chennai, Hyderabad, and Pune. These results appear in Tables A18−A19 in the Appendix. Table A18 replicates the R&D results, showing that workers from smaller towns are no more likely to engage in R&D when assigned to Chennai, Hyderabad, or Pune than when assigned elsewhere. Table A19 repeats the same exercise for additional coursework; here, workers from smaller towns enroll in more coursework in Chennai and Pune, and pass more of their courses in Chennai, than large-city workers. They do not take more English courses in any of the three locations; nor, conditional on taking English courses, do they pursue them at a higher level than large-city workers.

Jointly, these results indicate that our findings are unlikely to be driven by a pure technology-cluster effect. Bangalore does enjoy a disproportionate concentration of technology firms and MNCs, but we cannot rule out that at least part of its effect is driven by its headquarters status, which provides employees greater access to resources than other locations.

8. Discussion

This study posits that the firm-induced migration path is an important determinant of post-relocation human-capital outcomes. We exploit a natural experiment stemming from a randomized talent-allocation protocol at an Indian technology firm to show that workers from smaller towns exhibit superior on-the-job performance when assigned to Bangalore, India's largest technology cluster and the focal firm's headquarters, than when assigned to production centers elsewhere; however, such workers are also more likely to depart for competing firms. In contrast, workers hired from large cities exhibit lower on-the-job performance when assigned to Bangalore than when relocated elsewhere, but depart for competing firms at a lower rate than workers from smaller towns. These outcomes may have significant economic implications: back-of-the-envelope calculations of the net returns of moving workers along different migration paths, presented in Appendix D, suggest that such returns are highest when relocating workers from smaller towns to production centers other than Bangalore and lowest when relocating large-city workers to Bangalore. The net difference varies from $719 to $2,742 per worker per year, or from 12.2 percent to 46.7 percent of the annual salary of workers in this sample.[25]

With the aim of ruling in a plausible abductive explanation, as urged by King et al. (2019), we provide empirical evidence that differences in labor-market opportunities between workers' places of origin and destinations, and intrafirm skill-development opportunities at the destination, plausibly explain heterogeneous human-capital outcomes. We also provide evidence that smaller-town origins are a proxy for scarce socioeconomic opportunities and worse quality of life, and document that workers enjoy

[25] As Appendix D explains, these estimates are obtained by calculating the upper and lower bounds of the revenues that INDTECH probably generates from each worker. We derive the upper bound by dividing total INDTECH revenues in 2008 by its number of employees, and the lower bound by combining workers' starting salaries and replacement costs. We collect salary and replacement-cost data from management and from our turnover results.

disproportionately greater external opportunities and intrafirm human-capital-augmentation opportunities when relocated to Bangalore.

Our study's limitations and boundary conditions present a rich agenda for future work. In the tradition of insider econometrics (Baker et al. 1994; Bartel et al. 2004), our data is collected from a single firm. Several specific conditions at INDTECH potentially limit the generalizability of our findings. One limitation is that Bangalore was both the leading Indian IT cluster at the time of our study and INDTECH's headquarters. Thus, we are unable to cleanly distinguish between the performance and turnover effects of moving workers to the largest cluster and of moving workers to headquarters. We attempt to disentangle these two effects by studying the effects of moving INDTECH workers to Chennai, Hyderabad, and Pune, the next-largest technology clusters in our sample during our period of study. We observe no statistically significant performance and turnover effects when smaller-town workers are assigned to those non-headquarters cluster locations. Our results suggest that the Bangalore effect is unlikely to be driven by a cluster effect alone; rather, workers appear to be utilizing the unique resources available at the headquarters to develop skills and to signal employability and/or discovering better matches. This possibility raises a question for future research: the relative importance in driving employee turnover of intrafirm resources at headquarters vs. external labor-market opportunities in a cluster, and how such resources and opportunities are utilized by workers from different places of origin.

Second, the generalizability of our findings could be constrained by such industry-level characteristics as cluster size and the characteristics of entry-level workers who choose the industry. In a qualitative study, Deshpandé and Raina (2011) document that the values of smaller-town workers in the Indian hospitality industry differ from those of large-city counterparts. This finding raises the possibility that smaller-town workers self-select to join different industries, and may exhibit patterns of loyalty and turnover that vary with the industry they join, a fruitful question for future research. Relatedly, we present human-capital augmentation as a plausible explanation for turnover to competitors, but cannot test for, or rule out, alternative and complementary explanations such as motivation. For example, Chiswick (1978) found that foreign-born migrants were often more highly motivated than the native-born. Future research may further disentangle the mechanisms underlying the smaller-town effect. Third, our study is agnostic about differences in the relocation costs of different migration pathways. Future research could explore how relocation costs impact the net returns of heterogenous firm-induced migration paths.

Another limitation is that, in the longer-run equilibrium, individual outcomes might be subject to changing migration patterns, changing geographic distribution of firms, scaling-up of production centers in non-cluster and/or non-headquarters locations, and increased adoption of remote work, especially work-from-anywhere (WFA). Within a WFA regime, workers might prefer to live in smaller towns for such reasons as a lower cost of living and proximity to family (Choudhury et al. 2021). Future research should explore the conditions under which firms offer workers from heterogenous origins the flexibility of WFA, and how doing so impacts firm-induced migration paths.

Finally, studying firms' policies on hiring and managing smaller-town workers in western contexts and in emerging markets like India would be an important extension of this research. Prior literature on the dynamics of interorganizational careers has argued (Bidwell and Briscoe 2010) that workers are apt early in their careers to choose larger firms that provide more training, and later to move to smaller organizations that reward their skills. Smaller-town workers assigned to Bangalore may eventually adopt that pattern. Future research might explore whether there exists a separating equilibrium whereby some firms assign smaller-town workers to clusters while others (e.g., MNCs) hire such workers once they have arrived in a cluster.

Our study contributes to several streams of the strategic-human-capital literature, notably that on firm-induced migration (Foley and Kerr 2013; Hernandez 2014; Wang 2015; Choudhury 2015, 2017; Choudhury and Kim 2019; Hernandez and Kulchina 2020). We conceptualize *firm-induced migration paths*, rather than workers' place of origin and destination alone, as relevant to human-capital outcomes. Firm-induced migration paths are apt to differ from those taken by individuals who chart their own paths. The migration literature in sociology (e.g., Portes 1995), for example, documents individual migrants' decisions to move to "ethnic enclaves" and to leverage community-based resources. In contrast, firm-induced migration paths offer new opportunities at the assigned production center and in the external labor market that workers are unlikely to be aware of prior to moving. Future research can explore how the heterogeneity of firm-induced migration paths—both cross-border and within-country—affects outcomes other than performance and turnover, such as knowledge sharing, knowledge recombination, and entrepreneurship.[26]

Our results also contribute to the literature on value creation and value capture from human capital. Prior literature in strategic human capital has outlined how labor-market imperfections—especially frictions related to workers' geographic preferences—might limit the geographic mobility and career mobility of human capital (Campbell et al. 2012) and turnover to competitors (Carnahan et al. 2012; Carnahan and Somaya 2013; Carnahan et al. 2017; Ton and Huckman 2008). By matching workers to intrafirm opportunities and resources at their destinations, firms might help mitigate geographic-mobility frictions (Choudhury, forthcoming) and incentivize workers to pursue intrafirm mobility. Our study underlines, however, that firms should consider both value creation and value capture from human capital when relocating workers. As Campbell et al. (2012, p. 377, emphasis added) observe, "human capital can be at the core of a resource-based advantage if it is valuable, rare *and can be kept from rivals*."

Our results also speak to the literature on agglomeration economies and on locating talent in clusters (Alcácer and Chung 2014; Mariotti et al. 2019; Shaver and Flyer 2000). In a recent review, Kerr and Robert-Nicoud (2020) subdivide this vast literature into two strands: (i) studies of clusters as "high-velocity labor markets," characterized by rapid turnover; and (ii) studies of "immigration, diversity, and tech talent" in clusters. We link the two strands by documenting heterogeneous patterns of turnover based on differences in workers' geographic origins and migration paths. In fact, we show that a longstanding finding in the literature on clusters—rapid turnover to competitors—is more salient in our setting for workers hired from smaller towns than for those from large cities. Future research might explore whether multi-unit firms can strategically relocate high-quality human capital away from clusters to avoid poaching; recent anecdotal evidence suggests that firms such as MobSquad have done so (Choudhury 2020), while still locating some operations in clusters to benefit from "listening in" (Monteiro 2015) on competitors.

This study contributes to the ongoing conversation on the role of firms in moving talent between geographies by conceptualizing how heterogeneous firm-induced migration pathways affect individual human-capital outcomes. We exploit unique data and a randomized talent-allocation protocol to provide

[26] Within-country migration has long been studied in economics (Bazzi et al. 2016; Bryan et al. 2014; Harris and Todaro 1970; Munshi and Rosenzweig 2016; Young 2013), but ours is among the first empirical studies of firm-induced within-country migration and of how heterogeneous pathways affect individual outcomes.

causal evidence and to rule in a plausible abductive explanation: across firm-induced migration paths, both differences in external labor-market opportunities between the place of origin and the destination and intrafirm skill-development opportunities at the destination are related to heterogeneous human-capital outcomes. Our insights have relevance for managerial decisions on intrafirm talent location and relocation and for firms' future geographic footprints.

References

Alcácer, Juan, and Wilbur Chung. "Location strategies for agglomeration economies." *Strategic Management Journal,* 35.12 (2014): 1749–1761.

Almazan, A., De Motta, A., & Titman, S. "Firm location and the creation and utilization of human capital." *The Review of Economic Studies*, 74.4 (2007): 1305–1327.

Andersson, Martin, and Per Thulin. "Does spatial employment density spur inter-firm job switching?" *The Annals of Regional Science,* 51.1 (2013): 245–272.

Andersson, U., M. Forsgren, and U. Holm. "The Strategic Impact of External Networks: Subsidiary Performance and Competence Development in the Multinational Corporation." *Strategic Management Journal* 23.11 (2002): 979–96.

Austin, Benjamin, Edward Glaeser, and Lawrence Summers. "Jobs for the Heartland: Place-Based Policies in 21st-Century America." *Brookings Papers on Economic Activity*, Spring (2018): 151–255.

Baker, George, Michael Gibbs, and Bengt Holmstrom, "The Internal Economics of the Firm: Evidence from Personnel Data," *The Quarterly Journal of Economics*, 109.4 (1994): 881–919.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. "Remedying education: Evidence from two randomized experiments in India." *The Quarterly Journal of Economics*, 122.3 (2007): 1235–1264.

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw, "Using 'Insider Econometrics' to Study Productivity," *The American Economic Review*, 94.2 (2004): 217–223.

Bartlett, C. A. Building and managing the transnational: The new organizational challenge. In Michael E. Porter (Ed.), *Competition in global industries* (pp. 367–404), Harvard Business School Press, 1986.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg, and Maisy Wong. "Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia." *The American Economic Review,* 106.9 (2016): 2658–2698.

Bidwell, M. and F. Briscoe. "The Dynamics of Interorganizational Careers." *Organization Science,* 21.5 (2010): 1034–1053.

Björkman, Ingmar, Wilhelm Barner-Rasmussen, and Li Li. "Managing Knowledge Transfer in MNCs: The Impact of Headquarters Control Mechanisms." *Journal of International Business Studies,* 35.5 (2004): 443–55.

Bleakley, Hoyt, and Jeffrey Lin. "Thick-market effects and churning in the labor market: Evidence from U.S. cities." *Journal of Urban Economics,* 72.2–3 (2012): 87–103.

Borjas, George J. "The economics of immigration." *Journal of Economic Literature* 32.4 (1994): 1667–1717.

Bower, Joseph L. *Managing the resource allocation process: A study of corporate planning and investment*. Homewood: Irwin, 1972.

Bresnahan, Timothy, and Alfonso Gambardella, eds. *Building high-tech clusters: Silicon Valley and beyond*. Cambridge University Press, 2004.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh." *Econometrica,* 82.5 (2014): 1671–1748.

Campbell, Benjamin A., Russell Coff, and David Kryscynski. "Rethinking sustained competitive advantage from human capital." *Academy of Management Review* 37.3 (2012): 376–395.

Carnahan, Seth, Rajshree Agarwal, and Benjamin A. Campbell. "Heterogeneity in Turnover: The Effect of Relative Compensation Dispersion of Firms on the Mobility and Entrepreneurship of Extreme Performers." *Strategic Management Journal,* 33.12 (2012): 1411–30.

Carnahan, Seth, David Kryscynski, and Daniel Olson. "When Does Corporate Social Responsibility Reduce Worker Turnover? Evidence from Attorneys Before and After 9/11." *Academy of Management Journal,* 60.5 (2017): 1932–62.

Carnahan, Seth, and Deepak Somaya. "Alumni Effects and Relational Advantage: The Impact on Outsourcing When a Buyer Hires Workers from a Supplier's Competitors." *Academy of Management Journal,* 56.6 (2013): 1578–1600.

Chandler, Alfred Dupont. *Strategy and structure: Chapters in the history of the industrial enterprise*. Cambridge: MIT Press, 1969.

Chattopadhyay, Shinjinee, and Prithwiraj Choudhury. "Sink or swim: The role of workplace context in shaping career advancement and human-capital development." *Organization Science* 28.2 (2017): 211–227.

Chaudhury, Nazmul, Hammer, Jeffrey, Kremer, Michael, Muralidharan, Karthik, & Rogers, F. Halsey. "Missing in action: teacher and health worker absence in developing countries." *Journal of Economic Perspectives* 20.1 (2006): 91–116.

Chiswick, Barry R. "The effect of Americanization on the earnings of foreign-born men." *Journal of Political Economy* 86.5 (1978): 897–921.

Choudhury, Prithwiraj. "Return migration and geography of innovation in MNEs: a natural experiment of knowledge production by local workers reporting to return migrants." *Journal of Economic Geography* 16, no. 3 (2016): 585-610.

Choudhury, Prithwiraj. "Innovation Outcomes in a Distributed Organization: Intrafirm Mobility and Access to Resources." *Organization Science,* 28.2 (2017): 339–54.

Choudhury, Prithwiraj. "Our work-from-anywhere future." *Harvard Business Review* 98, no. 6 (2020): 58-+.

Choudhury, Prithwiraj. "Geographic Mobility, Immobility, and Geographic Flexibility – A Review and Agenda for Research on the Changing Geography of Work." *Academy of Management Annals.* Forthcoming.

Choudhury, Prithwiraj, and Do Yoon Kim. "The Ethnic Migrant Inventor Effect: Codification and Recombination of Knowledge across Borders." *Strategic Management Journal,* 40.2 (2019): 203–29.

Choudhury, Prithwiraj, Cirrus Foroughi, and Barbara Larson. "Work‐from‐anywhere: The productivity effects of geographic flexibility." *Strategic Management Journal* 42, no. 4 (2021): 655-683.

Ciccone, A., & Hall, R. "Productivity and the Density of Economic Activity." *The American Economic Review,* 86.1 (1996): 54–70.

Collis, David J., and Cynthia A. Montgomery. "Corporate strategy: Resources and the scope of the firm." Irwin, 1997.

Collis, David, David Young, and Michael Goold. "The size, structure, and performance of corporate headquarters." *Strategic Management Journal* 28.4 (2007): 383–405.

Conway, Dennis. "Step-wise migration: Toward a clarification of the mechanism." *International Migration Review* 14, no. 1 (1980): 3-14.

Dacin, M. Tina, Marc J. Ventresca, and Brent D. Beal. "The Embeddedness of Organizations: Dialogue & Directions." *Journal of Management,* 25.3 (1999): 317–56.

Dahl, Michael S., and Olav Sorenson. "Home sweet home: Entrepreneurs' location choices and the performance of their ventures." *Management Science* 58, no. 6 (2012): 1059-1071.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum. "Matching in Cities." National Bureau of Economic Research Working Paper 25227, 2018.

Deshpandé, Rohit, and Anjali Raina. "The ordinary heroes of the Taj." *Harvard Business Review* 89.12 (2011): 119–123.

Eeckhout, Jan. "Gibrat's law for (all) cities." *American Economic Review* 94, no. 5 (2004): 1429-1451.

Fallick, Bruce, Charles A. Fleischman, and James B. Rebitzer. "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster." *The Review of Economics and Statistics,* 88.3 (2006): 472–81.

Florida, Richard. *Cities and the Creative Class*. London: Routledge, 2005.

Foley, C. Fritz, and William R. Kerr. "Ethnic Innovation and U.S. Multinational Firm Activity." *Management Science,* 59.7 (2013): 1529–44.

Gabaix, Xavier. "Zipf's Law and the Growth of Cities." *American Economic Review* 89, no. 2 (1999): 129-132.

Ghoshal, Sumantra, and Christopher A. Bartlett. "The Multinational Corporation as an Interorganizational Network." *Academy of Management Review,* 15.4 (1990): 603–26.

Ghoshal, Sumantra, and Nitin Nohria. "Internal Differentiation within Multinational Corporations." *Strategic Management Journal,* 10.4 (1989): 323–37.

Glaeser, Edward L. "Learning in cities." *Journal of Urban Economics,* 46.2 (1999): 254–277.

Gomila, Robin. "Logistic or linear? Estimating causal effects of binary outcomes using regression analysis." *Journal of Experimental Psychology: General,* (2020).

Gupta, Anil K., and Vijay Govindarajan. "Knowledge Flows and the Structure of Control Within Multinational Corporations." *Academy of Management Review,* 16.4 (1991): 768–92.

Harris, John R., and Michael P. Todaro. "Migration, unemployment and development: a two-sector analysis." *The American Economic Review,* 60.1 (1970): 126–142.

Hernandez, Exequiel. "Finding a Home Away from Home: Effects of Immigrants on Firms' Foreign Location Choice and Performance." *Administrative Science Quarterly,* 59.1 (2014): 73–108.
Hernandez, Exequiel, and Elena Kulchina. "Immigrants and firm performance: Effects on foreign subsidiaries versus foreign entrepreneurial firms. *Organization Science*, 31.4 (2020): 797–820.
Hoetker, Glenn. "The use of logit and probit models in strategic management research: Critical issues." *Strategic Management Journal,* 28.4 (2007): 331–343.
Jensen, Robert, "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *The Quarterly Journal of Economics*, 127 (2012): 753–792.
Jovanovic, Boyan. "Job Matching and the Theory of Turnover." *Journal of Political Economy,* 87.5 (1979): 972–90.
Karim, Samina, and Charles Williams. "Structural knowledge: How executive experience with structural composition affects intrafirm mobility and unit reconfiguration." *Strategic Management Journal* 33, no. 6 (2012): 681-709.
Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. "Moving to opportunity in Boston: Early results of a randomized mobility experiment." *The Quarterly Journal of Economics* 116.2 (2001): 607–654.
Kerr, Sari Pekkala, and William R. Kerr. *Immigration Policy Levers for US Innovation and Startups*. National Bureau of Economic Research Working Paper 27040, 2020.
Kerr, Sari Pekkala, William R. Kerr, and William F. Lincoln. "Skilled immigration and the employment structures of US firms." *Journal of Labor Economics,* 33.S1 (2015): S147–S186.
Kerr, William R., and Frederic Robert-Nicoud. "Tech Clusters." *Journal of Economic Perspectives*, 34.3 (2020): 50–76.
Khan, Sama. "The other JNNURM: What does it mean for small towns in India." Centre for Policy Research Working Paper, 2014.
King, Andrew A., Brent Goldfarb, and Timothy Simcoe. "Learning from testimony on quantitative research in management." *Academy of Management Review* (2019).
Kline, Patrick, and Enrico Moretti. "Place based policies with unemployment." *American Economic Review,* 103.3 (2013): 238–43.
Kline, Patrick, and Enrico Moretti. "People, places, and public policy: Some simple welfare economics of local economic development programs." *Annual Review of Economics,* 6.1 (2014): 629–662.
Lee, Thomas W., and Terence R. Mitchell. "An Alternative Approach: The Unfolding Model of Voluntary Worker Turnover." *Academy of Management Review,* 19.1 (1994): 51–89.
Lee, Tae Heon, Barry Gerhart, Ingo Weller, and Charlie O. Trevor. "Understanding Voluntary Turnover: Path-Specific Job Satisfaction Effects and the Importance of Unsolicited Job Offers." *Academy of Management Journal,* 51.4 (2008): 651–71.
March, James, and Herbert Simon. *Organizations*. New York: John Wiley & Sons, 1958.
Mariotti, Sergio, Rocco Mosconi, and Lucia Piscitello. "Location and survival of MNEs' subsidiaries: Agglomeration and heterogeneity of firms." *Strategic Management Journal,* 40.13 (2019): 2242–2270.
Ministry of Housing and Urban Affairs, Government of India. *Ease of Living Index 2018*, 2018. Accessed at: https://eol2019.org/Images/ease-of-living-national-report.pdf
Mitchell, W., and A. S. Tsui. "Introduction: Research in Emerging Economy Contexts - Selected Articles from Management and Organization Review and the Strategic Management Society journals (GSJ, SEJ, and SMJ), Virtual Special Issue, Wiley-Blackwell (2012).
Monteiro, L. Felipe. "Selective attention and the initiation of the global knowledge-sourcing process in multinational corporations." *Journal of International Business Studies,* 46.5 (2015): 505–527.
Monteiro, L. Felipe, Niklas Arvidsson, and Julian Birkinshaw. "Knowledge Flows within Multinational Corporations: Explaining Subsidiary Isolation and Its Performance Implications." *Organization Science,* 19.1 (2008): 90–107.
Munshi, Kaivan, and Mark Rosenzweig. "Networks and misallocation: Insurance, migration, and the rural-urban wage gap." *The American Economic Review,* 106.1 (2016): 46–98.
Muralidharan, Karthik, and Venkatesh Sundararaman. "The aggregate effect of school choice: Evidence from a two-stage experiment in India." *The Quarterly Journal of Economics,* 130.3 (2015): 1011–1066.
National Crime Records Bureau. *Crime in India,* 2008. Accessed at: https://ncrb.gov.in/hi/crime-india-year-2008
Nell, Phillip C., and Björn Ambos. "Parenting Advantage in the MNC: An Embeddedness Perspective on the Value Added by Headquarters." *Strategic Management Journal,* 34.9 (2013): 1086–1103.

Parthasarathy, Balaji. "India's Silicon Valley or Silicon Valley's India? Socially embedding the computer software industry in Bangalore." *International Journal of Urban and Regional Research,* 28.3 (2004): 664–685.
Paul, Anju Mary. "Stepwise international migration: A multistage migration pattern for the aspiring migrant." *American Journal of Sociology* 116, no. 6 (2011): 1842-86.
Portes, Alejandro, ed. *The economic sociology of immigration: Essays on networks, ethnicity, and entrepreneurship*. Russell Sage Foundation, 1995.
Ravenstein, Ernest George. "The laws of migration." *Journal of the statistical society of London* 48, no. 2 (1885): 167-235.
Roodman, David, et al. "Fast and wild: Bootstrap inference in Stata using boottest." *The Stata Journal,* 19.1 (2019): 4–60.
Rosenkopf, Lori, and Paul Almeida. "Overcoming local search through alliances and mobility." *Management Science* 49, no. 6 (2003): 751-766.
Roy, Andrew Donald. "Some thoughts on the distribution of earnings." *Oxford economic papers* 3, no. 2 (1951): 135-146.
Saxenian, AnnaLee. (1994). *Regional advantage: Culture and competition in Silicon Valley and Route 128*. Cambridge, Mass.: Harvard University Press.
Saxenian, AnnaLee. "Silicon Valley's new immigrant entrepreneurs." University of California–San Diego Working Paper No. 15, 2000.
Schleicher, David. Stuck! The Law and Economics of Residential Stagnation. *Yale Law Journal*, 127.1 (2017): 78–154.
Shaver, J. Myles. *Headquarters economy: Managers, mobility, and migration*. Oxford University Press, 2018.
Shaver, Myles J., and Fredrick Flyer. "Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States." *Strategic Management Journal,* 21.12 (2000): 1175–1193.
Singh, Namrata. "Search firms give push to small-town hiring." *Times of India*, August 6, 2018.
The United Nations, Department of Economic and Social Affairs, Population Division. The World's Cities in 2018—Data Booklet 2018 (ST/ESA/ SER.A/417).
Ton, Zeynep, and Robert S. Huckman. "Managing the Impact of Worker Turnover on Performance: The Role of Process Conformance." *Organization Science,* 19.1 (2008): 56–68.
Wang, Dan. "Activating Cross-Border Brokerage: Interorganizational Knowledge Transfer through Skilled Return Migration." *Administrative Science Quarterly,* 60.1 (2015): 133–76.
Wheeler, Christopher H. "Local market scale and the pattern of job changes among young men." *Regional Science and Urban Economics,* 38.2 (2008): 101–118.
Yap, Lorene. "Internal Migration and Economic Development in Brazil." *The Quarterly Journal of Economics*, 90.1 (1976): 119–137.
Young, Alwyn. "Inequality, the urban-rural gap, and migration." *The Quarterly Journal of Economics,* 128.4 (2013): 1727–1785.
Zhao, Liqiu, Shouying Liu, and Wei Zhang. "New trends in internal migration in China: Profiles of the New-generation migrants." *China & World Economy,* 26.1 (2018): 18–41.

Figure 1

Conceptual Schema of Main Results and Evidence in Support of Abductive Explanations

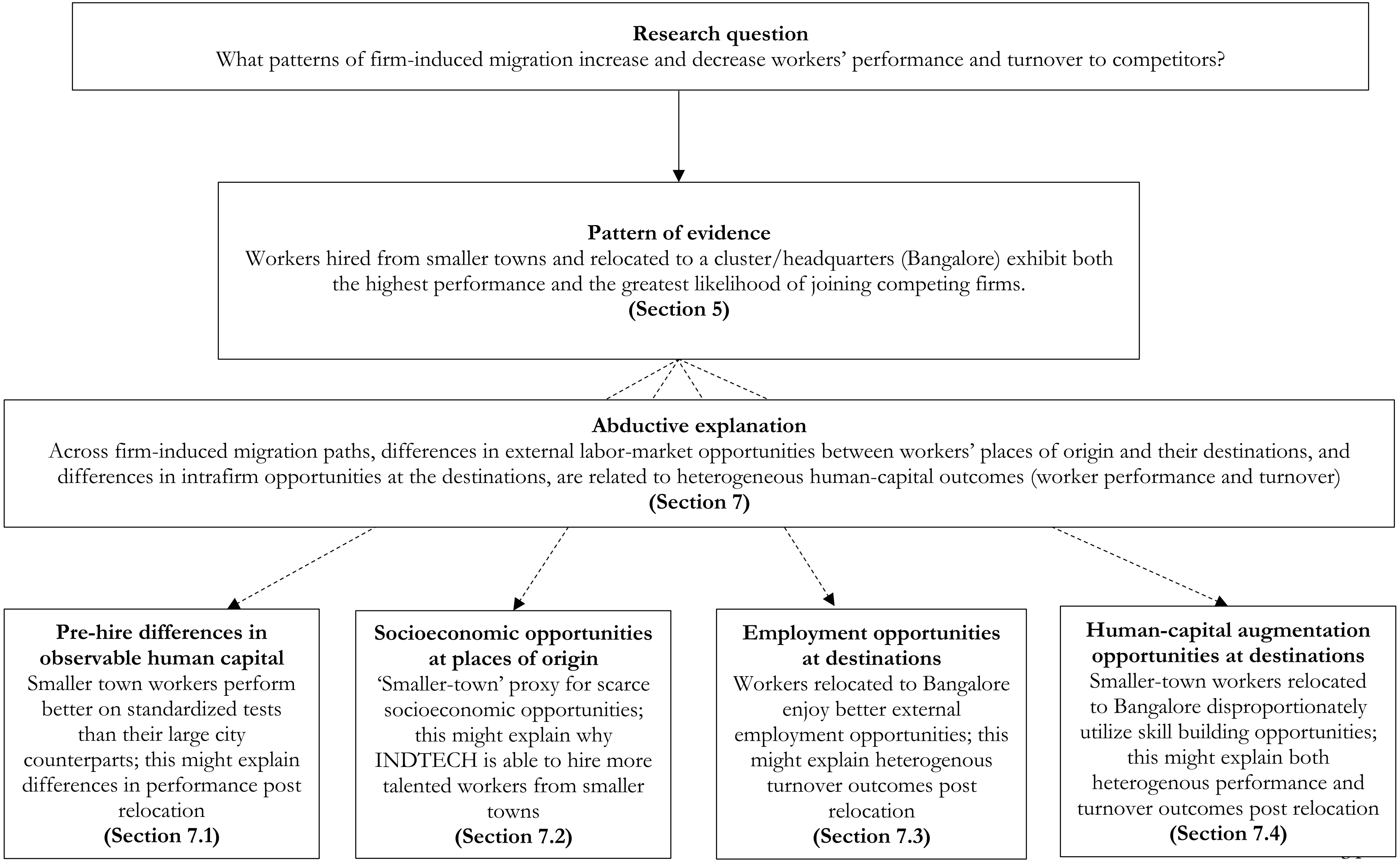
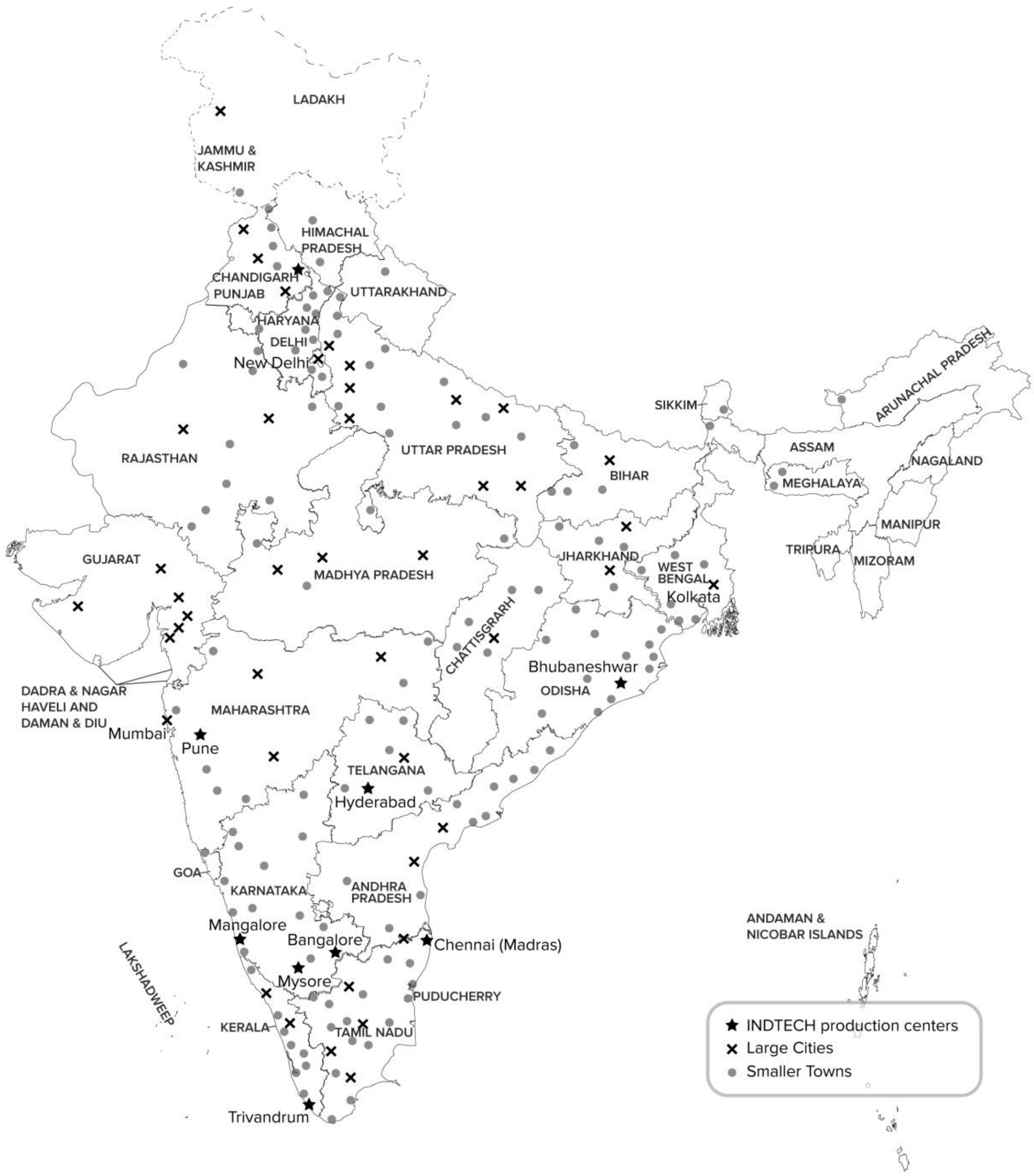


Figure 2

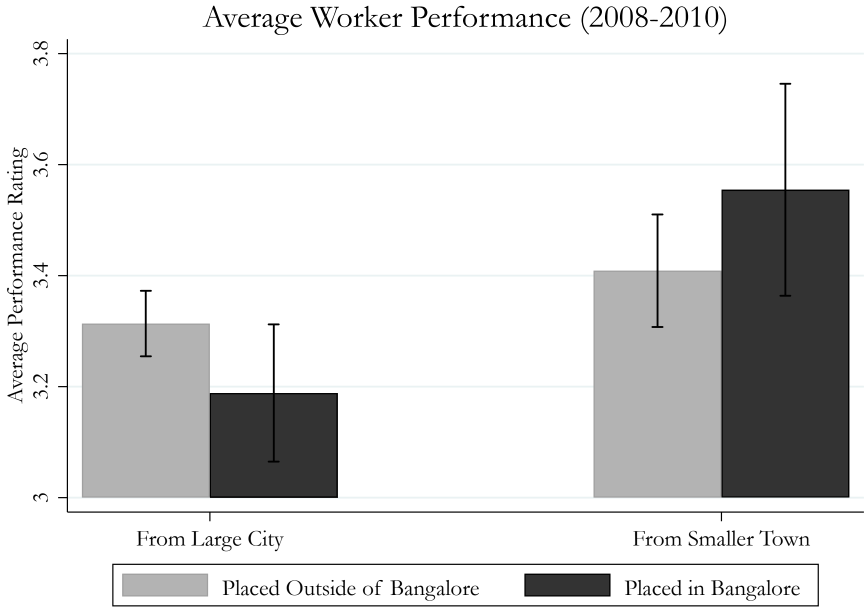
Location of Smaller Towns and Large Cities in Relation to INDTECH Production Centers



NOTES: This figure depicts the geographic distribution of INDTECH production centers, large cities, and smaller towns in India that appear in our sample. Note that prior to 2020, the border between Ladakh and Jammu & Kashmir did not exist; the dotted lines indicate that these borders continue to be disputed.

Figure 3

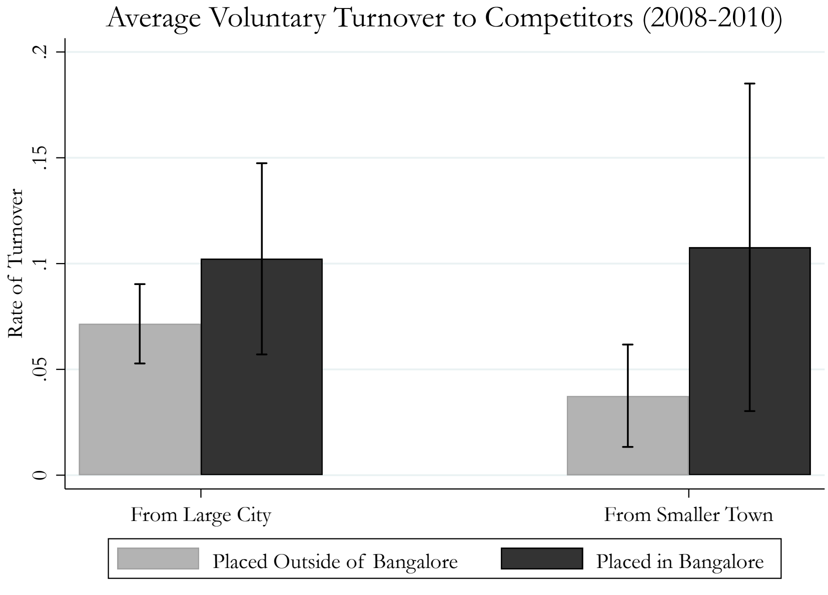
Average Worker Performance by Place of Origin and Assignment to Bangalore, 2008–2010



Note: This figure presents descriptive raw statistics on the average performance rating achieved by workers between 2008 and 2010, by place of origin and production-center placement. Bars indicate averages with 95% confidence intervals.

Figure 4

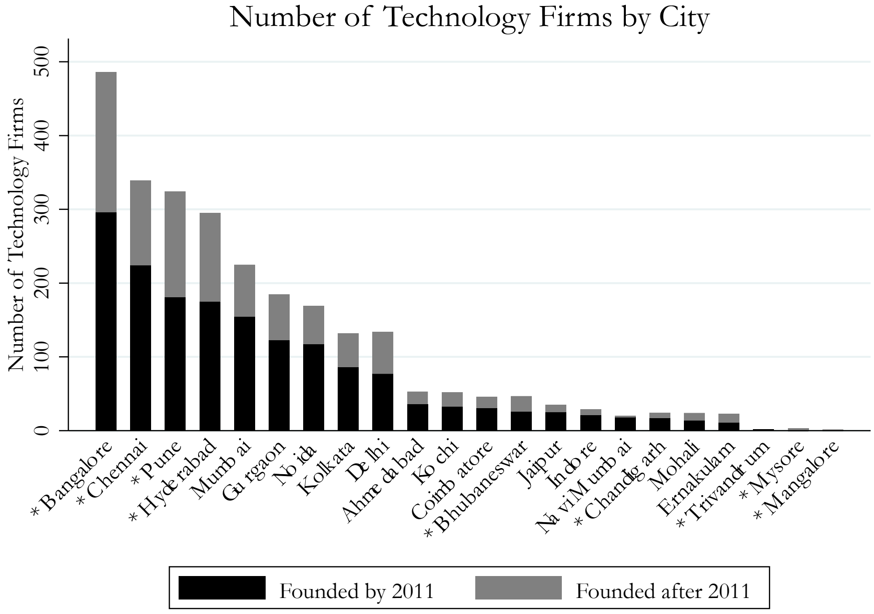
Average Turnover Rates of Workers by Place of Origin and Assignment to Bangalore, 2008–2010



Note: This figure presents descriptive raw statistics on workers' average rate of turnover to competing firms between 2008 and 2010, by place of location of origin and production-center placement. Bars indicate averages with 95% confidence intervals.

Figure 5

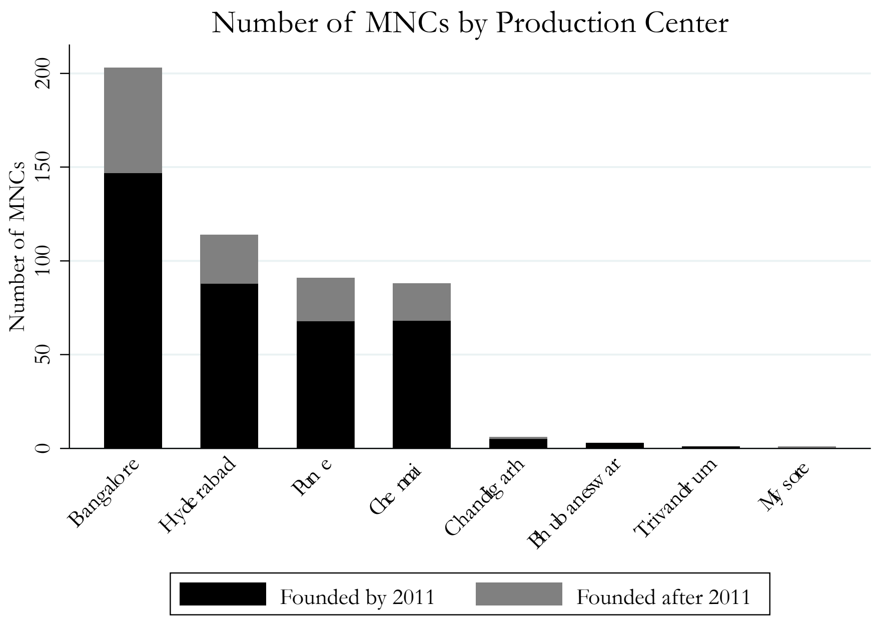
Number of Technology Firms Registered with NASSCOM, by City



Note: This figure represents the number of technology firms registered with NASSCOM, a non-profit trade association of Indian information-technology and business-process outsourcing companies, that were founded and active pre- and post-2011, by city, for the 22 cities in India with the largest number of technology firms. (For more information on NASSCOM, see https://nasscom.in/.) Cities with an asterisk are INDTECH production centers.

Figure 6

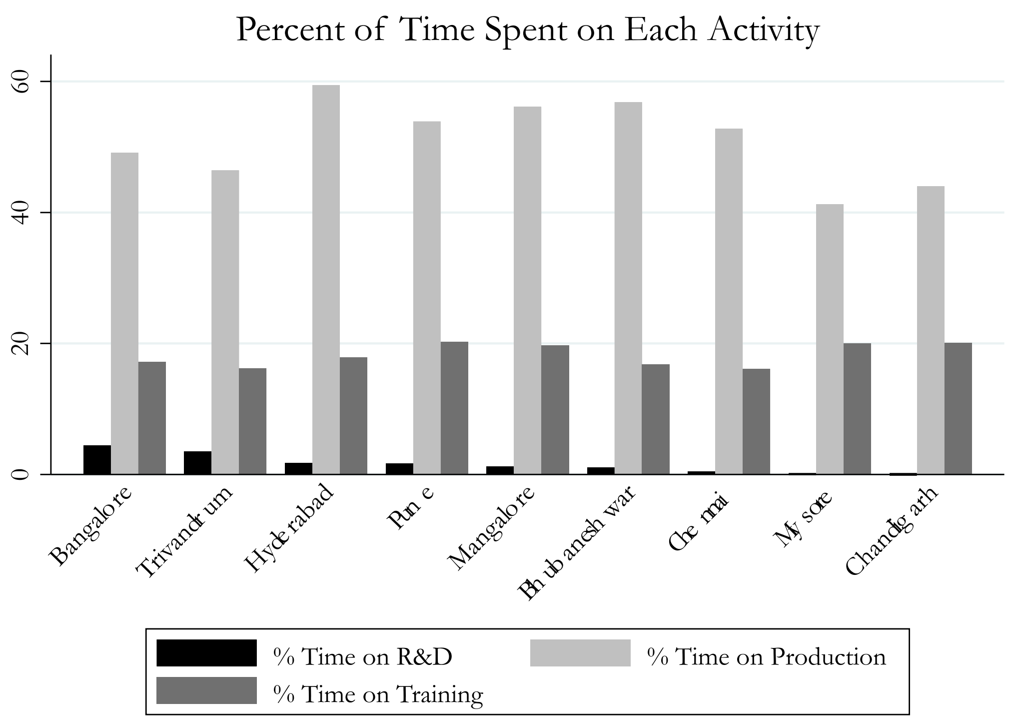
Number of Multinational Corporations Registered with NASSCOM, by INDTECH Production Center



Note: This figure represents the number of multinational corporations (MNCs) (firms with headquarters outside of India) in the technology sector, registered with NASSCOM, that were founded and active pre- and post-2011, by INDTECH production center. NASSCOM is a non-profit trade association of Indian information-technology and business-process outsourcing companies. (For more information, see https://nasscom.in/.)

Figure 7

Workers' Time Allocation by INDTECH Production Center



Note: This figure represents the average percentage of workers' total time dedicated to each of the three activities at each INDTECH production center in the period 2008–2010.

Table 1

City Tiers and Production Centers: Average Population and Distance to Bangalore

Panel A

	# Cities	Fraction of Sample who Attended School in Tier	Fraction of Sample who Attended University in Tier	Population (in thousands)				Driving Distance to Bangalore (in km)				Rail Distance to Bangalore (in km)			
				Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Tier 1	6	0.428	0.366	7,977.436	7,623.323	4,486.679	12,442.370	1,200.433	1,210.800	334.800	2,065.900	1,366.550	1,392.500	354.600	2,404.300
Tier 2	44	0.256	0.312	1,378.460	1,122.464	185.803	5,570.585	1,378.368	1,398.550	142.700	2,901.200	1,660.723	1,739.300	137.400	3,220.100
Tier 3	110	0.317	0.323	353.366	280.318	11.747	1,818.872	1,323.240	1,440.500	68.900	2,972.100	1,491.883	1,532.800	68.900	2,984.900

Panel B

	Production Center	Percent of Sample	Population in 2011	Driving Distance to Bangalore (in km)	Rail Distance to Bangalore (in km)
Tier 1	**Bangalore**	20.48	8,436.675	0.000	0.000
	Chennai	14.59	4,681.087	334.800	354.600
	Hyderabad	13.63	6,809.970	566.200	644.200
Tier 2	**Bhubaneshwar**	5.59	837.737	1,381.100	1,542.000
	Chandigarh	4.80	1,028.667	2,338.100	2,859.300
	Mangalore	2.64	499.486	352.300	363.200
	Mysore	14.89	920.550	142.700	137.400
	Pune	20.72	3,115.431	841.800	965.100
	Trivandrum	2.64	460.468	671.500	836.000

Note: Cities are assigned to tiers in accordance with the 2008 recommendation of the Sixth Central Pay Commission (see http://www.referencer.in/PayCommission/Reports/OM_Allowances.pdf for more details). Panel A includes all cities in our sample; all population and distance statistics are therefore specific to our sample. Panel B presents the raw population and distance statistics for each production center in our sample and the distribution of workers in our sample across those production centers. Note that the distances to Bangalore of cities in Panel A, Tier 1, exclude Bangalore itself, and thus represent the average distance of the five remaining cities.

Table 2

Summary Statistics: Comparison of Workers by Place of Origin and Placement Location

	Summary Statistics for Full Sample					Summary Statistics by Place of Origin			Summary Statistics by Placement Location		
						(1)	(2)	(3)	(4)	(5)	(6)
	Obs.	Mean	Std. dev.	Min	Max	From Smaller Town = 1 (From Smaller Town)	From Smaller Town = 0 (Not from Smaller Town)	Diff.	Placed in Bangalore = 1 (Placed in Bangalore)	Placed in Bangalore = 0 (Placed Elsewhere)	Diff.
Panel A: Employee characteristics											
From smaller town	1254	0.262	0.440	0.000	1.000	1.000	0.000	1.000	0.275	0.258	0.017
Placed in Bangalore	1665	0.205	0.404	0.000	1.000	0.207	0.193	0.014	1.000	0.000	1.000
Placed in hometown	1665	0.276	0.447	0.000	1.000	0.299	0.300	-0.001	0.284	0.274	0.010
Male	1665	0.656	0.475	0.000	1.000	0.643	0.656	-0.012	0.630	0.662	-0.032
Panel B: Recruitment and training scores											
Recruitment test score logical	1605	4.940	3.352	-4.000	9.000	5.531	4.988	0.543***	4.869	4.958	-0.089
Recruitment test score verbal	1605	4.295	3.983	-8.000	16.000	4.075	4.502	-0.426	4.474	4.249	0.225
CGPA training	1665	4.516	0.370	2.800	5.000	4.517	4.510	0.007	4.533	4.512	0.021
Panel C: Performance & turnover											
Performance in 2008	676	2.293	0.544	1.000	3.000	2.413	2.259	0.153***	2.240	2.308	-0.068
Performance in 2009	1283	3.081	0.804	1.000	5.000	3.175	3.029	0.146***	3.044	3.091	-0.046
Performance in 2010	1241	3.182	0.862	1.000	5.000	3.212	3.179	0.033	3.118	3.199	-0.082
Average performance (2008-2010)	1376	3.320	0.734	1.000	5.000	3.444	3.287	0.156***	3.264	3.336	-0.072
Moved to competitor	1665	0.077	0.267	0.000	1.000	0.058	0.078	-0.020	0.091	0.074	0.017

Note: The sample size for workers receiving performance ratings in 2008 is smaller than that for subsequent years because of workers' training schedules. INDTECH requires all workers to spend at least nine months on production prior to receiving a performance rating. However, the required four-month training program meant that workers whose training began after September 2007 were not placed on a project until April 2008, making them ineligible to receive a 2008 performance rating.

*p<0.1; **p<.05; ***p<.01

Table 3

Worker Performance and Turnover to Competitors by Production-Center Location (Bangalore or elsewhere) and Place of Origin

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Average Performance (z-score)	**Average Performance (z-score)**	**Average Performance (z-score)**	**Average Performance (z-score)**	**Average Performance (z-score)**	**Quit to Competitor**	**Quit to Competitor**	**Quit to Competitor**	**Quit to Competitor**	**Quit to Competitor**
VARIABLE	OLS	OLS	OLS with FEs	OLS	OLS	OLS	OLS	OLS with FEs	OLS	OLS
	Full Sample	*Full Sample*	*Full Sample*	*Large-City Sample*	*Smaller-Town Sample*	*Full Sample*	*Full Sample*	*Full Sample*	*Large-City Sample*	*Smaller-Town Sample*
From Smaller Town	0.178	0.116	0.113			-0.028	-0.036*	-0.033*		
	(0.108)	(0.126)	(0.126)			(0.015)	(0.016)	(0.018)		
	[0.136]	[0.398]	[0.413]			[0.094]	[0.053]	[0.088]		
Placed in Bangalore	-0.073	-0.155		-0.158*	0.156*	0.042***	0.032***		0.033***	0.070***
	(0.066)	(0.088)		(0.083)	(0.075)	(0.006)	(0.009)		(0.009)	(0.007)
	[0.351]	[0.104]		[0.082]	[0.022]	[0.000]	[0.000]		[0.000]	[0.000]
From Smaller Town * Bangalore		0.297**	0.301**				0.037**	0.035*		
		(0.125)	(0.126)				(0.014)	(0.016)		
		[0.051]	[0.066]				[0.025]	[0.047]		
Logical Score	-0.004	-0.004	-0.005	-0.014	0.022	0.000	0.000	-0.000	0.001	-0.003
	(0.015)	(0.015)	(0.014)	(0.018)	(0.033)	(0.004)	(0.004)	(0.004)	(0.005)	(0.003)
Verbal Score	0.006	0.006	0.009	0.006	0.005	-0.004**	-0.004**	-0.004**	-0.006**	0.006***
	(0.009)	(0.009)	(0.010)	(0.010)	(0.019)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
CGPA Training	0.914***	0.914***	0.945***	0.781***	1.189***	0.043	0.043	0.045	0.043	0.048
	(0.048)	(0.049)	(0.050)	(0.066)	(0.190)	(0.029)	(0.029)	(0.030)	(0.035)	(0.029)
Male	0.122	0.123	0.139*	0.052	0.307**	0.021	0.021	0.020	0.019*	0.020
	(0.073)	(0.073)	(0.068)	(0.078)	(0.119)	(0.015)	(0.015)	(0.016)	(0.008)	(0.040)
Placed in Hometown	-0.023	-0.022	-0.027	0.055	-0.191	-0.032	-0.032	-0.033*	-0.036	-0.024
	(0.077)	(0.077)	(0.068)	(0.065)	(0.189)	(0.017)	(0.017)	(0.017)	(0.023)	(0.020)
Constant	-4.251***	-4.236***	-4.427***	-3.555***	-5.584***	-0.115	-0.113	-0.112	-0.100	-0.191
	(0.287)	(0.280)	(0.278)	(0.371)	(0.916)	(0.143)	(0.142)	(0.147)	(0.159)	(0.149)
Observations	1,001	1,001	1,001	729	272	1,208	1,208	1,208	903	305
R-squared	0.089	0.091	0.110	0.061	0.184	0.018	0.018	0.023	0.019	0.037
Location FE	No	No	Yes	No	No	No	No	Yes	No	No

Note: Standard errors clustered at the production-center level appear in parentheses. Brackets signify p-values estimated using STATA's wild bootstrap-t procedure boottest, appropriate for data with a small number of clusters (see Roodman et al. (2019) for more details).
*p<0.1; **p<.05; ***p<.01

Table 4

Validity of Random Assignment

VARIABLE	(1) **Assigned to Bangalore** OLS	(2) **Assigned to Bangalore** OLS	(3) **Assigned to Bangalore** OLS	(4) **Assigned to Bangalore** OLS	(5) **Assigned to Bangalore** OLS	(6) **Assigned to Bangalore** OLS
From Smaller Town	0.014	0.020	0.023	0.023	0.022	0.022
	(0.023)	(0.023)	(0.024)	(0.025)	(0.025)	(0.025)
	[0.502]	[0.399]	[0.395]	[0.414]	[0.413]	[0.411]
Logical Score		-0.004	-0.005	-0.005	-0.005	-0.006
		(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
		[0.260]	[0.225]	[0.224]	[0.230]	[0.229]
Verbal Score			0.004	0.004	0.004	0.004
			(0.003)	(0.003)	(0.003)	(0.003)
			[0.292]	[0.294]	[0.299]	[0.300]
CGPA Training				0.033	0.034	0.034
				(0.035)	(0.037)	(0.036)
				[0.371]	[0.373]	[0.373]
Male					-0.035	-0.035
					(0.029)	(0.029)
					[0.325]	[0.324]
Placed in Hometown						-0.005
						(0.013)
						[0.625]
Constant	0.193	0.214	0.203	0.055	0.072	0.074
	(0.180)	(0.195)	(0.187)	(0.141)	(0.151)	(0.149)
Observations	1,254	1,208	1,208	1,208	1,208	1,208
R-squared	0.000	0.001	0.003	0.004	0.006	0.006
Location FE	No	No	No	No	No	No

Note: Standard errors clustered at the production-center level appear in parentheses. Brackets signify p-values estimated using STATA's wild bootstrap-t procedure boottest, appropriate for data with a small number of clusters (see Roodman et al. (2019) for more details).
*p<0.1; **p<.05; ***p<.01

Table 5

Replication of Table 3's Results (the placement in Bangalore effect) in smaller clusters (Chennai, Hyderabad, and Pune)

VARIABLE	(1) **Average Performance (z-score)** OLS *Large-City Sample*	(2) **Average Performance (z-score)** OLS *Smaller-Town Sample*	(3) **Average Performance (z-score)** OLS *Large-City Sample*	(4) **Average Performance (z-score)** OLS *Smaller-Town Sample*	(5) **Average Performance (z-score)** OLS *Large-City Sample*	(6) **Average Performance (z-score)** OLS *Smaller-Town Sample*	(7) **Quit to Competitor** OLS *Large-City Sample*	(8) **Quit to Competitor** OLS *Smaller-Town Sample*	(9) **Quit to Competitor** OLS *Large-City Sample*	(10) **Quit to Competitor** OLS *Smaller-Town Sample*	(11) **Quit to Competitor** OLS *Large-City Sample*	(12) **Quit to Competitor** OLS *Smaller-Town Sample*
Placed in Chennai	0.271**	-0.152+					0.005	-0.022+				
	(0.062)	(0.081)					(0.011)	(0.011)				
Placed in Hyderabad			0.204*	0.405***					-0.030**	0.008		
			(0.066)	(0.077)					(0.008)	(0.016)		
Placed in Pune					-0.177*	0.113					-0.014	-0.016
					(0.073)	(0.090)					(0.014)	(0.019)
Logical Score	-0.011	0.011	-0.009	0.009	-0.010	0.012	0.001	-0.004	0.001	-0.004	0.001	-0.004
	(0.018)	(0.041)	(0.018)	(0.041)	(0.018)	(0.040)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)
Verbal Score	0.009	-0.001	0.007	0.001	0.006	-0.001	-0.006*	0.005*	-0.006*	0.006*	-0.006*	0.006*
	(0.010)	(0.019)	(0.007)	(0.019)	(0.008)	(0.019)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CGPA Training	0.804***	1.184***	0.810***	1.196***	0.817***	1.197***	0.043	0.053	0.041	0.054+	0.044	0.054
	(0.061)	(0.213)	(0.056)	(0.212)	(0.055)	(0.208)	(0.035)	(0.030)	(0.036)	(0.029)	(0.036)	(0.030)
Male	0.081	0.264	0.063	0.268+	0.074	0.272	0.018*	0.017	0.018*	0.018	0.018+	0.019
	(0.060)	(0.153)	(0.063)	(0.143)	(0.067)	(0.153)	(0.007)	(0.041)	(0.008)	(0.042)	(0.008)	(0.042)
Placed in Hometown	0.059	-0.238	0.062	-0.236	0.056	-0.237	-0.034	-0.029	-0.034	-0.029	-0.035	-0.029
	(0.069)	(0.203)	(0.070)	(0.202)	(0.072)	(0.201)	(0.024)	(0.021)	(0.023)	(0.021)	(0.023)	(0.021)
Constant	-3.788***	-5.373**	-3.796***	-5.491***	-3.753***	-5.480***	-0.097	-0.182	-0.084	-0.193	-0.100	-0.190
	(0.400)	(1.073)	(0.343)	(1.051)	(0.347)	(1.019)	(0.156)	(0.140)	(0.159)	(0.135)	(0.159)	(0.141)
Observations	744	257	744	257	744	257	918	290	918	290	918	290
R-squared	0.069	0.181	0.064	0.193	0.064	0.180	0.015	0.024	0.017	0.023	0.016	0.024
Location FE	No	No	No	No	No	No	No	No	No	No	No	No

Parentheses contain standard errors clustered at the production center level. Square brackets contain p-values estimated using STATA's wild bootstrap-t procedure boottest appropriate for data with a small number of clusters (see Roodman et al. (2019) for more details).
*p<0.1; **p<.05; ***p<.01

Table 6

Correlation of Smaller-Town Origin with Place-of-Origin Characteristics

		Obs.	Mean	St. Dev.	Min	Max	1	2	3
1	From Smaller Town	1,254	0.262	0.440	0.000	1.000	1.000		
2	Overall Score (continuous), reversed	591	-41.246	8.471	-58.110	-17.000	0.169	1.000	
3	Crime Rate per 100,000 Inhabitants	661	293.242	136.778	98.400	941.400	0.214	0.100	1.000

Table 7

Interpreting 'Smaller town' construct: Place-of-Origin Characteristics (continuous measures) and Worker Performance and Turnover

VARIABLE	(1) **Average Performance (z-score)** OLS w/FE	(2) **Average Performance (z-score)** OLS w/FE	(3) **Quit to Competitor** OLS w/FE	(4) **Quit to Competitor** OLS w/FE	(5) **Average Performance (z-score)** OLS w/FE	(6) **Average Performance (z-score)** OLS w/FE	(7) **Quit to Competitor** OLS w/FE	(8) **Quit to Competitor** OLS w/FE
Overall Score Z-Score (Reversed)	-0.101	-0.082	0.002	-0.007				
	(0.071)	(0.075)	(0.016)	(0.014)				
Overall Score Z-Score * Bangalore		-0.117		0.057**				
		(0.078)		(0.017)				
Crime Rate					-0.001	-0.001	0.000	0.000
					(0.000)	(0.000)	(0.000)	(0.000)
Crime Rate * Bangalore						0.000		0.000
						(0.000)		(0.000)
Logical Score	-0.015	-0.014	-0.000	-0.000	-0.009	-0.009	-0.001	-0.001
	(0.018)	(0.018)	(0.004)	(0.004)	(0.019)	(0.020)	(0.003)	(0.003)
Verbal Score	0.009	0.008	-0.007**	-0.007*	0.006	0.006	-0.006*	-0.006*
	(0.011)	(0.011)	(0.003)	(0.003)	(0.008)	(0.008)	(0.003)	(0.003)
CGPA Training	1.040***	1.037***	0.047	0.050	0.898***	0.898***	0.035	0.035
	(0.089)	(0.087)	(0.029)	(0.031)	(0.126)	(0.126)	(0.031)	(0.031)
Male	0.276**	0.278**	0.024	0.022	0.129*	0.131*	0.000	0.001
	(0.101)	(0.100)	(0.013)	(0.013)	(0.066)	(0.066)	(0.015)	(0.016)
Hometown	0.110	0.113	-0.024	-0.027	0.053	0.050	-0.025	-0.026
	(0.089)	(0.087)	(0.025)	(0.024)	(0.108)	(0.109)	(0.025)	(0.025)
Constant	-4.915***	-4.906***	-0.113	-0.124	-4.063***	-4.057***	-0.077	-0.071
	(0.390)	(0.384)	(0.128)	(0.136)	(0.504)	(0.525)	(0.162)	(0.153)
Observations	470	470	579	579	513	513	648	648
R-squared	0.150	0.152	0.037	0.043	0.115	0.116	0.024	0.025
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered at the production-center level appear in parentheses.
*p<0.1; **p<.05; ***p<.01

Table 8

Interpreting 'Smaller town' construct: Place-of-Origin Characteristics (Top and Bottom 20%) and Worker Performance and Turnover

VARIABLE	(1) **Average Performance (z-score)** OLS w/FE	(2) **Average Performance (z-score)** OLS w/FE	(3) **Quit to Competitor** OLS w/FE	(4) **Quit to Competitor** OLS w/FE	(5) **Average Performance (z-score)** OLS w/FE	(6) **Average Performance (z-score)** OLS w/FE	(7) **Quit to Competitor** OLS w/FE	(8) **Quit to Competitor** OLS w/FE
Overall Score Bottom 20%	-0.242**	-0.278***	0.018	-0.029				
	(0.073)	(0.080)	(0.054)	(0.025)				
	[0.006]	[0.000]	[0.663]	[0.431]				
Overall Score Bottom 20% * Bangalore		0.192*		0.261***				
		(0.087)		(0.035)				
		[0.000]		[0.005]				
Crime Rate Top 20%					-0.355***	-0.391***	0.051	0.028
					(0.076)	(0.075)	(0.030)	(0.027)
					[0.000]	[0.000]	[0.159]	[0.337]
Crime Rate Top 20% * Bangalore						0.304**		0.178***
						(0.105)		(0.024)
						[0.002]		[0.000]
Logical Score	-0.015	-0.016	-0.000	-0.001	-0.007	-0.006	-0.001	-0.001
	(0.017)	(0.017)	(0.004)	(0.004)	(0.018)	(0.019)	(0.003)	(0.003)
Verbal Score	0.011	0.012	-0.007*	-0.006*	0.004	0.005	-0.005*	-0.005*
	(0.012)	(0.012)	(0.003)	(0.003)	(0.008)	(0.009)	(0.003)	(0.003)
CGPA Training	1.035***	1.030***	0.047	0.048	0.909***	0.898***	0.032	0.027
	(0.078)	(0.078)	(0.030)	(0.030)	(0.121)	(0.129)	(0.029)	(0.026)
Male	0.277**	0.278**	0.023	0.024	0.127*	0.131**	0.001	0.002
	(0.098)	(0.098)	(0.014)	(0.013)	(0.053)	(0.053)	(0.017)	(0.017)
Hometown	0.126	0.124	-0.025	-0.028	0.035	0.034	-0.024	-0.024
	(0.085)	(0.086)	(0.024)	(0.023)	(0.098)	(0.098)	(0.022)	(0.023)
Constant	-4.874***	-4.853***	-0.117	-0.120	-4.181***	-4.141***	-0.043	-0.023
	(0.346)	(0.346)	(0.136)	(0.131)	(0.536)	(0.570)	(0.137)	(0.120)
Observations	470	470	579	579	513	513	648	648
R-squared	0.148	0.148	0.038	0.054	0.130	0.131	0.027	0.034
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered at the production-center level appear in parentheses. Brackets signify p-values estimated using STATA's wild bootstrap-t procedure boottest, appropriate for data with a small number of clusters (see Roodman et al. (2019) for more details).
*p<0.1; **p<.05; ***p<.01

Table 9

Nature and Location of Departing Workers' Next Jobs: Bangalore and Elsewhere

Panel A. Placement Location

VARIABLE	(1) **Next Job in Tech**	(2) **Next Job in a Startup**	(3) **Next Job in a Public Firm**	(4) **Next Job in Bangalore**	(5) **Next Job in India excl. Bangalore**	(6) **Next Job Abroad**	(7) **HQ in Bangalore**	(8) **HQ in India excl. Bangalore**	(9) **HQ Abroad (MNC)**
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Placed in Bangalore	0.019	0.062*	0.006	0.422***	-0.399***	-0.023	0.003	-0.124***	0.121**
	(0.058)	(0.036)	(0.056)	(0.053)	(0.055)	(0.047)	(0.041)	(0.042)	(0.054)
Constant	0.630***	0.039**	0.695***	0.133***	0.673***	0.194***	0.133***	0.218***	0.648***
	(0.038)	(0.017)	(0.036)	(0.027)	(0.037)	(0.031)	(0.027)	(0.032)	(0.037)
Observations	282	216	281	282	282	282	282	282	282
R-squared	0.000	0.015	0.000	0.203	0.155	0.001	0.000	0.027	0.017

Panel B. Next Job Location

VARIABLE	(1) **Next Job in Tech**	(2) **Next Job in a Start-up**	(3) **Next Job in a Public Firm**	(4) **HQ in Bangalore**	(5) **HQ in India excl. Bangalore**	(6) **HQ Abroad (MNC)**
	OLS	OLS	OLS	OLS	OLS	OLS
Next Job in Bangalore	0.057*	0.069***	-0.017	0.122***	-0.222***	0.100***
	(0.033)	(0.022)	(0.031)	(0.025)	(0.026)	(0.033)
Constant	0.618***	0.036***	0.742***	0.081***	0.315***	0.604***
	(0.021)	(0.009)	(0.019)	(0.012)	(0.020)	(0.021)
Observations	877	659	875	877	877	877
R-squared	0.003	0.019	0.000	0.032	0.064	0.010

Robust standard errors appear in parentheses. Panel A, presents the results for the subsample of 282 workers whose INDTECH placement location we were able to determine. Panel B, presents the corresponding results for the full sample of 877 workers whose data we were able to collect from LinkedIn.
*p<0.1; **p<.05; ***p<.01

Table 10

Time Spent on R&D by Workers' Production-Center Location (Bangalore or elsewhere) and Place of Origin

VARIABLE	(1) **% Time on R&D** OLS *Large-City Sample*	(2) **% Time on R&D** OLS *Smaller-Town Sample*	(3) **% Time on R&D** OLS *Full Sample*	(4) **% Time on R&D** OLS w/FE *Full Sample*
From Smaller Town			-0.852**	-0.852**
			(0.329)	(0.306)
			[0.005]	[0.008]
Placed in Bangalore	1.431***	5.463***	1.489***	
	(0.404)	(0.535)	(0.415)	
	[0.000]	[0.000]	[0.000]	
From Smaller Town * Bangalore			3.465***	3.449***
			(0.346)	(0.325)
			[0.000]	[0.000]
Logical Score	-0.048	0.296	0.042	0.058
	(0.205)	(0.231)	(0.082)	(0.086)
Verbal Score	-0.038	0.308	0.049	0.047
	(0.135)	(0.205)	(0.141)	(0.141)
CGPA Training	1.479	1.770	1.523	1.467
	(0.965)	(1.542)	(1.054)	(1.092)
Male	-0.870	1.656	-0.184	-0.207
	(0.742)	(1.729)	(0.865)	(0.908)
Placed in Hometown	-0.950	-0.690	-0.992	-1.116
	(0.583)	(0.473)	(0.601)	(0.602)
Constant	-4.152	-11.410	-5.642	-5.159
	(4.008)	(10.007)	(5.348)	(5.488)
Observations	474	181	655	655
R-squared	0.013	0.104	0.024	0.030
Location FE	No	No	No	Yes

Standard errors clustered at the production-center level appear in parentheses. Brackets signify p-values estimated using STATA's wild bootstrap-t procedure boottest, appropriate for data with a small number of clusters (see Roodman et al. (2019) for more details).
*p<0.1; **p<.05; ***p<.01

Table 11

Enrollment in Voluntary Coursework by Workers' Production-Center Location (Bangalore or elsewhere) and Place of Origin

VARIABLE	(1) **Number of Courses** OLS with FEs	(2) **Number of Courses** OLS with FEs	(3) **% Courses Passed** OLS with FEs	(4) **% Courses Passed** OLS with FEs	(5) **Number of English Courses** OLS with FEs	(6) **Number of English Courses** OLS with FEs	(7) **Level of English Courses** OLS with FEs	(8) **Level of English Courses** OLS with FEs
From Smaller Town	0.858**	0.900**	2.103*	1.811	-0.047	-0.067**	0.475**	0.328
	(0.269)	(0.331)	(1.035)	(1.295)	(0.026)	(0.023)	(0.199)	(0.198)
	[0.003]	[0.017]	[0.070]	[0.194]	[0.079]	[0.014]	[0.054]	[0.165]
From Smaller Town * Bangalore		-0.200		1.477		0.101***		0.710***
		(0.316)		(1.458)		(0.020)		(0.171)
		[0.550]		[0.316]		[0.000]		[0.000]
Logical Score	0.043	0.043	-0.258	-0.256	0.001	0.001	-0.035	-0.037
	(0.035)	(0.035)	(0.348)	(0.348)	(0.005)	(0.005)	(0.044)	(0.044)
Verbal Score	-0.076***	-0.076***	0.288**	0.291*	-0.005	-0.005	0.084**	0.082**
	(0.019)	(0.019)	(0.125)	(0.127)	(0.003)	(0.003)	(0.029)	(0.030)
CGPA Training	0.630	0.629	16.664***	16.660***	-0.041	-0.042	-1.026*	-1.070*
	(0.446)	(0.447)	(3.622)	(3.632)	(0.040)	(0.040)	(0.527)	(0.523)
Male	0.239	0.239	-4.343**	-4.335**	0.011	0.012	-0.450	-0.430
	(0.142)	(0.142)	(1.772)	(1.773)	(0.031)	(0.031)	(0.243)	(0.232)
Placed in Hometown	0.108	0.108	-0.902	-0.905	0.006	0.006	-0.254	-0.268
	(0.233)	(0.233)	(1.932)	(1.937)	(0.031)	(0.031)	(0.240)	(0.236)
Constant	0.219	0.225	7.546	7.534	0.354*	0.353*	6.454**	6.655**
	(1.948)	(1.948)	(15.681)	(15.694)	(0.182)	(0.182)	(2.279)	(2.254)
Observations	1,208	1,208	687	687	687	687	96	96
R-squared	0.048	0.049	0.120	0.120	0.019	0.021	0.256	0.266
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

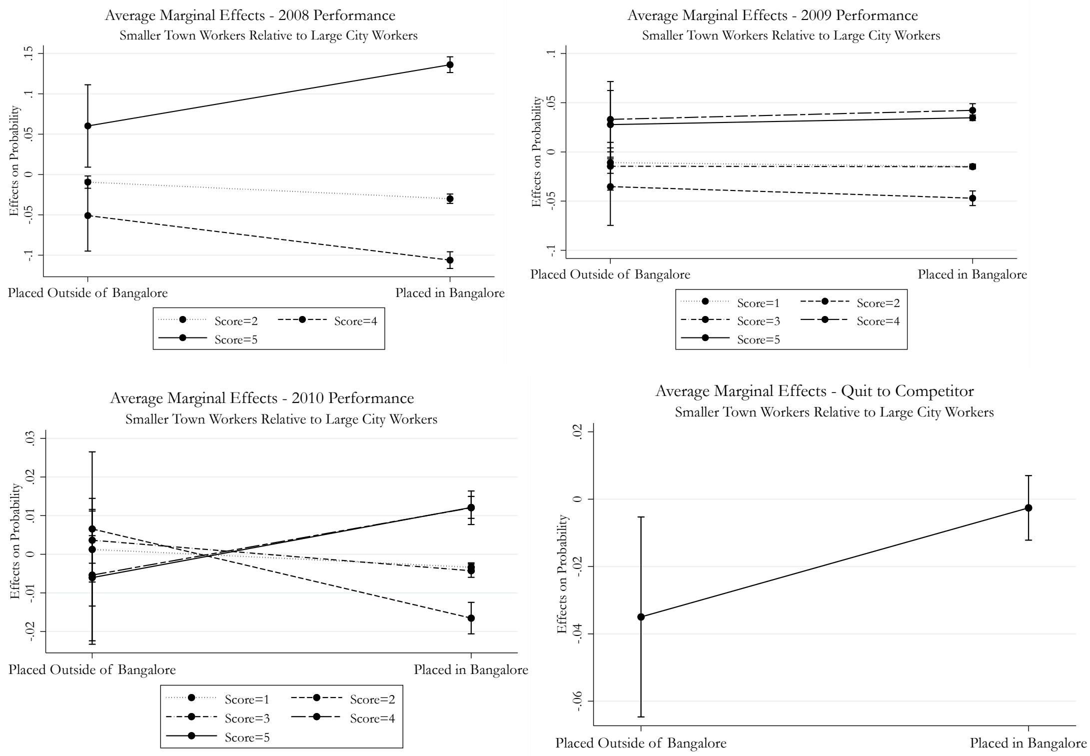
Standard errors clustered at the production-center level appear in parentheses. Brackets signify p-values estimated using STATA's wild bootstrap-t procedure boottest, appropriate for data with a small number of clusters (see Roodman et al. (2019) for more details).
*p<0.1; **p<.05; ***p<.01

APPENDIX

Table of Contents

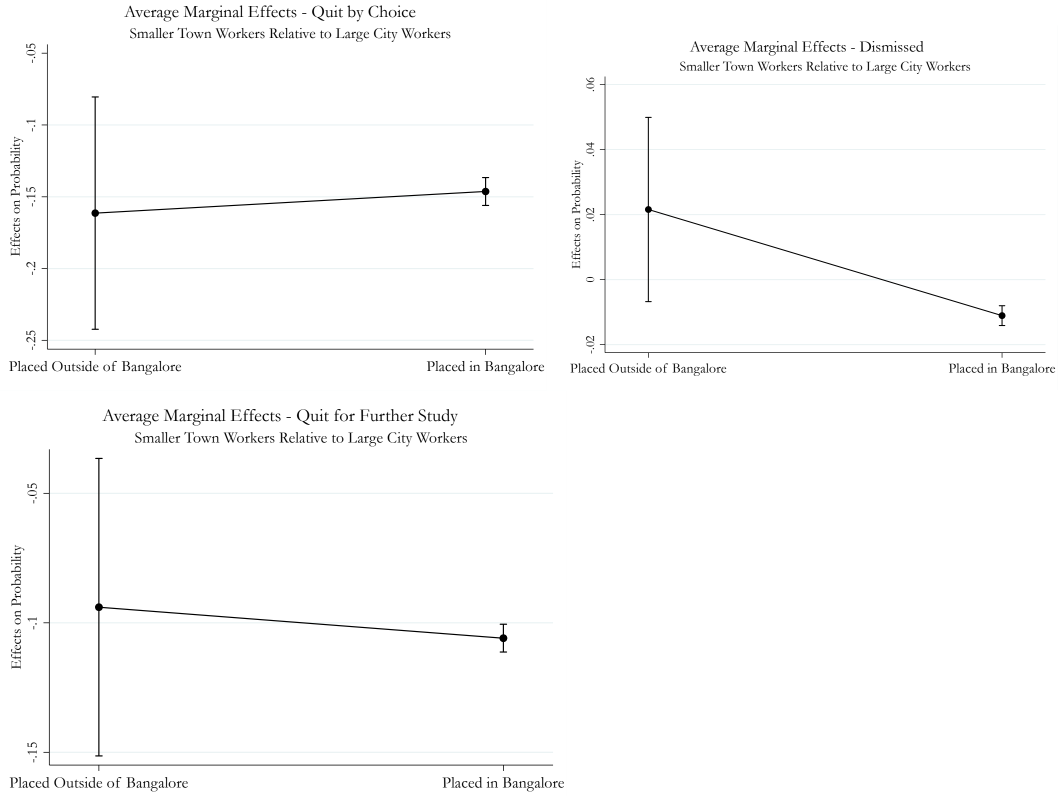
Figures

Figure A1
Marginal Effects Plots for Table A4 – Performance and Turnover



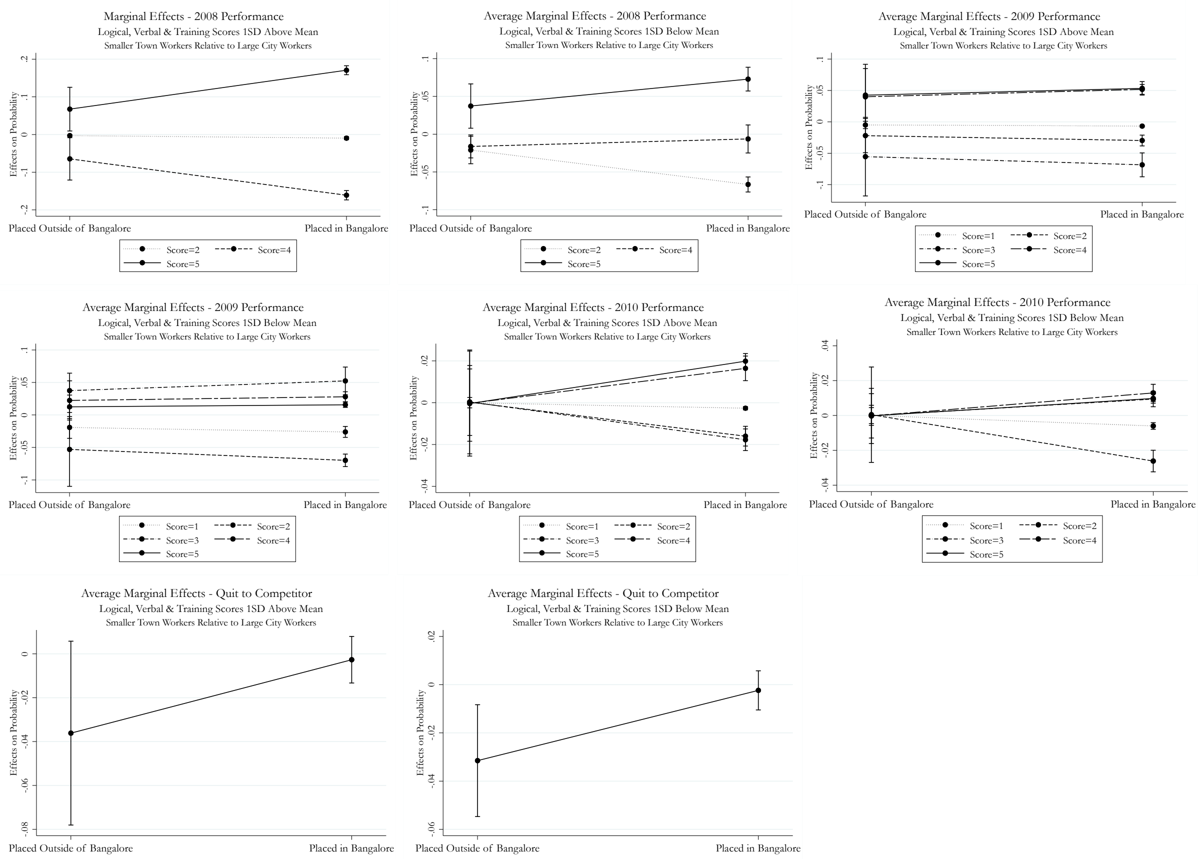
NOTES – Interpreting Marginal Effects*:* Here we plot the average marginal effects for each model interaction in Table A3. Each line represents the change in *relative* probability of a given outcome for workers from smaller towns compared to workers from large cities who are placed either outside of Bangalore or in Bangalore. Outcomes for *Performance* range between 1 and 5, with 5 being the highest. Outcomes for *Quit to Competitor* take the values of 0 or 1. We obtain these marginal effects using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*)" following each regression. Note that the results remain qualitatively the same if we instead use "margins *Bangalore # From Smaller Town*, atmeans". Different Number of Performance Levels in 2008: While outcomes range between 1 and 5, with 5 being the highest, for the 2007 incoming cohort that satisfied the "nine-month rule" and received a performance rating in 2008, no worker received a score of 1 or 3. We therefore show only the changes in probability of receiving each of the three other scores. For the ratings received in 2009 and 2010, all five score levels were given to workers.

Figure A2
Marginal Effects Plots for Table A6 – Attrition Types: Quit by Choice, Quit for Further Study and Dismissed



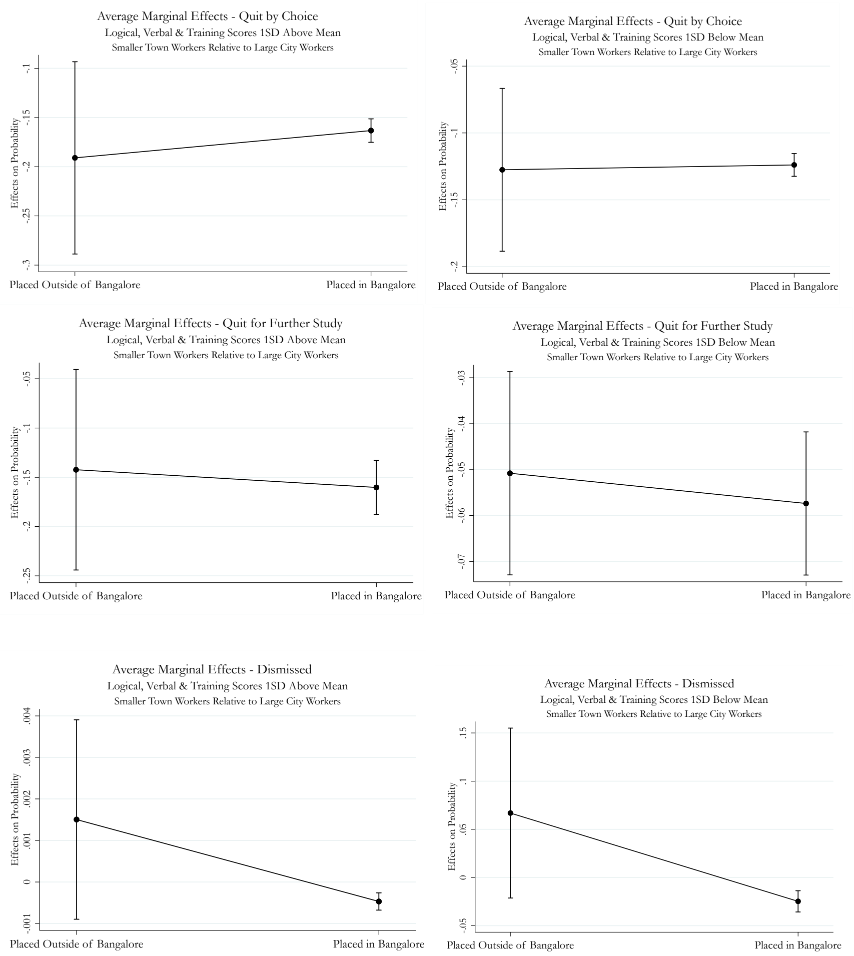
NOTES – <u>Interpreting Marginal Effects</u>*:* Here we plot the average marginal effects for each model interaction in Table A9. Each line represents the change in *relative* probability of a given outcome for workers from smaller towns compared to workers from large cities who are placed either outside of Bangalore or in Bangalore. We obtain these marginal effects using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*)" following each regression. Note that the results remain qualitatively the same if we instead use "margins *Bangalore # From Smaller Town*, atmeans".

Figure A3
Marginal Effects Plots for Table A4 – Performance and Turnover – At Alternative Estimates



NOTES – <u>Interpreting Marginal Effects</u>*:* Here we plot the average marginal effects for each model interaction in Table A3. Each line represents the change in *relative* probability of a given outcome for workers from smaller towns compared to workers from large cities who are placed either outside of Bangalore or in Bangalore. Outcomes for performance range between 1 and 5, with 5 being the highest. However, in contrast to Figure A1, here we condition all estimates at Logical Scores, Verbal Scores and CGPA training at one standard deviation above and below the mean, respectively, keeping gender and hometown location at their mean values (resetting *Male* and *Hometown* to either 0 or 1 does not change the qualitative conclusions of these graphs). We obtain the marginal effects for 1 standard deviation above the mean using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*) at(*Logical Score = 8.292 Verbal Score = 8.279 CGPA = 4.886*" following each regression. We obtain the marginal effects for 1 standard deviation below the mean using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*) at(*Logical Score = 1.588 Verbal Score = 0.312 CGPA = 4.146*" following each regression. The raw means and standard deviations of our control variables are available in Table 2 in the main manuscript. As the results in these figures show, our conclusions remain the same as the marginal effects conditioned at the means in Figure A1. <u>Different Number of Performance Levels in 2008</u>: While outcomes range between 1 and 5, with 5 being the highest, for the 2007 incoming cohort that satisfied the "nine-month rule" and received a performance rating in 2008, no worker received a score of 1 or 3. We therefore show only the changes in probability of receiving each of the three other scores. For the ratings received in 2009 and 2010, all five score levels were given to workers.

Figure A4
Marginal Effects Plots for Table A6 – Attrition Types: Quit by Choice, Quit for Further Study and Dismissed – Alternative Estimates



NOTES – Interpreting Marginal Effects*:* Here we plot the average marginal effects for each model interaction in Table A9. Each line represents the change in *relative* probability of a given outcome for workers from smaller towns compared to workers from large cities who are placed either outside of Bangalore or in Bangalore. We obtain the marginal effects for 1 standard deviation above the mean using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*) at(*Logical Score = 8.292 Verbal Score = 8.279 CGPA = 4.886*" following each regression. We obtain the marginal effects for 1 standard deviation below the mean using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*) at(*Logical Score = 1.588 Verbal Score = 0.312 CGPA = 4.146*" following each regression. The raw means and standard deviations of our control variables are available in Table 2 in the main manuscript. As the results in these figures show, our conclusions remain the same as the marginal effects conditioned at the means in Figure A2.

Tables

Table A1
University Rankings and Location

India Rank 2021	World University Rank 2021	University	City	City Tier	State
1	301–350	Indian Institute of Science	Bangalore	1	Karnataka
2	351–400	Indian Institute of Technology Ropar	Rupnagar	3	Punjab
3	401–500	Indian Institute of Technology Indore	Indore	2	Madhya Pradesh
4	601–800	Banaras Hindu University	Varanasi	2	Uttar Pradesh
4	601–800	Institute of Chemical Technology	Mumbai	1	Maharashtra
4	601–800	University of Delhi	Delhi	1	Delhi
4	601–800	Indian Institute of Science Education and Research, Pune	Pune	2	Maharashtra
4	601–800	Indian Institute of Science Education and Research Kolkata	Kolkata	1	West Bengal
4	601–800	Indian Institute of Technology Gandhinagar	Gandhinagar	2	Gujarat
4	601–800	Indian Institute of Technology Hyderabad	Hyderabad	1	Telangana
4	601–800	Indraprastha Institute of Information Technology Delhi	New Delhi	3	Delhi
4	601–800	Jamia Millia Islamia	Jamia Nagar	3	Delhi
4	601–800	Jawaharlal Nehru University	New Delhi	3	Delhi
4	601–800	King George's Medical University	Lucknow	2	Uttar Pradesh
4	601–800	Mahatma Gandhi University	Kottayam	3	Kerala
4	601–800	Panjab University	Chandigarh	2	Punjab
4	601–800	Savitribai Phule Pune University	Pune	2	Maharashtra
4	601–800	Thapar University	Patiala	2	Punjab
19	801–1000	Aligarh Muslim University	Aligarh	2	Uttar Pradesh
19	801–1000	Amrita Vishwa Vidyapeetham	Coimbatore	2	Tamil Nadu
19	801–1000	Bharathiar University	Coimbatore	2	Tamil Nadu
19	801–1000	Delhi Technological University	New Delhi	3	Delhi
19	801–1000	Guru Jambheshwar University of Science and Technology	Hisar	3	Haryana
19	801–1000	Indian Institute of Technology Bhubaneswar	Bhubaneswar	2	Odisha
19	801–1000	Indian Institute of Technology (Indian School of Mines) Dhanbad	Dhanbad	2	Jharkand
19	801–1000	Indian Institute of Science Education and Research Bhopal	Bhopal	2	Madhya Pradesh
19	801–1000	Jadavpur University	Kolkata	1	West Bengal
19	801–1000	Jamia Hamdard University	New Delhi	3	Delhi
19	801–1000	National Institute of Technology Rourkela	Rourkela	3	Odisha
19	801–1000	National Institute of Technology Silchar	Silchar	3	Assam
19	801–1000	Osmania University	Hyderabad	1	Telangana
19	801–1000	Sri Venkateswara University	Andhra	3	Andhra Pradesh
19	801–1000	VIT University	Vellore	2	Tamil Nadu
34	1001+	Acharya Nagarjuna University	Guntur	2	Andhra Pradesh
34	1001+	Amity University	Noida	3	National Capital Region
34	1001+	Andhra University	Visakhapatnam	2	Andhra Pradesh
34	1001+	Anna University	Chennai	1	Tamil Nadu
34	1001+	Annamalai University	Cuddalore	3	Tamil Nadu

34	1001+	Banasthali University	Vanasthali	3	Rajasthan
34	1001+	Birla Institute of Technology and Science, Pilani	Pilani	3	Rajasthan
34	1001+	University of Calcutta	Kolkata	1	West Bengal
34	1001+	Cochin University of Science and Technology	Cochin	3	Kerala
34	1001+	GITAM University	Visakhapatnam	2	Andhra Pradesh
34	1001+	Indian Institute of Technology Patna	Bihta	3	Bihar
34	1001+	Jawaharlal Nehru Technological University Anantapur (JNTUA)	Ananthapuramu	3	Andhra Pradesh
34	1001+	KIIT University	Bhubaneswar	2	Odisha
34	1001+	KL University	Guntur	2	Andhra Pradesh
34	1001+	University of Lucknow	Lucknow	2	Uttar Pradesh
34	1001+	Maharaja Sayajirao University of Baroda	Vadodara	2	Gujarat
34	1001+	Manipal Academy of Higher Education	Manipal	3	Karnataka
34	1001+	University of Mumbai	Mumbai	1	Maharashtra
34	1001+	University of Mysore	Mysuru	3	Karnataka
34	1001+	National Institute of Technology, Tiruchirappalli	Tiruchirappalli	3	Tamil Nadu
34	1001+	Pondicherry University	Puducherry	3	Puducherry
34	1001+	PSG College of Technology	Coimbatore	2	Tamil Nadu
34	1001+	SASTRA University	Thanjavur	3	Tamil Nadu
34	1001+	Sathyabama Institute of Science and Technology	Chennai	1	Tamil Nadu
34	1001+	Saveetha University	Chennai	1	Tamil Nadu
34	1001+	Siksha 'O' Anusandhan	Bhubaneswar	2	Odisha
34	1001+	SRM Institute of Science and Technology	Kattankulathur	3	Tamil Nadu
34	1001+	Tamil Nadu Agricultural University	Coimbatore	2	Tamil Nadu
34	1001+	Tezpur University	Sonitpur	3	Assam
34	1001+	Visvesvaraya National Institute of Technology, Nagpur	Nagpur	3	Maharashtra

NOTES – While unfortunately no comprehensive ranking is available for all top universities in India at the time of our study, given the slow evolution of university rankings over time (Bejan, 2007), we feel confident relying on a 2021 ranking to examine the geographic distribution of universities. This ranking is provided by the Times Higher Education and is available for the top 63 internationally ranked universities and colleges India. This table reproduces the 2021 India and World University Ranking produced by Times Higher Education. The original data can be obtained at: https://www.timeshighereducation.com/student/best-universities/best-universities-india.

Table A2
Origin Town Size - Continuous Population Measures

VARIABLES	(1) **Average Performance** OLS	(2) **Average Performance** OLS	(3) **Average Performance** OLS w/FE	(4) **Average Performance** OLS	(5) **Average Performance** OLS	(6) **Average Performance** OLS w/FE	(7) **Quit to Competitor** OLS	(8) **Quit to Competitor** OLS	(9) **Quit to Competitor** OLS w/FE	(10) **Quit to Competitor** OLS	(11) **Quit to Competitor** OLS	(12) **Quit to Competitor** OLS w/FE
Placed in Bangalore	-0.043	0.135*		-0.063	0.028		0.045***	0.087***		0.032***	0.051***	
	(0.050)	(0.059)		(0.045)	(0.042)		(0.008)	(0.016)		(0.004)	(0.010)	
Avg Origin Population	-0.023*	-0.013	-0.009				0.002	0.004	0.003			
	(0.011)	(0.010)	(0.011)				(0.003)	(0.003)	(0.003)			
Avg Origin Population * Bangalore		-0.044***	-0.049***					-0.010***	-0.009***			
		(0.011)	(0.012)					(0.002)	(0.003)			
Avg Origin Population (Relaxed)				-0.012	-0.006	-0.004				0.004	0.005	0.004
				(0.008)	(0.010)	(0.010)				(0.002)	(0.003)	(0.003)
Avg Origin Population (Relaxed) * Bangalore					-0.025**	-0.028**					-0.005*	-0.005
					(0.010)	(0.011)					(0.003)	(0.003)
Score Logical	-0.002	-0.003	-0.004	-0.004	-0.004	-0.005	0.001	0.001	0.000	0.001	0.001	0.001
	(0.011)	(0.010)	(0.009)	(0.008)	(0.008)	(0.007)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Score Verbal	0.007	0.008	0.009	0.001	0.001	0.003	-0.004*	-0.004*	-0.004*	-0.005**	-0.005**	-0.005**
	(0.008)	(0.007)	(0.008)	(0.006)	(0.006)	(0.006)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CGPA	0.690***	0.687***	0.713***	0.619***	0.617***	0.637***	0.050	0.049	0.051	0.053	0.053	0.053
	(0.037)	(0.036)	(0.041)	(0.050)	(0.050)	(0.048)	(0.038)	(0.038)	(0.039)	(0.030)	(0.029)	(0.030)
Male	0.116**	0.113**	0.123**	0.096*	0.097*	0.103*	0.015	0.014	0.012	0.026	0.026	0.025
	(0.040)	(0.039)	(0.037)	(0.050)	(0.050)	(0.048)	(0.015)	(0.014)	(0.016)	(0.014)	(0.015)	(0.015)
Constant	0.181	0.158	0.045	0.505*	0.494	0.395	-0.162	-0.166	-0.152	-0.184	-0.187	-0.177
	(0.188)	(0.201)	(0.233)	(0.266)	(0.273)	(0.276)	(0.182)	(0.188)	(0.193)	(0.142)	(0.146)	(0.148)
Observations	842	842	842	1,187	1,187	1,187	1,032	1,032	1,032	1,448	1,448	1,448
R-squared	0.091	0.098	0.118	0.078	0.080	0.098	0.015	0.017	0.024	0.017	0.018	0.019
Location FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

NOTES – Standard errors in parentheses are clustered at the production centre level. This table replicates Table 3 in the main Results section with two alternative measures of small-town worker origin. The first measure, *Avg. Origin Population*, calculates the average population size across the worker's school, high school, and university towns. The second measure, *Avg. Origin Population (Relaxed)*, also calculates the average population size across the worker's school, high school, and university towns, but allows for missing data for one of the worker's origin locations.

*p<0.1; **p<.05; ***p<.01

Table A3
Non-linear Estimation of Performance and Turnover

VARIABLES	(1) **Performance 2008** Ordered Logit	(2) **Performance 2008** Ordered Logit	(3) **Performance 2008** BUC Ordered Logit	(4) **Performance 2009** Ordered Logit	(5) **Performance 2009** Ordered Logit	(6) **Performance 2009** BUC Ordered Logit	(7) **Performance 2010** Ordered Logit	(8) **Performance 2010** Ordered Logit	(9) **Performance 2010** BUC Ordered Logit	(10) **Quit to Competitor** Logit	(11) **Quit to Competitor** Logit	(12) **Quit to Competitor** Conditional Logit
Placed in Bangalore	-0.367***	-0.519***		-0.097	-0.128*		-0.224***	-0.276***		0.576***	0.426***	
	(0.104)	(0.120)		(0.072)	(0.075)		(0.072)	(0.080)		(0.115)	(0.150)	
From Smaller Town	0.382***	0.284**	0.267	0.358**	0.335*	0.287	-0.020	-0.060	-0.101	-0.479*	-0.712**	-0.670**
	(0.123)	(0.126)	(0.180)	(0.159)	(0.201)	(0.237)	(0.080)	(0.092)	(0.075)	(0.275)	(0.307)	(0.341)
From Smaller Town * Bangalore		0.410***	0.481***		0.107	0.063		0.199*	0.232***		0.684**	0.640**
		(0.125)	(0.183)		(0.201)	(0.234)		(0.103)	(0.082)		(0.272)	(0.303)
Received 2008 Rating				0.256***	0.256***	0.226**	0.266***	0.265***	0.276***			
				(0.083)	(0.084)	(0.093)	(0.057)	(0.057)	(0.057)			
Received 2009 Rating							1.392***	1.416***	1.465***			
							(0.314)	(0.324)	(0.276)			
Logical Score	0.083***	0.084***	0.086***	-0.000	-0.000	-0.009	0.002	0.002	-0.004	0.002	0.002	-0.001
	(0.025)	(0.026)	(0.027)	(0.029)	(0.029)	(0.032)	(0.025)	(0.025)	(0.025)	(0.058)	(0.058)	(0.057)
Verbal Score	0.002	0.003	0.001	-0.000	-0.000	0.009	-0.010	-0.010	-0.010	-0.056**	-0.055**	-0.056**
	(0.015)	(0.014)	(0.015)	(0.023)	(0.023)	(0.024)	(0.016)	(0.016)	(0.016)	(0.023)	(0.023)	(0.022)
CGPA Training	2.109***	2.114***	2.122***	1.946***	1.947***	2.115***	1.270***	1.272***	1.347***	0.780	0.779	0.822
	(0.278)	(0.276)	(0.248)	(0.257)	(0.257)	(0.246)	(0.182)	(0.181)	(0.159)	(0.573)	(0.574)	(0.586)
Male	0.219	0.221	0.147	0.173	0.173	0.222	0.462***	0.463***	0.518***	0.338	0.335	0.324
	(0.208)	(0.208)	(0.212)	(0.185)	(0.185)	(0.174)	(0.085)	(0.084)	(0.095)	(0.220)	(0.219)	(0.233)
Placed in Hometown	0.061	0.060	0.064	-0.296**	-0.296**	-0.321**	0.006	0.007	0.000	-0.562	-0.558	-0.582
	(0.176)	(0.175)	(0.178)	(0.142)	(0.142)	(0.160)	(0.108)	(0.109)	(0.107)	(0.358)	(0.357)	(0.355)
Constant										-6.028**	-5.988**	
										(2.819)	(2.797)	
Observations	511	511	1,014	933	933	3,681	903	903	3,563	1,208	1,208	1,177
Location FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

NOTES – Standard errors in parentheses are clustered at the production centre level. Note on the differences in the number of observations for performance regressions: For new hires, workers' training schedule affected whether they satisfied the "nine-month rule" in their first year on the job. For instance, in the 2007 sample, workers who started their training after September 2007 would not finish until early 2008. Most of those workers were not assigned to a project prior to March 2008, making them ineligible to receive a 2008 performance rating. As a result, the number of workers receiving the 2008 performance rating is lower than the number of workers receiving the 2009 performance rating. However, INDTECH's decision to deploy a worker to a project did not depend on superior ability or observable and/or unobservable characteristics.[27] Interpreting Marginal Effects: Figure A1 in the Appendix plots the average marginal effects for each model interaction, obtained using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*)". Note that the results remain qualitatively the same if we instead use "margins *Bangalore* # *From Smaller Town*, atmeans". All results reported here are robust to re-running all models with OLS with and without production center fixed effects.

*p<0.1; **p<.05; ***p<.01

[27] We also empirically validate that controlling for the training batch, logical and verbal scores, gender, and smaller town origin were not significant predictors of receiving a performance rating in 2008; only CGPA training was a positive predictor of receiving a performance rating in 2008, so we include this control in all further specifications.

Table A4
Worker Performance & Turnover to Competitors – Matched Sample Analysis – Coarsened Exact Matching – Automatic Coarsening

	(1)	(2)	(3)	(4)	(5)	(6)
	Average Performance (z-score)	**Average Performance (z-score)**	**Average Performance (z-score)**	**Quit to Competitor**	**Quit to Competitor**	**Quit to Competitor**
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS
	CEM Automatic Bins	CEM Automatic Bins	CEM Automatic Bins	CEM Automatic Bins	CEM Automatic Bins	CEM Automatic Bins
From Smaller Town	0.271	0.236	0.178	-0.030	-0.049***	-0.049***
	(0.148)	(0.185)	(0.178)	(0.020)	(0.010)	(0.010)
Placed in Bangalore	-0.005	-0.067		0.058***	0.027**	
	(0.110)	(0.172)		(0.007)	(0.009)	
From Smaller Town * Bangalore		0.174	0.231		0.090***	0.090***
		(0.194)	(0.186)		(0.011)	(0.011)
Logical Score	0.004	0.003	0.004	-0.001	-0.001	-0.001
	(0.019)	(0.019)	(0.019)	(0.008)	(0.008)	(0.008)
Verbal Score	-0.001	-0.000	0.006	-0.005**	-0.005**	-0.005**
	(0.021)	(0.021)	(0.021)	(0.002)	(0.002)	(0.002)
CGPA Training	0.847***	0.845***	0.910***	0.038	0.037	0.038
	(0.111)	(0.112)	(0.105)	(0.058)	(0.058)	(0.061)
Male	0.218	0.214	0.218	0.039	0.037	0.039
	(0.225)	(0.223)	(0.222)	(0.034)	(0.033)	(0.034)
Placed in Hometown	0.094	0.094	0.067	-0.026	-0.025	-0.028
	(0.108)	(0.108)	(0.101)	(0.036)	(0.035)	(0.035)
Constant	-4.165***	-4.136***	-4.458***	-0.088	-0.075	-0.074
	(0.463)	(0.442)	(0.495)	(0.311)	(0.305)	(0.318)
Observations	569	569	569	651	651	651
R-squared	0.088	0.089	0.117	0.022	0.026	0.031
Location FE	No	No	Yes	No	No	Yes

NOTES – Standard errors in parentheses are clustered at the location level. Coarsened Exact Matching (CEM) matches workers on their pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender. However, the results are also generally robust to matching on Logical and Verbal Scores only. CEM Bins are determined with STATA's in-built automatic identification of appropriate bins using Sturge's Rule (i.e., the number of bins is equal to 1+3.322log(n) where n is the number of observations), and results in a multivariate L1 distance equal to 0. 581 with 136 matched strata comprising of 232 treated and 419 untreated observations.

*p<0.1; **p<.05; ***p<.01

Table A5
Worker Performance & Turnover to Competitors – Matched Sample Analysis – Coarsened Exact Matching – Natural Breakpoints

VARIABLES	(1) **Average Performance (z-score)** OLS CEM Manual Natural Bins	(2) **Average Performance (z-score)** OLS CEM Manual Natural Bins	(3) **Average Performance (z-score)** OLS CEM Manual Natural Bins	(4) **Quit to Competitor** OLS CEM Manual Natural Bins	(5) **Quit to Competitor** OLS CEM Manual Natural Bins	(6) **Quit to Competitor** OLS CEM Manual Natural Bins
From Smaller Town	0.171	0.093	0.094	-0.052*	-0.069**	-0.071**
	(0.139)	(0.148)	(0.152)	(0.024)	(0.021)	(0.023)
Placed in Bangalore	-0.033	-0.173		0.064***	0.034	
	(0.060)	(0.105)		(0.012)	(0.019)	
From Smaller Town * Bangalore		0.381**	0.379**		0.082***	0.084***
		(0.137)	(0.140)		(0.022)	(0.024)
Logical Score	0.002	0.001	0.000	-0.002	-0.002	-0.002
	(0.019)	(0.019)	(0.019)	(0.009)	(0.009)	(0.009)
Verbal Score	0.015	0.015	0.017	-0.004	-0.004	-0.004
	(0.013)	(0.013)	(0.013)	(0.003)	(0.003)	(0.003)
CGPA Training	0.988***	0.986***	1.018***	0.079	0.078	0.067
	(0.114)	(0.116)	(0.113)	(0.055)	(0.055)	(0.059)
Male	0.180*	0.189*	0.208*	0.038	0.040	0.039
	(0.097)	(0.098)	(0.105)	(0.031)	(0.032)	(0.033)
Placed in Hometown	-0.068	-0.058	-0.065	-0.046	-0.044	-0.045
	(0.113)	(0.111)	(0.109)	(0.036)	(0.036)	(0.035)
Constant	-4.683***	-4.639***	-4.848***	-0.254	-0.244	-0.184
	(0.477)	(0.481)	(0.446)	(0.285)	(0.281)	(0.300)
Observations	454	454	454	526	526	526
R-squared	0.115	0.121	0.136	0.041	0.044	0.054
Location FE	No	No	Yes	No	No	Yes

NOTES – Standard errors in parentheses are clustered at the production center level. All models match workers on their pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender. However, the results are also generally robust to matching on Logical and Verbal Scores only. Coarsened Exact Matching (CEM) bins are determined based on natural breakpoints in the data within Logical and Verbal scores. Breakpoints for CGPA Training scores are determined with STATA's in-built Sturge's Rule (i.e., the number of bins is equal to 1+3.322log(n) where n is the number of observations). This approach leads to a multivariate L1 distance of 0.468, with155 matched strata containing 209 treated and 317 untreated observations.

*p<0.1; **p<.05; ***p<.01

Table A6
Worker Performance & Turnover to Competitors – Matched Sample Analysis – Coarsened Exact Matching – Decile Breakpoints

VARIABLES	(1) **Average Performance (z-score)** OLS CEM Manual Decile Bins	(2) **Average Performance (z-score)** OLS CEM Manual Decile Bins	(3) **Average Performance (z-score)** OLS CEM Manual Decile Bins	(4) **Quit to Competitor** OLS CEM Manual Decile Bins	(5) **Quit to Competitor** OLS CEM Manual Decile Bins	(6) **Quit to Competitor** OLS CEM Manual Decile Bins
From Smaller Town	0.138	0.010	0.038	-0.040	-0.054*	-0.058*
	(0.159)	(0.130)	(0.127)	(0.025)	(0.028)	(0.031)
Placed in Bangalore	-0.099	-0.354***		0.075***	0.050*	
	(0.077)	(0.105)		(0.015)	(0.023)	
From Smaller Town * Bangalore		0.618***	0.592***		0.064*	0.068*
		(0.126)	(0.123)		(0.028)	(0.030)
Logical Score	0.010	0.010	0.014	-0.005	-0.005	-0.006
	(0.034)	(0.034)	(0.038)	(0.005)	(0.005)	(0.006)
Verbal Score	-0.014	-0.014	-0.018	-0.001	-0.001	-0.001
	(0.027)	(0.027)	(0.027)	(0.004)	(0.004)	(0.004)
CGPA Training	1.007***	0.982***	0.964***	0.062*	0.062*	0.068*
	(0.100)	(0.110)	(0.117)	(0.029)	(0.028)	(0.030)
Male	0.210*	0.219*	0.235*	0.032	0.033	0.026
	(0.111)	(0.111)	(0.115)	(0.032)	(0.032)	(0.032)
Placed in Hometown	-0.046	-0.047	-0.059	-0.020	-0.020	-0.019
	(0.140)	(0.141)	(0.139)	(0.029)	(0.029)	(0.024)
Constant	-4.730***	-4.560***	-4.585***	-0.192	-0.184	-0.190
	(0.465)	(0.530)	(0.540)	(0.126)	(0.121)	(0.127)
Observations	281	281	281	332	332	332
R-squared	0.129	0.144	0.178	0.040	0.042	0.057
Location FE	No	No	Yes	No	No	Yes

NOTES – Standard errors in parentheses are clustered at the production center level. All models match workers on their pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender. However, the results are also generally robust to matching on Logical and Verbal Scores only. Coarsened Exact Matching (CEM) bins are determined based on deciles within Logical, Verbal and CGPA Training scores. This approach leads to a multivariate L1 distance of 0.358, with 118 matched strata containing 141 treated and 191 untreated observations.

*p<0.1; **p<.05; ***p<.01

Table A7
Additional Measures of Performance and Turnover

VARIABLES	(1) **Dismissed** OLS	(2) **Dismissed** OLS	(3) **Dismissed** OLS w/FE	(4) **Quit for Further Study** OLS	(5) **Quit for Further Study** OLS	(6) **Quit for Further Study** OLS w/FE	(7) **Quit by Choice** OLS	(8) **Quit by Choice** OLS	(9) **Quit by Choice** OLS w/FE
From Smaller Town	0.024	0.029	0.026	-0.094***	-0.091**	-0.080**	-0.157***	-0.160***	-0.146***
	(0.017)	(0.021)	(0.021)	(0.025)	(0.031)	(0.028)	(0.035)	(0.043)	(0.041)
Placed in Bangalore	-0.024***	-0.018***		0.006	0.009		0.061**	0.057*	
	(0.004)	(0.005)		(0.019)	(0.025)		(0.021)	(0.030)	
From Smaller Town * Bangalore		-0.022	-0.019		-0.012	-0.024		0.015	0.001
		(0.020)	(0.021)		(0.030)	(0.026)		(0.040)	(0.039)
Logical Score	-0.003	-0.003	-0.003	-0.002	-0.002	-0.002	-0.003	-0.003	-0.002
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
Verbal Score	0.001	0.001	0.001	0.013***	0.013***	0.012***	0.009***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
CGPA Training	-0.296***	-0.296***	-0.293***	0.123**	0.123**	0.113**	0.195***	0.196***	0.184***
	(0.044)	(0.044)	(0.044)	(0.040)	(0.041)	(0.037)	(0.045)	(0.045)	(0.045)
Male	0.011	0.011	0.012	0.027**	0.027**	0.020	-0.027	-0.027	-0.036
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.011)	(0.026)	(0.026)	(0.028)
Placed in Hometown	-0.007	-0.007	-0.009	0.000	0.000	0.006	-0.025	-0.025	-0.020
	(0.013)	(0.013)	(0.013)	(0.031)	(0.031)	(0.031)	(0.027)	(0.027)	(0.026)
Constant	1.388***	1.387***	1.374***	-0.424**	-0.424**	-0.377**	-0.505**	-0.504**	-0.440*
	(0.201)	(0.202)	(0.203)	(0.165)	(0.165)	(0.161)	(0.185)	(0.184)	(0.201)
Observations	1,208	1,208	1,208	1,208	1,208	1,208	1,208	1,208	1,208
R-squared	0.259	0.260	0.262	0.049	0.049	0.061	0.056	0.056	0.068
Location FE	No	No	Yes	No	No	Yes	No	No	Yes

NOTES – Standard errors in parentheses are clustered at the production center level. The variable *Dismissed* takes the value of 1 if INDTECH dismissed the focal worker from his/her job within the first three years of their employment and 0 otherwise. Dismissals are driven almost entirely by low performance. For instance, no workers receiving at least the middle performance rating were dismissed, but 70% of those receiving the lowest performance rating were dismissed. However, the dismissal outcomes are also available for workers who left INDTECH before they were eligible to receive their first performance rating. Overall, across the full sample of 1,665 workers, 5.3% were dismissed. *Quit by Choice* takes the value of 1 if the worker departed INDTECH by 2011 of their own volition and 0 otherwise. *Quit for Further Study* takes the value of 1 if the worker departed INDTECH by 2011 of their own volition and listed further study as the reason and 0 otherwise. Furthermore, all results are robust to re-estimating the modes with logit.

*p<0.1; **p<.05; ***p<.01

Table A8
Additional Measures of Performance and Turnover – Non-Linear Models

VARIABLES	(1) **Dismissed** Logit	(2) **Dismissed** Logit	(3) **Dismissed** Conditional Logit	(4) **Quit for Further Study** Logit	(5) **Quit for Further Study** Logit	(6) **Quit for Further Study** Conditional Logit	(7) **Quit by Choice** Logit	(8) **Quit by Choice** Logit	(9) **Quit by Choice** Conditional Logit
From Smaller Town	0.520	0.662	0.598	-0.831***	-0.809***	-0.722***	-0.774***	-0.808***	-0.751***
	(0.396)	(0.405)	(0.421)	(0.211)	(0.261)	(0.232)	(0.168)	(0.209)	(0.200)
Placed in Bangalore	-0.962***	-0.528**		0.030	0.047		0.276***	0.243*	
	(0.202)	(0.232)		(0.135)	(0.153)		(0.096)	(0.127)	
From Smaller Town * Bangalore		-1.344***	-1.263***		-0.105	-0.185		0.147	0.092
		(0.442)	(0.445)		(0.235)	(0.208)		(0.190)	(0.183)
Logical Score	-0.072**	-0.073**	-0.088***	-0.010	-0.010	-0.009	-0.011	-0.011	-0.010
	(0.032)	(0.033)	(0.031)	(0.020)	(0.020)	(0.019)	(0.026)	(0.026)	(0.025)
Verbal Score	-0.012	-0.013	-0.012	0.087***	0.087***	0.084***	0.041***	0.041***	0.038***
	(0.041)	(0.041)	(0.042)	(0.014)	(0.014)	(0.013)	(0.013)	(0.012)	(0.012)
CGPA Training	-4.615***	-4.672***	-4.514***	1.098***	1.098***	1.011***	0.999***	0.999***	0.953***
	(0.390)	(0.385)	(0.384)	(0.368)	(0.369)	(0.337)	(0.234)	(0.233)	(0.235)
Male	0.257	0.253	0.239	0.204**	0.204**	0.153	-0.126	-0.127	-0.170
	(0.196)	(0.196)	(0.177)	(0.094)	(0.093)	(0.098)	(0.121)	(0.121)	(0.127)
Placed in Hometown	0.196	0.204	0.124	0.019	0.019	0.065	-0.109	-0.108	-0.086
	(0.322)	(0.319)	(0.323)	(0.236)	(0.236)	(0.239)	(0.126)	(0.125)	(0.126)
Constant	16.675***	16.870***		-6.948***	-6.951***		-5.037***	-5.032***	
	(1.616)	(1.610)		(1.587)	(1.585)		(0.997)	(0.988)	
Observations	1,208	1,208	1,157	1,208	1,208	1,208	1,208	1,208	1,208
Location FE	No	No	Yes	No	No	Yes	No	No	Yes

NOTES – Standard errors in parentheses are clustered at the production center level. Interpreting Marginal Effects*:* Figure A2 in the Appendix plots the average marginal effects for each model interaction, obtained using the STATA command "margins *Bangalore*, dydx(*From Smaller Town*)". Note that the results remain qualitatively the same if we instead use "margins *Bangalore # From Smaller Town*, atmeans". All results reported here are robust to re-running all models with OLS with and without production center fixed effects.
*p<0.1; **p<.05; ***p<.01

Table A9
Robustness Test for Table 3 - Error Clustering Validation

	Average Performance (z-score)						Quit to Competitor					
	OLS		**OLS**		**OLS with FEs**		**OLS**		**OLS**		**OLS with FEs**	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
From Smaller Town	0.178	0.170	0.116	0.420	0.113	0.420	-0.028	0.130	-0.036	0.120	-0.033	0.140
Placed in Bangalore	-0.073	0.210	-0.155	0.140	-0.410	0.210	0.042	0.000	0.032	0.030	0.051	0.040
From Smaller Town * Bangalore			0.297	0.070	0.301	0.060			0.037	0.060	0.035	0.080
Logical Score	-0.004	0.800	-0.004	0.800	-0.005	0.780	0.000	0.980	0.000	0.970	0.000	0.980
Verbal Score	0.006	0.570	0.006	0.550	0.009	0.480	-0.004	0.060	-0.004	0.060	-0.004	0.060
CGPA Training	0.914	0.000	0.914	0.000	0.945	0.000	0.043	0.260	0.043	0.270	0.045	0.260
Male	0.122	0.150	0.123	0.150	0.139	0.070	0.021	0.330	0.021	0.330	0.020	0.350
Placed in Hometown	-0.023	0.790	-0.022	0.810	-0.027	0.730	-0.032	0.130	-0.032	0.130	-0.033	0.120
Constant	-4.251	0.000	-4.236	0.000	-4.140	0.000	-0.115	0.490	-0.113	0.490	-0.137	0.430
Observations	1,001		1,001		1,001		1,208		1,208		1,208	
R-squared	0.078		0.078		0.095		0.254		0.254		0.257	
Location FE	No		No		Yes		No		No		Yes	

NOTES – In addition to the robustness tests listed in the Results section, we specifically address the potential problem of our relatively small number of clusters (the number of INDTECH placement locations) in Table 3 by re-running these results with an OLS model with placement location fixed effects and standard errors estimated using the paired bootstrap-t clustered estimator with 1000 reps. As described in Cameron, Gelbach, and Miller (2008), this estimator is specifically designed to address the small number of clusters problem, by estimating p-values with the pairs cluster bootstrap-t procedure appropriate for data with small number of clusters (see Cameron, Gelbach & Miller (2008) for more details). As the results in this table indicate, our qualitative conclusions in Table 3 continue to hold. As before, the results show that the difference in smaller town workers' performance across placements relative to their large city counterparts is larger and significant at the 10% level. Similarly, smaller town workers continue to be significantly more likely to move to competitors when placed in Bangalore, relative to their counterparts outside of Bangalore and relative to large city workers in Bangalore. Control variables also continue to have the same signs and similar levels of significance.

Table A10
Worker Enrollment in Additional Coursework – Matched Sample Analysis – Coarsened Exact Matching – Automatic Coarsening

VARIABLES	(1) **Number of Courses** OLS CEM Automatic Bins	(2) **Number of Courses** OLS CEM Automatic Bins	(3) **% Courses Passed** OLS CEM Automatic Bins	(4) **% Courses Passed** OLS CEM Automatic Bins	(5) **Number of English Courses** OLS CEM Automatic Bins	(6) **Number of English Courses** OLS CEM Automatic Bins	(7) **Level of English Courses** OLS CEM Automatic Bins	(8) **Level of English Courses** OLS CEM Automatic Bins
From Smaller Town	1.095***	1.162***	1.603	0.971	-0.025	-0.047	0.220	0.284
	(0.253)	(0.303)	(1.357)	(1.642)	(0.035)	(0.030)	(0.274)	(0.344)
Placed in Bangalore	-0.584***	-0.469**	-0.330	-1.829	-0.040	-0.092***	0.709**	1.013***
	(0.136)	(0.159)	(2.483)	(3.405)	(0.022)	(0.024)	(0.273)	(0.261)
From Smaller Town * Bangalore		-0.332		3.683		0.127***		-0.461
		(0.302)		(2.462)		(0.025)		(0.530)
Logical Score	-0.005	-0.004	-0.557	-0.569*	0.007	0.006	0.047	0.056
	(0.065)	(0.066)	(0.303)	(0.300)	(0.007)	(0.007)	(0.031)	(0.038)
Verbal Score	-0.060	-0.061	0.528*	0.549*	-0.010	-0.009	0.081	0.084
	(0.039)	(0.039)	(0.275)	(0.277)	(0.007)	(0.008)	(0.069)	(0.070)
CGPA Training	-0.875	-0.871	16.386*	16.285*	0.069	0.066	0.429**	0.471**
	(0.742)	(0.738)	(7.357)	(7.435)	(0.061)	(0.061)	(0.167)	(0.151)
Male	0.251	0.259	-6.027*	-5.984*	0.010	0.011	-0.416	-0.436
	(0.229)	(0.228)	(2.765)	(2.765)	(0.044)	(0.045)	(0.373)	(0.385)
Placed in Hometown	0.205	0.203	-0.445	-0.377	0.012	0.015	0.047	0.065
	(0.354)	(0.351)	(2.816)	(2.850)	(0.042)	(0.041)	(0.412)	(0.400)
Constant	7.401*	7.352*	10.597	11.273	-0.176	-0.153	-0.826	-1.098
	(3.244)	(3.210)	(33.842)	(34.234)	(0.286)	(0.280)	(1.194)	(1.190)
Observations	651	651	410	410	410	410	55	55
R-squared	0.054	0.054	0.093	0.094	0.019	0.024	0.203	0.206
Location FE	No	No	No	No	No	No	No	No

NOTES – Standard errors in parentheses are clustered at the production center level. All models match workers on their pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender. However, the results are also generally robust to matching on Logical and Verbal Scores only. Coarsened Exact Matching (CEM) bins are calculated using STATA's in-built identification of appropriate bins using Sturge's Rule (i.e., the number of bins is equal to 1+3.322log(n) where n is the number of observations). These modes result in a multivariate L1 distance equal to 0. 581 with 136 matched strata comprising of 232 treated and 419 untreated observations.

*p<0.1; **p<.05; ***p<.01

Table A11
Worker Enrollment in Additional Coursework – Matched Sample Analysis – Coarsened Exact Matching – Natural Breakpoints

VARIABLES	(1) **Number of Courses** OLS CEM Manual Natural Bins	(2) **Number of Courses** OLS CEM Manual Natural Bins	(3) **% Courses Passed** OLS CEM Manual Natural Bins	(4) **% Courses Passed** OLS CEM Manual Natural Bins	(5) **Number of English Courses** OLS CEM Manual Natural Bins	(6) **Number of English Courses** OLS CEM Manual Natural Bins	(7) **Level of English Courses** OLS CEM Manual Natural Bins	(8) **Level of English Courses** OLS CEM Manual Natural Bins
From Smaller Town	0.921*	1.081*	1.324	0.190	-0.048	-0.087*	0.633	0.584
	(0.416)	(0.479)	(1.885)	(1.885)	(0.052)	(0.044)	(0.418)	(0.400)
Placed in Bangalore	-0.950***	-0.673*	-0.182	-2.927	-0.050	-0.144**	0.280	-0.853
	(0.207)	(0.294)	(2.154)	(3.184)	(0.033)	(0.045)	(0.589)	(0.862)
From Smaller Town * Bangalore		-0.767		6.841**		0.235***		1.252**
		(0.489)		(2.740)		(0.033)		(0.491)
Logical Score	0.020	0.022	-1.150*	-1.182**	0.005	0.004	0.035	0.033
	(0.067)	(0.068)	(0.516)	(0.507)	(0.006)	(0.006)	(0.036)	(0.035)
Verbal Score	-0.071	-0.072	0.681***	0.716***	-0.007	-0.005	0.073	0.081
	(0.051)	(0.051)	(0.173)	(0.158)	(0.007)	(0.008)	(0.062)	(0.063)
CGPA Training	-0.410	-0.402	15.761*	15.685*	0.136	0.133	-0.023	-0.151
	(0.648)	(0.642)	(7.161)	(7.235)	(0.077)	(0.081)	(0.498)	(0.578)
Male	0.151	0.137	-4.674	-4.423	-0.066*	-0.058*	-0.684**	-0.706*
	(0.233)	(0.232)	(3.332)	(3.370)	(0.029)	(0.026)	(0.282)	(0.307)
Placed in Hometown	0.054	0.038	-2.311	-2.196	0.017	0.021	-0.505	-0.545
	(0.196)	(0.194)	(4.339)	(4.399)	(0.049)	(0.049)	(0.387)	(0.436)
Constant	5.413*	5.320	15.237	15.953	-0.425	-0.401	1.602	2.224
	(2.908)	(2.872)	(31.197)	(31.544)	(0.296)	(0.306)	(2.319)	(2.744)
Observations	526	526	318	318	318	318	39	39
R-squared	0.051	0.053	0.097	0.101	0.040	0.057	0.257	0.265
Location FE	No	No	No	No	No	No	No	No

NOTES – Standard errors in parentheses are clustered at the production center level. All models match workers on their pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender. However, the results are also generally robust to matching on Logical and Verbal Scores only. Coarsened Exact Matching (CEM) bins are determined based on natural breakpoints in the data within Logical and Verbal scores. Breakpoints for CGPA Training scores are determined with STATA's in-built Sturge's Rule (i.e., the number of bins is equal to 1+3.322log(n) where n is the number of observations). This approach leads to a multivariate L1 distance of 0. 468, but is also the most conservative of the three, with only 155 matched strata containing 209 treated and 317 untreated observations.

*p<0.1; **p<.05; ***p<.01

Table A12
Worker Enrollment in Additional Coursework – Matched Sample Analysis – Coarsened Exact Matching – Decile Breakpoints

	(1)	(2)	(3)	(4)	(5)	(6)
	Number of Courses	**Number of Courses**	**% Courses Passed**	**% Courses Passed**	**Number of English Courses**	**Number of English Courses**
	OLS	OLS	OLS	OLS	OLS	OLS
VARIABLES	CEM Manual Decile Bins	CEM Manual Decile Bins	CEM Manual Decile Bins	CEM Manual Decile Bins	CEM Manual Decile Bins	CEM Manual Decile Bins
From Smaller Town	1.140*	1.231*	1.224	-0.585	-0.039	-0.079*
	(0.509)	(0.619)	(2.584)	(2.104)	(0.046)	(0.035)
Placed in Bangalore	-1.064***	-0.895**	1.570	-3.824	-0.025	-0.145**
	(0.258)	(0.374)	(2.026)	(2.906)	(0.034)	(0.044)
From Smaller Town * Bangalore		-0.434		11.183***		0.250***
		(0.579)		(2.480)		(0.029)
Logical Score	0.181*	0.182*	-0.856	-0.872	-0.006	-0.006
	(0.086)	(0.086)	(0.476)	(0.472)	(0.012)	(0.011)
Verbal Score	-0.161**	-0.162**	0.425	0.499	0.005	0.006
	(0.066)	(0.066)	(0.285)	(0.277)	(0.009)	(0.009)
CGPA Training	0.700	0.704	17.104**	16.289**	0.180**	0.162**
	(0.719)	(0.712)	(5.250)	(5.026)	(0.054)	(0.052)
Male	-0.435	-0.438	-2.779	-2.647	-0.092*	-0.089*
	(0.336)	(0.338)	(4.518)	(4.469)	(0.048)	(0.047)
Placed in Hometown	-0.535*	-0.533*	-2.816	-2.978	0.046	0.043
	(0.257)	(0.256)	(4.564)	(4.658)	(0.050)	(0.049)
Constant	0.143	0.091	8.378	12.802	-0.604**	-0.505*
	(2.942)	(2.865)	(22.114)	(20.943)	(0.240)	(0.222)
Observations	332	332	200	200	200	200
R-squared	0.119	0.120	0.099	0.110	0.050	0.069
Location FE	No	No	No	No	No	No

NOTES – Standard errors in parentheses are clustered at the production center level. All models match workers on their pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender. However, the results are also generally robust to matching on Logical and Verbal Scores only. Coarsened Exact Matching (CEM) bins are determined based on deciles within Logical, Verbal and CGPA Training scores. This approach leads to the lowest multivariate L1 distance of 0. 358, but is also the most conservative of the three, with only 118 matched strata containing 141 treated and 191 untreated observations.

*p<0.1; **p<.05; ***p<.01

Table A13
2008 Crime Rates, National Crime Records Bureau

	City	Incidence of total cognizable crimes	Population (in 100,000's) as per 2011 census	Rate of total cognizable crimes per 100,000 inhabitants
1	AGRA	4826	13.21	365.3
2	AHMEDABAD	18544	45.19	410.4
3	ALLAHABAD	2068	10.5	197
4	AMRITSAR	2327	10.11	230.2
5	ASANSOL	1676	10.91	153.6
6	BENGALURU	29664	56.87	521.6
7	BHOPAL	11515	14.55	791.4
8	CHENNAI	11829	64.25	184.1
9	COIMBATORE	4180	14.46	289.1
10	DELHI (CITY)	44573	127.91	348.5
11	DHANBAD	1302	10.64	122.4
12	FARIDABAD	4516	10.55	428.1
13	HYDERABAD	18567	55.34	335.5
14	INDORE	15430	16.39	941.4
15	JABALPUR	5128	11.17	459.1
16	JAIPUR	15407	23.24	663
17	JAMSHEDPUR	2685	11.02	243.6
18	KANPUR	8885	26.9	330.3
19	KOCHI	7956	13.55	587.2
20	KOLKATA	13005	132.17	98.4
21	LUCKNOW	11735	22.67	517.6
22	LUDHIANA	2847	13.95	204.1
23	MADURAI	2470	11.95	206.7
24	MEERUT	2765	11.67	236.9
25	MUMBAI	32770	163.68	200.2
26	NAGPUR	8661	21.23	408
27	NASIK	3813	11.52	331
28	PATNA	9014	17.07	528.1
29	PUNE	14467	37.56	385.2
30	RAJKOT	5525	10.02	551.4
31	SURAT	10741	28.11	382.1
32	VADODARA	5386	14.92	361
33	VARANASI	2734	12.12	225.6
34	VIJAYAWADA	5127	10.11	507.1
35	VISHAKHAPATNAM	5015	13.29	377.4
	TOTAL (CITIES)	**347153**	**1078.8**	**321.8**

NOTES – This table reproduces the publicly available 2008 crime rate statistics for 35 cities in India provided by the National Crime Records Bureau, part of the Ministry of Home Affairs. These data are available at https://ncrb.gov.in/en/crime-in-india-table-addtional-table-and-chapter-contents?page=1.

Table A14
Origin Location Characteristics – 30% Cut-off

VARIABLES	(1) **Average Performance (z-score)** OLS w/FE	(2) **Average Performance (z-score)** OLS w/FE	(3) **Quit to Competitor** OLS w/FE	(4) **Quit to Competitor** OLS w/FE	(5) **Average Performance (z-score)** OLS w/FE	(6) **Average Performance (z-score)** OLS w/FE	(7) **Quit to Competitor** OLS w/FE	(8) **Quit to Competitor** OLS w/FE
Overall Score Bottom 30%	-0.293**	-0.300*	0.022	-0.007				
	(0.101)	(0.132)	(0.047)	(0.041)				
Overall Score bottom 30% * Bangalore		0.033		0.147**				
		(0.144)		(0.047)				
Crime Rate Top 30%					-0.007	-0.096	0.046*	0.035
					(0.159)	(0.174)	(0.024)	(0.027)
Crime Rate Top 30% * Bangalore						0.480**		0.066**
						(0.191)		(0.025)
Logical Score	-0.015	-0.015	-0.000	-0.000	-0.009	-0.008	-0.001	-0.001
	(0.018)	(0.018)	(0.004)	(0.004)	(0.019)	(0.020)	(0.003)	(0.003)
Verbal Score	0.010	0.011	-0.007**	-0.007*	0.007	0.008	-0.006*	-0.005*
	(0.011)	(0.012)	(0.003)	(0.003)	(0.008)	(0.009)	(0.003)	(0.003)
CGPA Training	1.022***	1.022***	0.047	0.051	0.895***	0.905***	0.038	0.038
	(0.078)	(0.079)	(0.030)	(0.032)	(0.115)	(0.110)	(0.032)	(0.032)
Male	0.286**	0.286**	0.023	0.022	0.145**	0.158**	0.001	0.003
	(0.099)	(0.100)	(0.014)	(0.013)	(0.058)	(0.060)	(0.017)	(0.017)
Hometown	0.080	0.081	-0.022	-0.023	0.068	0.061	-0.023	-0.024
	(0.091)	(0.091)	(0.025)	(0.025)	(0.100)	(0.101)	(0.023)	(0.023)
Constant	-4.757***	-4.758***	-0.122	-0.136	-4.212***	-4.270***	-0.070	-0.073
	(0.325)	(0.327)	(0.138)	(0.142)	(0.465)	(0.458)	(0.151)	(0.149)
Observations	470	470	579	579	513	513	648	648
R-squared	0.155	0.155	0.038	0.046	0.112	0.118	0.027	0.029
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTES – This table replicates the analyses in Table 9 in the main manuscript but replaces the binary variables capturing the top (bottom) 20% of the distribution for workers' origin town characteristics with an alternative classification based on a cut-off of 30%. Standard errors in parentheses are clustered at the production centre level.

*p<0.1; **p<.05; ***p<.01

Table A15
Origin Location Characteristics – 10% Cut-off

VARIABLES	(1) **Average Performance (z-score)** OLS w/FE	(2) **Average Performance (z-score)** OLS w/FE	(3) **Quit to Competitor** OLS w/FE	(4) **Quit to Competitor** OLS w/FE	(5) **Average Performance (z-score)** OLS w/FE	(6) **Average Performance (z-score)** OLS w/FE	(7) **Quit to Competitor** OLS w/FE	(8) **Quit to Competitor** OLS w/FE
Overall Score Bottom 10%	-0.192**	-0.221**	0.034	-0.021				
	(0.067)	(0.077)	(0.061)	(0.025)				
Overall Score bottom 10% * Bangalore		0.133		0.274***				
		(0.095)		(0.036)				
Crime Rate Top 10%					-0.301***	-0.344**	0.100	0.034
					(0.087)	(0.107)	(0.084)	(0.059)
Crime Rate Top 10% * Bangalore						0.166		0.292***
						(0.105)		(0.065)
Logical Score	-0.015	-0.016	-0.000	-0.001	-0.009	-0.009	-0.001	-0.001
	(0.017)	(0.017)	(0.004)	(0.004)	(0.019)	(0.019)	(0.003)	(0.003)
Verbal Score	0.010	0.011	-0.007**	-0.006*	0.007	0.007	-0.006*	-0.005*
	(0.012)	(0.012)	(0.003)	(0.003)	(0.009)	(0.009)	(0.003)	(0.003)
CGPA Training	1.040***	1.037***	0.048	0.048	0.895***	0.889***	0.038	0.030
	(0.077)	(0.077)	(0.030)	(0.030)	(0.111)	(0.117)	(0.031)	(0.027)
Male	0.268**	0.268**	0.023	0.023	0.152*	0.155*	-0.005	-0.001
	(0.093)	(0.093)	(0.014)	(0.014)	(0.068)	(0.068)	(0.017)	(0.018)
Hometown	0.128	0.128	-0.025	-0.027	0.059	0.059	-0.026	-0.027
	(0.087)	(0.087)	(0.024)	(0.023)	(0.112)	(0.113)	(0.024)	(0.023)
Constant	-4.898***	-4.884***	-0.120	-0.123	-4.194***	-4.168***	-0.061	-0.029
	(0.343)	(0.342)	(0.139)	(0.133)	(0.497)	(0.521)	(0.149)	(0.126)
Observations	470	470	579	579	513	513	648	648
R-squared	0.145	0.145	0.039	0.055	0.117	0.118	0.031	0.044
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTES – Standard errors in parentheses are clustered at the production centre level. This table replicates the analyses in Table 9 in the main manuscript but replaces the binary variables capturing the top (bottom) 20% of the distribution for workers' origin town characteristics with an alternative classification based on a cut-off of 10%.

*p<0.1; **p<.05; ***p<.01

Table A16
Origin Location Characteristics – All Sub-components of Overall Score

VARIABLES	(1) Average Performance (z-score) OLS w/FE	(2) Average Performance (z-score) OLS w/FE	(3) Quit to Competitor OLS w/FE	(4) Quit to Competitor OLS w/FE	(5) Average Performance (z-score) OLS w/FE	(6) Average Performance (z-score) OLS w/FE	(7) Quit to Competitor OLS w/FE	(8) Quit to Competitor OLS w/FE	(9) Average Performance (z-score) OLS w/FE	(10) Average Performance (z-score) OLS w/FE	(11) Quit to Competitor OLS w/FE	(12) Quit to Competitor OLS w/FE	(13) Average Performance (z-score) OLS w/FE	(14) Average Performance (z-score) OLS w/FE	(15) Quit to Competitor OLS w/FE	(16) Quit to Competitor OLS w/FE
Econ. Score Bottom 20%	-0.243*	-0.219	0.004	0.011												
	(0.118)	(0.132)	(0.031)	(0.039)												
Econ. Score Bottom 20% * Bangalore		-0.127		-0.040												
		(0.130)		(0.042)												
Social Score Bottom 20%					-0.328***	-0.387***	0.035	-0.013								
					(0.064)	(0.074)	(0.051)	(0.011)								
Social Score Bottom 20% * Bangalore						0.298**		0.286***								
						(0.097)		(0.022)								
Institutional Score Bottom 20%									-0.104	-0.127	-0.043**	-0.045**				
									(0.157)	(0.177)	(0.014)	(0.017)				
Institutional Score Bottom 20% * Bangalore										0.151		0.020				
										(0.167)		(0.025)				
Physical Score Bottom 20%													-0.195	-0.181	-0.016	-0.011
													(0.167)	(0.204)	(0.032)	(0.042)
Physical Score Bottom 20% * Bangalore														-0.070		-0.027
														(0.201)		(0.046)
Logical Score	-0.017	-0.016	-0.000	0.000	-0.013	-0.014	-0.000	-0.001	-0.016	-0.016	0.000	0.000	-0.016	-0.016	-0.000	-0.000
	(0.018)	(0.018)	(0.004)	(0.004)	(0.017)	(0.017)	(0.004)	(0.004)	(0.017)	(0.017)	(0.004)	(0.004)	(0.018)	(0.018)	(0.004)	(0.004)
Verbal Score	0.011	0.011	-0.007*	-0.007*	0.010	0.011	-0.007**	-0.006*	0.010	0.010	-0.007*	-0.007*	0.011	0.011	-0.007*	-0.007*
	(0.012)	(0.012)	(0.003)	(0.003)	(0.012)	(0.012)	(0.003)	(0.003)	(0.011)	(0.011)	(0.003)	(0.003)	(0.012)	(0.012)	(0.003)	(0.003)
CGPA Training	1.024***	1.021***	0.047	0.045	1.027***	1.019***	0.048	0.044	1.050***	1.049***	0.047	0.047	1.031***	1.030***	0.046	0.045
	(0.079)	(0.081)	(0.028)	(0.028)	(0.084)	(0.085)	(0.029)	(0.026)	(0.079)	(0.079)	(0.028)	(0.029)	(0.073)	(0.075)	(0.027)	(0.027)
Male	0.268**	0.269**	0.024	0.024	0.281**	0.282**	0.022	0.022	0.257**	0.257**	0.024*	0.024*	0.267**	0.267**	0.024	0.024
	(0.094)	(0.094)	(0.014)	(0.015)	(0.092)	(0.091)	(0.014)	(0.014)	(0.095)	(0.096)	(0.013)	(0.013)	(0.092)	(0.093)	(0.014)	(0.014)
Hometown	0.092	0.092	-0.024	-0.024	0.110	0.106	-0.023	-0.026	0.127	0.127	-0.025	-0.025	0.106	0.106	-0.026	-0.026
	(0.095)	(0.095)	(0.025)	(0.025)	(0.079)	(0.082)	(0.024)	(0.024)	(0.087)	(0.087)	(0.024)	(0.024)	(0.093)	(0.093)	(0.023)	(0.024)
Constant	-4.788***	-4.775***	-0.114	-0.108	-4.832***	-4.795***	-0.120	-0.102	-4.939***	-4.929***	-0.107	-0.107	-4.835***	-4.832***	-0.106	-0.103
	(0.340)	(0.349)	(0.123)	(0.121)	(0.372)	(0.373)	(0.134)	(0.115)	(0.348)	(0.349)	(0.126)	(0.125)	(0.305)	(0.311)	(0.119)	(0.119)
Observations	470	470	579	579	470	470	579	579	470	470	579	579	470	470	579	579
R-squared	0.149	0.149	0.037	0.038	0.153	0.154	0.039	0.056	0.143	0.143	0.041	0.041	0.146	0.146	0.038	0.038
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTES – This table replicates the analyses in Columns (1)-(4) in Table 9 in the main manuscript by disaggregating the *Overall Score* capturing the ease-of-living in each worker's origin location with its four constituent pillars: *Economy*, *Social*, *Institutional* and *Physical Scores*. For more information on the methodology behind the construction. Standard errors in parentheses are clustered at the production centre level.

*p<0.1; **p<.05; ***p<.01

Table A17
Time Spent on R&D – Matched Sample Analysis – Coarsened Exact Matching

VARIABLES	(1) **% Time on R&D** OLS CEM Automatic Bins	(2) **% Time on R&D** OLS CEM Automatic Bins	(3) **% Time on R&D** OLS CEM Manual Natural Bins	(4) **% Time on R&D** OLS CEM Manual Natural Bins	(5) **% Time on R&D** OLS CEM Manual Decile Bins	(6) **% Time on R&D** OLS CEM Manual Deciles Bins
From Smaller Town	0.405	-0.772*	1.547	-0.056	4.116	0.228
	(1.188)	(0.402)	(1.349)	(0.532)	(3.487)	(0.221)
Placed in Bangalore	4.520***	1.886**	7.791***	4.071***	9.009***	0.603
	(0.681)	(0.759)	(0.374)	(0.368)	(0.631)	(0.412)
From Smaller Town * Bangalore		7.682***		11.216***		24.828***
		(0.742)		(0.487)		(1.157)
Logical Score	0.116*	0.106	-0.054	-0.103	0.142	0.113
	(0.059)	(0.057)	(0.125)	(0.157)	(0.188)	(0.140)
Verbal Score	0.369	0.332	0.539	0.541	0.423	0.146
	(0.377)	(0.352)	(0.394)	(0.393)	(0.281)	(0.101)
CGPA Training	2.543	2.396	4.201	4.241	2.336	2.422
	(2.944)	(2.812)	(3.492)	(3.554)	(1.568)	(1.574)
Male	0.212	0.454	1.661	2.174	1.332	0.997
	(1.399)	(1.518)	(1.904)	(2.294)	(1.662)	(1.296)
Placed in Hometown	-2.365	-1.954	-2.654	-2.010	-3.395	-1.559
	(1.327)	(1.059)	(2.022)	(1.521)	(2.291)	(1.137)
Constant	-12.337	-11.280	-22.106	-21.855	-15.144	-12.408
	(15.652)	(14.687)	(18.528)	(18.305)	(10.900)	(8.837)
Observations	303	303	195	195	103	103
R-squared	0.056	0.077	0.143	0.176	0.175	0.329
Location FE	No	No	No	No	No	No

NOTES – Standard errors in parentheses are clustered at the location level except for Columns 7 and 8 where the errors are robust. Three versions of Coarsened Exact Matching (CEM) are presented, all of which match workers on their pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender. However, the results are also robust to matching on Logical and Verbal Scores only. The first in Columns 1 and 2 labelled Automatic, uses STATA's in-built identification of appropriate bins using Sturge's Rule (i.e., the number of bins is equal to 1+3.322log(n) where n is the number of observations), and results in a multivariate L1 distance equal to 0.526 with 81 matched strata comprising of 118 treated and 185 untreated observations. In the second approach in Columns 3 and 4 we use is a set of natural bins created based on the common Logical and Verbal scores achieved on the pre-entry exams and allowing STATA to select appropriate bins for CGPA training based on Sturge's rule. This approach leads to an improved multivariate L1 distance equal to 0.433, but fewer matched strata – 71 – and fewer treated (89) and untreated (106) observations. Finally, the third, and most conservative approach in Columns 5 and 6 creates bins based on deciles within Logical, Verbal and CGPA Training scores. This approach leads to the lowest multivariate L1 distance of 0.227, but is also the most conservative of the three, with only 41 matched strata containing 47 treated and 56 untreated observations. Columns 7 and 8 present our results with Propensity Score Matching with the nearest neighbour algorithm with a single nearest neighbour, using the same pre-location assignment characteristics of Logical Scores, Verbal Scores, CGPA Training and Gender for the initial matching. However, note that since we are interested in the difference in the Bangalore placement effect between workers from smaller towns and large cities we perform the matching separately for workers from smaller towns and large cities and present the results for the treatment dummy on From Smaller Town. All observations end up on common support after this procedure. In unreported results we also confirm that balance is achieved on all variables in the matching process.

*p<0.1; **p<.05; ***p<.01

Table A18
Time Spent on R&D by Workers Located in Chennai, Hyderabad, and Pune

VARIABLES	(1) **% Time on R&D** OLS *Large City Sample*	(2) **% Time on R&D** OLS *Smaller Town Sample*	(3) **% Time on R&D** OLS *Large City Sample*	(4) **% Time on R&D** OLS *Smaller Town Sample*	(5) **% Time on R&D** OLS *Large City Sample*	(6) **% Time on R&D** OLS *Smaller Town Sample*
Placed in Chennai	-1.071**	-1.191				
	(0.344)	(0.850)				
Placed in Hyderabad			1.110*	-1.056		
			(0.498)	(1.326)		
Placed in Pune					-0.356	-1.024
					(0.552)	(1.383)
Logical Score	-0.047	0.256	-0.037	0.251	-0.050	0.245
	(0.207)	(0.201)	(0.210)	(0.201)	(0.207)	(0.194)
Verbal Score	-0.043	0.299	-0.042	0.304	-0.042	0.314
	(0.132)	(0.194)	(0.133)	(0.198)	(0.133)	(0.210)
CGPA Training	1.413	1.131	1.440	1.197	1.502	1.245
	(0.976)	(1.048)	(0.980)	(1.100)	(0.997)	(1.163)
Male	-0.926	1.177	-0.966	1.302	-0.919	1.281
	(0.713)	(1.362)	(0.715)	(1.477)	(0.704)	(1.457)
Placed in Hometown	-0.880	-1.168	-0.870	-1.158	-0.875	-1.106
	(0.532)	(0.795)	(0.509)	(0.779)	(0.531)	(0.697)
Constant	-3.451	-6.817	-3.941	-7.144	-3.890	-7.342
	(3.834)	(6.201)	(3.715)	(6.421)	(3.809)	(6.684)
Observations	474	181	474	181	474	181
R-squared	0.011	0.049	0.012	0.050	0.009	0.050
Location FE	No	No	No	No	No	No

NOTES – Parentheses contain standard errors clustered at the production centre level. This table replicates the analyses in Table 5 in the main manuscript for each of the next three largest technology hubs where INDTECH has a production center: Chennai, Hyderabad and Pune.
*p<0.1; **p<.05; ***p<.01

Table A19
Worker Enrolment in Additional Coursework in Chennai, Hyderabad, and Pune

VARIABLES	(1) **Number of Courses** OLS with FEs	(2) **% Courses Passed** OLS with FEs	(3) **Number of English Courses** OLS with FEs	(4) **Level of English Courses** OLS with FEs	(5) **Number of Courses** OLS with FEs	(6) **% Courses Passed** OLS with FEs	(7) **Number of English Courses** OLS with FEs	(8) **Level of English Courses** OLS with FEs	(9) **Number of Courses** OLS with FEs	(10) **% Courses Passed** OLS with FEs	(11) **Number of English Courses** OLS with FEs	(12) **Level of English Courses** OLS with FEs
From Smaller Town	0.771**	1.223	-0.051	0.386	0.960***	2.434*	-0.048	0.548**	0.631***	2.649*	-0.047	0.502**
	(0.301)	(0.933)	(0.031)	(0.256)	(0.274)	(1.153)	(0.030)	(0.199)	(0.182)	(1.165)	(0.032)	(0.212)
From Smaller Town * Chennai	0.555*	5.365***	0.029	0.347								
	(0.261)	(1.013)	(0.030)	(0.326)								
From Smaller Town * Hyderabad					-0.795**	-2.562*	0.007	-0.737**				
					(0.278)	(1.256)	(0.031)	(0.308)				
From Smaller Town * Pune									1.424***	-3.102**	0.001	-0.196
									(0.196)	(1.257)	(0.034)	(0.185)
Logical Score	0.043	-0.261	0.001	-0.035	0.044	-0.254	0.001	-0.034	0.043	-0.255	0.001	-0.035
	(0.035)	(0.347)	(0.005)	(0.045)	(0.035)	(0.347)	(0.005)	(0.045)	(0.035)	(0.348)	(0.005)	(0.044)
Verbal Score	-0.076***	0.289**	-0.005	0.086**	-0.075***	0.290**	-0.005	0.084**	-0.078***	0.296**	-0.005	0.084**
	(0.019)	(0.123)	(0.003)	(0.030)	(0.019)	(0.126)	(0.003)	(0.029)	(0.018)	(0.121)	(0.003)	(0.029)
CGPA Training	0.626	16.695***	-0.041	-0.984	0.625	16.647***	-0.041	-1.055*	0.653	16.498***	-0.041	-1.024*
	(0.446)	(3.650)	(0.040)	(0.530)	(0.443)	(3.608)	(0.040)	(0.517)	(0.444)	(3.758)	(0.040)	(0.528)
Male	0.245	-4.289**	0.011	-0.429	0.240	-4.311**	0.011	-0.423	0.231	-4.313**	0.011	-0.451
	(0.141)	(1.795)	(0.031)	(0.254)	(0.143)	(1.758)	(0.031)	(0.245)	(0.137)	(1.801)	(0.031)	(0.246)
Placed in Hometown	0.114	-0.910	0.006	-0.272	0.106	-0.912	0.006	-0.265	0.088	-0.885	0.006	-0.241
	(0.230)	(1.948)	(0.031)	(0.240)	(0.233)	(1.927)	(0.031)	(0.251)	(0.227)	(1.936)	(0.031)	(0.243)
Constant	0.232	7.387	0.353*	6.249**	0.234	7.561	0.354*	6.562**	0.148	8.211	0.354*	6.439**
	(1.949)	(15.692)	(0.183)	(2.321)	(1.952)	(15.682)	(0.183)	(2.235)	(1.960)	(16.116)	(0.185)	(2.289)
Observations	1,208	687	687	96	1,208	687	687	96	1,208	687	687	96
R-squared	0.049	0.122	0.019	0.258	0.050	0.120	0.019	0.262	0.054	0.121	0.019	0.256
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTES – This table replicates the analyses in Table 6 in the main manuscript for each of the next three largest technology hubs where INDTECH has a production center: Chennai, Hyderabad, and Pune. Standard errors in parentheses are clustered at the production centre level.

*p<0.1; **p<.05; ***p<.01

References

Bejan, Adrian. "Why university rankings do not change: education as a natural hierarchical flow architecture." *International Journal of Design & Nature and Ecodynamics* 2.4 (2007): 319-327.

Appendix B

Additional Evidence for our Interpretation of the "Smaller Town" Construct

We scrape data from publicly available profiles of employees for the three largest Indian technology firms (TCS, Infosys, and Wipro), available on the social networking site LinkedIn. We reconcile differences in naming conventions for each of these firms on the LinkedIn profiles (e.g., Wipro was listed as "*Wipro," "Wipro technologies," "Wipro Ltd.,"* etc.) and clean and code the graduate school of each employee for 30,131 employees in this sample. We follow the same coding principle as in our definition of the *From smaller town* variable and code the graduate school as being located within a top six metropolitan city in India (i.e., New Delhi, Mumbai, Kolkata, Chennai, Bangalore, and Hyderabad) or elsewhere (i.e., in smaller towns). The fraction of employees who graduated from colleges located in smaller towns was 0.59, 0.62, and 0.62, respectively, for employees from TCS, Infosys, and Wipro.

To follow up, we conduct another survey at one of the top three Indian technology firms (not INDTECH). We hired a professional survey company to survey 1,054 employees at the firm's Bangalore production center. 84 percent of employees surveyed indicated they had gone to school in a smaller town in India. We also find that employees from smaller towns had shorter tenures at the firm (difference in means = -1.12 years, t=-4.11) and were more likely to be employed as a contractor (difference in means = 0.22, t=4.99).

In the final step, we collected the CVs of the top 593 scientists in the government-owned research labs of India. India's 42 state-owned national laboratories are organized under an autonomous umbrella organization, The Council of Scientific and Industrial Research (CSIR); collectively they have around 12,500 scientific and technical employees. The laboratories, covering all major scientific and engineering disciplines, were created in the 1940s and 1950s. We collected the CVs of the top 593 scientists across all 42 laboratories at the top three tiers of the organizational hierarchy. We then coded whether or not the individual was educated in a college located in a smaller town. The mean fraction of top CSIR scientists educated in smaller towns is 0.74, 0.71, and 0.57, respectively, for the top three organizational hierarchy levels of 'Scientist-F,' 'Scientist-G,' and 'Scientist-H.'[28]

[28] Our secondary analyses benefit from looking at employees across the organizational hierarchy of these firms (by contrast, our main analyses look at entry-level employees at INDTECH) but have several limitations. For this broader sample of employees, we are unable to compare the performance of individuals educated in smaller towns vs. larger cities. We also do not know whether the Indian firms hired employees educated in smaller towns directly from their colleges or whether the individuals migrated to larger cities prior to being hired by these firms.

Table B1
Survey of Large City and Smaller Town Engineering Colleges

	Large city colleges	Smaller town colleges
Average size of graduating class in computer science/IT (undergraduate and master's)	342	458
Average percentage of graduating class in computer science/IT hired by INDTECH (in 2011, 2012)	0.17%	0.06%
Average percentage of graduating class in computer science/IT hired by multinational technology firms IBM and Cognizant (in 2011, 2012)	9%	1%
Mean annual salary (Rupees Lakhs, 2011 and 2012 average)	6.20	2.70
N	7	4

NOTES – The researchers randomly selected 10 large city and 10 smaller town engineering colleges from the list of colleges from which INDTECH hires and contacted the colleges' representatives to ask them to participate in a telephone survey. The researchers were able to conduct interviews with representatives at seven out of the 10 large city colleges. These were the R.V. College of Engineering, Bangalore; M.S. Ramaiah Institute of Technology, Bangalore; MLR Institute of Technology, Hyderabad; Muffakham Jah College of Engineering and Technology, Hyderabad; Vasavi College of Engineering, Hyderabad; G. Narayanamma Institute of Technology & Science (GNITS), Hyderabad; and Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad. The researchers were also able to conduct interviews with representatives of four out of the 10 selected smaller town colleges. These included M.J.P. Rohilkhand University in Bareilly, Uttar Pradesh; Majhighariani Institute of Technology & Science, Rayagada Orissa; Bapatla Engineering College, Guntur, Andhra Pradesh; and Jaya Prakash Narayan College of Engineering, Mahabubnagar, Dharmapur, Telangana. The survey results indicated that the mean salaries for 2011 and 2012 were significantly higher for individuals hired from the large city colleges as compared to those from the smaller town colleges. We found this difference to be statistically significant, based on a t-test comparison of means. In addition, the survey revealed that multinational technology firms predominantly hire from large city colleges, while INDTECH follows the distinctive policy of hiring from both large city and smaller town colleges. Note that the results from the survey might be upward biased, given the small sample of colleges that participated in the survey. Rs.1 Lakh = Rs. 100,000

Appendix C
Employee Random Assignment Protocol

This text is based on field interviews and INDTECH internal documents. (Part of it is copied from INDTECH internal documents.)

INDTECH assigns its software engineer trainees to production centres based on a computer application called "Talent Planning," which is part of the firm's enterprise resource allocation software system. This application allocates trainees to a location based on the quarterly manpower budget released by Corporate Planning.

The "process life cycle steps" are:

- Collating the manpower budget and unit-wise requirements
- Trainee assignment (location)
- Communication with stakeholders

Talent Planning bases the assignment of employees on the following:

- Production center requirements: HR at each production center provides data on requirement for trainees trained in various technologies.
- Data from HR stationed at the training location: Two weeks prior to the completion of training batches, HR at the training location releases data on which employees are expected to complete training.

The two variables that the Talent Planning team considers while assigning employees to production centres, using the automated system, include the stream of training for the trainee and the estimated date of training completion. The prior background of the employee and the test scores of the employee are **not** considered in this decision. INDTECH communicates employee assignments through a centralized portal.

Appendix D
Human Capital Rents by Worker Origin and Placement Location

Our results raise the question of whether the decision to hire employees from smaller towns and place them in Bangalore rather than an alternative location creates economic value for INDTECH. In particular, on net, do workers from smaller towns who are placed in Bangalore create less value than their large city counterparts when we account for their higher mobility to competing firms? Similarly, what is the relative net return on placing workers from smaller towns and large cities outside of Bangalore? Unfortunately, answering this question requires detailed employee-level data comparing INDTECH's costs associated with hiring individuals from smaller towns versus large cities and the monetary returns to their employment, which we do not have access to. As a result, we could not run regressions using hiring costs. However, based on a set of assumptions gathered during field interviews, we are able to estimate at least the approximate net payoff of hiring individuals from smaller towns and large cities and placing them in different locations in their first year of employment at INDTECH. We outline here the steps in our estimation process and the results.

In the first step, we estimate the "revenues" generated by different types of workers. To do so, we begin by estimating the dollar value of productivity gains associated with hiring a worker from a smaller town and also of placing this individual in Bangalore. Our interviews suggest that at least in their first year on the job, as compared to those who achieve the highest performance rating in a given year, other workers need 35% more man-days to correct coding/testing/documentation errors. This is based on rough calculations with INDTECH HR managers on error rates and lost man-days due to coding/testing/documentation errors, and implies that workers achieving the top performance rating can complete the same amount of work as those receiving lower performance ratings in just under three quarters (0.741) the amount of time as lower-performing workers. We use the predicted probabilities of achieving the highest performance rating in 2008 for small-town versus large-city workers reported in Column (1) of Table A3 in the Appendix. Average marginal effects (estimated with the STATA command "margins i.from_smaller_town") indicate that small-town workers receive the highest performance rating at a rate of 40.9 percent, while workers hired from large cities do so at a rate of 33.3 percent.

To arrive at a dollar-value estimate of the total value generated by each type of worker, however, we need to estimate each worker's contribution to firm revenues. While direct measures of the revenue contributions of workers in our sample are not available, we are able to obtain an estimate of the likely upper and lower bounds for the average revenues that a worker at INDTECH generates. We obtain the upper bound by dividing publicly available data on INDTECH's total firm revenues in 2008 by the total number of workers at INDTECH in 2008. This calculation yields an average revenue of about $50,000 per worker. Our reasoning is that while the workers in our sample work on the organization's core products and services, and are therefore likely to generate more revenue than workers in support functions at the firm, they are nonetheless entry-level workers and therefore unlikely to be contributing at a level above the mean revenue per worker within their first year on the job. We therefore use the average revenue of $50,000 as our likely upper bound on the worker's productivity. We estimate the lower bound of productivity conservatively as the sum of the worker's annual wages and training costs, discussed below. However, please note, that while these figures ignore the substantial variation across workers and job types, our calculations of the relative rents from different types of workers are not sensitive (in relative rank) to the absolute value of average revenues chosen. We start by calculating the net returns using only the upper bound and then discuss the changes in the results with the lower bound.

Combining the relative performance differences across workers and the average upper bound of revenue contributions by all workers at INDTECH, we arrive at the following formula to calculate the total contribution to revenues by each worker type:

*Total Contribution to Firm Revenues = Probability of Achieving Top Performance Rating * Extra Revenue From Top Performers * $50,000 + (1 – Probability of Achieving Top Performance Rating) * $50,000*

Applying this formula to workers hired from smaller towns, we calculate that their total value generated for INDTECH in their first year of employment amounts to $57,158 (=0.409*1.35*50,000+0.591*50,000) and the total value generated by workers from large cities amounts to $55,828 using the same formula. However, since we know that the *average* employee at INDTECH generates about $50,000 in value, we rescale these figures to preserve this average for all workers, and arrive at the final value generated by smaller town workers of $50,756 and large city workers of $49,575 (the rescaling is achieved by assuming an equal number of workers from smaller towns and large cities and setting their average contributions to equal $50,000; the resulting scaling factor is equal to 0.888, and we multiple this factor by each of the workers' contributions).

However, these figures are the averages for workers hired from smaller towns and large cities, regardless of placement location. If these workers are instead posted to headquarters in Bangalore, the relative performance differences grow between large city and smaller town workers, to 25.5% and 39.1% chances of receiving the top performance rating, respectively. Therefore, the relative re-scaled productivities of employees from smaller towns and large cities become $50,920 and $48,363, respectively.

In the second step, we estimate the costs of recruiting workers of different types. INDTECH's entry-level salaries are about $5,870 per year (at 2008 U.S. Dollar to Rupee exchange rates) regardless of placement and worker productivity. Therefore, workers from smaller towns and large cities all receive the same salary of $5,870. However, recruiting workers from smaller towns requires additional expenditures. Based on our discussions with INDTECH's recruiting managers, we estimate that there is a $21 incremental cost of hiring a remote worker. This is based on several criteria: incremental travel costs for INDTECH executives involved in hiring from smaller towns, the additional search costs associated with trips to screen colleges and students from smaller towns, and the larger number of candidates who need to be interviewed in smaller towns as compared to large cities.

In addition, differences in attrition rates across different types of workers and placement locations create significant differences in the cost of training, turnover, and replacement. Specifically, INDTECH spends about $3,500 to train each newly hired worker in the four-month training program. Since all hires go through this training, regardless of origin, the figure only enters our calculations through attrition rates. We therefore compare the overall attrition rates for workers from smaller towns and large cities, as well as their differential rates by placement location, using Columns (10) and (11) in Table A3. Average marginal analyses indicate that workers from smaller towns exit INDTECH at an average rate of 7.8% per year relative to workers from large cities, who do so at a rate of 13.0% per year.[29] This difference is reduced but remains large even when both types of workers are posted to Bangalore (9.7% annual attrition rate for workers from smaller towns and 14.6% annual attrition rate for workers from large cities). Multiplying these figures by the training ($3,500) and recruitment ($21) costs of smaller town workers yields a total recruitment, training, and replacement cost of $275 each for workers from smaller towns and $455 each for workers from large cities, regardless of placement location. When posted to Bangalore, these figures grow to $342 and $511 for smaller town and large city workers, respectively.

In sum, the expected costs of employing workers from smaller towns regardless of location amount to $6,145 ($5,870 in salary and $275 in recruitment, training, and replacement costs) and employing workers from large cities amounts to $6,325. In Bangalore, the respective estimated costs are $6,212 and $6,381 for smaller town and large city workers in their first year of employment.

Bringing together the estimates of revenue and costs for each worker type, we arrive at our back-of-the-envelope estimates for human capital rents captured by INDTECH each year from workers in our sample. Depicted as the "rent rectangles" in Figure D1, they show that workers from smaller towns placed outside of Bangalore generate the greatest amount of rents ($44,724), followed closely by workers from smaller towns placed in Bangalore ($44,708). The lowest rents derive from workers from large cities posted to Bangalore, whose higher rates of attrition and lower productivity generate only $41,982 in rents, a difference of $2,742 as compared to workers from smaller towns posted outside of Bangalore, which amounts to about 46.7% of the workers' annual salary in 2008 of $5,870.

Similarly, we calculate the lower bound of productivity using the same formula as above, but instead of scaling the returns by the average worker revenue of $50,000, we scale it by the average lower bound of revenue in the worker's first year on the job – the sum of the worker's annual salary ($5,870) and training costs ($3,500) as the first year revenue, or $9,370. Applying the same formula, we arrive at:

[29] Note that the attrition results in Table A3 are based on worker attrition after three years at INDTECH from the date of entry. We annualize these figures by assuming a constant rate of attrition each year, using the size of the entry cohort as the base with respect to which we calculate rates of departure. The annualized rates are therefore one-third of the total rates for the three years for each group – 39.1% for large city workers and 23.4% for workers from smaller towns.

*Total Contribution to Firm Revenues = Probability of Achieving Top Performance Rating * Extra Revenue From Top Performers * $9,370 + (1 – Probability of Achieving Top Performance Rating) * $9,370*

For workers hired from smaller towns and placed in Bangalore, this calculation amounts to $10,652 (=0.391*1.35*9,370+0.609*9,370) and the total value generated by workers from large cities placed in Bangalore amounts to $10,206 using the same formula. Rescaling these values to preserve the average yields $9,459 and $9,063, respectively. For workers placed outside of Bangalore, these figures are $9,351 for workers from large cities and $9,529 for workers from smaller towns. Subtracting from these values the average replacement and re-training costs for workers placed in Bangalore from smaller towns ($340) and large cities ($510), and outside of Bangalore from smaller towns ($257) and large cities ($443) as well as the $5,870 in wages, we arrive at net figures of $3,249 for workers from smaller towns and $2,683 for workers from large cities, placed in Bangalore. For large city workers outside of Bangalore, this figure is $3,038 and for small town workers outside of Bangalore this figure is $3,402. Therefore, as depicted in Figure D2, the difference between large city workers in Bangalore and small town workers in Bangalore amounts to $719 per worker per year or 12.2% of annual worker salary.

However, given our data constraints, these calculations provide only *a very rough estimate* of the net payoff associated with hiring from smaller towns and have at least two important limitations. First, we do not have an estimate of sunk costs of investments that INDTECH made related to hiring from smaller towns. Nor do we have an estimate of operating costs per worker related to real estate, utilities, etc. Second, we cannot distinguish between the costs of losing workers to voluntary turnover for pursuing higher studies versus turnover to competing firms. The latter may have important observed and unobserved costs that affect the trade-off as estimated. For instance, additional (unreported) results indicate that top-performing workers from smaller towns are even more likely to leave to join competing firms and less likely to leave to take up further education when posted to Bangalore as compared to elsewhere and as compared to their large city counterparts, suggesting that such turnover may be even more costly than estimated here. However, while our data limit more accurate calculations, we hope that our approach will spur future research on the role of worker characteristics, not only in terms of value creation and capture, but also net rent generation for firms, since the latter is likely to be the true underlying driver of recruitment decisions.

Figure D1
'Rent Rectangles': INDTECH's Value Creation and Appropriation by Employee Origin and Placement Location (in U.S. dollars) – Upper Bound

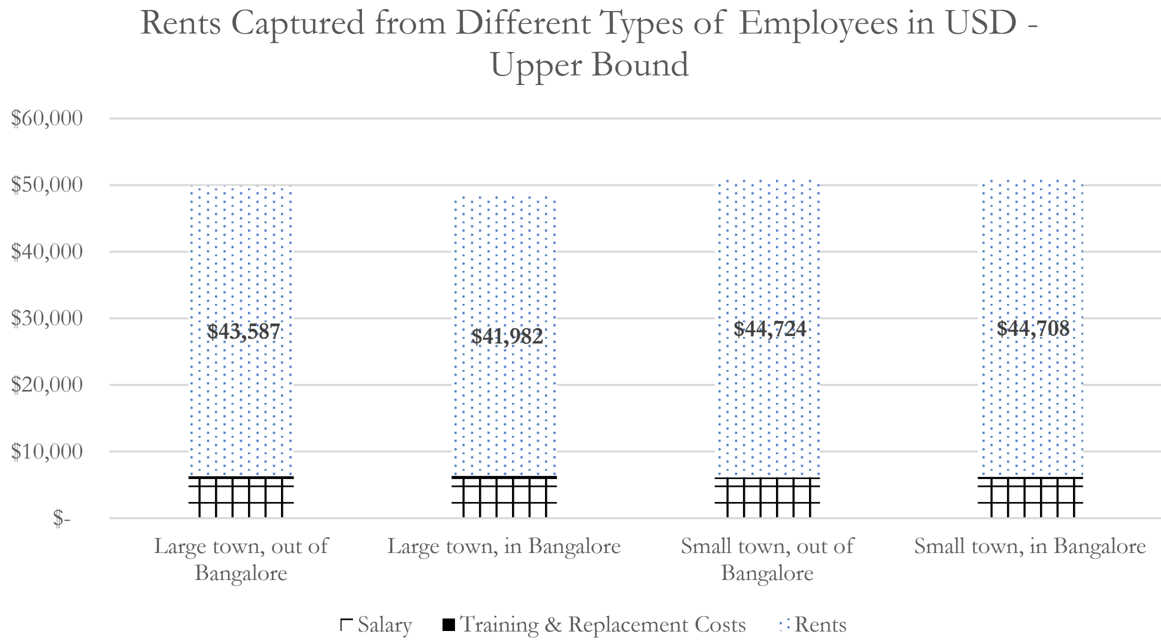
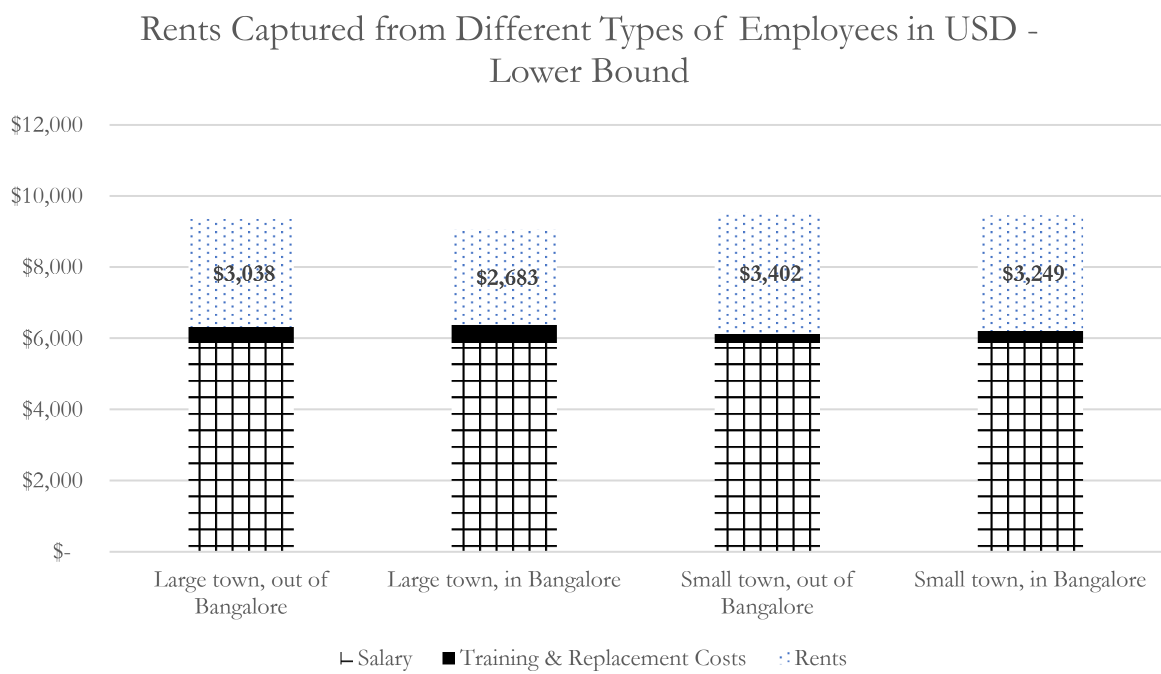


Figure D2
'Rent Rectangles': INDTECH's Value Creation and Appropriation by Employee Origin and Placement Location (in U.S. dollars) – Lower Bound


Rents Captured from Different Types of Employees in USD - Lower Bound
$12,000
$10,000
$8,000
$6,000
$4,000
$2,000
$-
$3,038
$2,683
$3,402
$3,249
Large town, out of Bangalore
Large town, in Bangalore
Small town, out of Bangalore
Small town, in Bangalore
Salary
Training & Replacement Costs
Rents