7-1-2018

Internal Borders and Migration in India

Zovanga L. Kone
University of Oxford

Maggie Y. Liu
Smith College, yliu77@smith.edu

Aaditya Mattoo
The World Bank, USA

Caglar Ozden
The World Bank, USA

Siddharth Sharma
The World Bank, USA

Recommended Citation

Internal Borders and Migration in India[*]

Zovanga L. Kone[b], Maggie Y. Liu[a,c], Aaditya Mattoo[a], Caglar Ozden[a] and Siddharth
Sharma[a]

[a]World Bank Group
[b]University of Oxford
[c]Smith College

November 30, 2017

Abstract

Internal mobility is a critical component of economic growth and development, as it enables the reallocation of labor to more productive opportunities across sectors and regions. Using detailed district-to-district migration data from the 2001 Census of India, the paper highlights the role of state borders as significant impediments to internal mobility. The analysis finds that average migration between neighboring districts in the same state is at least 50 percent larger than neighboring districts on different sides of a state border, even after accounting for linguistic differences. Although the impact of state borders differs by education, age, and reason for migration, it is always large and significant. The paper suggests that inter-state mobility is inhibited by state-level entitlement schemes, ranging from access to subsidized goods through the public distribution system to the bias for states' own residents in access to tertiary education and public sector employment.

Keywords: Internal migration; internal borders; immigration; emigration
JEL codes: J6; O15; F22

1 Introduction

Development and economic growth take place through the more efficient allocation of inputs among alternative productive uses. Labor is a key input since it is the main asset of the majority of the population, especially of the poor, in developing countries. The reallocation of labor can take place across sectors, occupations and, most importantly, geographic regions. Thus, it is no surprise that every successful development experience and growth episode is accompanied by large labor movements, especially from rural to urban areas, and from low to higher productivity sectors and occupations. In this regard, India presents a paradox and daunting challenge.

As of 2001, internal migrants represented 30 percent of India's population, but this number is deceptively large. A closer inspection of the data reveals that two-thirds are intra-district migrants, more than half of whom are women migrating for marriage. Comparing India's migration rates with those of Brazil, China and the US reveals that they are relatively low. As seen in the last column of Table 1, India has the lowest cross-district migration rate at 2.8 percent while the rate is over 9 percent in Brazil, almost 10 percent in China and 20 percent in the U.S. Internal migrants in India are less likely to move across major administrative units (states or provinces) compared to those in the other three countries. Inter-state migration is slightly above 1 percent in India, while it is 3.6 percent in Brazil, 4.7 percent in China and almost 10 percent in the U.S.[1] In fact, a cross national comparison of internal migration rates over a five year interval between the years 2000 and 2010, (Bell, Charles-Edwards, Ueffing, Stillwell, Kupiszewski, and Kupiszewska 2015) shows that India ranks last in a sample of 80 countries.[2]

This paper makes several contributions in exploring internal migration patterns

---

[1]These data are broadly consistent with another study of the U.S. which finds that those who moved from one state to another within a given five-year period accounted for 12 percent of the population in 2005 (Molloy, Smith, and Wozniak 2011).

[2]Government of India (2017), using provisional tables from the 2011 census, suggests that the share of migrants for economic reasons rose from 8.1 percent of the workforce in 2001 to 10.5 percent in 2011. Given the large differences in migration rates between India and other countries shown in Table 1, growth of this magnitude would not change the characterization of India as a country with relatively low internal migration.

and their determinants in India. The first is the presentation of internal migration patterns in India in greater detail by using district-to-district census-based migration data, disaggregated by age, education, duration of stay and reason for migration. Most existing studies in India use household survey data that suffer from sampling and aggregation biases and are rarely bilateral. Our data allow us to control for origin and destination specific factors (such as natural endowments, economic and social conditions, and climate) through fixed effects in a gravity model. Thus, we are able to focus on the bilateral variables emphasized in the literature. Among these are the critical contiguity variables – being in the same state and/or being neighbors – in addition to the standard physical distance and linguistic overlap measures. Furthermore, by using bilateral migration data between 585 districts, instead of the standard state-to-state analysis in other papers, we are able to solve many of the aggregation problems that arise in large countries like India. For example, Uttar Pradesh would rank as the fifth most populous country in the world if it were independent, and treating it as a single observation creates many biases.

The second and more substantive contribution is to demonstrate the role played by administrative barriers, particularly state borders, in limiting internal migration in India. Our empirical analysis shows that, even when we control for numerous barriers to internal mobility, such as physical distance, linguistic differences and economic and social features of origin and destination districts (through district fixed effects), state borders continue to be important impediments. Migration between neighboring districts in the same state is at least 50 percent larger than migration between districts which are on different sides of a state border. This gap varies by education level, age and reason for migration, yet it is always large and significant.

The low level of internal mobility in India – including the role of state borders – cannot be attributed to restrictions imposed by the state or federal governments. In China, for example, federal government policies have constrained migration through measures such as the *hukou* system. No such administrative measures exist in India, and anyone is legally free to move from one district or one state to another. Moreover,

3

federal laws in India protect migrant workers from exploitation in destination regions. One such provision is the Inter State Migrant Workmen Act, 1979, which requires that migrants are paid timely wages equal to or higher than the minimum wage.[3]

We provide preliminary evidence that mobility is India is inhibited by explicit and implicit entitlement programs implemented at the state level. First, many social benefits are not portable across state boundaries. For example, access to subsidized food through the Public Distribution System (PDS), with a coverage of over half of the population (Government of India 2016), and even admission to public hospitals is administered on the basis of "ration cards", issued and accepted only by the home state government. While non-portability of such benefits inhibits the movement of the poor and the unskilled, two other factors contribute to the inertia of the skilled. Many universities and technical institutes are under the control of state governments, and state residents get preferential admission. Furthermore, government jobs account for more than half of the employment opportunities for individuals with secondary education and above. State domicile is required for employment in such government entities. We show patterns that suggest these state-level policies inhibit inter-state mobility for both low and high-skilled people. Specifically, the relative share of unskilled migrants moving out-of-state is lower precisely in the states with higher levels of participation in the public distributions system. The relative share of skilled migrants moving out-of-state is lower in states with higher rates of public employment. And the relative share of migrants moving out-of-state to seek education is lower in states with higher rates of access to tertiary education.

The limited labor mobility in India has been documented since the early 1960s (Srivastava and McGee 1998; Singh 1998; Lusome and Bhagat 2006 and Srivastava and Sasikumar 2003). In spite of these observations, there have been few attempts to empirically investigate the causes (Rajan and Mishra 2012). Most studies on the topic have been concerned with identifying patterns of migration and the general

---

[3]Menon (2012) questions the effectiveness and implementation of this provision. Other legal provisions that migrants can benefit from are the Minimum Wage Act, 1948; the Contract Labour Act, 1970; the Equal Remuneration Act, 1976; and the Building and Other Construction Workers' Act, 1996 (Srivastava and Sasikumar 2003).

characteristics of migrants (Singh 1998; Lusome and Bhagat 2006; Hnatkovska and Lahiri 2015). A more recent study (Pandey 2014) documents a slight upward trend in the overall level of migration from the early 1990s, primarily driven by increased intra-district and intra-state movements.[4]

Studies by Bhattacharyya (1985), Munshi and Rosenzweig (2016) and Viswanathan and Kumar (2015) are exceptions that move beyond descriptive analyses. While the last paper examines how migration responds to environmental changes, the first two papers provide an explanation for the low levels of rural to urban migration in India. Bhattacharyya (1985) presents a theoretical framework for developing countries, in which migration decisions are more likely to be taken at the (extended) family level as opposed to the individual level, with the objective of increasing overall family income. Closely related, Munshi and Rosenzweig (2016) explore the linkages between the caste networks in rural areas and migration incentives. They argue that emigration of an income-earning individual reduces the family's access to the caste network as a social safety net. This reduces the incentives for internal migration considerably. While this explanation addresses low rural to urban migration, where community networks exert a strong influence on the decisions of members, it does not explain why urban to urban migration is also low or why we observe differences in migration patterns across state borders.

Mobility across certain administrative boundaries can be costly, especially if these boundaries reflect differences in societal characteristics such as language, culture, laws and institutions or geographic barriers (Belot and Ederveen 2012). The first study to point out such a cost was by McCallum (1995) using trade as an example. McCallum showed that Canadian provinces adjacent to the United States trade more with their neighboring provinces than with the states in the U.S.[5] Subsequent studies have confirmed McCallum's findings and unearthed evidence of a border cost in the case of

---

[4]Government of India (2017) also finds an upward trend in migration using estimates based not on actual migration but on railway passenger data and changes in the population within state and district level age cohorts.

[5]There is now an extensive literature on the role of national borders in trade, as reviewed in Anderson and van Wincoop (2003; 2004).

migration (Helliwell 1997; and Poncet 2006). Helliwell (1997), for example, suggests that inter-provincial migration in Canadian provinces is almost 100 times more likely than migration to Canadian provinces from the United States. But these studies explore the role of international borders rather than the internal ones.

Migration costs naturally impede internal migration flows in a country. Bayer and Juessen (2012) suggest that inter-state migration in the U.S. can cost a potential migrant up to two-thirds of an average household annual income. In her study of internal migration in China, Poncet (2006) suggests migration flows between two localities are negatively related to distance but positively related to contiguity (as well as with wage levels at the destination). More relevant to our paper, they too find that there is more intra-province migration in comparison to inter-province migration.

The next section of the paper presents the internal migration data, the geographic and linguistic distance variables as well as several empirical observations that motivate the analysis. Section 3 introduces the gravity model and our empirical specification, followed by empirical results in Section 4. We then discuss the results in Section 5 and end with the conclusions.

2 Data

2.1 Data source and empirical observations

The National Census of India for 2001 is the main data source in this paper.[6] The census has been conducted every decade since 1871 and is the responsibility of the Office of the Registrar General and Census Commissioner in the Ministry of Home Affairs. The national census, like those in many other countries, collects individual and household level information on various demographic and labor market characteristics for the entire population. We supplement the census with additional household and labor force data from the 55th Round (1999-2000) and the 61st Round (2004-2005)

---

[6] As of the date of drafting of this paper, the migration related sections of the 2011 Census have not been processed.

of the National Sample Survey (NSS), which cover over 100 thousand households. In addition to standard household modules on consumption, health, education and employment, it includes specialized surveys that rotate each year. The NSS has a significantly larger set of questions and therefore provides more detailed data in comparison to the Census, but for a much smaller sample of the population.

The census asks two different questions pertaining to the migration status of the respondents – one based on birthplace and one on place of last residence. The last residence question is less common in censuses, but it is more relevant for economic analysis of internal mobility (Carletto, Larrison, and Ozden 2014). We define an individual as a migrant "if the place in which he is enumerated during the census is other than his place of immediate last residence" (Census, 2001). The Census includes additional questions based on the last residence criteria. These questions include reason for migration (marriage, education, employment etc.), the urban/rural status of the location of last residence and the duration of stay in the current residence since migration. Such information sheds additional light on the patterns and determinants of internal mobility.

While the census questionnaire asks these questions to each respondent, the resulting individual level data are not made publicly available. Instead, the data are aggregated up to the geographic units – depending on the purpose – and are disseminated through tables. For example, we can find the number of people living in a given district whose previous residence was in a different state or another district within the same state. In some cases, the publicly available tables include additional variables on gender, education or reason for migration. However, these datasets do not present bilateral migrant stocks at the district level, and therefore, they do not lend themselves to empirical analysis, especially to gravity type estimation. We therefore requested detailed bilateral (district-to-district) migration data from the Census Bureau, which provided us with a series of tables under a special administration agreement. These tables contained the following for all pairs of districts in India: (i) migration stocks by gender and educational attainment levels, (ii) migration stocks by gender and age

groups, (iii) migration stocks by gender and reason for migrating, and (iv) migration stocks by gender and duration of stay at the destination.[7]

Using the compiled data, we distinguish four subgroups of the population: (i) non-migrants, (ii) intra-district migrants, i.e. those who moved from one enumeration area to another one within the *same* district, (iii) inter-district migrants within the same state, i.e. people who moved *across* districts within the *same* state, and (iv) inter-state migrants, i.e. those who moved *across* states. Table 2 presents the sizes of these groups by gender. Migrants account for close to 30 percent of the population in 2001, albeit with considerable divergence in patterns across genders. The share of migrants among females (43.3 percent) is almost three times larger than among males (16.3 percent). This gap is due to the well-known migration of women within the same or neighboring districts for marriage. The share of intra-district migrants among women is 29.5 percent, over three times the level among men. Inter-district (but intra-state) migration among women is 9.8 percent, over twice the level among men. Finally, inter-state migration among women is 4 percent, slightly higher than among men.

The low level of internal migration in India, its spatial variation and gender gaps are further illustrated by district-level heat maps of Central India in Figure 1. In each map, state boundaries are outlined with thick lines, and districts are color-coded so that darker-shaded districts have relatively higher shares of the relevant migration measure.

Figures 1a and 1b plot the share of all inter-district migrants (the sum of intra-state-inter-district migrants and inter-state migrants) by gender among the existing population in each district. Figure 1b is much "darker" in color, indicating that inter-district migration is higher among women. In 337 districts, over ten percent of the current female population are inter-district migrants, while only 101 districts have

---

[7]Each table has over 350,000 rows and between 10 and 16 columns. The 2001 administrative division of India has 593 districts, 9 of which are districts in Delhi. In our analysis, we combined the nine districts in Delhi, and treat Delhi as one single district. This leaves us with 585 districts in the empirical analysis.

the same share among males. Furthermore, we observe more migration to the West coast, especially to districts in Maharashtra, and to Northwestern states, especially to Punjab, Haryana and Delhi.

The data allow us to compare those who stay within the same state (intra-state migrants) with those who move to another state (inter-state migrants) as presented in Figure 2. More specifically, Figures 2a and 2b present the share of inter-state migrants among all inter-district migrants in destination districts for males and females, respectively. Even though the number of female migrants far exceeds that of male migrants, female migration is mostly within the same state while male migrants are more likely to cross state borders. That is why most districts in Figure 2a (for men) are darker in color compared to Figure 2b (for women). On average, 43 percent of male inter-district migrants are from another state, compared to 29 percent of female inter-district migrants. Furthermore, districts that receive higher shares of migrants from other states are located along state borders, an issue which we will explore in detail in the empirical section.

The key feature of our dataset is its bilateral nature at the district level. To highlight the role of state borders on internal migration, we take the district of Nagpur in Maharashtra as an example. We chose Nagpur since it is geographically located at the center of India and close to three other states – Andra Pradesh, Madhya Pradesh and Chattisgarh. Figures 3a and **??** plot the color-coded distribution of the origin districts of the migrants coming to Nagpur. The vast majority of these migrants come from other districts in Maharastra or from districts in neighboring states. In fact, four out of the top five origin districts are in Maharashtra, and six out of the seven districts that share a border with Nagpur are among the top ten senders. The four neighboring districts in Maharashtra (Bhandara, Wardha, Amravati, and Chandrapur) send a total of 31 percent of Nagpur's immigrants. The remaining three neighboring districts in Madhya Pradesh (Balaghat, Chhindwara, and Seoni) send a total of 13 percent. The prohibitive role of state borders becomes more clear when we note there are more migrants from several distant districts in Maharashtra than

from neighboring districts in other states. Similar patterns are observed when we look at out-migration from Nagpur to other districts in Figures 3c and 3d. The most popular destinations of Nagpur's emigrants are neighboring districts in Maharashtra (Bhandara, Wardha, Amravati, and Chandrapur) which receive a total of 32 percent of emigrants from Nagpur. Neighboring districts in other states receive much fewer migrants when compared to distant coastal districts of Maharashtra (see Figure 3d).

Table 4 presents summary statistics by gender and various other dimensions. The first disaggregation is by age groups. The share of migrants are highest among those between 25 and 65 years of age. The gap is especially stark for women where the migrant ratio dramatically increases from 23.2 percent for 14-19 year olds to 69.1 percent for the 25-34 year olds, highlighting the role of marriage in migration. The corresponding increase is less drastic among men.

The second set of rows in Table 4 illustrates patterns of migration for four education levels: (i) illiterate, (ii) primary school education, (iii) secondary school education, and (iv) tertiary education. People with higher educational levels appear to be more mobile. This holds for both aggregate levels of migration as well as for movements across geographical boundaries. For instance, migrants account for 35.6 percent of the tertiary educated male population compared with 11.5 percent for the illiterate males, and inter-state migrants represent 8.4 percent among the former but only 2.1 among the latter. The patterns are similar for females.

The reason for migration (third set of rows in Table 4) is one of the most important questions in the census. We aggregated the answers into five main categories: (i) work or business, (ii) marriage, (iii) move with the family, (iv) education, and (v) other reasons. For men, work/business, move with the family and others are the main reasons (around 30 percent each) while marriage dominates all other categories for women (70 percent). Unfortunately, the format of the data does not allow us to construct cross-tabulations, such as by education and reason for migration, which would provide further insights.

Closely linked to the propensity of moving across geographical boundaries is the

duration of stay at the destination. The bottom set of rows in Table 4 reports summary statistics on the origin distribution of migrants across four intervals of duration of stay at their destinations. The data suggest that most migrants (i.e. about 50 percent) have lived at their destination for over 10 years, although this is driven by female migrants. Regardless of the duration of stay considered, there is very little variation in the distribution of migrants by origin (e.g. inter-state versus intra-state), especially among males.

2.2   Migration measures and other controls

The key dependent variables, bilateral migration stocks, are based on the Census data described above. In addition, we construct several explanatory variables needed for the gravity estimation. These are standard bilateral distance, linguistic overlap and other geographic proximity variables. They are described and discussed in detail below.

Bilateral Migration Stocks

We define m_{ij} as the stock of migrants who moved from origin (or previous) district i to destination (or current) district j as of 2001.[8] We also amalgamate intra-district migrants with non-migrants in the empirical analysis. Lastly, we disaggregate the migrant numbers also by education, age, reason for migration and duration in later sections and m_{ij} represents the relevant bilateral migrant stock in each regression.

Following the approach in gravity models of international migration, we control for dyadic factors that influence migration costs: physical distance, linguistic proximity,[9] contiguity, and state borders.[10] The construction of these control variables is explained below.

State Borders and Contiguity

Borders, either physical or institutional, could impose costs on mobility. To cap-

---

[8]Since we only measure the migrant stock at 2001, we do not observe return or circular migration.

[9]We include cultural proximity variables using caste information in a robustness check.

[10]We should note that origin and destination specific factors are not included since we control for them with origin and destination fixed effects in our empirical analysis.

ture the effects of state borders on mobility, we first construct a contiguity variable which takes a value of 1 if two districts share a common land border. Empirical studies on international migration (Mayda 2010; Artuc, Docquier, Ozden, and Parsons 2015) have documented higher migration flows between countries with common border relative to noncontiguous ones and the same properties arguably hold for internal migration. Next, we construct a dummy variable to indicate whether the origin and destination districts are located in the same state. These two variables allow us to categorize district-pairs into 4 distinct groups: (i) different states and not neighbors, (ii) different states and neighbors, (iii) same state and not neighbors, and (iv) same state and neighbors.

We note that three of the states were in fact newly created in November 2000 by splitting existing states. Chhattisgarh was created out of eastern Madhya Pradesh; Uttaranchal (renamed Uttarakhand in 2007) was created out of the mountainous districts of northwest Uttar Pradesh; and Jharkhand was created out of the southern districts of Bihar. In other words, new state borders were created *within* Madhya Pradesh, Uttar Pradesh, and Bihar (see Figure A1 in the Online Appendix). Since their creation predates 2001, we treat these three new states as "different" states throughout most of our analysis. However, as discussed in the results section, we confirm that our analysis is robust to ignoring the state division of November 2000 and using only the boundaries of the original, undivided states.

The first column in Table 3 tabulates the number of district-pairs that fall into each contiguity category. We have a total of 341,640 district-pairs in our dataset. Among these, for example, 323,906 (95 percent) are in different states and they are not neighbors while 14,576 are in the same state and not neighbors.

Distance

The physical distance between two districts is expected to influence migration through its effect on transportation costs and the degree of uncertainty about earnings at the prospective destination. For bilateral distance between any two districts, we calculate geodesic distances – the length of the shortest curve between two points along the

surface of a mathematical model of the earth – between the districts' geographical centers.[11] In robustness checks, we include several other distance variables. These are (i) geodesic distances between largest cities in each district, (ii) driving distance between these cities using the transport network and (iii) driving time between these cities. [12]

Linguistic Proximity

Another important component of bilateral migration costs is the linguistic differences (Belot and Ederveen 2012; Adsera and Pytlikova 2015); linguistic proximity facilitates communication and skill transferability, especially for the less skilled. First, we measure linguistic distance between any two districts (i, j) following the commonly used ethnolinguistic fractionalization (EFL) index (Mira 1964), which measures the probability of two randomly chosen individuals from different districts speaking the same language. We concentrate on the mother tongue, which is "the language spoken in childhood to the person by the person's mother", as reported in the 2001 Census of India. In addition to data availability, we argue that there are two advantages in using the mother tongue. First, each individual has a unique mother tongue even if they are multilingual. Second, mother tongue relates more closely to an individual's birth place, family background, and social networks. In the 2001 Census of India, there are 122 separate mother tongues, and all districts have multiple mother tongues spoken by the native population.

We construct two different measures of linguistic proximity between two districts: *Common Language*$_{ij}$ and *Language Overlap*$_{ij}$. Let s_i^l and s_j^l be the share of individuals speaking mother tongue l in districts i and j, respectively. Then $s_i^l * s_j^l$ is the probability that an individual from i can speak to an individual from j in language l. Summing over all possible mother tongues, *Common Language*$_{ij}$ measures the likelihood of any two individuals being able to communicate in a common language. This

---

[11] We restrict centroids to be inside the boundaries of a polygon.

[12] See Alder, Roberts, and Tewari (2017) for more details on how these distances and travel times are calculated using the road network data.

13

is given by:

$$Common\ Language_{ij} = \sum_l s_i^l \cdot s_j^l$$

Similarly, *Language Overlap$_{ij}$* measures the degree of overlap in languages spoken at any pair of districts. $\min\{s_i^l, s_j^l\}$ is the intersection of people from each district who speak the same language l. Since each person has only one mother tongue, summing over all possible mother tongues, we have the overlap of people from two districts that can understand each other. This is calculated as:

$$Language\ Overlap_{ij} = \sum_l \min\ \{s_i^l, s_j^l\}$$

Our linguistic proximity measures do not take into account the genealogical relations (linguistic distance) between languages,[13] and thus can be considered a lower bound of the linguistic proximity across districts.

Table 3 summarizes the language and distance measures by contiguity of district-pairs. Overall, neighboring districts are closer to each other in terms of distance and linguistic proximity relative to non-neighboring districts. The average district-to-district log distance is 6.8. For neighbors, regardless of whether they are in the same state or not, the log distance is 4.3, which is 12 times smaller. Districts that are in the same state have greater linguistic proximity than district-pairs from two different states. This confirms that language was an important consideration in the drawing of state borders. Consistent with this, even though neighboring districts in different states have higher linguistic overlap than non-neighboring districts in different states, this overlap is lower than that among districts in the same state.

---

[13]Several studies use language trees from Ethnologue and use number of shared nodes between two languages to construct a linguistic proximity measure. Such studies include Adsera and Pytlikova (2015); Belot and Hatton (2012); Desmet, Weber, and Ortuño-Ortín (2009) and Desmet, Ortuno-Ortin, and Wacziarg (2012).

3 Empirical specification

In the empirical analysis to follow, we adopt a gravity specification, which is based on a random utility maximization model. This specification has been extensively used in the analysis of migration patterns.[14] Our specification is given by:

$$m_{ij} = \alpha + \beta_1 \cdot lnDIST_{ij} + \beta_2 \cdot LANG_{ij} + \gamma_1 \cdot D_{ij}^{\mathrm{diff}-NBR} + \gamma_2 \cdot D_{ij}^{\mathrm{same}-NBR} + \gamma_3 \cdot D_{ij}^{\mathrm{same}-notNBR} + \delta_i + \delta_j + \epsilon_{ij}$$

(1)

The dependent variable, m_{ij}, measures migration from origin i to destination j. In our case, it is the size of the inter-district migration stock. The bilateral independent variables introduced previously are: $lnDIST_{ij}$, log geodesic distance between districts i and j; $LANG_{ij}$, linguistic proximity between districts i and j. There are three contiguity variables: $D_{ij}^{\mathrm{diff}-NBR}$ is a dummy variable that takes the value of 1 if districts i and j are in different states but are neighbors; $D_{ij}^{\mathrm{same}-NBR}$ is dummy variable that is equal to 1 if the districts i and j are in the same state and are neighbors; $D_{ij}^{\mathrm{same}-notNBR}$ is dummy variable that is equal to 1 if the districts i and j are in the same state but are not neighbors. The base group is "not in the same state and not neighbors". The difference between γ_2 and γ_1 gauges the role of the state borders.

Multilateral resistance, in the context of bilateral migration decisions, is the influence exerted by the attractiveness of other destinations (Bertoli and Moraga 2013), and can introduce bias in the estimation if not properly addressed. We include origin and destination fixed effects, δ_i and δ_j, to account for the multilateral resistance as well as for unobserved heterogeneity in sending and receiving districts in our cross-sectional data.

We estimate the above specified gravity model using Poisson Pseudo-Maximum Likelihood, or PPML (Silva and Tenreyro 2006). As thoroughly discussed by Beine, Bertoli, and Fernandez-Huertas Moraga (2015), PPML is a more reliable estimator

---

[14]Beine, Bertoli, and Fernandez-Huertas Moraga (2015); Beine, Docquier, and Ozden (2011); Beine and Parsons (2015); Bertoli and Moraga (2013); Grogger and Hanson (2011); Mayda (2010).

since, (1) OLS estimates are biased and inconsistent in the presence of heteroskedasticity of ϵ_{ij}, and (2) PPML performs well in the presence of a large share of zeros, which is slightly over 40 percent of observations in our data.

4 Empirical results

4.1 Main results

Our first set of results explores the determinants of bilateral migration patterns, and more specifically the role of district and state borders. As discussed earlier, the dependent variable is the stock of migrants currently living in district j and whose previous residence was in district i. Since we have fixed effects for both origin and destination districts, we include only bilateral variables in the estimation – distance, language overlap and dummy variables for the contiguity relationships. Each pair of districts can have one of the four possible relationships: (i) different states and not neighbors, (ii) different states and neighbors, (iii) same state and not neighbors, (iv) same state and neighbors. In the estimations that follow, "different states and not neighbors" is the base category, and hence dropped from the regression.

Table 5 presents our main gravity estimates. The first set of three columns relates to total migration, the next set of three columns pertains to men, and the last set of three columns to women. The first and second columns in each set have different linguistic proximity variables. The third column presents the results when the newly split states are included as a separate group.

The distance variable has a negative coefficient in all specifications, as expected, and all the estimates are quantitatively close to each other. The language variables all have a positive sign, again as expected, with higher coefficients for men, indicating linguistic proximity is a more important pull factor for them. The most important variables are the contiguity dummy variables. We see that relative to the base category of "different states and not neighbors", being in the same state and being neighbors both increase migration. For example, being in the same state but not

16

neighbors increases migration (in Column 1) by almost twice ($e^{1.097} - 1$). The impact of being in the same state is higher for men than women. Being in different states but neighbors also has a large positive effect. In column 1, we see that total migration is around 4.5 times ($e^{1.730} - 1$) larger in this case and this effect is stronger for women.

The most important observation is that the coefficient for same-state-neighbor dummy variable is larger than the different-state-neighbor coefficient in every column. This difference is statistically significant. For example, in the first column, being neighbors and in the same state increases total migration by almost eight times ($e^{2.177} - 1$), indicating that the state borders have a large negative effect on internal migration in India. To put it differently, migration between neighboring districts in the *same* state is around at least 50 percent larger than migration between neighboring districts in *different* states ($e^{2.177-1.730} - 1$). The state border effect is almost identical for men and women when we compare the differences between the relevant coefficients in columns 4 and 7.

As noted earlier, we treat the recently created states of Chattisgarh, Uttaranchal and Jharkhand as "different states" in most of our analysis. To confirm that our analysis is robust to this event, in the third column of each set, we create a separate state border category called "split states" to indicate two districts that were in the same state before 2000, but now belong to different states post 2000 due to the state split. For example, Godda and Banka used to be in the Bihar before 2000. After Bihar was split, Godda went to Jharkhand while Banka remained in Bihar. Thus Godda and Banka are coded as districts from "split states". We see that the coefficient for the "split states and neighbors" dummy is never statistically different from the coefficient for the "same state and neighbors" dummy (columns 3, 6 and 9). This is consistent with the fact that the migration observed in the 2001 census largely predates the creation of the new states. If the state borders represented natural mobility barriers, the coefficient of the "split states and neighbors" dummy would have been closer to the "different states and neighbors" dummy rather than the "same state and neighbor" dummy. More convincing evidence will come from the 2011 census, when we will

be able to see what happens to migration flows after the new state boundaries were imposed.[15]

The next set of tables presents the results of the gravity estimation for different sub-groups of migrants, by age, education, reason for migration and duration of migration. Estimates when the sample only comprises males are reported on the left, and those for females are on the right. We only use the *share of common language* variable since the choice of the linguistic overlap variable does not seem to affect the results.

In Table 6, we explore the impact of the distance and contiguity variables on different age groups. The signs on distance and language variables are as expected, and similar for all age groups. Being in the same state and being neighbors increase migration, with the *same state* effect being higher for men and the *neighbor* effect being higher for women. Most importantly, there does not seem to be much difference across age groups. The state border effect – the difference between the "different states and neighbors" and "same state and neighbors" coefficients – are slightly higher for younger men of working age and younger women in the marrying age group, relative to older people (above age 65).

The next disaggregation is by education level, as presented in Table 7. In this case, as education levels increase, distance becomes less of an impediment while linguistic proximity becomes more important. Furthermore, the changes in these coefficients are larger for women relative to men. With respect to the contiguity variables, we observe interesting patterns. Being in the same state is significantly more important for more educated people while being neighbors is less important for them. As a result, the state border effect between neighboring districts is rapidly increasing in education levels. For example, for illiterate men, the state border effect is only 17 percent ($e^{1.482-1.325} - 1$) as seen in Column 1. On the other hand, for college educated

---

[15]In all specifications except for Table 5, we group "Split States" with "Different States" because at least some migration in our dataset took place after the split, and their inclusion in the "Different States" category mitigates the risk of creating a bias in favor of finding a significant border effect. Tables in the Online Appendix show that our results are robust to how we treat district pairs from split states.

men, being in the same state increases migration between neighboring districts by about 149 percent ($e^{1.852-0.939} - 1$) as seen in column 4.

Table 8 splits the population by reason for migration, revealing large differences between men and women. As mentioned earlier, women migrate predominantly for marriage reasons to nearby districts while men migrate for employment reasons to more distant areas. As a result, distance is a large impediment for women migrating for marriage (column 5) relative to other reasons. The importance of distance for women migrating for marriage appears again in the neighborhood coefficients, which are significantly higher for this group. For men migrating for work, common language and being neighbors seem to be less important (column 2). The negative state border effect for males is significant for migration motivated by movement with the family, work and education, but not marriage.

4.2 Robustness checks

The labor mobility between two districts could depend on the *relative* level of attributes such as income levels, extent of urbanization or literacy rates, in addition to distance, contiguity and linguistic overlap. To account for this possibility, we include several relative "attraction" metrics as controls. Using 2001 census tables and 2004 National Sample Survey (NSS) data, we calculate the following variables at the district level: (i) the percentage of non ST/SC population,[16] (ii) the literacy rate, (iii) the urbanization rate, (iv) the share of private employment in the labor force, (v) the share of formal employment in the labor force, and (vi) average income. Districts with higher values of these metrics are likely to attract more migrants from districts with lower values. The attraction between an origin district i and destination district j due to an attribute a is then measured by $s_{ij}^a = \frac{a_j}{a_i}$. Since these bilateral variables are correlated across district-pairs, we do not insert them separately into the regression. Instead, we calculate the overall "attraction index" between i and j which is a

---

[16]"ST/SC" refers to "scheduled tribes and scheduled castes."

simple average of the six attributes: $s_{ij} = \frac{1}{6} \sum_a s_{ij}^a$.[17]

Table 9 presents the PPML regression results when the bilateral attraction variable s_{ij} is included in the gravity regression. Column 1 has the original results and column 2 includes the attraction index. Comparing columns (1) and (2), the coefficients of distance and contiguity variables barely change and are robust to the inclusion of the attraction variable. More importantly, the state border effect remains strong. The attraction index is significant, suggesting that the listed pull factors lead to higher migration flows.

Column 3 presents the results when we interact the attraction index with each one of the contiguity variables. We first note that the gap between 'same state neighbor' and 'different state neighbor' dummies disappears and they are no longer statistically different, indicating that the state border effect is zero when $s_{ij} = 0$ and the destination is not attractive at all relative to the origin. The coefficient of the interaction term between s_{ij} and the 'different state neighbor' is greater than the coefficient of the interaction term with the 'same state neighbor' dummy. In other words, the state border effect becomes stronger as the bilateral attractiveness of the destination district increases. [18]

Our next extension introduces other measures of distance which are highly relevant in the context of low-income countries with poor infrastructure. The analysis above relied on the flight distance between the geographic centers of origin and destination districts as a measure of distance and traveling cost. This measure may suffer from two measurement errors. First, flight distance does not account for the transport network across India, and thus distorts the actual cost of travel. For two districts that are not connected by highways, the flight distance underestimates the relative traveling time. If this measurement error is more relevant among district pairs that are in different states, the gravity estimation could overstate the state border effect. Second, the

---

[17]See Table A5 in the Online Appendix for the summary statistics of district characteristics included in the attraction index.

[18]Specifically, the state border effect is given by $(2.420 - 2.345) + [-0.286 - (-0.576)] \cdot s_{ij}$, and therefore increasing in s_{ij}.

geographic centers are not necessarily the economic or population centers that send and receive most migrants. Thus distance measures using geographic centers might not accurately reflect traveling cost between the more relevant economic centers.

Table 10 replicates the original results from Table 5 and confirms that the earlier results are robust to alternative measures of distance. '$l.dist_{ij}$, geo. centroids' is the geodesic (flight) distance between the geographic centers of districts i and j – it is the same distance measure used in the previous tables. Columns (1), (5), and (9) repeat the results from Table 5 for ease of comparison. We use three additional measures of distance: (i) '$l.dist_{ij}$ in columns (2), (6), and (10) is the flight distance between the economic centers of districts i and j, (ii) '$l.TravelTime_{ij}$' in columns (4), (8), and (12) takes into account India's transport network of national highways and measures the driving time on the shortest path between the economic centers of i and j,[19] and (iii) '$l.TravelTime_{ij}$, flat' in columns (3), (7), and (11) assumes the same driving speed on and off the roads – this measure is similar to the flight distance between economic centers. The coefficients of all of these distance variables are negative. Furthermore in each case, the state border effect remains significant.

5   Discussion: Some explanations for the invisible wall at the border

Why do state borders inhibit migration? In this section, we highlight a number of policies implemented at the state level which act as inhibitors, either explicitly or implicitly, of mobility across state boundaries. Three key inhibitors of inter-state migration will be discussed: inadequate portability of social welfare benefits and a significant home bias in access to education and public employment.

---

[19]See Alder, Roberts, and Tewari (2017) for more details on the method of computing these shortest paths.

5.1 Inadequate portability of social welfare benefits

Social welfare entitlements in India, like any country, require proper identification of the recipients. When the recently launched "Unique Identity Documentation" project reaches completion, India will possess a unified system of national identity documentation. Until then, the de facto identity document for most Indian households is the "ration card" issued by state governments. The basic purpose of this card is to enable access to the "Public Distribution System" (PDS), a program of subsidized food for poor households, but because there is no national identity documentation system and the PDS covers the majority of the population, it also serves as the proof of identity and address when requesting public services such as hospital care and education. It is also needed for purposes such as initiating telephone service or opening a bank account (Zelazny 2012; Abbas and Varma 2014).

Ration cards are not portable across states; that is, they are accepted only by the issuing state. This has to do with the design of the PDS system for which these cards were designed. Even though most of the PDS subsidy cost is borne by the central government, the program is administered by state governments on the basis of their own poverty lines and lists of poor households. Further, some states add subsidies of their own to the central subsidy amount, or have a more inclusive subsidy entitlement policy than the central government. In Tamil Nadu, for example, every person is entitled to receive subsidized food. In Andhra Pradesh and Chhattisgarh, more than 70 percent of the population is entitled to subsidized ration. The differences in cost are borne by the state government. As a result, state governments generally do not extend PDS benefits to migrants who hold ration cards from other states (Srivastava 2012).

In order to get access to subsidized food and other public services in their destination state, inter-state migrants need to surrender the ration card issued by their origin state, and obtain a new ration card from their destination state. However this process is fraught with difficulties, particularly for poor and less educated people who are not familiar with the bureaucratic processes and lack social or political connec-

tions in the destination state. Procedures for issuing documentation for the PDS are complicated and vary by state. They are also prone to corruption and administrative errors. For example, issuing officials in the destination state may refuse to accept prior identity documentation provided by poor migrants because they are looking for bribes (Government of India 2008; Abbas and Varma 2014).

Individuals moving across state boundaries risk losing access to the PDS, and a host of other public services linked to the PDS for a substantial period until their destination state issues them a new ration card. The loss of access to subsidized PDS food could be a significant issue for most households. According to household survey data, 27 percent of all rural households and 15 percent of all urban households were fully dependent on PDS grain, and most households in the country were eligible in 2004-05 (Kumar, Parappurathu, Babu, and Betne 2014). Despite widespread leakage to non-eligible households, the PDS subsidy is a particularly important source of calories for poor households. One study estimates that in 2004-05, access to PDS lowered the rate of nutritional deficiency in households officially categorized as "Below Poverty Line" (BPL) from 49 percent to 37 percent (Kumar, Parappurathu, Babu, and Betne 2014). Using survey data from 2009, another study estimates that the PDS reduced the poverty-gap index of rural poverty in Indian states by 18 to 22 percent (Dreze 2013). Therefore, the low inter-state portability of PDS cards and a host of other associated welfare benefits could act as an indirect barrier to migration in India.

A survey of seasonal migrant workers in the construction industry in Delhi suggests that the lack of identity documents also makes it difficult for low-skilled interstate migrants to claim the benefits that they are entitled to under labor laws (Srivastava and Sutradhar 2016). For example, the migrant workers surveyed were not registered under the Building and Construction Workers' Welfare Act, a law that regulates social welfare, health care, and safety for construction workers. Lacking formal protection, the workers had to work long hours under poor health and safety conditions. Thus, poor inter-state portability of identity documentation leads to asymmetric enforce-

ment of labor regulation across inter-state migrants, further reducing incentives to move even if wage gains are substantial.

Recognizing these issues, the central government passed a law, called the 1979 Inter State Migrant Workmen Act, specifically to regulate practices associated with the recruitment and employment of interstate migrant workers. The law requires middlemen who recruit interstate migrant workers and the firms that hire them to get a special license. It requires that migrant workers be paid in accordance with local minimum wage laws, be issued a passbook recording their identity, nature of work and remuneration, and be provided with accommodation and health care. However, as pointed out in Section 2, studies suggest that this law is not enforced: most firms hiring migrant workers do not carry the proper license and most migrant workers do not possess the required passbooks (Srivastava and Sasikumar 2003; Srivastava 2012).

We expect the lack of portability of PDS benefits and cards to contribute to the inertia of the unskilled who are likely to be most dependent on it. In Figure 4a, we plot the partial regression of the share of in-state unskilled emigration on participation in the PDS. The dependent variable (on the y-axis) is the number of unskilled emigrants who moved to destinations within the state of their origin, divided by the total number of unskilled migrants from the said state. This measure comes from the bilateral migration data in the 2001 Census, aggregated to the state level. The explanatory variable (on the x-axis) is the share of the unskilled population participating in the Public Distribution System (PDS).[20] The regression controls for the log average household income per capita and the share of agricultural households at the state level, both of which are also calculated from the NSS data. We find a positive and significant relationship between the two variables, i.e. the larger the share of unskilled population who rely on PDS, the higher the tendency for potential emigrants to choose home-state destinations over out-of-state destinations. This finding is con-

---

[20]We calculate this measure from the consumption module of the 55th round of National Sample Survey (1999-2000). The unskilled population refers to all members from households with a male household head who has completed primary education or below. Any household that reported a positive amount of PDS purchase is considered participating in the PDS, and consequently, so are all individuals from such households.

sistent with, and preliminary evidence for, the argument that inadequate portability of social welfare programs such as PDS tends to deter households who rely on these benefits from moving across state borders.

5.2 State government employment policies

The state domicile requirements for employment in government entities could act as a disincentive to move across states. Under India's policy of affirmative action, a sizable proportion of jobs in central and state government entities are reserved for individuals belonging to disadvantaged minority groups, principally the "Scheduled Castes" (SCs) and "Scheduled Tribes" (STs). According to the Constitution of India, the percentage of employment quota for SCs and STs in state government jobs must be equal to their respective shares of a state's total population. In 1999, on average 25 percent of employment in state-level government jobs was reserved for SCs and STs (Howard and Prakash 2012). In order to be eligible for the SC/ST employment quota in a particular state, an individual has to belong to an SC/ST community and be domiciled in that state. Thus, individuals belonging to an SC/ST group would lose access to reserved government jobs in their home state if they were to migrate to another state.

This disincentive for inter-state migration is likely to matter the most for highly educated individuals belonging to SC/ST communities but is reportedly also relevant for non-SC/ST individuals. While the public sector accounts for only about 5 percent of total employment in India, it is a major employer for educated individuals. On average, 51 percent of wage-earning individuals with secondary education and above in 2000 were employed in government jobs (Schundeln and Playforth 2014). Moreover, the majority of government jobs are with state government entities. In 2001, 76 percent of government jobs in the median state were with the state government. Taken together, these numbers suggests that, on average, state government jobs account for more than 25 percent of employment among individuals with secondary education and above. Thus, educated individuals, especially but not only SC and ST individuals,

would care about remaining eligible for the employment opportunities in their home state government.

While all states reserve some government jobs for resident SC/STs and are reported to de facto prefer residents of that state, some states even have explicit "jobs for natives" policies that cut across communities. For example, the state of Karnataka announced a policy in 2016 under which both private and public sector firms would have to reserve 70 percent of their jobs for state residents to be eligible for any state government industrial policy benefits. Orissa, Maharashtra and Himachal Pradesh have similar quotas for state residents in factory jobs.[21] To our knowledge, there is no systematic quantitative evidence on the extent, enforcement and impact of such policies. Potentially, such policies can create yet another disincentive to migrate across state boundaries.

In Figure 4b, we plot the partial regression of in-state skilled emigration on public sector employment at the state level. The dependent variable (on the y-axis) is number of high-skilled (i.e. those who completed at least secondary education) emigrants who moved to destinations within the state of their origin, divided by the total number of high-skilled migrants from that state. The explanatory variable on the x-axis is the share of high-skilled workers who are employed by the public sector. This variable comes from the employment module of the National Sample Survey (1999-2000). Log average household income per capita is also calculated from the NSS data, and controlled for in the regression. The positive relationship shown in the graph suggests that the higher the share of government job opportunities for the high-skilled, the stronger the incentive for potential migrants to stay in their home states. The argument that state domicile requirements for public sector employment inhibit high-skilled workers from moving across state borders is novel and requires more careful analysis.

---

[21] See newspaper article published in the Economic Times: "Karnataka's 70% jobs quota for locals faces criticism; Phenomenon not limited to the state". November 9, 2014 Edition.

5.3 State government policies for access to higher education

Many universities and technical institutes in India are public and under the control of the government of the state in which they are located. For example, in 2003-04, state-level engineering and "polytechnic" colleges in the state of Tamil Nadu (TN) had a total entering class size of about 120,000 students (Government of Tamil Nadu, 2004). State residents get preferential access to state-level colleges and institutes of higher education through "state quota seats". The size of the state quota varies by state and by whether the university in question is public or private, but in general, it is a substantial proportion of the total class size. [22]

"Domicile certificates" are proofs of residence in a state that are issued by state governments and are necessary to be eligible for the state quota in educational institutes. The certificate is issued upon proof of continuous residence in the state. The duration of continuous residence that qualifies an individual for this certificate varies from 3 to 10 years, depending on the state. For example, the state of Rajasthan issues domicile certificates to individuals who have resided continuously in the state for at least 10 years, while the state of Uttar Pradesh (UP) requires continuous residence for at least 3 years (Government of India, 2016).

Domicile requirements for state quota eligibility provide clear and strong disincentives for inter-state migration. For example, a 16 year old who was born and attended high school in TN would lose eligibility for state quota seats in state-level universities in Tamil Nadu if his family were to move to another state, say Uttar Pradesh. Moreover, because of the three-year wait period for domicile certification in Uttar Pradesh, he would not be eligible for quota seats in state-level universities there for at least three years.

In Figure 4c, we examine the effect of state government policies determining ac-

---

[22]For example, in 2004, 50 percent of the seats in all state-level engineering colleges and medical colleges in Tamil Nadu were under the state quota (Government of Tamil Nadu, 2005). In the state of Maharashtra, the current state quota in state-level medical colleges varies from 70 percent to as high as 100 percent (Government of Maharashtra, 2015). In the state of Madhya Pradesh, 38 percent of seats in private medical and dental institutes are in the state quota (Government of Madhya Pradesh, 2014).

cess to higher education on emigration for the purpose of education. The dependent variable (on the y-axes) is the share of the migrants who chose home-state destinations among all migrants who moved for education related reasons. This variable is constructed from the bilateral data from the 2001 Census as discussed earlier. The explanatory variable on the x-axis comes from the employment module of the National Sample Survey (1999-2000). It measures the share of college attending students among all 18-22 year old state-natives in each state. Log average household income per capita is also calculated from the NSS data, and controlled for in the regression. The positive slope in the graph is consistent with the argument that state government policies granting preferential access to higher education to in-state students tend to induce potential migrants moving for education to choose home-state institutions.

6   Conclusion

That international borders limit migration is obvious. More surprising is the role of provincial or state borders in inhibiting mobility within a country. We are able to demonstrate the existence of these "invisible walls" by putting together, with the help of the Indian census authorities, detailed district-to-district migration data from the 2001 Census. Even after controlling for key bilateral barriers to mobility, such as physical distance and linguistic differences, and for origin and destination specific factors through district fixed effects, we find that average migration between neighboring districts in the same state is at least 50 percent larger than between neighboring districts on different sides of a state border. This gap varies by education level, age and the reason for migration, but is always large and significant. The evidence from the recent creation of three new states in 2000 provides additional evidence that these state borders are not natural barriers.

There are no barriers at state borders or explicit legal restrictions on people's mobility between states in India, and we control for distance and difference in language. Then the question is what other reasons can explain the presence of these

invisible walls. We argue that interstate mobility is inhibited by the existence of state level entitlement schemes. The non-portability across state borders of social welfare benefits, such as access to subsidized food or issuance of PDS ration cards, weakens the incentive to move for the poor and the unskilled. People are deterred from seeking education in other states because state residents get preferential access in the numerous universities and technical institutes that are under state government control. Finally, the skilled are reluctant to move to other states to seek employment because state governments are still major employers and grant de facto preferences to their own residents. We provide preliminary evidence that that the relative share of migrants moving out-of-state is linked to the importance of these entitlement schemes in each state.

This research can be taken forward in at least three ways. First, the data can be updated when the Census Bureau releases the data for 2011 and enriched in several ways. The data tables that were made available to us are two dimensional, for example, we can observe either the skill composition or the motive for migration in bilateral flows between districts but not both dimensions simultaneously. Multidimensional data would facilitate richer analysis of the determinants and consequences of internal migration in India. Second, our analysis of the reasons why state borders restrict mobility is both selective and preliminary at this stage. A fuller analysis would examine the role of other factors, e.g. such as the National Rural Employment Guarantee scheme, and for finer evidence of their relative impact. Finally, we motivate this study by noting that labor mobility enables the reallocation of labor to more productive opportunities across sectors and regions and hence promotes growth. Future analysis should assess how far India's "fragmented entitlements" – i.e. state-level administration of welfare benefits, as well as education and employment preferences – dampen growth by preventing the efficient allocation of labor. It may also be possible to assess the impact of the implementation of a unique national identification system which will lower but not eliminate the costs of moving.

References

Abbas, R. and D. Varma (2014). Internal labor migration in india raises integration challenges for migrants. *Migration Information Source, Migration Policy Institute, Washington, DC*.

Adsera, A. and M. Pytlikova (2015). The role of language in shaping international migration. *The Economic Journal 125*(586), F49–F81.

Alder, S., M. Roberts, and M. Tewari (2017). The effect of transport infrastructure on india's urban and rural development. *Working Paper*.

Anderson, J. E. and E. Van Wincoop (2003). Gravity with gravitas: a solution to the border puzzle. *the american economic review 93*(1), 170–192.

Anderson, J. E. and E. Van Wincoop (2004). Trade costs. *Journal of Economic literature 42*(3), 691–751.

Artuc, E., F. Docquier, C. Ozden, and C. Parsons (2015). A global assessment of human capital mobility: the role of non-oecd destinations. *World Development 65*, 6–26.

Bayer, C. and F. Juessen (2012). On the dynamics of interstate migration: Migration costs and self-selection. *Review of Economic Dynamics 15*(3), 377–401.

Beine, M., S. Bertoli, and J. Fernandez-Huertas Moraga (2015). A practitioner's guide to gravity models of international migration. *The World Economy*.

Beine, M., F. Docquier, and C. Ozden (2011). Diasporas. *Journal of Development Economics 95*(1), 30–41.

Beine, M. and C. Parsons (2015). Climatic factors as determinants of international migration. *The Scandinavian Journal of Economics 117*(2), 723–767.

Bell, M., E. Charles-Edwards, P. Ueffing, J. Stillwell, M. Kupiszewski, and D. Kupiszewska (2015). Internal migration and development: comparing migration intensities around the world. *Population and Development Review 41*(1), 33–58.

Belot, M. and S. Ederveen (2012). Cultural barriers in migration between oecd countries. *Journal of Population Economics 25*(3), 1077–1105.

Belot, M. V. and T. J. Hatton (2012). Immigrant selection in the oecd*. *The Scandinavian Journal of Economics 114*(4), 1105–1128.

Bertoli, S. and J. F.-H. Moraga (2013). Multilateral resistance to migration. *Journal of Development Economics 102*, 79–100.

Bhattacharyya, B. (1985). The role of family decision in internal migration: The case of india. *Journal of Development Economics 18*(1), 51–66.

Carletto, C., J. Larrison, and C. Ozden (2014). Informing migration policies: a data primer. *International Handbook on Migration and Economic Development, Robert E. B. Lucas (ed.).*

Desmet, K., I. Ortuno-Ortin, and R. Wacziarg (2012). The political economy of linguistic cleavages. *Journal of development Economics 97*(2), 322–338.

Desmet, K., S. Weber, and I. Ortuño-Ortín (2009). Linguistic diversity and redistribution. *Journal of the European Economic Association 7*(6), 1291–1318.

Dreze, J. (2013). *Rural poverty and the public distribution system*. Ph. D. thesis, Department of Economics, Delhi School of Economics.

Government of India (2008). Nutrition and social safety net. In *Eleventh Five Year Plan 2007-2012*, Volume 2. Planning Commission, Government of India.

Government of India (2016). Evaluation study on role of public distribution system in shaping household and nutritional security in India. Development Monitoring and Evaluation Office, Government of India, New Delhi.

Government of India (2017). India on the move and churning: New evidence. In *Economic Survey*, Chapter 12. Economic Division, Department of Economic Affairs, Ministry of Finance.

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

Helliwell, J. F. (1997). National borders, trade and migration. *Pacific Economic Review 2*(3), 165–185.

Hnatkovska, V. and A. Lahiri (2015). Rural and urban migrants in india: 1983–2008. *The World Bank Economic Review 29*(suppl 1), S257–S270.

Howard, L. L. and N. Prakash (2012). Do employment quotas explain the occupational choices of disadvantaged minorities in india? *International Review of Applied Economics 26*(4), 489–513.

Kumar, A., S. Parappurathu, S. Babu, and R. Betne (2014). Public distribution system in india: Implications for food security. *International Food Policy Research Institute, India, Working Paper*. Paper presented at 97th Indian Economic Association Conference in Udaipur.

Lusome, R. and R. Bhagat (2006). Trends and patterns of internal migration in india, 1971-2001. In *Paper presented at the Annual Conference of Indian Association for the Study of Population (IASP) during*, Volume 7, pp. 9.

Mayda, A. M. (2010). International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics 23*(4), 1249–1274.

Menon, N. M. (2012). Can the licensing–inspection mechanism deliver justice to interstate migrant workmen? *India Migration Report 2011: Migration, Identity and Conflict*, 102.

Mira, A. N. (1964). Moscow: Miklukho-maklai ethnological institute at the department of geodesy and cartography of the state geological committee of the soviet union.

Molloy, R., C. L. Smith, and A. Wozniak (2011). Internal migration in the united states. *The Journal of Economic Perspectives 25*(3), 173–196.

Munshi, K. and M. Rosenzweig (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *The American Economic Review 106*(1), 46–98.

Pandey, A. K. (2014). Spatio-temporal changes in internal migration in india during post reform period. *Journal of Economic & Social Development 10*(1).

Poncet, S. (2006). Provincial migration dynamics in china: Borders, costs and economic motivations. *Regional Science and Urban Economics 36*(3), 385–398.

Rajan, S. I. and U. Mishra (2012). Facets of indian mobility: An update. *India Migration Report 2011: Migration, Identity and Conflict*, 1.

Schundeln, M. and J. Playforth (2014). Private versus social returns to human capital: Education and economic growth in india. *European Economic Review 66*, 266–283.

Silva, J. S. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and statistics 88*(4), 641–658.

Singh, D. (1998). Internal migration in india: 1961-1991. *Demography India 27*(1), 245–61.

Srivastava, R. (2012). Internal migrants and social protection in india. *Human Development in India.*

Srivastava, R. and T. McGee (1998). Migration and the labour market in india. *Indian Journal of Labour Economics 41*(4), 583–616.

Srivastava, R. and S. Sasikumar (2003). An overview of migration in india, its impacts and key issues. In *Regional Conference on Migration, Development and Pro-Poor Policy Choices in Asia*, pp. 22–24.

Srivastava, R. and R. Sutradhar (2016). Labour migration to the construction sector in india and its impact on rural poverty. *Indian Journal of Human Development 10*(1), 27–48.

Viswanathan, B. and K. K. Kumar (2015). Weather, agriculture and rural migration: evidence from state and district level migration in india. *Environment and Development Economics 20*(04), 469–492.

Zelazny, F. (2012). The evolution of india?s uid program: Lessons learned and implications for other developing countries. *CGD Policy Paper 8.*

Figure 1: Share of inter-district in-migrants in population at destination districts (%)

(a) Male

(b) Female
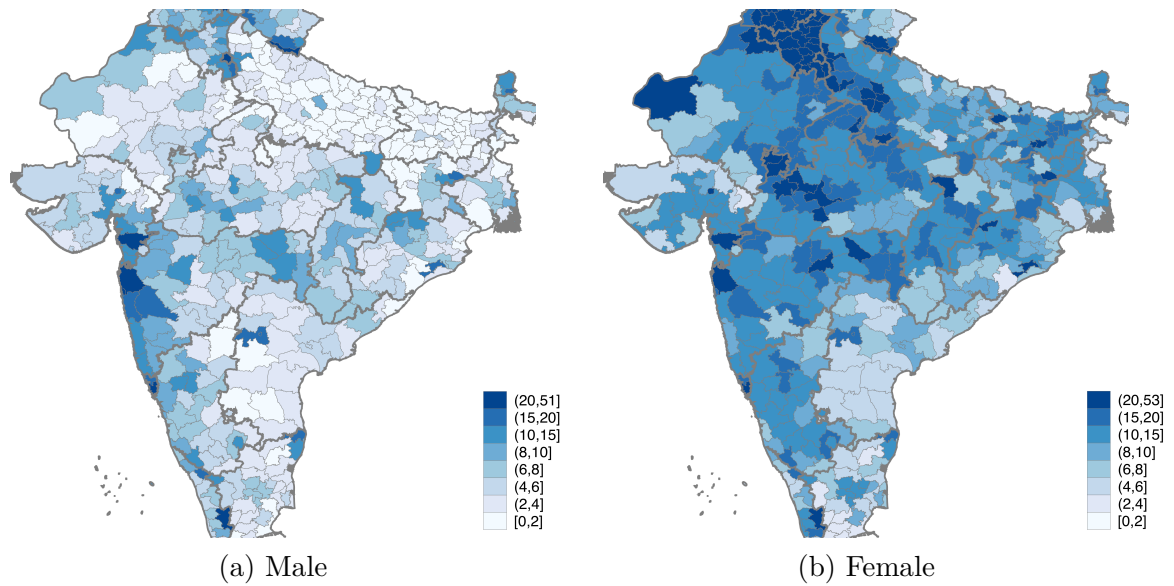
Figure 2: Share of inter-state in-migrants in inter-district in-migrants at destination districts (%)
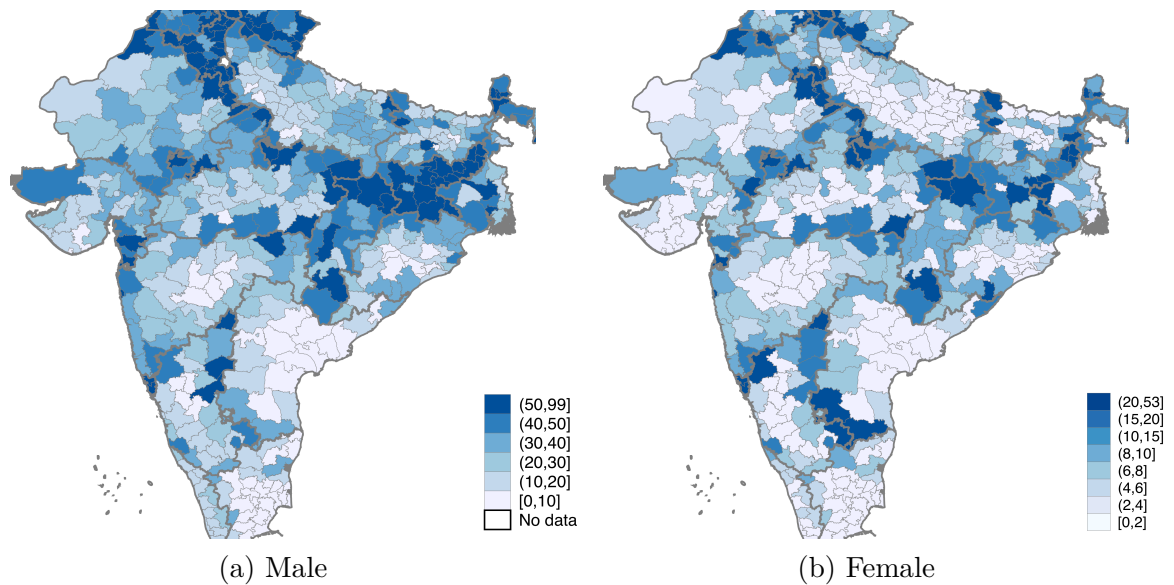
(a) Male

(b) Female

Source: Prepared by the authors based on migration data from 2001 census provided by Registrar General and Census Commissioner, Government of India. Base map is provided by the World Bank. *Notes:* Figures 1(a) and 1(b) plot each district's share of inter-district in-migrants out of total observed population in 2001. Figures 2(a) and 2(b) plot each district's share of inter-state in-migrants among observed inter-district in-migrants in 2001. Each polygon represents a district, and state borders are outlined in thick lines.

Figure 3: Nagpur, Maharashtra
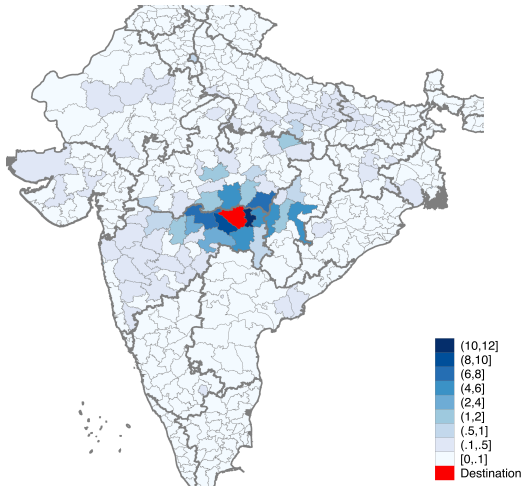
(a) Origins of in-migrants in Nagpur
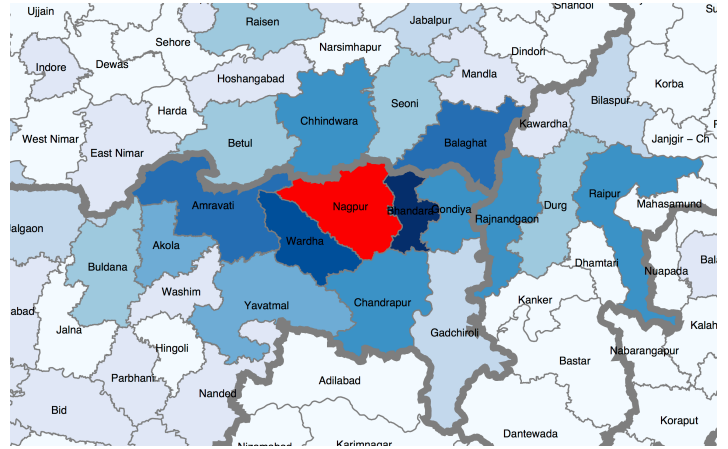
(b) Origins of in-migrants (zoomed in)
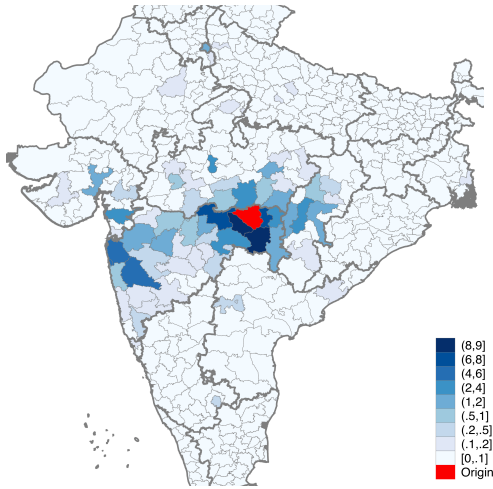
(c) Destinations of out-migrants in Nagpur
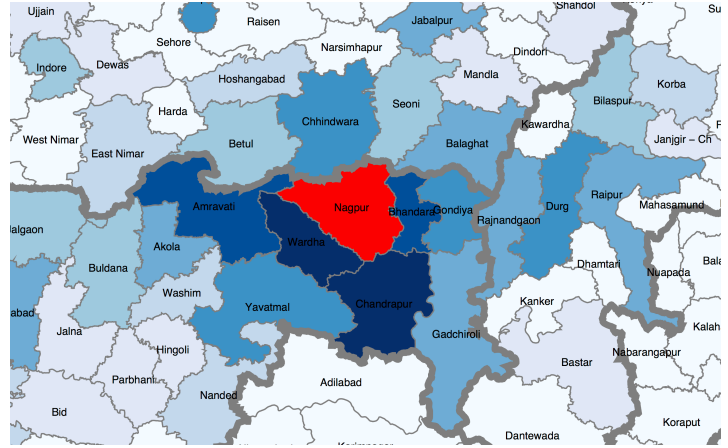
(d) Destinations of out-migrants (zoomed in)

Source: Prepared by the authors based on migration data from 2001 census provided by Registrar General and Census Commissioner, Government of India. Base map is provided by the World Bank. *Notes:* In this figure, we focus on Nagpur (in the state of Maharashtra) as a destination district in (a) and (b) and an origin district in (c) and (d). Nagpur is highlighted in red in the middle of the maps, and all other districts are in ascending shades of blue depending on the share of migrants they send to Nagpur or receive from Nagpur. In (a) and (b), we plot the origin districts of migrants coming to Nagpur; in (c) and (d), we plot the destination districts of migrants from Nagpur.

Figure 4: Institutional Barriers and Migration Inertia
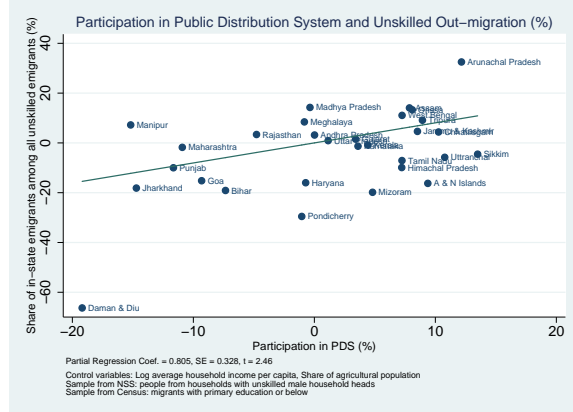
(a) Public Distribution System
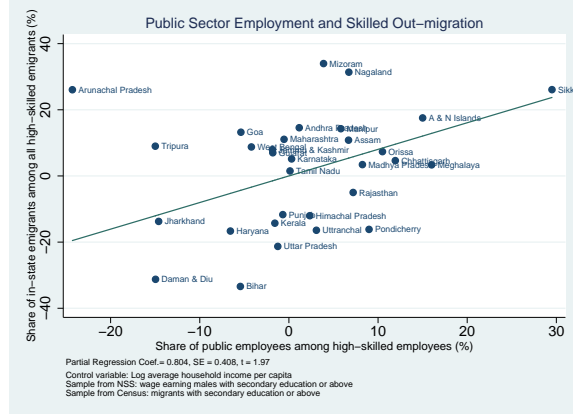
(b) Public employment of the high-skilled
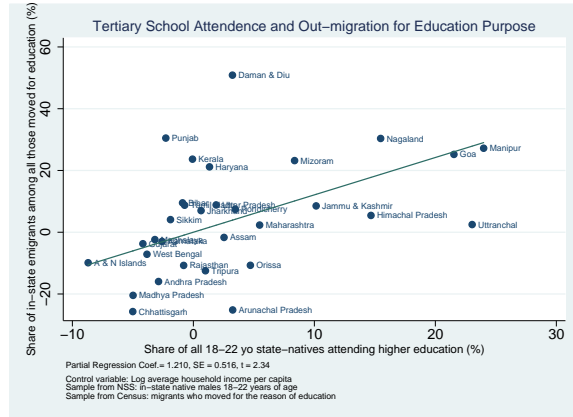
(c) Tertiary education enrollment

Source: Prepared by the authors based on migration data from 2001 census and 1999-2000 National Sample Survey (NSS 55th round).

Notes: This figure plots partial regression results of the effect of different entitlement policies (e.g., participation in PDS, share of public employment among the high-skilled, and share of tertiary enrollment among 18-22.) on out-migration shares at the state level.

36

Table 1: Internal migration flows in 2001 (or 2000)

India: 585 districts; 35 states		**Within District**	**Within State, Across Districts**	**Across States**	**Total Cross District**
Population (thousands)	1,028,610				
Last 5 years migrant flow (thousands)		36,482	18,126	10,870	28,996
Last 5 years migration rate (%)		3.55	1.76	1.06	2.82

Brazil: 2376 municipalities; 27 states		**Within Municipality**	**Within State; Across Municipalities**	**Across States**	**Total Cross Municipality**
Population (thousands)	169,077				
Last 5 years migrant flow (thousands)		51,589	9,211	6,057	15,268
Last 5 years migration rate (%)		30.51	5.45	3.58	9.03

China: 340 prefecture; 31 provinces		**Within Prefecture**	**Within Province; Across Prefectures**	**Across Provinces**	**Total Cross Prefecture**
Population (*16-65 yo; thousands)	825,544				
Last 5 years migrant flow (thousands)			43,518	38,364	81,882
Last 5 years migration rate (%)			5.27	4.65	9.92

| **United States**: 1024 PUMAs; 51 states[a]		**Within PUMA**	**Within State; Across PUMAs**	**Across States**	**Total Cross PUMA**
Population (*16-64 yo; thousands)	154,435				
Last 5 years migrant flow (thousands)			16,062	15,283	31,345
Last 5 years migration rate (%)			10.40	9.90	20.30

Source: Prepared by the authors based on migration data from 2001 Indian census (provided by Registrar General and Census Commissioner, Government of India), 2000 Brazilian census, 2000 Chinese census, and 2000 American Community Survey.

Notes: This table lists the five-year internal migration in India, Brazil, China, and the United States. First column reports the total population count, and the ensuing columns reports internal mobility at different administrative boundaries. Second Column reports mobility within secondary administrative units – district (India), municipality (Brazil), prefecture (China), or PUMA (Public Use Microdata Areas in the U.S); third column reports mobility across secondary unites but within first administrative units – states (India, Brazil, U.S.) , or provinces (China); fourth column reports mobility across first administrative units within each country.

---

[a]We count District of Columbia as a state level entity.

Table 2: Population distribution by gender and resident type

	MALE					FEMALE				
Resident Type	Native (non-migrant)	Intra-district migrant	Intra-state-inter-district migrant	Inter-state migrant	Total	Native (non-migrant)	Intra-district migrant	Intra-state-inter-district migrant	Inter-state migrant	Total
Total # (thousand)	445,373	47,338	22,468	16,978	532,157	281,735	146,255	48,639	19,825	496,454
Share (%) in population	83.7	8.9	4.2	3.2	100.0	56.7	29.5	9.8	4.0	100.0

Source: Prepared by the authors based on migration data from 2001 census provided by Registrar General and Census Commissioner, Government of India.

Notes: This table describes the population distribution of India by gender and resident type in 2001. First row reports the total count, and the second row reports the share of a group in total population. "Native (non-migrant)" refers to those who didn't move; "Intra-district migrant" to one who moved within the district; "Intra-state-inter-district migrant" to those who moved to a different district within the state; and "Inter-state migrant" to those who moved to a different state. Our sample excludes those who reported last usual residence as "unknown".

Table 3: Bilateral migration costs between origin and destination by border and contiguity

		Language		Distance
CONTIGUITY and BORDER	N	Share of common language	Language overlap	log distance (km)
Different states; Neighbor	814	0.40	0.50	4.41
Different states; Non neighbor	323,906	0.16	0.19	6.91
Same state; Neighbor	2,344	0.70	0.83	4.30
Same state; Non neighbor	14,576	0.70	0.79	5.54
Total	341,640	0.18	0.22	6.83

Notes: This table reports the mean values of linguistic proximity and physical distance between district pairs by contiguity and border. First column reports the number of district pairs that fall into each contiguity/border group. With 585 districts in the 2001 census, there are in total 341,640 (= 585*584) pairs of origin and destination districts. We use two measures for linguistic proximity: Share of common language and Language overlap. See Section 3.2 for details. Physical distance is measured as the geodesic distance between the geographic centers of two districts.

38

Table 4: Demographic distribution of population by gender, resident type, and age/education/reason/duration

	MALE					FEMALE				
Resident Type	Native (non-migrant)	Intra-district migrant	Intra-state inter-district migrant	Inter-state migrant	Total	Native (non-migrant)	Intra-district migrant	Intra-state inter-district migrant	Inter-state migrant	Total
	(%)	(%)	(%)	(%)	(thousand)	(%)	(%)	(%)	(%)	(thousand)
AGE GROUP										
0-13	89.6	6.9	2.3	1.2	177,675	89.6	7.0	2.2	1.2	163,348
14-19	85.7	8.4	3.4	2.5	65,753	76.8	15.9	5.2	2.1	57,051
20-24	82.5	8.5	4.5	4.4	46,321	42.0	39.4	13.4	5.2	43,443
25-34	80.0	9.7	5.4	4.9	78,919	30.9	46.3	16.2	6.6	78,777
35-44	77.7	11.2	6.3	4.9	65,917	29.7	47.3	16.3	6.6	60,395
45-54	77.1	11.5	6.6	4.8	44,719	30.8	47.4	15.6	6.2	39,277
55-64	79.7	10.6	5.7	4.0	27,169	32.2	48.2	14.4	5.3	28,001
65+	82.2	9.9	4.8	3.2	24,182	35.9	45.4	13.7	5.0	24,924
Age not stated	85.1	9.1	3.8	2.0	1,501	68.6	21.9	7.2	2.4	1,238
EDUCATION LEVEL										
Illiterate	88.5	7.1	2.3	2.1	195,623	54.2	33.6	8.7	3.5	272,299
Primary	85.2	8.9	3.4	2.5	176,035	63.8	24.6	8.5	3.1	135,560
Secondary	78.4	10.7	6.3	4.7	134,898	54.9	25.3	14.1	5.7	76,428
College +	64.4	13.5	13.7	8.4	25,533	47.2	18.3	21.7	12.8	12,137
Education level unknown	100.0	0.0	0.0	0.0	68	100.0	0.0	0.0	0.0	30
REASON FOR MIGRATION										
Work or business		30.1	34.0	35.9	26,867		43.8	34.6	21.6	3,902
Marriage		70.6	22.0	7.4	2,125		71.2	21.5	7.3	151,656
Move with family		53.9	29.6	16.5	25,590		48.5	31.5	20.0	29,402
Education		49.7	34.9	15.4	2,266		54.9	32.8	12.4	939
Other reason		76.3	15.0	8.7	29,935		75.5	17.6	6.9	28,819
DURATION OF MIGRATION										
0-1 years		43.0	31.2	25.8	3,976		53.4	29.3	17.3	4,579
1-5 years		45.8	30.2	24.0	19,324		62.4	25.8	11.7	37,599
6-10 years		43.8	31.5	24.7	12,176		64.6	24.9	10.5	31,508
10 + years		44.3	31.4	24.3	29,050		69.5	22.1	8.4	120,360
Duration unknown		83.4	11.0	5.6	22,258		78.9	15.3	5.8	20,674

Source: Prepared by the authors based on migration data from 2001 census provided by Registrar General and Census Commissioner, Government of India.

Notes: This table describes the demographic distribution of 2001 India population by gender, resident type, and demographic groups including age, education level, reason for migration, or duration of migration. Definitions of 4 types of residents are introduced in Table 2. Education level is the highest degree that an individual has completed. "Secondary" includes Lower Secondary, High Secondary (or Senior Secondary) degrees, and vocational/professional diplomas. "College +" includes undergraduate degrees and above. Row percentages are reported, as well as total counts of migrants for each demographic group.

Table 5: PPML gravity estimation on district-to-district migration by gender, 2001

Sample	All			Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log Distance	-1.510	-1.492	-1.479	-1.436	-1.412	-1.396	-1.603	-1.590	-1.579
	(0.091)***	(0.104)***	(0.104)***	(0.101)***	(0.116)***	(0.117)***	(0.082)***	(0.092)***	(0.092)***
Share of Common Language		0.690	0.575		0.758	0.621		0.690	0.591
		(0.128)***	(0.132)***		(0.173)***	(0.180)***		(0.104)***	(0.107)***
Language Overlap	0.391			0.405			0.421		
	(0.107)***			(0.114)***			(0.107)***		
Different states, neighbors	1.730	1.729	1.765	1.300	1.305	1.356	1.853	1.849	1.879
	(0.149)***	(0.149)***	(0.155)***	(0.154)***	(0.156)***	(0.161)***	(0.138)***	(0.136)***	(0.143)***
Same state; neighbors	2.177	2.125	2.242	1.780	1.703	1.848	2.259	2.218	2.317
	(0.107)***	(0.078)***	(0.077)***	(0.089)***	(0.074)***	(0.073)***	(0.110)***	(0.085)***	(0.085)***
Same state; not neighbors	1.097	1.029	1.126	1.294	1.198	1.316	0.968	0.913	0.996
	(0.144)***	(0.095)***	(0.092)***	(0.156)***	(0.092)***	(0.088)***	(0.129)***	(0.091)***	(0.089)***
Split states, neighbors			2.306			2.044			2.314
			(0.147)***			(0.141)***			(0.142)***
Split states, not neighbors			0.793			0.988			0.662
			(0.086)***			(0.089)***			(0.095)***
p-value: Same.nbr = Split.nbr			0.58			0.16			0.98
p-value: Same.nbr = Diff.nbr	0	0	0	0	0.01	0	0	0	0
R2	0.32	0.32	0.32	0.25	0.26	0.26	0.43	0.43	0.43
N	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640

$* \ p < 0.1; \ ** \ p < 0.05; \ *** \ p < 0.01$

Table 6: PPML gravity estimation on district-to-district migration by gender and age, 2001

Ages	Males			Females		
	25-34	35-64	65 +	25-34	35-64	65+
log Distance	-1.407	-1.489	-1.507	-1.590	-1.643	-1.722
	(0.122)***	(0.112)***	(0.128)***	(0.087)***	(0.092)***	(0.091)***
Share of Common Language	0.719	0.700	0.873	0.679	0.665	0.675
	(0.191)***	(0.161)***	(0.171)***	(0.100)***	(0.106)***	(0.119)***
Different state, neighbors	1.295	1.161	1.234	1.897	1.812	1.891
	(0.166)***	(0.149)***	(0.157)***	(0.137)***	(0.136)***	(0.130)***
Same state; neighbors	1.683	1.541	1.430	2.282	2.163	2.205
	(0.083)***	(0.077)***	(0.078)***	(0.086)***	(0.088)***	(0.096)***
Same state; not neighbors	1.262	1.175	0.957	0.907	0.839	0.777
	(0.093)***	(0.090)***	(0.102)***	(0.084)***	(0.090)***	(0.092)***
p-value: Same.nbr = Diff.nbr	0.03	0.02	0.21	0	0	0
R2	0.27	0.30	0.36	0.48	0.49	0.58
N	341,640	341,640	341,640	341,640	341,640	341,640

$* p < 0.1; ** p < 0.05; *** p < 0.01$

Notes: Huber-White robust standard errors (clustered by origin district) are reported in parentheses. All specifications include origin fixed effect and destination fixed effect. The sample includes 341,640 observations, which accounts for bilateral migration stock between 585 origin and 584 destination districts in 2001. Dependent variable is the bilateral migration stock, m_{ij}, of inter-district male migrants by age group, and of inter-district female migrants by age group. See Table 4 for the age composition of males and females. Definition and construction of distance and language measures are described in text. All district pairs fall into four mutually exclusive categories regarding contiguity ({Neighbors; not neighbors}) and state borders ({Different states; Same state}). We include three dummy variables in the estimation to indicate the dyadic relationship of district pairs regarding contiguity and state borders. For example, "Different states, neighbors" takes the value 1 when the sending and receiving districts are from different districts and share a common border; 0 otherwise. p-values from t-tests comparing border coefficients are reported under coefficients.

Table 7: PPML gravity estimation on district-to-district migration by gender and education attainment, 2001

Education Level	Males				Females			
	Illiterate	Primary	Secondary	College +	Illiterate	Primary	Secondary	College +
log Distance	-1.653	-1.510	-1.388	-1.167	-1.710	-1.644	-1.539	-1.202
	(0.102)***	(0.139)***	(0.120)***	(0.094)***	(0.077)***	(0.112)***	(0.084)***	(0.083)***
Share of Common Language	0.601	0.705	0.746	1.167	0.442	0.768	0.859	1.184
	(0.174)***	(0.192)***	(0.183)***	(0.148)***	(0.108)***	(0.115)***	(0.109)***	(0.149)***
Different state, neighbors	1.325	1.451	1.137	0.939	2.058	1.823	1.320	0.844
	(0.137)***	(0.176)***	(0.167)***	(0.138)***	(0.116)***	(0.153)***	(0.130)***	(0.127)***
Same state; neighbors	1.482	1.745	1.717	1.852	2.359	2.188	1.894	1.632
	(0.099)***	(0.084)***	(0.084)***	(0.096)***	(0.093)***	(0.097)***	(0.070)***	(0.097)***
Same state; not neighbors	0.807	1.122	1.336	1.527	0.604	1.038	1.130	1.265
	(0.102)***	(0.115)***	(0.101)***	(0.058)***	(0.076)***	(0.117)***	(0.062)***	(0.060)***
p-value: Same.nbr = Diff.nbr	0.18	0.06	0.03	0	0	0	0	0
R2	0.40	0.25	0.22	0.42	0.66	0.38	0.36	0.44
N	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Huber-White robust standard errors (clustered by origin district) are reported in parentheses. All specifications include origin fixed effect and destination fixed effect. The sample includes 341,640 observations, which accounts for bilateral migration stock between 585 origin and 584 destination districts in 2001. Dependent variable is the bilateral migration stock, m_{ij}, of inter-district male migrants by education attainment, and of inter-district female migrants by education attainment. Education attainment is the highest degree that an individual has completed. "Secondary" includes Lower Secondary (or Senior Secondary) degrees, and vocational/professional diplomas. "College +" includes undergraduate degrees and above. See Table 4 for the education attainment composition of males and females. Definition and construction of distance and language measures are described in text. All district pairs fall into four mutually exclusive categories regarding contiguity ({Neighbors; not neighbors}) and state borders ({Different states; Same state}). We include three dummy variables in the estimation to indicate the dyadic relationship of district pairs regarding contiguity and state borders. For example, "Different states, neighbors" takes the value 1 when the sending and receiving districts are from different districts and share a common border; 0 otherwise. p-values from t-tests comparing border coefficients are reported under coefficients.

Table 8: PPML gravity estimation on district-to-district migration by gender and reason for migration, 2001

Reason for Migration	Males				Females			
	Marriage	Work or Business	Move with Family	Education	Marriage	Work or Business	Move with Family	Education
log Distance	-1.639	-1.48	-1.454	-1.206	-1.767	-1.578	-1.426	-1.222
	(0.077)***	(0.105)***	(0.080)***	(0.089)***	(0.082)***	(0.089)***	(0.078)***	(0.091)***
Share of Common Language	1.062	0.496	0.939	1.334	0.587	0.789	0.881	1.316
	(0.116)***	(0.171)***	(0.131)***	(0.134)***	(0.098)***	(0.153)***	(0.123)***	(0.151)***
Different state, neighbors	2.145	1.052	1.368	1.031	1.996	1.047	1.255	0.896
	(0.111)***	(0.149)***	(0.116)***	(0.151)***	(0.122)***	(0.139)***	(0.118)***	(0.147)***
Same state; neighbors	2.257	1.511	1.684	2.365	2.317	1.376	1.668	2.486
	(0.080)***	(0.084)***	(0.077)***	(0.089)***	(0.095)***	(0.109)***	(0.077)***	(0.107)***
Same state; not neighbors	0.877	1.227	1.148	1.778	0.717	1.049	1.173	1.806
	(0.066)***	(0.083)***	(0.072)***	(0.084)***	(0.078)***	(0.086)***	(0.067)***	(0.091)***
p-value: Same.nbr = Diff.nbr	0.11	0.01	0.01	0	0	0.01	0	0
R2	0.82	0.40	0.30	0.49	0.67	0.49	0.32	0.44
N	341,640	341,640	341,640	341,640	341,640	341,640	341,640	340,472

$* p < 0.1; ** p < 0.05; *** p < 0.01$

Notes: Huber-White robust standard errors (clustered by origin district) are reported in parentheses. All specifications include origin fixed effect and destination fixed effect. The sample includes 341,640 observations, which accounts for bilateral migration stock between 585 origin and 584 destination districts in 2001. Dependent variable is the bilateral migration stock, m_{ij}, of inter-district by gender and reason for migration. See Table 4 for the composition of reasons for migration. Definition and construction of distance and language measures are described in text. All district pairs fall into four mutually exclusive categories regarding contiguity ({Neighbors; not neighbors}) and state borders ({Different states; Same state}). We include three dummy variables in the estimation to indicate the dyadic relationship of district pairs regarding contiguity and state borders. For example, "Different states, neighbors" takes the value 1 when the sending and receiving districts are from different districts and share a common border; 0 otherwise. p-values from t-tests comparing border coefficients are reported under coefficients.

Table 9: District migration gravity estimation– attraction index

	(1)	(2)	(3)
$l.dist_{ij}$, geo. centroids	-1.600	-1.616	-1.611
	(0.050)***	(0.051)***	(0.050)***
Share of Common Language	0.596	0.593	0.590
	(0.092)***	(0.094)***	(0.093)***
Different state, neighbors	1.605	1.591	2.345
	(0.105)***	(0.102)***	(0.222)***
Same state; neighbors	2.076	2.057	2.420
	(0.068)***	(0.068)***	(0.102)***
Same state; not neighbors	0.930	0.927	0.906
	(0.053)***	(0.053)***	(0.098)***
Attraction Index		0.182	0.161
		(0.051)***	(0.049)***
Attraction Index * Different states, neighbors			-0.576
			(0.157)***
Attraction Index * Same state, neighbors			-0.286
			(0.072)***
Attraction Index * Different state, not neighbors			0.023
			(0.067)
R^2	0.72	0.72	0.73
N	329,460	329,460	329,460

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Huber-White robust standard errors (clustered by origin district) are reported in parentheses. All specifications include origin fixed effect and destination fixed effect. The sample is restricted to district pairs with non-missing district attributes including percentage non-ST/SC in population, literacy rate, urban population share, share of private employment, share of formal sector, and average income. Dependent variable is the bilateral migration stock, m_{ij}, of inter-district migration of both males and females. See text for definition of 'Attraction Index'. Construction of other variables follow Tables 5 - 9.

Table 10: District migration gravity estimation – alternative distance measures

Sample	All				Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share of Common Language	0.690	0.929	0.789	1.023	0.758	0.979	0.841	1.088	0.690	0.943	0.798	1.021
	(0.128)***	(0.150)***	(0.123)***	(0.140)***	(0.173)***	(0.209)***	(0.167)***	(0.193)***	(0.104)***	(0.118)***	(0.101)***	(0.112)***
Different state, neighbors	1.729	2.187	1.856	2.155	1.305	1.628	1.401	1.603	1.849	2.419	2.005	2.356
	(0.149)***	(0.210)***	(0.144)***	(0.221)***	(0.156)***	(0.250)***	(0.150)***	(0.256)***	(0.136)***	(0.186)***	(0.133)***	(0.194)***
Same state; neighbors	2.125	2.572	2.235	2.590	1.703	1.993	1.767	2.004	2.218	2.799	2.367	2.793
	(0.078)***	(0.118)***	(0.076)***	(0.128)***	(0.074)***	(0.104)***	(0.078)***	(0.112)***	(0.085)***	(0.119)***	(0.082)***	(0.126)***
Same state; not neighbors	1.029	1.197	1.067	1.092	1.198	1.291	1.221	1.182	0.913	1.141	0.965	1.020
	(0.095)***	(0.118)***	(0.091)***	(0.139)***	(0.092)***	(0.121)***	(0.090)***	(0.147)***	(0.091)***	(0.109)***	(0.087)***	(0.127)***
$l.dist_{ij}$, geo. centroids	-1.492				-1.412				-1.590			
	(0.104)***				(0.116)***				(0.092)***			
$l.dist_{ij}$, economic centers		-1.171				-1.168				-1.199		
		(0.126)***				(0.156)***				(0.106)***		
$l.TravelTime_{ij}$, flat			-1.413				-1.359				-1.489	
			(0.097)***				(0.111)***				(0.084)***	
$l.TravelTime_{ij}$				-1.403				-1.389				-1.465
				(0.151)***				(0.182)***				(0.126)***
p-value: Same.nbr = Diff.nbr	0	0	0	0	.01	.06	.02	.03	0	0	0	0
R^2	0.32	0.31	0.32	0.29	0.26	0.25	0.26	0.21	0.43	0.40	0.43	0.40
N	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640	341,640

$* \ p < 0.1; ** \ p < 0.05; *** \ p < 0.01$

Notes: Huber-White robust standard errors (clustered by origin district) are reported in parentheses. All specifications include origin fixed effect and destination fixed effect. The sample includes 329,460 observations, which accounts for bilateral migration stock between 585 origin and 584 destination districts in 2001. Dependent variable is the bilateral migration stock, m_{ij}, of inter-district by gender. Construction of other variables follow Tables 5 - 9. Four measures of distance are used for this robustness check. '$l.dist_{ij}$, geo. centroids' is the geodesic (flight) distance between the **geographic** centers of district i and j – it is the same distance measure used in Tables 5 - 9. Alternatively, '$l.dist_{ij}$, economic centers' calculates the flight distance between the **economic** centers of district i and j. '$l.TravelTime_{ij}$' takes into account India's transport network (national highways and the GQ), and measures the driving time on the shortest path between the economic centers of i and j. See Alder, Roberts, and Tewari (2017) for more details on the method of computing the shortest paths. '$l.TravelTime_{ij}$, flat' assumes the same driving speed on and off the roads.