

Costs and Benefits of Rural-Urban Migration : Evidence from India

Clément Imbert and John Papp

[This paper also appears as CAGE Discussion Paper No: 459](#)

March 2020

No: 1253

[Published in the Journal of Development Economics](#)

Warwick Economics Research Papers

ISSN 2059-4283 (online)

ISSN 0083-7350 (print)

Costs and Benefits of Rural-Urban Migration: Evidence from India *

Clément Imbert[†] and John Papp[‡]

This version: August 2019. First version: June 2012.

Abstract

This paper provides new evidence on rural-urban migration decisions in developing countries. Using original survey data from rural India, we show that seasonal migrants prefer to earn 35 percent less on local public works rather than incur the cost of migrating. Structural estimates suggest that the fixed cost of migration is small, and can be entirely explained by travel costs and income risk. In contrast, the flow cost of migration is very high. We argue that higher living costs in the city explain only a small part of the flow cost of migration, and that most of it is non-monetary.

Keywords: Internal Migration, Workfare Programs, India, Urban, Rural.

JEL Classification: H53, J22, J61, O15, R23

*Thanks to Sam Asher, Sam Bazzi, Gharad Bryan, Robin Burgess, Stefano Caria, Anne Case, Angus Deaton, Taryn Dinkelman, Dave Donaldson, Esther Duflo, James Fenske, Erica Field, Jessica Goldberg, Doug Gollin, Catherine Guirkinger, Viktoria Hnatkovska, Seema Jayachandran, Michael Keane, Reetika Khera, Marieke Kleemans, Julien Labonne, Karen Macours, Mushfiq Mobarak, Melanie Morten, Kaivan Munshi, Rossa O’Keeffe-O’Donovan, Rohini Pande, Simon Quinn, Mark Rosenzweig, Chris Woodruff, as well as seminar and conference participants at the Barcelona Summer Forum, Migration and Development Conference, NEUDC, Paris School of Economics, University of Kent, University of Oxford, University of Warwick, and World Bank for very helpful comments. John Papp gratefully acknowledges financial support from the Fellowship of Woodrow Wilson Scholars, the National Science Foundation, the Industrial Relations Section at Princeton, the Russell Sage Foundation, and the Center for Economic Policy Studies at Princeton.

[†]University of Warwick, Department of Economics, Coventry, UK, c.imbert@warwick.ac.uk.

[‡]R.I.C.E., johnhpapp@gmail.com

1 Introduction

The recent economic literature has documented large average labor productivity gaps between rural and urban areas of developing countries (Gollin et al., 2014; Young, 2013). The source of these gaps, which has important implications for economic growth and development policy, is hotly debated. One possible explanation is that mobility frictions, such as transportation costs, prevent workers from being optimally allocated across sectors (Gollin et al., 2014; Bryan and Morten, 2017). Another is that the most productive workers sort into urban areas, so that despite large average wage gaps returns to migrating for the marginal worker are close to zero (Young, 2013; Hicks et al., 2017).

Providing evidence on the source of rural-urban labor productivity gaps is challenging. First, rural and urban workers are different along observable and unobservable characteristics, so that rural workers may not gain from moving to the city even if urban workers earn more. Second, rural and urban workers do different types of jobs (e.g. farming and factory work), which makes it hard to compare labor productivity in the village and in the city. Third, living costs are usually higher in urban than in rural areas, so that nominal comparisons may be misleading. Finally, rural workers who move to urban areas may have to pay a risk premium, e.g. they face the risk of being unemployed (Harris and Todaro, 1970).

To shed light on the nature of rural-urban wage gaps, we exploit original survey data on seasonal migrants in India who choose between construction work in the city and employment on public works in the village. This unique context allows us to observe the same person doing the same type of work in the same season either in the city or in the village. We measure the value of time in the village and in the city by comparing daily migration earnings and daily wages on public works. We build deflators based on migrants' actual living conditions at destination. And we compute the distribution of potential earnings and the implied risk premium, based on variation in seasonal migration earnings within worker across years.

We first provide reduced form evidence that when offered employment on local public works, most seasonal migrants choose to stay in the village for a 35 percent lower wage. This suggests rural-urban wage gaps may be large for a given worker, but migration costs are larger. We then use a structural approach to quantify these costs. We show that the fixed cost of migration is small: neither transportation costs nor unemployment risk are significant barriers to migration. Higher living costs only have a small impact on migrants, who consume little in the city. We argue that a large part of migration costs is non-monetary: the disutility of harsh living and working conditions in the city is the main barrier to migration.

The study of seasonal migration requires dedicated survey data: we use original data collected in a high out-migration area located at the border of three Indian states (Rajasthan, Madhya Pradesh and Gujarat).¹ Our analysis proceeds in two steps. In the first step, we exploit variation in employment provision under India’s rural workfare program, the National Rural Employment Guarantee Act (NREGA) across seasons and states to estimate its effect on seasonal migration. We argue that this variation reflects exogenous constraints on employment provision, rather than differences in demand for NREGA work. We find that availability of NREGA work has a strong negative effect on seasonal migration: in Rajasthan during the summer, the average adult worked eight more days on local public works and migrated seven fewer days. The migration response is present at the extensive margin (workers are 5.5 percentage points less likely to migrate) and the intensive margin (migration trips are shorter by seven days).

Utility costs associated with migration need to be large for migrants to prefer NREGA work which pays 35 percent less than daily earnings outside of the village. In the second step, we estimate a structural model which includes both a flow cost and a fixed cost of migration to fit the reduced form evidence. The estimated fixed cost of migration, on the one hand, is relatively small, about 7 percent of total migration earnings. Travel costs reported by migrants make for a quarter of this cost. We compute a risk premium, based on the variance of earnings within migrant across years, and find that income risk may explain the rest of the fixed cost. The estimated flow cost of migration, on the other hand, is large, about 80 percent of daily migration earnings. Using detailed information on migrants’ trips, we find that higher living costs at destination can only explain a small share of the flow cost. We argue that non-monetary costs of living and working in the city must play an important role in migrants’ decision to stay in the village. These costs seem to be higher for more educated and older migrants, and for migrants who went to destinations where crime is high.

This paper contributes to the literature in three ways. First, we measure net benefits of migration for the same worker performing a similar task in and outside of the village. This helps overcome selection issues which plague the debate on rural-urban wage gaps in developing countries. Some authors interpret average differences in real wages, or productivity per worker between rural and urban areas as evidence of significant barriers to migration ([Gollin et al., 2014](#); [Bryan and Morten, 2017](#)). In contrast, [Young \(2013\)](#) argues that the entire gap can be explained by the fact that production in urban areas is more skill intensive and attracts more skilled workers. [Hicks et al. \(2017\)](#) also find little income gains for workers who

¹The data was collected by Diane Coffey, John Papp and Dean Spears ([Coffey et al., 2015](#)).

settle down in urban areas. Our contribution is to show that the same worker in the same season can earn 35 percent more on urban private construction sites than on local public works but still prefers to stay in the village. We find that the disutility of migration is high, and that differences in living costs and income risk are a small part of it. This suggests that rural-urban wage gaps do exist in expectation and in real terms, and rural workers choose not to take advantage of them.

Second, we use demand for employment on public works among migrants to shed light on the determinants of migration decisions in developing countries. The literature highlights the importance of financial constraints, and reliance on village-based insurance networks (Angelucci, 2015; Bazzi, 2017; Kleemans, 2015; Munshi and Rosenzweig, 2016; Morten, 2019). Bryan et al. (2014) argue that the risk of failed migration and the lack of information prevent rural workers from taking full advantage of seasonal migration opportunities, which could bring significant income and consumption gains. Our contribution is to show that in a context where information on migration opportunities is widespread, migrants prefer to stay back and shorten their trips for a much lower wage on local public works. A structural estimation shows that migration costs need to be large to rationalize this finding, and that these costs are variable rather than fixed. We find that higher living costs at destination, travel costs, and even the risk associated with migration earnings are a small part of migration costs in this context. We argue that leaving the village has a high non-monetary cost for migrants, which could reflect a preference for staying in the village (“home bias”) or difficult living and working conditions at destination.²

Third, we present new causal evidence on the effect of workfare programs on private sector employment. Workfare programs are a popular form of safety nets, present in 94 countries in 2014 (The World Bank, 2015). The existing evidence is mixed and focuses on local impacts (Zimmermann, 2012; Imbert and Papp, 2015; Muralidharan et al., 2017; Berg et al., 2018).³ Some studies have argued that India’s NREGA might provide an alternative to seasonal migration (Jacob, 2008; Ashish and Bhatia, 2009; Morten, 2019). This paper provides the first causal evidence of NREGA’s impact on rural-to-urban migration using a dedicated survey and a border discontinuity design. We show that workfare programs operating during the agricultural off-season may have a significant negative impact on employment outside of the village. In a follow-up paper, we build on this finding and study the spatial equilibrium

²In a recent re-evaluation of the Bryan et al. (2014) experiment, Lagakos et al. (2018) also argue that the welfare gains from migration may be limited by a large non-monetary cost of migration.

³India’s NREGA being the largest workfare program in the world, has attracted more attention, but (Beegle et al., 2017) and Alik-Lagrange et al. (2017) study public works programs in other contexts.

effects of the NREGA on urban labor markets across India (Imbert and Papp, 2016).⁴ Since migrants forgo much higher earnings in the city, our results imply that the program *reduces* household income. Hence, a complete welfare analysis of the program needs to include not only forgone earning opportunities in the private sector (Imbert and Papp, 2015, for example), but also the non-monetary costs of these alternatives.

The following section describes the workfare program and presents the data set used in the paper. Section 3 uses variation in public employment provision across states and seasons to estimate the impact of the program on short-term migration. Section 4 uses a structural model of migration decisions and detailed information on migration trips to provide evidence on the costs and benefits of seasonal migration.

2 Context and data

2.1 NREGA

This paper studies India’s National Rural Employment Guarantee Act (NREGA), which entitles every household in rural India to 100 days of work per year at a state-level minimum wage. The NREGA is the largest workfare program in the world: in 2016-17 it provided 2.36 billion person-days of employment to 51 million households.⁵ For identification, this paper will use variation in NREGA work across states and seasons to estimate its effect on seasonal migration. We provide here some general information about the source of this variation.

NREGA implementation is highly heterogeneous across states (Dreze and Khera, 2009; Dreze and Oldiges, 2009). As Figure 1 shows, the number of days spent on public works by the average rural adult ranges from almost zero in Haryana (HR) to 12 in Andhra Pradesh (AP). Days spent on public works also vary widely across the three states of our study: Rajasthan (RJ) provides 11 days of public works employment per adult, Madhya Pradesh (MP) 2.6 days, and Gujarat (GJ) 1.4 days.⁶ Imbert and Papp (2016) show that cross-state differences in NREGA implementation cannot be explained by differences in socio-economic characteristics. Dutta et al. (2012) argue that they do not reflect underlying demand for NREGA work, but some combination of political will, administrative capacity, and previous

⁴As compared to this paper, Imbert and Papp (2016) rely on NSS data, which has the advantage of being nationally representative, but lacks important information on migration trips, including duration, location, unemployment and earnings, and hence does not allow one to study migration costs.

⁵Figures are from the official NREGA website nrega.nic.in.

⁶Authors’ calculations based on the National Sample Survey Organization (NSS) Employment-Unemployment survey Round 66.

experience in providing public works.

Employment provision under the NREGA also varies within the year. Public works are often closed at the time of the monsoon (July) and reopen after the main harvest (December). As a result, most NREGA employment is provided during the first half of the year. The seasonality of NREGA works is driven by both practical and political considerations. Most NREGA works are construction projects, which are difficult to carry out during the heavy monsoon rains. Also, local governments in charge of NREGA implementation tend to avoid competing with demand for work in agriculture ([Association for Indian Development, 2009](#)). As Figure 2 shows, in 2009-10, public employment per rural household is the lowest between July to September, which is the peak agricultural season, and highest between April and June, which is the agricultural off-season.⁷

Work under the act is short-term, often on the order of a few weeks per adult. Households with at least one member employed under the act during agricultural year 2009-10 report a mean of only 38 days of work and a median of 30 days for *all* members of the household during that year, which is well below the guaranteed 100 days. Work under the program is rationed ([Dutta et al., 2012](#)): during the agricultural year 2009-10, an estimated 19 percent of Indian households reported attempting to get work under the act without success.⁸ The rationing rule is at the discretion of local officials: workers are recruited rather than applying for work ([The World Bank, 2011](#)). Also, since work is provided only to households who are registered in the village council (Gram Panchayat), workers cannot migrate to another village - let alone another state - to participate in the NREGA.

2.2 Survey Design

Our analysis draws from a survey collected in 2010 by Diane Coffey, John Papp and Dean Spears ([Coffey et al., 2015](#)). Appendix Figure 3 shows the location of the 70 sample villages. The selection of sample villages proceeded in three steps. First, we selected four neighboring districts: one in Rajasthan, one in Gujarat and two in Madhya Pradesh. The survey location was chosen because previous studies in the area reported high rates of out-migration and poverty ([Mosse et al., 2002](#)), and because surveying along the border of the three states provided variation in state-level policies. Second, we matched villages in Rajasthan with villages across the border in Gujarat and Madhya Pradesh based on seven criteria measured in the 2001 census: distance, fraction of Scheduled Castes (SC), fraction of Scheduled Tribes

⁷ Authors' calculations based on the NSS Employment-Unemployment survey Round 66.

⁸ Author's calculations based on the NSS Employment-Unemployment Survey Round 66.

(ST), cultivated area, irrigated and non-irrigated cultivated area and population per cultivated area.⁹ Finally, we selected the 25 best matches along the Madhya Pradesh border and the 10 best matches along the Gujarat border to be part of the survey sample.¹⁰ As Panel A of Table 1 shows, this procedure ensured that village pairs were well balanced along these dimensions.

The survey itself consisted of three modules: village, household, and adult modules. The household module was completed by the household head or other knowledgeable member. One-on-one interviews were attempted with each adult aged 14 to 69 in each household. The analysis in this paper focuses mostly on those adults who completed the full one-on-one interviews. In order to maximize response rates, the survey was carried out between the end of June and the beginning of September, which is the time when migrants come back for the start of the agricultural peak season, and multiple visits were made to households whose members were away. Out of 2,722 adults aged 14-69, we were able to complete face-to-face interviews with 2,224 (81.7 percent). For the 498 remaining individuals, the head of the household provided the information about their employment and migration spells in the last year. The adults that we were unable to survey personally are different from adults that were interviewed: they were much more likely to have migrated all three seasons of the year, and much less likely to have ever done NREGA work (see Appendix Table A.1). We choose not to use the information on these 498 individuals in our main specification to maximize the accuracy of our estimates, but include it later as a robustness check. The results on the full 2,722 sample are similar both qualitatively and quantitatively (see Appendix Table A.7).

To assess how the adults in our sample compare with the rural population in India, Column 5 in appendix Table A.1 presents means from the rural sample of the nationally representative NSS Employment-Unemployment Survey. Literacy rates are substantially lower in the study sample compared with India as a whole, reflecting the fact that the study area is a particularly poor area of rural India. The NSS asks only one question about short-term migration, which is whether an individual spent between 30 and 180 days away from the village for work within the past year. Based on this measure, adults in our sample are 28 percentage points more likely to be short-term migrants than adults in India as a whole. Part of this difference may be due to the fact that the survey instrument was specifically designed to pick up short-term migration, though most of the difference is more likely due to the fact that the sample is drawn from a high out-migration area. Column 6 in Table A.1 shows the

⁹Village characteristics used for matching were measured in the 2001 census, before the NREGA.

¹⁰The best matches were pairs with the lowest sum of squared distances over the seven criteria.

short-term migration rate is 16 percent for the four districts chosen for the migration survey according to NSS, which is half the mean in sample villages (30 percent) but well above the all-India average (2 percent).

The survey instrument was specifically designed to measure migration, cultivation, and participation in the NREGA, which are all highly seasonal. Respondents were much more familiar with seasons than calendar months: summer is mid-March through mid-July, the monsoon season is mid-July through mid-November, and winter is mid-November through mid-March. The survey was implemented at the end of the summer and beginning of the monsoon 2010, when most seasonal migrants come back to work in agriculture. Surveyors asked retrospective questions to each household member about each activity separately for summer 2010, winter 2009-10, monsoon 2009, and summer 2009.

2.3 Demand and Supply of NREGA Work

Two important variables for the following analysis are whether an individual worked for the NREGA, and whether they would have liked to do more NREGA work during a particular season. Specifically, we asked “if more NREGA work were available during [season] would you work more?” for individuals who had worked for the NREGA. For individuals who did not work for the NREGA, we asked “did you want to work for the NREGA during [season]?” One should be skeptical that the answer to these questions truly indicates a person’s willingness to work. Column 1 in Appendix Table A.2 shows that the correlations between the response to the resulting measure of demand and respondent characteristics are sensible. Across all seasons, demand for NREGA was lower for adults with secondary education and for older respondents, but higher for adults who were married and came from poorer families (kuccha roof). Members of families who depend on farming were less likely to declare that they wanted NREGA work during the peak agricultural season (monsoon). Other work opportunities, studies, and illness were the the main reasons given by respondents who did not want NREGA work.

Table 2 presents NREGA demand and participation across states and seasons. As expected, demand for NREGA work was higher during the summer (80 percent) and the winter season (75 percent), when work opportunities are scarce, and lower during the peak agricultural season (54 percent). Demand was uniformly high across states: 79 percent of respondents in Gujarat, 82 percent in Madhya Pradesh and 83 percent in Rajasthan said they would have liked to do NREGA work during the summer 2009. Strikingly, actual NREGA work provision did not follow demand patterns. Most NREGA work was provided in the

summer (11 days per adult), and very little in the winter (1 day per adult), when demand was equally high. During the summer 2009, Rajasthan villages provided 16 NREGA days per adult, as compared to seven days in villages just across the border, where demand was just as high. Hence, as in the rest of India, demand for NREGA work in the survey sample was heavily rationed: the proportion of respondents who said they wanted work but did not work ranged from 33 percent in Rajasthan to 69 percent in Gujarat in the summer, and was highest in the winter (75 percent in all states). Survey respondents mentioned the closure of worksites and the inaction of village officials as the main reasons for why they they did not work for NREGA when they wanted to, suggesting the restriction was supply-related rather than demand-related (see Appendix Table A.3). The assumption that variation in NREGA work is exogenous to local economic conditions will be crucial for our identification strategy.

2.4 Migration Patterns

In the survey, migration is defined as spending more than two nights away from the village for work or looking for work. Table 2 presents migration rates across states and seasons. Migration is high during the summer, from mid-March to mid-July (35 percent), low during the monsoon, from mid-July to mid-November, which is also the peak agricultural season (10 percent), and high again in the winter, from mid-November onward (29 percent). Column 2 in Appendix Table A.2 presents the correlates of migration. Migrants are more likely to be male, they are younger and have lower education than the average respondent. They are also more likely to come from families that are poorer, have a house with earthen floor, have less land, and whose main income source is not farming.

In order to assess the costs of migration, we require reliable information on what migrants do and how much they earn. Given the short-term nature of most migrant jobs, the same migrant might work for multiple employers for different wages and make multiple trips within the same season. For this reason, the survey instrument included questions about earnings, wages, and jobs for each trip within the past four seasons up to a maximum of four trips. In total, this yields detailed observations for 2,749 trips taken by 1,125 adults.¹¹ In 95 percent of the cases, migrants report only one trip per season. For those who have done multiple trips, we calculate the average trip characteristic for each migrant for each season that the migrant was away, using trip duration as weight. Finally, we take into account the possibility

¹¹One important downside of asking respondents about their four most recent trips, is that we do not have detailed information on trips done in earlier seasons for people who migrated more than four times in the last year. This is the case of 200 adults (out of 768) who migrated in the summer 2009. For them we only know that they migrated, with whom and for how long.

that migrants do not always find work at destination by using earnings per day away, rather than earnings per day worked as our main measure of migration returns.¹²

Columns 1 to 3 in Table 3 present descriptive information about short-term migration trips based on the survey. Most short-term migrants travel with a family member (71 percent in the summer). They travel relatively long distances (300km on average during the summer) for a relatively cheap fare (Rs. 116 for the summer), and a large majority works in construction in urban areas in a different state. Figure 4 illustrates the remarkable geographical spread of their trips. The figure also shows that migrants from Rajasthan travel to the same destinations as migrants from the other two states: this will be important in our analysis, because it suggests that migrants from different states face the same labor demand (in the private sector). Employer-employee relationships are often short-term: only 37 percent of summer migrants knew their employer or labor contractor before leaving the village. Living arrangements at destination are rudimentary, with about 85 percent of migrants reporting having no formal shelter (often a bivouac on the work-site itself). Most migrants do only one trip, and they stay on average 50 to 60 days - half of the season - at destination. Table 3 shows that migrants are close to full employment. They work on average six days per week spent at destination. They earn on average Rs. 5,666 in total, that is Rs. 118 per day worked or Rs. 101 per day away, which means that their travel cost is recouped in one day of work. Column 4 presents national averages and Column 5 presents averages in the four districts of the survey sample according to NSS data. Migration patterns are similar along the few dimensions measured in both surveys.

2.5 Effect of NREGA on Migration: Descriptive Evidence

Before providing causal evidence of the program effect on migration, we describe the correlation between demand for NREGA work, program participation and short-term migration (Table 2). Adults who migrated during the summer 2009 were less likely to have worked for NREGA than the whole sample (35 percent against 40 percent), but they were also more likely to say that they would have liked to do NREGA work during that season without being able to (55 percent against 42 percent). Hence, on the whole, demand for NREGA work was higher among migrants (89 percent against 82 percent), which is consistent with the idea that the NREGA competes with seasonal migration. Table 2 also provides descriptive evidence that higher NREGA work provision is associated with lower migration. The proportion of adults who declared they stopped migrating because of NREGA in the sum-

¹²Appendix A describes the construction of the earnings measures in more detail.

mer increases from 3 percent in Gujarat to 8 percent in Madhya Pradesh and 10 percent in Rajasthan (Panel A). This fraction is negligible in the other seasons (Panel B and C). In the following sections, we use variation in NREGA employment provision across states and seasons to estimate the causal effect of the program on seasonal migration and to quantify migration costs.

3 Program effect on migration: reduced form evidence

3.1 Empirical Strategy

In order to estimate the impact of the NREGA on days spent outside the village we exploit the variation in program implementation across states and compare Rajasthan with Gujarat and Madhya Pradesh. We also take advantage of the seasonality of public employment provision and compare the summer months, when most employment is provided, to the rest of the year. The estimating equation is:

$$Y_{is} = \alpha + \beta_0 Raj_i + \beta_1 Sum_s + \beta_2 Raj_i * Sum_s + \gamma_0 \mathbf{X}_i + \gamma_1 \mathbf{X}_i * Sum_s + \varepsilon_{is} \quad (1)$$

where Y_{is} is the outcome for adult i in season s , Raj_i is a dummy variable equal to one if the adult lives in Rajasthan, Sum_s is a dummy variable equal to one for the summer season (mid-March to mid-July) and X_i are controls. The vector X_i includes all worker characteristics (gender, age, education, marital status, language skills), households characteristics (size, religion, caste, wealth), and village controls (population and connectivity) listed in Table 1, as well as village pair fixed effects. We interact the vector of controls X_i with the summer season dummy, to allow worker, household or village characteristics to affect outcomes differently in the summer and in the other seasons. Standard errors are clustered at the village level.

In order for β_2 to identify the impact of the NREGA, villages in Rajasthan need to be comparable with their match on the other side of the border in all respects other than NREGA implementation. Potential threats to our identification strategy include differences in socio-economic conditions or state policies other than NREGA. Table 1 compares village, households and individual characteristics in villages in Rajasthan with matched villages in Gujarat and Madhya Pradesh and village pairs in Rajasthan. SC/ST population, land use are balanced by design (Panel A). As Panel B shows, other village characteristics, such as population size, access with paved road and distance to town are also balanced. However, villages in Rajasthan are less likely to have education facilities, or benefit from bus services

(Panel B), and households are less likely to have electricity (Panel C). This suggests that the policy mix in Rajasthan may include better NREGA implementation but lower infrastructure provision. We will control for these variables in the analysis. Socio-economic characteristics are very comparable otherwise on either side of the Rajasthan border, a half depends on farming for subsistence, with many households living in houses with dirt floor and thatch roof (Panel C). Literacy rates (Panel A) and education levels (Panel D) are also very similar.

3.2 Results

Table 4 presents our results on the causal effect of the program on migration. We first use days worked for the NREGA in each season as an outcome and estimate Equation 1. Column 1 confirms that no public employment was provided outside of the summer months, and there is no difference between Rajasthan and the other states. During the summer 2009, adults in Madhya Pradesh and Gujarat, worked about seven days for the NREGA. The coefficient on the interaction of Rajasthan and summer suggests that in Rajasthan nine more days of public employment were provided per adult. The estimated coefficient declines slightly after including controls and village pair fixed effects (Column 2).

Columns 3 and 4 in Table 4 repeat the same analysis with days spent outside the village for work as the dependent variable. The average adult spent 11 days away for work during the monsoon and the winter 2009, and there is again no difference between Rajasthan and the other states. In contrast, in the summer 2009 adults in Rajasthan villages spent five fewer days on average working outside the village than their counterparts in Gujarat and Madhya Pradesh, who were away for 24 days on average. The estimated coefficient becomes more negative (-6.6 days) with the inclusion of village pair fixed effects, of village, household and worker controls. Taken together, the estimates suggest that one additional day of NREGA work reduces migration by 0.6 to 0.8 days.

Columns 5-6 and 7-8 in Table 4 decompose the migration response along the extensive and intensive margin, respectively. The probability of migration during the summer 2009 was 6.4p.p. lower in Rajasthan villages, a 16 percent reduction as compared to the 40 percent migration rate in villages across the border. Conditional on migrating, adults in Rajasthan villages migrated 5.5 fewer days than adults in Madhya Pradesh and Gujarat villages, who migrated 60 days on average. As before, in the seasons when NREGA provision is low, we observe no differences in migration rates (20 percent) or migration duration (53 days) across states. The estimates are very similar once we include controls: the effect on the extensive margin decreases (5.5p.p) and the effect on the intensive margin increases (7 days).

We next perform a number of robustness checks. First, to make sure the identifying variation comes from exogenous variation in supply of NREGA work and not from endogenous variation in demand, we add to X_i a dummy variable for whether the adult did or wanted work during the summer 2009. The estimates remain perfectly stable (see Appendix Table A.4). Second, to ensure that workers in Rajasthan are comparable to workers on the other side of the border, we estimate a propensity score for living in Rajasthan based on all the controls from Table 1 and restrict the sample to observations on the common support. As Appendix Table A.5 shows, differences in NREGA days worked become slightly smaller, and as a result the negative effect on migration duration is a bit smaller, but the results overall are very similar. Third, we estimate the same specification without the pairs that include Gujarat villages, which have starkly different levels of NREGA provision and may hence be less comparable with Rajasthan villages. The magnitude of the coefficients of interest increases slightly (see Appendix Table A.6). As a fourth robustness check, we repeat the same analysis including adults who were not interviewed personally but about whom information was collected from the household head. The coefficients in Appendix Table A.7 are again similar, just slightly smaller: adults who were not interviewed personally are less likely to participate in NREGA, and hence their migration behavior is less affected.

Despite the robustness of our results, one may still worry that differences we observe in migration patterns between Rajasthan villages and the others might be due to preexisting differences unrelated to the NREGA. To alleviate this concern, we provide two additional pieces of evidence. First, we compare the number of long-term migrants across states, i.e. individuals who changed residence and left the household in the last five years. The results in Appendix Table A.8 show that there is no difference in permanent migration levels between Rajasthan and the other states. Second, the survey included retrospective questions about migration trips in previous years. Unfortunately, the fraction of respondents who forgot whether they migrated five and six years ago is high: 22 percent for 2005 and 47 percent for 2004. Using non-missing responses however, we find no significant difference in migration levels in 2004 and 2005, i.e. before the NREGA was implemented (see Appendix Figure A.1).

3.3 Nominal Wage Gaps

Our empirical results suggest that migrants shorten their trips, and sometimes stop migrating altogether to work for the NREGA. One would hence expect NREGA work to be more profitable than migration. Table 5 compares migrants' earnings per day spent outside the village with daily wages on NREGA work sites for adults who worked outside of the village

in the summer 2009. The construction of these variables is described in detail in Section 2.4 and Appendix A, but we will emphasize here two important points. First, we use earnings per day migrated, and not per day worked at destination to account for the possibility that migrants do not find work. In practice, however, as we saw in Table 3 migrants in our sample are close to full employment, and work on average six days a week. Second, for migrants who did not actually work for the NREGA in that season, we predict NREGA wages using a linear regression of NREGA daily earnings of those who work on individual, household and village characteristics (see Appendix Table A.9). The mean of predicted and actual NREGA wages are extremely close, which is due to the fact that the NREGA wage is for the most part fixed by state law, and that NREGA participation is driven by supply of work rather than demand factors.

For the average migrant in the summer 2009, earnings outside of the village were 37 percent higher than earnings on NREGA work sites (Column 1). Column 2 presents the wage gap for the minority of migrants who did not want NREGA work in that season. For them the difference between migration and NREGA earnings is much higher than the average (46%), which is consistent with the idea that migrants arbitrage the benefits of migration against the opportunity to do NREGA work in the village. In contrast, the wage gap for migrants who did NREGA work in that season is slightly lower than the average (35 percent), but still very substantial (Column 3). In Column 4, we show the wage gap for migrants who would have likely preferred to stay in the village and do NREGA work rather than migrate, either because they did NREGA work or they said they would have wanted to do NREGA work. The wage gap for them is also 35 percent, the same as for migrants who actually did NREGA work, which is consistent with our argument that NREGA participation is supply and not demand-driven. Overall, these wage patterns suggest that earnings outside of the village are much higher than earnings from NREGA work in nominal terms. For migrants to prefer NREGA work to migrating, migration costs need to be high. The following section quantifies these costs and explore their monetary and non-monetary components.

4 Migration Costs

The empirical evidence presented in the previous section suggests that rural workers prefer to stay and earn a much lower wage rather than go and earn in the city. In this section we develop a simple model of migration choice and NREGA participation. We then estimate migration costs structurally to match observed migration patterns and the reduced form

evidence of the program impact from the previous section.

4.1 Theoretical framework

Let us consider a rural worker during the off-peak season of agriculture (summer), who splits their time T between time spent migrating L_m and time spent at home $T - L_m$. Time spent at home has value $f(T - L_m)$, with $f(\cdot)$ increasing and concave. The value of each day spent migrating is $(w_m - c_v)$ where w_m denotes daily earnings from migration and c_v is a flow cost of migration. c_v represents the costs of living and working at destination. To migrate, workers pay a fixed cost c_f , which represents the cost of travel and migration arrangements (e.g. transportation). Time spent outside the village L_m maximizes:

$$\max_{L_m \in [0, T]} \mathcal{U} = f(T - L_m) + (w_m - c_v)L_m - c_f \mathbb{1}\{L_m > 0\}$$

For any interior solution $L_m > 0$, the optimal period of time spent migrating is L_m^* such that:

$$f'(T - L_m^*) = w_m - c_v$$

The model assumes that the utility function is linear in earnings and that there is no leisure choice. More generally, one could think of $f(T - L_m^*)$ as capturing utility from time spent in the village after the individual has optimally chosen work outside of the village L_m and leisure given a time constraint of T . The model also assumes that migration earnings are a linear function of time spent at destination. One may expect migrants who commit to go for longer periods to earn more per day, or migrants who go for shorter periods to spend time without pay looking for jobs. We check that daily earnings are independent of migration duration in the data: Appendix Figure A.2 shows that this is indeed the case.

There is no explicit role for risk in the model, as earning at home or at destination are perfectly anticipated. This may be a plausible assumption for time spent at home, since we focus on the summer months, in which climatic conditions or other hazards (such as pests) which influence agricultural productivity during the rest of the year are less prevalent. It is however highly unlikely that migration earnings are fixed in advance, given that two thirds of migrants do not even know their employer before leaving the village (see Table 3). A simple way to include risk considerations in the model is to assume that the decision to migrate is based on expected earnings and that actual earnings depend on an idiosyncratic shock realized once migrants are at destination. In that case the utility cost of uncertainty can be

captured by a risk premium, which is part of the fixed migration cost c_f .¹³

We next introduce the employment guarantee scheme, which offers L_g days of employment on public works paying w_g per day. We assume that L_g is small relative to the usual duration of migration trips ($L_g < L_m^*$) and fixed, i.e. workers may choose whether or not to participate in the program but not the number of days they work. These assumptions are consistent with the fact that demand for NREGA work is rationed and that in our survey, adults who worked for NREGA worked on average 28 days, while adults who migrated were away on average 60 days. Under these assumptions, the marginal value of time for workers who migrate and participate in the program remains $(w_m - c_v)$ and their migration duration is $(L_m^* - L_g)$. Workers can now be in one of four cases:

1. Stay at home and do not participate: $\mathcal{U}_{00} = f(T)$
2. Migrate and do not participate: $\mathcal{U}_{10} = f(T - L_m^*) + (w_m - c_v)L_m^* - c_f$
3. Stay at home and participate: $\mathcal{U}_{01} = f(T - L_g) + w_g L_g$
4. Migrate and participate: $\mathcal{U}_{11} = f(T - L_m^*) + (w_m - c_v)(L_m^* - L_g) + w_g L_g - c_f$

As we saw in the previous section, the employment guarantee may affect migration on the intensive and the extensive margin. The intensive margin is driven by workers who both migrate and participate in the program during the same season. For them the daily wage on the program must be higher than the opportunity cost of time ($w_m - c_v$):

$$w_m - c_v < w_g$$

The extensive margin response is driven by workers who would otherwise migrate but decide to stay home and work for the NREGA. For them, NREGA participation reduces migration duration so much that it is no longer profitable to pay the fixed cost of migrating:

$$f(T - L_m^*) + (w_m - c_v)(L_m^* - L_g) - f(T - L_g) < c_f < f(T - L_m^*) + (w_m - c_v)L_m^* - f(T)$$

¹³A more sophisticated model would allow migrants to choose how long they migrate based on actual, and not expected earnings. It would be the case for example if migrants cut short migration trips in which they earn less than expected. This would lead to a positive relationship between migration duration and earnings, which we do not see in the data (Appendix Figure A.2). For this reason, we keep to a simpler formulation and assume that both decisions of whether and how long to migrate are made at the same time.

This is the opportunity cost effect of income shocks on migration documented in the literature (Bazzi, 2017; Kleemans, 2015). Theoretically, there could also be a wealth effect, with NREGA participation allowing workers to pay the fixed cost of migration (Angelucci, 2015). The reduced form evidence suggests that the opportunity cost effect dominates.

4.2 Structural Estimation

We next use our model to estimate the fixed and the variable migration costs. We assume that

$$f(L) = A \log(L)$$

where A is a constant. This implies that the optimal migration duration is

$$L_m^* = \frac{A}{(w_m - c_v)}$$

We also assume that both migration costs c_f and c_v follow a normal distribution.

$$c_f \sim \mathcal{N}(\mu_f, \sigma_f) \quad ; \quad c_v \sim \mathcal{N}(\mu_v, \sigma_v)$$

The mean and standard deviations of these two variables as well as the constant A are the five structural parameters that we estimate. We use the method of simulated moments, in which the five targeted moments are:

- (i) the migration rate in Gujarat and Madhya Pradesh villages.
- (ii) the mean of the migration duration in Gujarat and Madhya Pradesh villages.
- (iii) the standard deviation of the migration duration in these villages.
- (iv) the difference in the migration rate between Rajasthan villages and the others.
- (v) the difference in the mean migration duration between Rajasthan villages and the others.

We also test the robustness of our model and consider a sixth, untargeted moment: the proportion of migrants who do or want NREGA work.

For the purpose of the simulation, we need to draw a migration wage w_m and a NREGA wage w_g for each worker. As migration wage, we use earnings per day spent outside of the village, to account for the possibility that migrants do not find work every day, although

in practice migrants work on average six days a week (Table 3). As NREGA wage, we use earnings per day worked in the program. To predict the migration wage of workers who did not migrate, and the NREGA wage of workers who did not participate in the program we use a linear regression of daily migration earnings on individual, household and village characteristics (see Appendix Table A.9).¹⁴

We also need to simulate the rationing rule, i.e. the way in which NREGA work (L_g) is allocated across workers. Given a draw of w_m , w_g , c_f , c_v and A , the model predicts whether a worker would migrate in the absence of the program. The model also predicts whether they would like to participate in the NREGA, which is whenever the marginal value of their time ($w_m - c_v$ for migrants and $f'(T)$ for non-migrants) is lower than the NREGA wage w_g . We assign to each worker who would like NREGA work a participation probability equal to the participation rate we observe among workers who worked or would have liked to do NREGA work. If the draw is favorable, we assign them a number of days equal to the observed average number of days worked by program participants. We do these steps separately for Rajasthan villages and villages in other states.

For each value of the parameters, the simulated moments are computed by averaging across 100 simulations of the whole sample. We then find the value of the parameters that minimize the squared distance between population moments and simulated moments. To increase precision, moments are weighted by the inverse of their variance across 1,000 bootstraps (each composed of 70 independent draws of entire villages). The estimates are the values that minimize the weighted distance after 100 iterations of the genetic algorithm written for R by (Scrucca, 2013). For inference, we use 100 bootstrapped samples, each composed of 70 independent draws of entire villages. We report standard errors equal to the standard deviation of the estimates over the bootstrapped samples adjusted for the number of bootstraps. We also report 5 percent confidence intervals equal to the 2.5th percentile and the 97.5th percentile of the distribution of bootstrapped estimates.

We estimate the model on simulated data to provide some sense on how and how well the model is identified. First, Appendix Figure A.3 displays the objective function that is minimized, the squared distance between simulated and actual moments weighted by the inverse of their variance, and how it changes when one parameter changes holding the others constant. For each parameter, the objective function is convex, and the model estimates are the global minimum. Second, Appendix Figure A.3 shows how the simulated moments respond to changes in each parameter, holding the others constant. There is not a one-to-one

¹⁴The construction of these variables is described in detail in Section 2.4 and Appendix A.

mapping between moments and parameters, but every moment endogeneously respond when each parameter varies. The migration rate and duration are mostly related to the average costs of migration (flow and fixed) and home productivity. Interestingly, the migration rate is a decreasing function of the flow and the fixed costs (and the home productivity), whereas the migration duration is decreasing in the flow cost but increasing in the fixed cost, due to a selection effect. As expected, the key moments for the identification of all parameters are the difference in the migration rate and the average migration duration between Rajasthan and non-Rajasthan villages.

4.3 Results

Table 6 presents the results of the estimation. The estimated mean of the flow cost of migration is large, Rs. 80, which represents 80 percent of earnings per day spent at destination (Rs. 101 on average). It is precisely estimated: the 5 percent confidence interval ranges from Rs. 76 to 86. In contrast, our estimate of the mean of the fixed cost of migration seems relatively small, Rs. 480 or 8.5 percent of average migration earnings for a whole trip (the average trip duration is 60 days). The confidence interval for the fixed cost is wider, from Rs. 302 to 804, but even the higher bound is only 14 percent of average migration earnings. The estimated standard deviation of these costs are large, which suggests substantial individual heterogeneity: the estimated standard deviation for the fixed cost of migration is 1733, which is more than three times the mean. Finally the constant A in the production function is equal to 1602.

In order to test the goodness-of-fit of the model, columns 2 and 3 of Table 7 compare the empirical moments with the simulated moments based on our main estimates. The only noticeable difference is that the simulated migration rate in Rajasthan villages is slightly higher than the actual migration rate, so that the simulated difference with the other states (-0.61 p.p) is lower than the actual difference (-0.84 p.p). It could be that a richer model that would take into account complementarities in migration decisions between individuals would predict a larger decrease in the migration rate: 71 percent of migrants in the summer 2009 migrated with a household member (Table 3). Overall, however, the fit of the model is very good for the targeted moments as well as for the sixth (untargeted) moment, the high demand for NREGA work among migrants.

Based on our structural estimates, we can also simulate migration rates and migration duration in two counterfactual scenarios. In the first scenario, workers who want NREGA work are offered their legal entitlement of 100 days per household (we divide 100 days by

the number of household members who want NREGA work). As column 4 in table 7 shows, as compared to actual migration patterns, the migration rate declines from 39 percent to 27 percent and the migration duration drops from 61 to 46 days. As a benchmark, we simulate a second scenario in which no NREGA work is provided (column 5). The migration rate is essentially the same as the one observed in non-Rajasthan villages (39 percent), and the migration duration is only slightly higher (66 days). Comparing columns 4 and 5, our counterfactual suggests that if it had been implemented as intended, NREGA would have reduced both migration prevalence and migration duration by a bit less than a third, i.e. it would have reduced total migration days by half.

The identification of our model relies on the choice of a home production function, which we have assumed to be $f(L) = A \log(L)$. To assess the robustness of our estimates to this particular choice, we estimate an alternative model in which assume instead that $f(L) = A\sqrt{L}$. The estimates based on the alternative model are presented in Appendix Table A.10. The point estimates are indeed different: the estimated flow cost mean is Rs. 63 (against Rs. 80 with our main model) and the estimated fixed cost mean is Rs. 280 (against Rs. 480 with our main model). They do not however affect qualitatively our conclusions: the flow cost is large (as compared to daily earnings) and the fixed cost is not (as compared to total earnings). Appendix Table A.10 presents the moments simulated based on the alternative model: the fit is not good for the targeted nor the untargeted moments. Hence, our preferred model seems better specified, and its results seem robust to the choice of the production function.

4.4 Interpretation of the fixed cost of migration

The first surprising finding from our structural analysis is that the fixed cost of migration need not be very large: the mean estimate is Rs. 480, about 7 percent of total migration earnings. We can quantify two components of this fixed cost: transportation costs and the risk premium. Survey responses suggest that transportation costs are indeed small: the average migrant paid Rs. 116 to travel to destination, less than 2 percent of total migration earnings (Table 3). Another source of utility cost associated with migration is income risk: migrants may not find work at destination or may have to work for lower wages than expected. Bryan et al. (2014) argue the risk of failed migration is an important barrier to seasonal migration during the hunger season in Bangladesh.

We can again leverage the richness of the survey data and use information on migration earnings from repeated trips to estimate the idiosyncratic risk migrants are exposed to.

Earnings are defined as earnings per day away, to account for both the employment and wage risks. We restrict the analysis to 435 migrants for whom we have earnings per day away for both summers 2009 and 2010. Average daily earnings in the Summer 2009 are Rs. 100. We then run a regression of earnings per season on season and migrant fixed effects and estimate the standard deviation of the residuals, which is Rs. 25.¹⁵ We next use the estimated mean and variance of migration earnings to compute the relative risk premium, i.e. the amount one would need to give migrants at home to make them indifferent between migrating and not migrating, expressed as a fraction of daily migration earnings. For the sake of the calibration, we assume a simple static relationship between consumption and income, i.e. that migrants consume their income and cannot insure their consumption against income shocks. If we further assume migrants' utility has constant relative risk aversion ρ then the relative risk premium (RPP) can be approximated as a simple function of the mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ of daily migration earnings:

$$RRP \approx \frac{\rho \hat{\sigma}^2}{2 \hat{\mu}^2} \approx \frac{\rho}{32}$$

The risk premium depends on the level of relative risk aversion. Even with very high levels of relative risk aversion $\rho = 5$, the risk premium is only 15 percent of migration earnings. For more moderate levels of risk aversion $\rho = 2$, which [Bryan et al. \(2014\)](#) argue match the evidence on migration decisions relatively well, the risk premium is 6 percent. The risk aversion parameters we consider here usually apply to variations in consumption, which for the sake of the exercise we have assumed to be exactly equal to variations in income. Since in reality migrants have different ways of insuring their consumption against bad income shocks (e.g. using informal risk sharing networks), the risk premium we compute should be considered as an upper bound. Hence, our back-of-the-envelope calculations suggest that the combination of low travel costs and moderate income risk may rationalize the small estimated fixed cost of migration.

4.5 Interpretation of the flow cost of migration

The second striking finding of the structural estimation is that the flow cost of migration is very large. The flow cost represents the utility cost of living and working one day in the

¹⁵Alternatively, we can use cross-sectional variation only and estimate idiosyncratic risk as the standard deviation of the residuals of a regression of daily migration earnings in the Summer 2009 on workers characteristics, migration history and village fixed effects. The estimated standard deviation is Rs. 29, close to, but higher than our preferred estimate.

city relative to one day at home. In particular, it includes differences in living costs: living in urban areas is more expensive than living in the village, and migrants may need to pay for goods they would get for free or cheaply at home. In order to evaluate what fraction of the estimated flow cost of migration can be explained by differences in living costs, we consider three possible deflators for migration earnings. We first follow [Hnatkovska and Lahiri \(2013\)](#) and use the ratio of the urban poverty line in the state of destination to the average rural poverty line in the three states of origin: Gujarat, Madhya Pradesh, and Rajasthan ([Planning Commission, 2009](#)).¹⁶ This deflator is only valid if migrants have the same consumption basket as poor urban residents, which is not true in our context, since migrants expenditures at destination do not include housing (less than 2 percent declare renting a room and 86 percent declare sleeping rough) or durable goods.

We next use the consumer expenditure module of the 2009-10 NSS Employment Unemployment Survey to estimate the share of these expenditures in urban and rural areas of each state. Specifically, we compute the share of food expenditures (S^1), as well as the share of non-durable and non-housing expenditures (S^2) for the average rural and urban household of each state.¹⁷ Let P_r denote the poverty line for households in rural areas of the states of origin and P_u the urban poverty line of the states of destination. The three deflators are:

$$D_1 = \frac{P_u}{P_r}, \quad D_2 = \frac{P_u * S_u^1 + P_r * (1 - S_r^1)}{P_r} \quad \text{and} \quad D_3 = \frac{P_u * S_u^2 + P_r * (1 - S_r^2)}{P_r}$$

For the migrants in our sample, the first deflator is equal to 1.3 (sd 0.12), the second to 1.04 (sd 0.04) and the third to 1.13 (sd 0.09). According to the third deflator, the monetary cost of living in the city is hence equal to 13 percent of migration earnings, or Rs. 13 per day away, which is much smaller than the estimated flow cost of migration (Rs. 80). This suggests that the utility cost of one day away is mostly non-monetary. In the case of housing, migrants seem to trade the monetary cost of accommodation for the non-monetary cost of sleeping rough. The next section will try to provide more evidence on non-monetary costs of migration.

¹⁶The consumer price index for agricultural laborer and for industrial workers collected by the Labour Bureau, which are the official deflators used for rural and urban areas respectively are useful to track inflation but cannot be used for cross-sectional comparison.

¹⁷Appendix Table [A.12](#) provides the full list of items included in these expenditures: non-durable and non-housing expenditures include entertainment, medical expenditures, personal care, conveyance etc.

4.6 Non-monetary costs of migration

Taken together, our findings suggest that migrating and living in the city imply high utility costs, but that a only a small fraction of these costs is monetary. The disutility cost of bivouacking for months in the city and leaving family behind is presumably also important, but harder to quantify. In order to provide evidence on the non-monetary dimension of migration costs, we regress actual or desired NREGA participation on migrant and trip characteristics, controlling for migration and predicted NREGA earnings. Specifically, let D_{id} be a dummy variable set equal to one if individual i who migrated to destination d in the summer 2009 did NREGA work or declared they would have liked to do NREGA work in that season. Let w_m^i and w_g^i denote migration and NREGA earnings respectively. Let \mathbf{X}_i denote a vector of migrant characteristics, including worker and household controls from Table 1. Let \mathbf{Z}_d denote a vector of trip characteristics, including those from Table 3, as well as crime and pollution levels at destination (see Appendix A for more detail). Let μ_S denote state fixed effects. We estimate the following regression through probit, with standard errors clustered at the village level:

$$D_{id} = \alpha_m \log(w_m^i) + \alpha_g \log(w_g^i) + \beta \mathbf{X}_i + \delta \mathbf{Z}_d + \mu_S + \varepsilon_{id}$$

Table 8 presents the probit estimates (marginal effects at the mean). As expected, we find that migrants who had higher earnings were less likely to want NREGA work. Two individual characteristics have significant effects: age and education. We find that migrants with some education were more likely to want NREGA work than workers with no education, which may be due to the fact that the non-monetary costs of migration are more important than the monetary gains for more educated workers. The disutility of migration also increases with age: migrants above 30 were more likely to want NREGA work. Only one household characteristic has a significant effect: the disutility of migration was higher for migrants from households who have electricity at home, which may make time in the village more productive / attractive. Finally, we find some evidence that amenities at destination matter: the utility cost of migration was lower among migrants who went to an urban area, but was higher among those who went to a destination where crime levels were relatively high.

5 Conclusion

This paper provides new evidence on the costs and benefits of rural to urban migration in developing countries. We study seasonal migrants who choose between construction work in the city and on local public works, which provides a unique opportunity to observe the same worker doing similar work in the village and in the city in the same season.

Using supply-side variation in employment provision on local public works across states and seasons, we show that when more employment is available migrants are less likely to leave the village, and make shorter trips to the city. This is despite the fact that daily earnings from the program are 35 percent lower than migrants' earnings per day outside of the village. We estimate migration costs structurally to fit the reduced form. We find that the fixed costs of migration are small (7 percent of total migration earnings): neither transportation costs nor the risk of unemployment are large barriers to migration. The flow cost of migration on the other hand is large (80 percent of daily migration earnings). We show that higher living costs in the city only explain a small share of the flow costs, because migrants consume very little at destination. We argue that the main barrier to migration in this context are non-monetary costs of harsh living and working conditions in the city and the non-monetary value of staying in the village.

Our findings have important implications for development policy. On the one hand, many governments consider that migration has undesirable effects for the migrants themselves, for their communities of origin or their city (or country) of destination, and promote development at origin as an alternative to migration. For example the new European Fund for Sustainable Development's main objective is to "tackle the root causes of migration" from Africa ([European Parliament, 2017](#)). Our results suggest that policies aimed at reducing migration flows do not need to fully compensate workers for the loss in potential migration earnings to convince them to stay at home. On the other hand, economists tend to argue that governments should in fact encourage mobility to urban areas, which would improve the allocation of labor in the economy ([Gollin et al., 2014](#); [Kraay and McKenzie, 2014](#); [Bryan and Morten, 2017](#)). Our results suggest that improving working and living conditions for migrants in urban areas may go a long way towards this aim.

References

- Alik-Lagrange, A., O. Attanasio, C. Meghir, S. Polana-Reyes, and M. Vera-Hernandes (2017). Work pays: Different benefits of a workfare program in colombia.
- Angelucci, M. (2015, March). Migration and Financial Constraints: Evidence from Mexico. *The Review of Economics and Statistics* 97(1), 224–228.
- Ashish, R. and K. Bhatia (2009). Alternative to Migration. *Frontline* 16.
- Association for Indian Development (2009). Key Observations on NREGA work in Andhra Pradesh. Available at <http://aidindia.org/main/content/view/788/272/1/1/>.
- Bazzi, S. (2017). Wealth heterogeneity and the income elasticity of migration. *American Economic Journal: Applied Economics* 9(2), 219–55.
- Beegle, K., E. Galasso, and J. Goldberg (2017). Direct and indirect effects of Malawi’s public works program on food security. *Journal of Development Economics* 128(C), 1–23.
- Berg, E., S. Bhattacharyya, R. Durgam, and M. Ramachandra (2018, 03). Can public works increase equilibrium wages? evidence from india’s national rural employment guarantee. 103, 239–254.
- Bryan, G., S. Chowdhury, and A. M. Mobarak (2014, 09). Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica* 82, 1671–1748.
- Bryan, G. and M. Morten (2017, June). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. NBER Working Papers 23540, National Bureau of Economic Research, Inc.
- Coffey, D., J. Papp, and D. Spears (2015, June). Short-Term Labor Migration from Rural North India: Evidence from New Survey Data. *Population Research and Policy Review* 34(3), 361–380.
- Dreze, J. and R. Khera (2009). The Battle for Employment Guarantee. *Frontline* 26(1).
- Dreze, J. and C. Oldiges (2009). Work in Progress. *Frontline* 26(4).
- Dutta, P., R. Murgai, M. Ravallion, and D. van de Walle (2012). Does India’s employment guarantee scheme guarantee employment ? Policy Research Discussion Paper 6003, The World Bank.

- European Parliament (2017). Eu investment plan on track to tackle root causes of migration.
- Gollin, D., D. Lagakos, and M. E. Waugh (2014). The agricultural productivity gap. *The Quarterly Journal of Economics* 129(2), 939–993.
- Greenstone, M., J. Nilekani, R. Pande, N. Ryan, A. Sudarshan, and A. Sugathan (2015). Lower pollution, longer lives. *Economic and Political Weekly* Vol. 50(8).
- Harris, J. R. and M. P. Todaro (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review* 60(1), pp. 126–142.
- Hicks, J. H., M. Kleemans, N. Li, and E. Miguel (2017). Reevaluating agricultural productivity gaps with longitudinal microdata. Manuscript.
- Hnatkovska, V. and A. Lahiri (2013). Structural transformation and the rural-urban divide. University of British Columbia, typescript.
- Imbert, C. and J. Papp (2015). Labor market effects of social programs: Evidence from india’s employment guarantee. *American Economic Journal: Applied Economics* 7(2), 233–63.
- Imbert, C. and J. Papp (2016). Short-term Migration, Rural Workfare Programs and Urban Labor Markets - Evidence from India. The Warwick Economics Research Paper Series (TWERPS) 1116, University of Warwick, Department of Economics.
- Jacob, N. (2008). The Impact of NREGA on Rural-Urban Migration: Field Survey of Villupuram District, Tamil Nadu. CCS Working Paper No. 202.
- Kleemans, M. (2015). Migration Choice under Risk and Liquidity Constraints. 2015 AAEA & WAEA Joint Annual Meeting, July 26-28, San Francisco, California 200702, Agricultural and Applied Economics Association; Western Agricultural Economics Association.
- Kraay, A. and D. McKenzie (2014). Do poverty traps exist? assessing the evidence. *Journal of Economic Perspectives* 28(3), 127–48.
- Lagakos, D., A. M. Mobarak, and M. E. Waugh (2018, January). The Welfare Effects of Encouraging Rural-Urban Migration. NBER Working Papers 24193, National Bureau of Economic Research, Inc.

- Morten, M. (2019). Temporary migration and endogenous risk sharing in village india. *Journal of Political Economy* 127(1), 1–46.
- Mosse, D., S. Gupta, M. Mehta, V. Shah, and J. Rees (2002). Brokered Livelihoods: Debt, Labour Migration and Development in Tribal Western India. *Journal of Development Studies* 38(5), 59–88.
- Munshi, K. and M. Rosenzweig (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review* 106(1), 46–98.
- Muralidharan, K., P. Niehaus, and S. Sukhtankar (2017, September). General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India. NBER Working Papers 23838, National Bureau of Economic Research, Inc.
- Planning Commission (2009). Report of the expert group to review the methodology for estimation of poverty. Technical report, Government of India.
- Prakash, N., M. Rockmore, and Y. Uppal (2014, November). Do Criminally Accused Politicians Affect Economic Outcomes? Evidence from India. HiCN Working Papers 192, Households in Conflict Network.
- Scrucca, L. (2013). Ga: A package for genetic algorithms in r. *Journal of Statistical Software, Articles* 53(4), 1–37.
- The World Bank (2011). Social Protection for a Changing India. Washington, DC.
- The World Bank (2015). The state of safety nets 2015. Technical report, The World Bank.
- Young, A. (2013). Inequality, the Urban-Rural Gap, and Migration. *The Quarterly Journal of Economics* 128(4), 1727–1785.
- Zimmermann, L. (2012). Labor market impacts of a large-scale public works program: Evidence from the indian employment guarantee scheme. Technical Report 6858, IZA Discussion Papers.

Figure 1: Cross-state variation in public employment provision

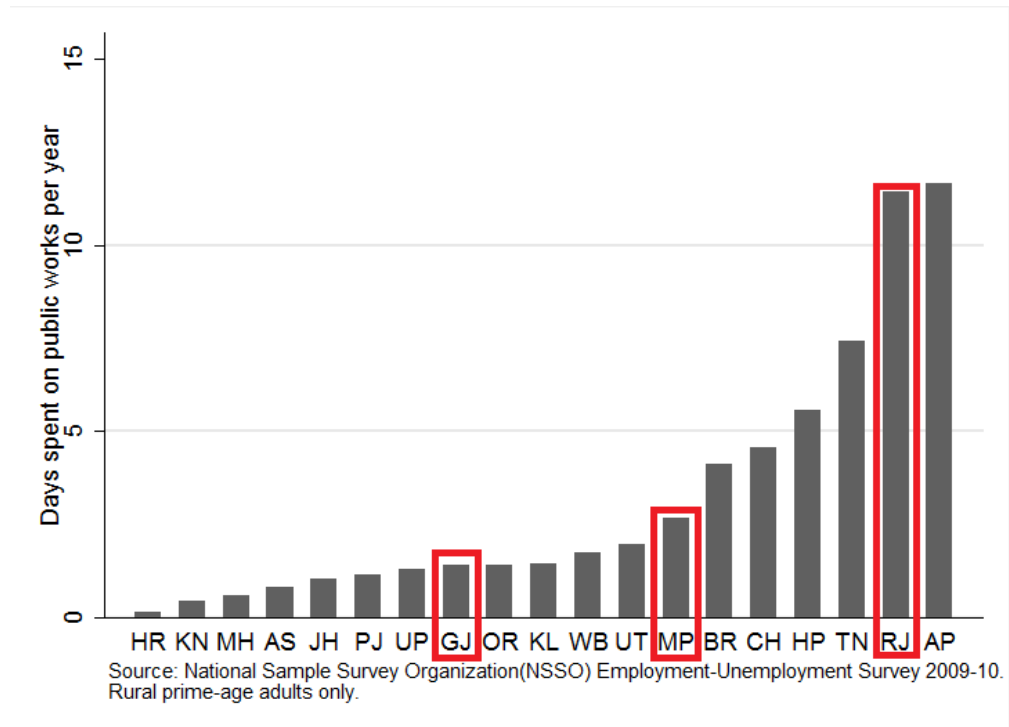


Figure 2: Seasonality of public employment provision

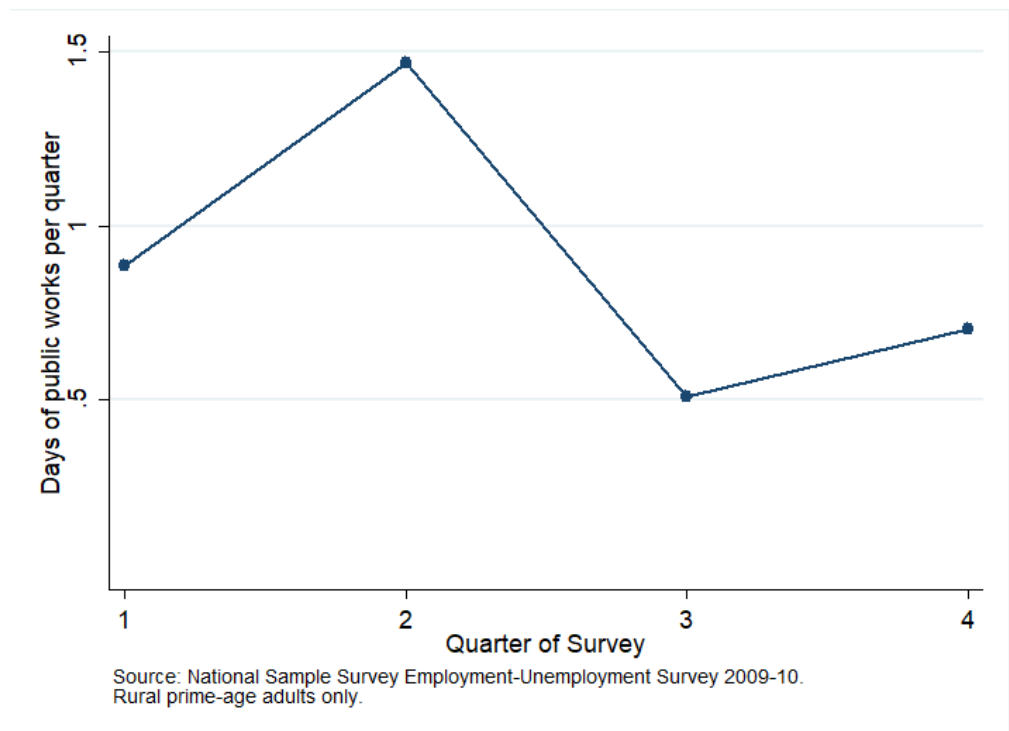


Figure 3: Seasonal out-migration rates in India and in the survey sample

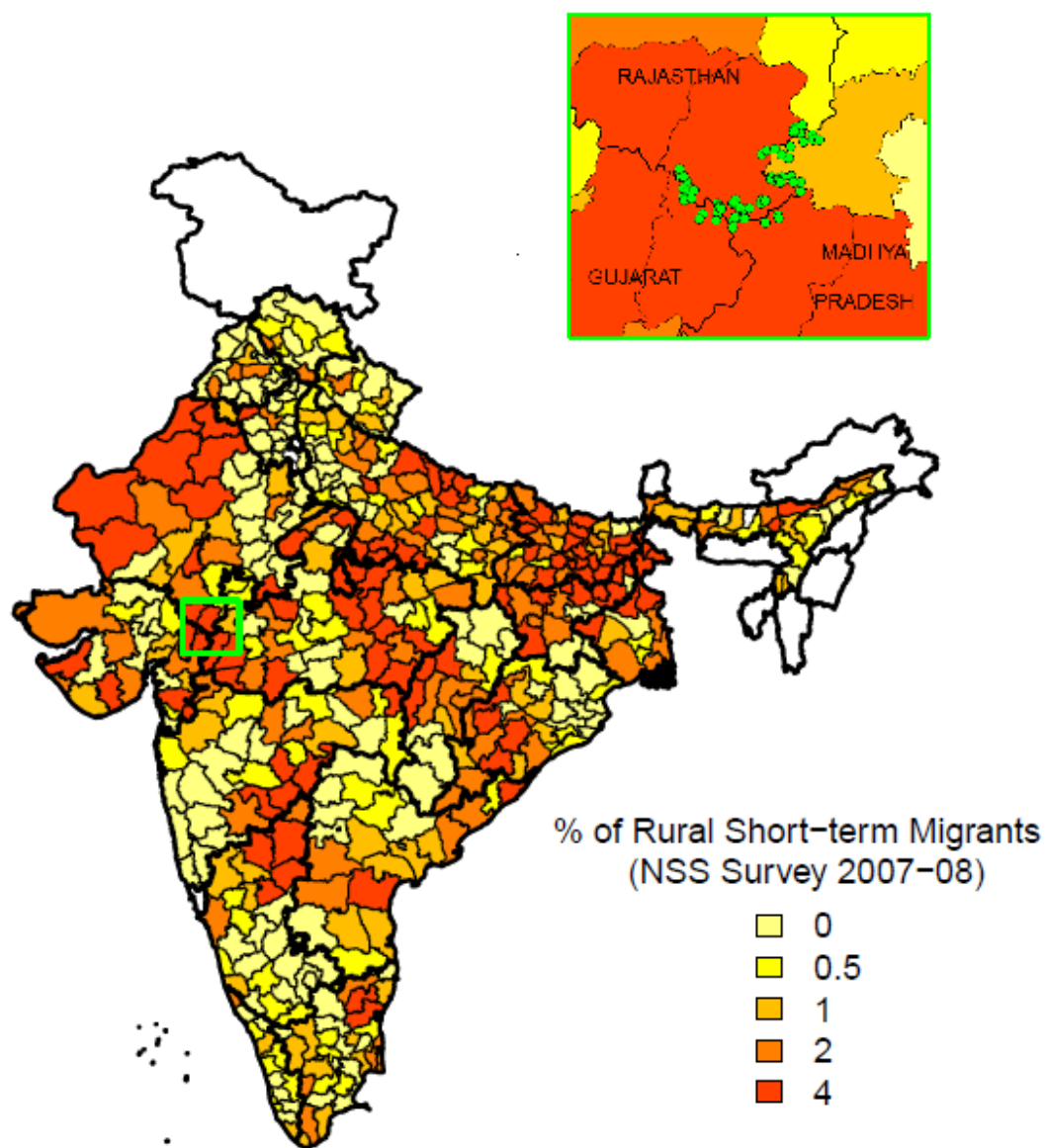
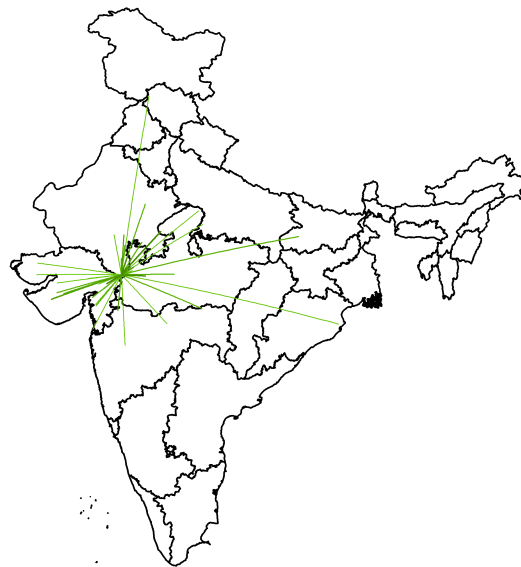


Figure 4: Destinations of seasonal migrants in the summer 2009

(a) Rajasthan Villages



(b) Madhya Pradesh Villages



(c) Gujarat Villages



(d) All Villages

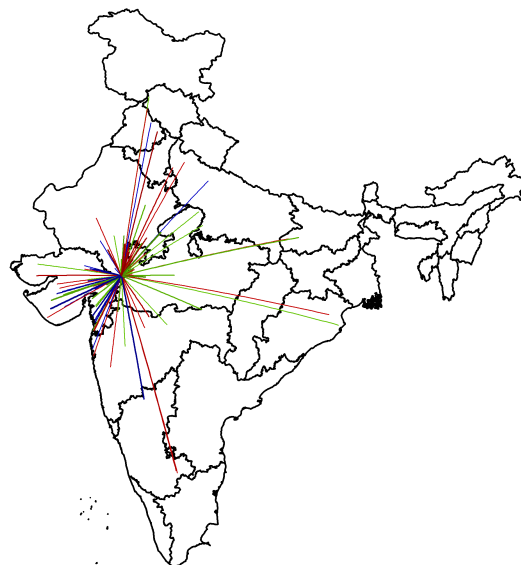


Table 1: Village Characteristics

	MP or Gujarat		Rajasthan		P-value
	Mean	St. Dev.	Mean	St. Dev.	
	(1)	(2)	(3)	(4)	
Panel A: Matching variables					
Population Share Scheduled Castes	0.01	(0.02)	0.01	(0.02)	0.62
Population Share Scheduled Tribes	0.97	(0.15)	0.96	(0.13)	0.83
Total Culturable Land	185	(115)	189	(133)	0.92
Share of Culturable Land Irrigated	0.26	(0.19)	0.26	(0.2)	0.94
Share of Culturable Land Non-irrigated	0.56	(0.2)	0.59	(0.2)	0.59
Population per ha of Culturable Land	4.24	(1.42)	4.16	(1.42)	0.81
Panel B: Village controls					
Total Population	1090	(786)	1071	(890)	0.93
Population Literate	0.36	(0.13)	0.37	(0.09)	0.54
Education Facility	1.00	(0)	0.77	(0.42)	0.00
Drinking Facility	1.00	(0)	0.95	(0.23)	0.16
Medical Facility	0.15	(0.35)	0.17	(0.37)	0.81
Paved Road	0.84	(0.37)	0.89	(0.32)	0.53
Bus Service?	0.47	(0.5)	0.27	(0.44)	0.09
Post and Telecommunication Facility?	0.11	(0.31)	0.19	(0.39)	0.37
Distance to Town (km)	30.93	(17.83)	31.74	(19.57)	0.86
Panel C: Household controls					
Number of Adults	5.64	(3.48)	4.84	(1.95)	0.09
Number of Children (below 12)	3.26	(2.45)	2.86	(2.03)	0.14
Hindu	0.78	(0.42)	0.90	(0.3)	0.00
Scheduled Tribes	0.91	(0.29)	0.88	(0.33)	0.56
House with Kuccha Roof	0.35	(0.48)	0.32	(0.47)	0.52
House with Dirt Floor	0.93	(0.26)	0.88	(0.33)	0.20
Owns Cell Phone	0.43	(0.5)	0.37	(0.48)	0.33
House with Electricity	0.39	(0.49)	0.23	(0.42)	0.03
Main Income Source is Farming	0.48	(0.5)	0.52	(0.5)	0.46
Panel D: Worker controls					
Female	0.51	(0.5)	0.52	(0.5)	0.27
Primary Education	0.12	(0.33)	0.15	(0.35)	0.09
Education Beyond Primary	0.20	(0.4)	0.18	(0.39)	0.58
Age between 30 and 45	0.24	(0.43)	0.25	(0.44)	0.46
Age higher than 45	0.26	(0.44)	0.26	(0.44)	0.67
Married	0.71	(0.46)	0.70	(0.46)	0.83
Speaks Gujarati	0.37	(0.48)	0.08	(0.27)	0.00
Speaks Hindi	0.35	(0.48)	0.35	(0.48)	0.95
Number of villages	35		35		

Note: Matching variables come from the Census 2001, village characteristics from the Census 2011, and household and worker characteristics from the 2010 survey. MP stands for Madhya Pradesh. All village, household and worker characteristics listed in this table are included as control in our main specification.

Table 2: NREGA work across states and seasons

	Gujarat	Madhya Pradesh	Rajasthan	Whole Sample
	(1)	(2)	(3)	(4)
Panel A: Summer (March-July 2009)				
Worked for NREGA	10%	39%	50%	40%
NREGA days worked if participated	25.3	21.7	31.7	28.1
Wanted work but did not work	69%	42%	33%	42%
Did or wanted NREGA work	79%	82%	83%	82%
Migrated	34%	41%	30%	35%
Of which worked for NREGA	7%	36%	43%	35%
Of which wanted but did not work	80%	55%	46%	55%
Of which did or wanted NREGA work	87%	91%	89%	89%
Would have migrated if no NREGA work	3%	8%	10%	8%
Panel B: Monsoon (July-November 2009)				
Worked for NREGA	0%	0%	1%	0%
NREGA days worked if participated	0.0	13.5	29.7	26.1
Wanted work but did not work	63%	50%	53%	53%
Did or wanted NREGA work	63%	51%	53%	54%
Migrated	18%	7%	9%	10%
Of which worked for NREGA	0%	0%	0%	0%
Of which wanted but did not work	72%	70%	74%	72%
Of which did or wanted NREGA work	72%	70%	74%	72%
Would have migrated if no NREGA work	0%	0%	0%	0%
Panel C: Winter (November 2009-March 2010)				
Worked for NREGA	2%	10%	5%	6%
NREGA days worked if participated	21.5	16.1	20.1	18.0
Wanted work but did not work	73%	64%	71%	69%
Did or wanted NREGA work	75%	74%	76%	75%
Migrated	35%	28%	28%	29%
Of which worked for NREGA	2%	11%	4%	6%
Of which wanted but did not work	83%	78%	84%	82%
Of which did or wanted NREGA work	85%	88%	88%	87%
Would have migrated if no NREGA work	1%	2%	1%	2%
Adults	330	749	1145	2224

Source: Retrospective questions from the 2010 survey. The unit of observation is an adult.

Table 3: Migration patterns

	Own survey (2009-10)			NSS 2007-08	
	Summer	Monsoon	Winter	All India	Sample districts
	(1)	(2)	(3)	(4)	(5)
Migrated?	35%	10%	29%	3%	16%
Observations (whole sample)	2224	2224	2224	188324	1937
Migrant is female	40%	33%	43%	14%	32%
Migrated with household member	71%	63%	74%	42%	81%
Destination is in same state	22%	33%	28%	53%	82%
Destination is urban	84%	88%	73%	68%	71%
Worked in construction	70%	70%	56%	42%	69%
Distance (km)	300	445	286	-	-
Transportation cost (Rs)	116	144	107	-	-
Number of trips	1.01	1.01	1.04	-	-
Migration duration (days)	60	50	52	-	-
Found employer after leaving	63%	64%	54%	-	-
No formal shelter in destination	86%	85%	83%	-	-
Days worked per week at destination	5666	4821	5109	-	-
Total Migration Earnings (Rs)	6	6	6	-	-
Earnings per day worked (Rs)	118	127	123	-	-
Earnings per day away (Rs)	101	107	109	-	-
Observations (migrants only)	768	218	646	13411	327

Notes: Columns 1 to 3 present means based on the migration survey described in Section 2. The unit of observation is a prime-age adult. Each column restricts the sample to responses for a particular season. Seasons are defined as follows: summer from April to June, monsoon from July to November, winter from December to March. Columns 4 and 5 present means based on the National Sample Survey (NSS). In Column 4 the sample includes all rural adults. In Column 5 the sample is restricted to adults living in the four districts of the migration survey sample. Information on distance, migration duration, transportation cost, job search and accommodation at destination is not collected by the NSS.

Table 4: Effect of NREGA on seasonal migration

	NREGA Days		Migration Days	
	(1)	(2)	(3)	(4)
Rajasthan	-0.124 (0.121)	-0.301 (0.135)	-1.012 (1.014)	1.067 (1.229)
Rajasthan x Summer	8.850 (0.642)	8.013 (0.922)	-5.255 (1.257)	-6.603 (1.196)
Village Pair FE Season FE	Yes	Yes	Yes	Yes
Controls Season FE	No	Yes	No	Yes
Observations	6,650	6,650	6,666	6,666
Other States Outside Summer	.67	.67	10.72	10.72
Other States in Summer	6.63	6.63	23.93	23.93
	Migrated?		Migration Days if Migrated	
	(5)	(6)	(7)	(8)
Rajasthan	-0.0123 (0.0139)	0.00365 (0.0160)	-1.669 (2.537)	2.472 (3.526)
Rajasthan x Summer	-0.0637 (0.0147)	-0.0555 (0.0163)	-5.526 (2.450)	-7.142 (2.495)
Village Pair FE Season FE	Yes	Yes	Yes	Yes
Controls Season FE	No	Yes	No	Yes
Observations	6,672	6,672	1,626	1,626
Other States Outside Summer	.2	.2	53.17	53.17
Other States in Summer	.39	.39	62.07	62.07

Notes: The unit of observation is an adult in a given season. Results in Panel B are based on pairs of villages in Madhya Pradesh and Rajasthan only. Column 1 and 2 present results from a regression of days spent working on the NREGA during a particular season on a set of explanatory variables. The outcome is missing for 22 adults who did NREGA work in the season. In Column 3 and 4 the outcome is the number of days spent away for work. It is missing for six individuals who migrated in this season. In Columns 5 and 6 the outcome is a dummy variable equal to one for adults who migrated during the season. In Columns 7 and 8 the outcome is the number of days spent away for work for adults who migrated. It is missing for adults who did not migrate and for six individuals who did migrate. Rajasthan is a dummy variable for whether the adult lives within a village in Rajasthan. Summer is a dummy variable for the summer months (mid-March to mid-July) Other States Outside Summer is the mean of the outcome variable for the monsoon and winter season in Madhya Pradesh and Gujarat villages. Other States in Summer is the outcome mean for the summer months in Madhya Pradesh and Gujarat villages. Controls are the variables shown in Table 1. Standard errors are clustered at the village-level.

Table 5: Wage Gaps

	All Migrants	Did not want NREGA work	Did NREGA work	Did or wanted NREGA work
	(1)	(2)	(3)	(4)
Daily Migration Earnings	100.9 (40.9)	118.0 (48.5)	98.8 (38.3)	98.9 (39.5)
Daily NREGA Wage	63.7 (12.4)	64.6 (11.4)	63.0 (12.2)	63.6 (12.5)
Difference	37%	46%	35%	35%
Observations	568	58	205	510

Note: The unit of observation is an adult. Daily Migration Earnings are equal to the ratio of earnings from migration in the summer 2009 divided by the total number of days spent migrating in that season. Daily NREGA Wage is equal to the ratio of NREGA earnings in the summer 2009 divided by the total number of days spent on NREGA works in that season. For migrants who did not work for NREGA, Daily NREGA Wage is predicted using worker, village and household controls described in Table 1 (see Appendix Table A5 for regression results). Column 1 uses the full sample of adults who left the village during the summer 2009 and for whom migration earnings are known. Column 2 includes only migrants who have done NREGA work during the summer 2009. Column 3 includes only migrants who have not done nor wanted to do NREGA work during the summer 2009. Column 4 includes only migrants who have done or wanted to do NREGA work during the summer 2009.

Table 6: Structural Estimates

	Flow Cost of Migration		Fixed Cost of Migration		Home Productivity
	Mean	SD	Mean	SD	Constant
	(1)	(2)	(3)	(4)	(5)
Estimate	80.4	14.7	480	1733	1602
Average Estimate	80.2	17.9	537	1622	1716
Standard Error	(2.7)	(1.7)	(127.8)	(121.2)	(217.7)
5% Conf. Interval	[75.5;85.8]	[14.6;21.5]	[302;804]	[1408;1902]	[1313;2105]

Note: This table presents estimates from the method of simulated moments described in section 4.3. Average estimates, standard errors, and confidence intervals (2.5th and 97.5th percentiles) are computed using 100 bootstrapped samples composed of independent village draws.

Table 7: Model Fit and Counterfactuals

		Targeted	Actual	Fitted	100 Days	No NREGA
		(1)	(2)	(3)	(4)	(5)
Non-Rajasthan villages						
(i)	Migration rate	Yes	38.8%	39.9%	35.0%	29.9%
(ii)	Average migration duration	Yes	61.3	61.3	63.3	62.4
(iii)	Standard deviation of duration	Yes	32.4	28.6	30.2	32.9
Rajasthan vs other states						
(iv)	Difference in migration rate	Yes	-0.084	-0.061	-0.021	-0.023
(v)	Difference in migration duration	Yes	-4.92	-5.57	-3.69	-1.63
All Migrants						
(vi)	% did or wanted NREGA work	No	89.3%	87.6%	83.8%	89.1%

Note: This table presents six empirical moments. (i) is the migration rate in Madhya Pradesh or Gujarat villages. (ii) is the average migration duration and (iii) is the standard deviation of the migration duration in these villages. (iv) is the difference in migration rate and (v) is the difference in average duration between Rajasthan villages and the other villages. (vi) is the proportion of migrants who work or want to work for NREGA. Column 1 indicates which moments were targeted in the estimation. Column 2 displays the moments observed in the data. Column 3 displays the simulated moments based on the main estimates presented in Table 5. Column 4 presents counterfactual moments if all workers who wanted work were given up to the legal entitlement of 100 days per household. Column 5 presents the counterfactual moments if no NREGA work was provided.

Table 8: Determinants of demand for NREGA work among migrants

	Did or wanted NREGA work?	
	Coefficient	Standard Error
Log real daily migration earnings	-0.0836	(0.0321)
Log NREGA wage	-0.0487	(0.0686)
Female	-0.00540	(0.0209)
Primary education	0.0536	(0.0152)
Education beyond primary	-0.100	(0.0675)
Age 30 to 45	0.0631	(0.0168)
Age higher than 45	0.0379	(0.0304)
Married	-0.0112	(0.0241)
Speaks Gujarati	-0.00829	(0.0354)
Speaks Hindi	-0.0296	(0.0348)
Number of adults	0.0141	(0.00674)
Number of children (below 12)	-0.00397	(0.00609)
Hindu	-0.00164	(0.0327)
Scheduled Tribes	0.0915	(0.0572)
HH has dirt floor	0.00371	(0.0392)
HH has cell phone	-0.0177	(0.0224)
HH has electricity	0.0498	(0.0216)
HH main income source is farming	0.0196	(0.0218)
Migrated with household member	-0.0267	(0.0214)
Destination is in same state	-0.00632	(0.0347)
Destination is urban	-0.0485	(0.0219)
Worked in construction	-0.0111	(0.0269)
Found employer after leaving	0.0450	(0.0310)
No formal shelter at destination	-0.0159	(0.0309)
Total crime per 1000	0.148	(0.0637)
Pollution SPM μg per m ³	-0.00569	(0.0364)
Observations	546	
State of Origin FE	Yes	

Note: The unit of observation is an adult. The sample includes only adults who were interviewed personally and migrated in the summer 2009. We report marginal effects at the mean from a separate probit estimation. The outcome is a dummy variable equal to one if the respondent has done or said they would have liked to do more NREGA work during the summer 2009. Worker controls and Household controls are described in Table 1. Trip characteristics are described in Table 2. The city deflator is the ratio of the urban poverty line in the state of destination to the average of rural poverty lines in Rajasthan, Madhya Pradesh and Gujarat. Crime per 1000 is computed using the National Crime Records Bureau report in the state of destination for 2009. Pollution is measured as Suspended Particle Matter per cubic meter, which is reported at the city level by the Central Pollution Control Board for the year 2010. Standard errors are clustered at the village-level.

FOR ONLINE PUBLICATION ONLY

A Appendix

A.1 Construction of Earnings Variables

Migration Earnings The survey instrument included questions about the frequency of payment and the typical amount per pay period. In most cases (74 percent), respondents were paid daily and in these cases we used the typical daily payment as earnings per day worked. We also asked respondents how many days per week they typically worked. Respondents worked on average six days per week and the median respondent worked six days. For respondents who were paid weekly, fortnightly, or monthly, we used the reported payment adjusted by the typical number of days per week worked. For example, a migrant paid 800 rupees weekly and working six days per week earns $800/6 = 133$ rupees per day worked. For migrants that were paid irregularly or in one lump sum at the end of work, we used the total earnings from the trip divided by the number of days worked. For migrants with missing values of days worked per week, we assumed they worked six days. Five percent of respondents received payment in-kind for their work, being paid in wheat for example. We leave these daily earnings observations as missing. For respondents with non-missing total earnings (62 percent), earnings per day away was computed using total earnings divided by days away. For respondents with missing total earnings, we used earnings per day worked adjusted downwards using days worked per week away.

As discussed in Section 2.4, the survey recorded detailed information on the last four trips only. Hence for adults who migrated in the summer 2009 but took more than four trips afterward, we did not record information for any of the trips taken during summer 2009. Out of 768 migrants, we have non-missing earnings for 593 migrants (77 percent). For migrants with missing earnings, we construct linear predictions by projecting summer 2009 earnings onto migration earnings in the following seasons (monsoon 2009, winter 2009 and summer 2010), individual and household controls (the list of controls is the same as in Panel C and D of Table 1. The regression coefficients are shown in appendix Table A.9. The mean of the observed and the predicted migration earnings are Rs. 101 and Rs. 102 respectively, which provides reassurance that the migrants with missing earnings are very similar to the other migrants in terms of observable characteristics.

NREGA Earnings Out of the 895 adults who worked for the NREGA during summer 2009,

32 (3.6 percent) report not having been paid in full at the time of the survey. Assuming a wage of zero for those who were not paid yields a wage of 64.4 rupees per day compared with 67 for only those who were paid. For our estimation of migration costs, we need a measure of daily earnings on NREGA that non-NREGA participants would expect to receive. For this, we restrict our sample to the 238 adults who both migrated and did NREGA work in the summer 2009 and we regress NREGA earnings on NREGA daily earnings for the following season, individual and worker controls (see Panel C and D in Table 1). The regression estimates are shown in appendix Table A.9. Interestingly, none of the predictors except summer 2010 NREGA daily earnings are statistically significant, suggesting that the NREGA wage does not vary with productivity or observable characteristics. The R-square of the regression is low, about 6 percent. In contrast, the R-square for the prediction of migration earnings is 10 percent. We use these estimates to predict NREGA earnings for migrants who did not participate to the program in the summer 2009. Again, mean predicted earnings are extremely close to observed earnings, Rs. 96 and Rs. 101 respectively.

A.2 Destination Characteristics

In order to better understand the determinants of migrants' demand for NREGA work, we use several sources of information on their destination.

Poverty lines First, we use poverty lines from (Planning Commission, 2009) to compute deflators of migration earnings. Specifically, we compute the ratio of the urban poverty line in the state of destination divided by the average of the rural poverty line in the three sample states (Rajasthan, Madhya Pradesh, Gujarat).

Crime Second, we use information on total crime per 1000 inhabitants in the state of destination in 2009 from the National Crime Records Bureau. We thank Nishith Prakash for sharing the data (Prakash et al., 2014).

Pollution Third, we use an index of urban air pollution at the state level in 2010, Suspended Particle Matter per cubic meter reported by the Central Pollution Control Board. We thank Anant Sudarshan for sharing the data (Greenstone et al., 2015).

Figure A.1: Differences in seasonal migration across states using retrospective questions

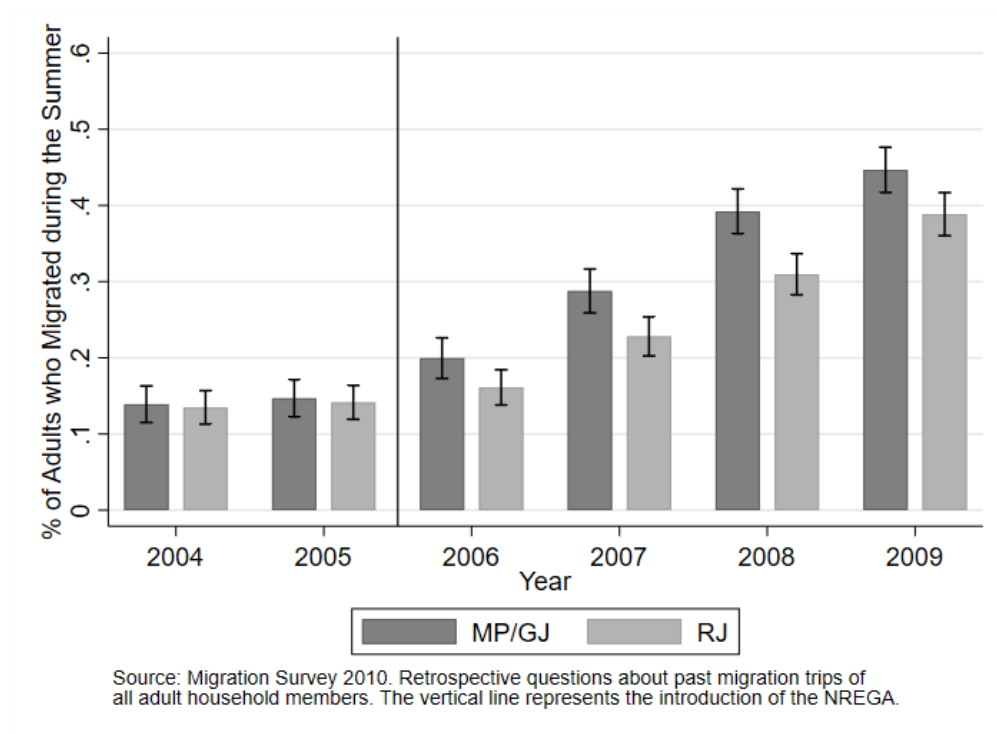


Figure A.2: Daily migration earnings as a function of migration duration

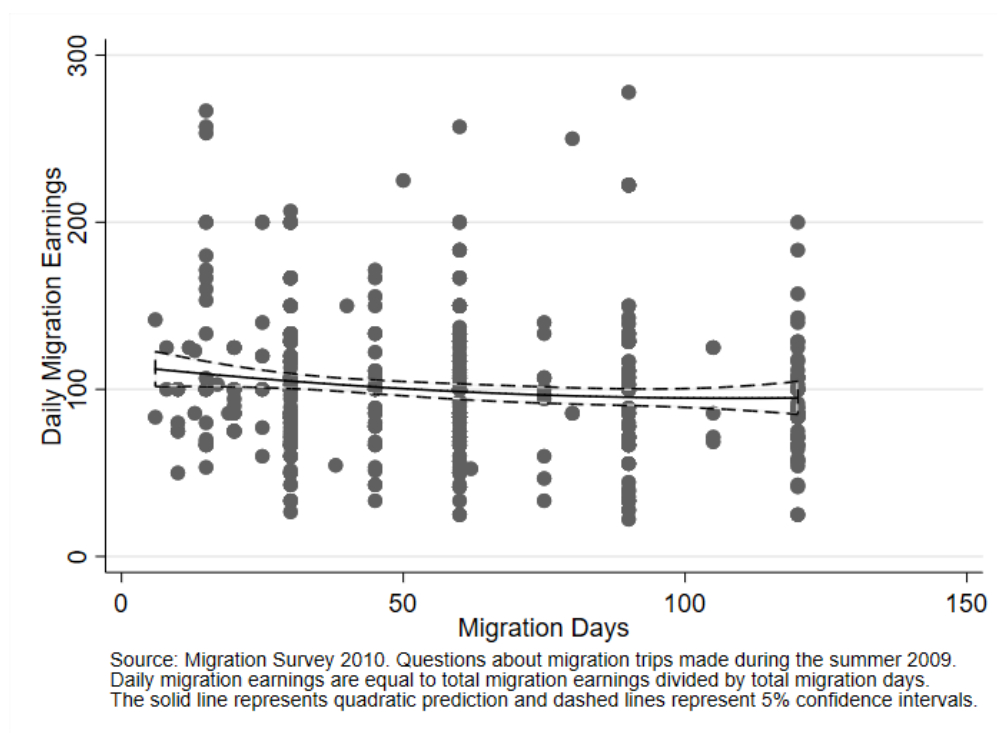
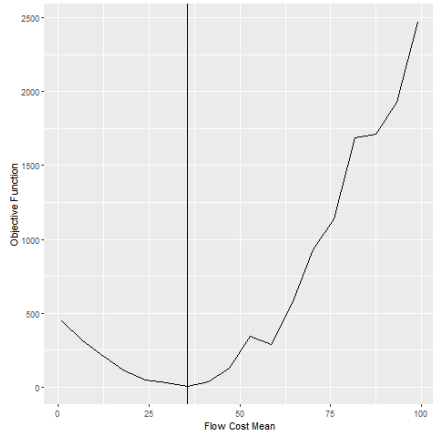
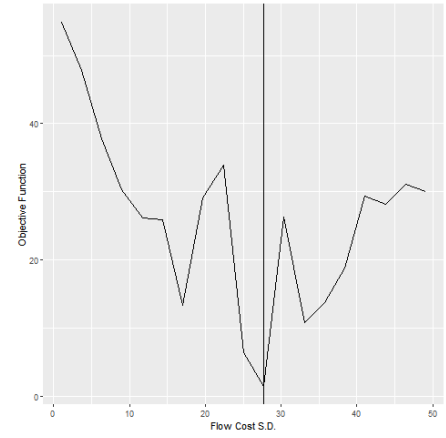


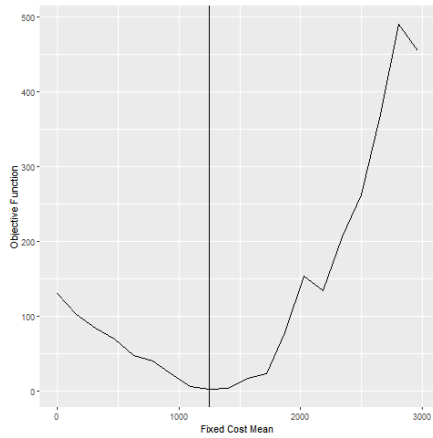
Figure A.3: Model Identification: Objective Function as a Function of the Parameters



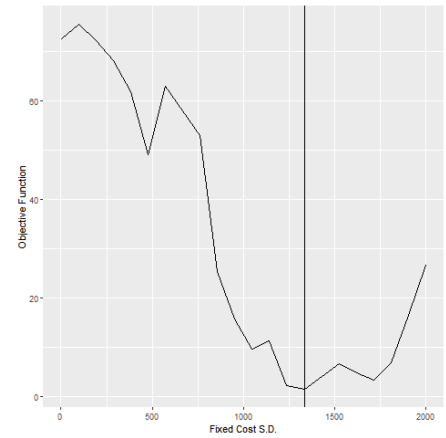
(a) Flow Cost Mean



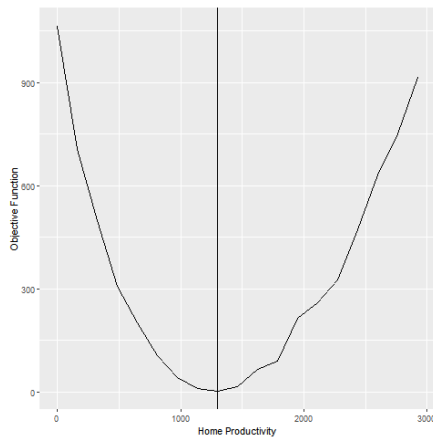
(b) Flow Cost S.D.



(c) Fixed Cost Mean



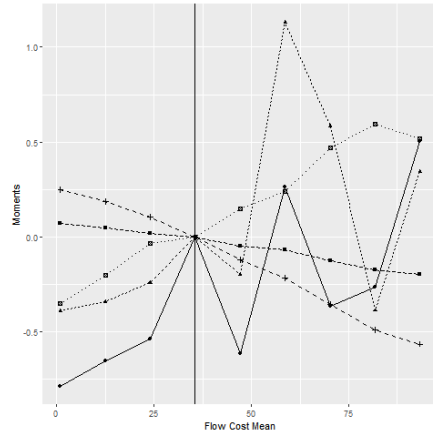
(d) Fixed Cost S.D.



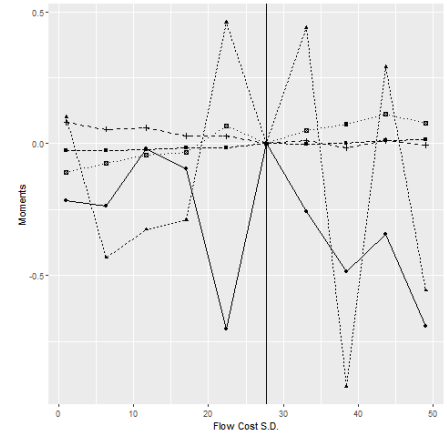
(e) Home Productivity

Note: Simulated data. Each figure shows the objective function as a function of each parameter, holding the others constant at the value of the model estimates. The objective function is the square of the distance between simulated and actual moments weighted by the inverse variance of the moments. Vertical lines denote model estimates.

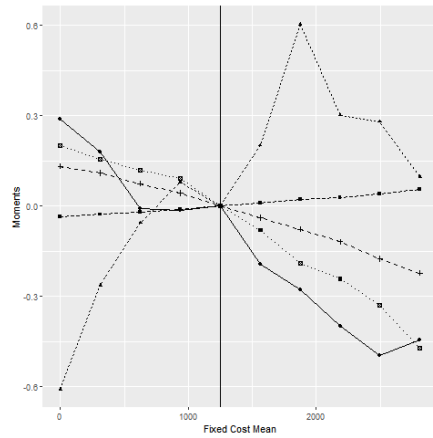
Figure A.3: Model Identification: Simulated Moments as a Function of the Parameters



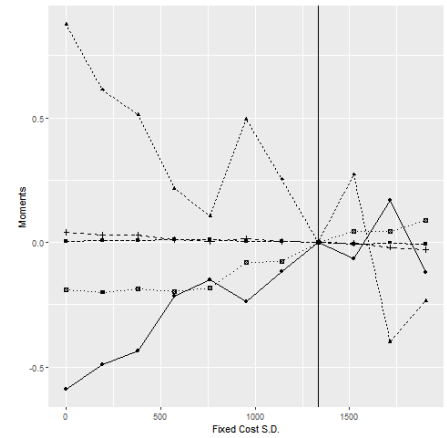
(a) Flow Cost Mean



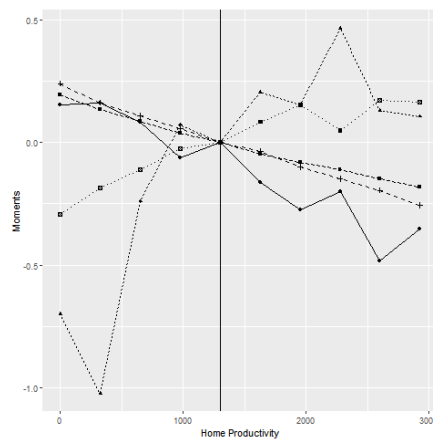
(b) Flow Cost S.D.



(c) Fixed Cost Mean



(d) Fixed Cost S.D.



(e) Home Productivity



Note: Simulated data. Each figure shows the simulated moments as a function of each parameter, holding the others constant. Each moment is standardized by subtracting and dividing by the moment simulated at the value of the model estimates. Vertical lines denote model estimates.

Table A.1: Survey Sample Selection

	All Adults	Own Survey		Difference (3) - (2)	NSS Survey 2007-08	
		Adult Survey Completed	Adult Survey not Completed		All Adults (India)	All Adults (Sample Districts)
	(1)	(2)	(3)	(4)	(5)	(5)
Female	0.511 (0.0056)	0.525 (0.0166)	0.448 (0.0067)	-0.077 (0.019)	0.499 (0.0011)	0.492 (0.0076)
Married	0.704 (0.0091)	0.729 (0.021)	0.594 (0.0105)	-0.134 (0.0233)	0.720 (0.0019)	0.753 (0.0185)
Illiterate	0.672 (0.0187)	0.690 (0.0327)	0.593 (0.019)	-0.097 (0.03)	0.372 (0.003)	0.491 (0.0313)
Scheduled Tribe	0.897 (0.0272)	0.894 (0.0278)	0.910 (0.0287)	0.016 (0.0225)	0.106 (0.0033)	0.660 (0.0601)
Age	32.8 (0.248)	34.1 (0.484)	27.0 (0.301)	-7.11 (0.592)	32.7 (0.038)	31.6 (0.3508)
Spent 2-330 Days away for Work	0.433 (0.0179)	0.422 (0.0394)	0.482 (0.0187)	0.060 (0.0412)	-	-
Migrated for Work all Three Seasons	0.119 (0.011)	0.080 (0.0318)	0.295 (0.0101)	0.215 (0.0324)	-	-
Ever Worked for NREGA	0.528 (0.0253)	0.581 (0.0354)	0.291 (0.0259)	-0.290 (0.0332)	-	-
Spent 30-180 Days away for Work	0.301 (0.0159)	0.312 (0.0351)	0.251 (0.0166)	-0.061 (0.0362)	0.028 (0.0009)	0.170 (0.0366)
Adults	2,722	2,224	498		188,324	1,937

Notes: The unit of observation is an adult. Standard errors computed assuming correlation of errors at the village level in parentheses. The first four columns present means based on subsets of the adults aged 14 to 69 from the main data set discussed in the paper. The first column includes the full sample of persons aged 14 to 69 for whom the adult survey was attempted. The second column includes all persons aged 14 to 69 for which the full adult survey was completed. The third column includes all persons aged 14 to 69 for which the full adult survey was not completed. The fourth column presents the difference between the third and second columns. The fifth and sixth columns present means computed using all adults aged 14 to 69 in the rural sample of the NSS Employment and Unemployment survey Round 64 conducted between July 2007 and June 2008 for all of India and for the six sample districts respectively. Means from the NSS survey are constructed using sampling weights. "-" denotes not available.

Table A.2: Determinants of demand for NREGA work and migration

	Did or wanted NREGA Work?	Migrated?
	(1)	(2)
Female	-0.0331 (0.0148)	-0.204 (0.0200)
Primary Education	-0.0418 (0.0236)	-0.0788 (0.0332)
Education Beyond Primary	-0.152 (0.0266)	-0.220 (0.0373)
Age 30 to 45	0.0356 (0.0191)	-0.168 (0.0297)
Age higher than 45	-0.0765 (0.0289)	-0.390 (0.0293)
Married	0.0913 (0.0229)	0.0938 (0.0225)
Speaks Gujarati	-0.00411 (0.0291)	0.146 (0.0342)
Speaks Hindi	-0.0383 (0.0194)	0.0221 (0.0237)
Number of Adults	0.00499 (0.00490)	-0.00326 (0.00556)
Number of Children (below 12)	0.000120 (0.00483)	0.0133 (0.00613)
Hindu	0.0423 (0.0265)	0.0915 (0.0267)
Scheduled Tribes	0.0332 (0.0361)	-0.0600 (0.0337)
HH with Kuccha roof	0.0621 (0.0206)	0.00371 (0.0255)
HH with dirt floor	0.0363 (0.0450)	0.0869 (0.0353)
HH with cell phone	-0.00854 (0.0215)	-0.0328 (0.0240)
HH with electricity	0.0193 (0.0227)	-0.00293 (0.0302)
HH whose main income source is farming	0.00385 (0.0163)	-0.111 (0.0248)
Observations	2,224	2,224
Mean	0.820	0.350

Note: The unit of observation is an adult. In Column 1 the dependent variable is a dummy variable for whether the individual has worked or would have wanted to work for the NREGA during the summer 2009. In Column 2 the dependent variable is a dummy variable set equal to one if the individual migrated for work during the summer 2009. Standard errors are clustered at the village level.

Table A.3: Stated reason for not getting more work or not wanting more work

	Summer 2009	Monsoon 2009	Winter 2009-10
	(1)	(2)	(3)
Panel A: Full Sample			
Wanted More NREGA Work	80%	75%	54%
Adults	2,224	2,224	2,224
Panel B: Reasons given by Adults Who Wanted More Work			
Family Worked Maximum 100 days	3%	1%	0%
Works Finished/No Work Available	56%	79%	87%
No Program ID Card/Name Not on ID Card	6%	4%	7%
Officials Would not Provide More Work	18%	12%	8%
Too Young	1%	1%	2%
Too Old / Sick	3%	2%	1%
Other	15%	9%	6%
Adults	1,779	1,194	1,673
Panel C: Reasons given by Adults Who Did Not Want More Work			
Working Outside the Village	18%	13%	5%
Other Work in Village	21%	30%	69%
Sick/injured/unable to work	29%	22%	11%
Pregnant / Needed to Take Care of Children	2%	3%	1%
Studying	21%	31%	17%
NREGA Does Not Pay Enough	4%	4%	1%
No Need for Work/Do Not Want to Do Manual Work	4%	2%	1%
Other	21%	11%	5%
Adults	445	1,030	551

Notes: The unit of observation is an adult. Each column restricts the sample to responses for a particular season. Panel A includes all adults who completed the adult survey. Panel B restricts the sample to adults who report wanting to work more for the NREGA during the season specified in the column heading. Panel C restricts the same to adults who report not wanting to work more for the NREGA during the season specified in the column heading.

Table A.4: Effect of the NREGA on seasonal migration including controls gradually

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: NREGA Days						
Rajasthan	-0.124 (0.121)	-0.124 (0.121)	-0.328 (0.129)	-0.235 (0.135)	-0.301 (0.135)	-0.323 (0.139)
Rajasthan x Summer	8.850 (0.642)	8.850 (0.642)	8.346 (0.798)	7.723 (0.810)	8.013 (0.922)	7.618 (0.770)
Observations	6,650	6,650	6,650	6,650	6,650	6,650
Other States Outside Summer	.67	.67	.67	.67	.67	.67
Other States in Summer	6.63	6.63	6.63	6.63	6.63	6.63
Panel B: Days Away						
Rajasthan	-1.012 (1.014)	-1.012 (1.014)	-0.905 (1.164)	-0.524 (1.212)	1.067 (1.229)	1.079 (1.229)
Rajasthan x Summer	-5.255 (1.257)	-5.255 (1.257)	-6.089 (0.990)	-5.792 (1.126)	-6.603 (1.196)	-6.726 (1.212)
Observations	6,666	6,666	6,666	6,666	6,666	6,666
Other States Outside Summer	10.72	10.72	10.72	10.72	10.72	10.72
Other States in Summer	23.93	23.93	23.93	23.93	23.93	23.93
Panel C: Migrated?						
Rajasthan	-0.0123 (0.0139)	-0.0123 (0.0139)	-0.0275 (0.0155)	-0.0218 (0.0171)	0.00365 (0.0160)	0.00204 (0.0157)
Rajasthan x Summer	-0.0637 (0.0147)	-0.0637 (0.0147)	-0.0622 (0.0123)	-0.0612 (0.0139)	-0.0555 (0.0163)	-0.0581 (0.0162)
Observations	6,672	6,672	6,672	6,672	6,672	6,672
Other States Outside Summer	.2	.2	.2	.2	.2	.2
Other States in Summer	.39	.39	.39	.39	.39	.39
Panel D: Migration Days if Migrated						
Rajasthan	-1.669 (2.537)	-1.669 (2.537)	1.551 (2.689)	1.240 (3.027)	2.472 (3.526)	1.710 (3.444)
Rajasthan x Summer	-5.526 (2.450)	-5.526 (2.450)	-6.068 (2.721)	-4.378 (2.698)	-7.142 (2.495)	-6.592 (2.470)
Observations	1,626	1,626	1,626	1,626	1,626	1,626
Other States Outside Summer	53.17	53.17	53.17	53.17	53.17	53.17
Other States in Summer	62.07	62.07	62.07	62.07	62.07	62.07
Village Pair FE	No	Yes	Yes	Yes	Yes	Yes
Village Controls	No	No	Yes	Yes	Yes	Yes
Household Controls	No	No	No	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes	Yes
Demand for NREGA	No	No	No	No	No	Yes

Notes: The unit of observation is an adult in a given season. Panel A present results from a regression of days spent working on the NREGA during a particular season on a set of explanatory variables. The outcome is missing for 22 observations. In Panel B the outcome is the number of days spent away for work. The outcome is missing for 6 observations. In Panel C, the outcome is a dummy equal to one for adults who migrated during the season. In Panel D the outcome is the number of days spent away from the village for adults who migrated. It is missing for adults who did not migrate and for 6 adults who did. Rajasthan is a dummy for whether the adult lives within a village in Rajasthan. Summer is a dummy for the summer months (mid-March to mid-July). Control Mean is the mean of the outcome variable in villages which are not in Rajasthan outside of the summer months. Village, Household and Worker Controls are described in Table 1. Demand for NREGA is a dummy variable equal to one for individuals who either worked or wanted work during the season. Standard errors are clustered at the village-level.

Table A.5: Effect of the NREGA on seasonal migration on common support

	NREGA Days	Migration Days	Migrated?	Migration Days if Migrated
	(1)	(2)	(3)	(4)
Rajasthan	-0.197 (0.132)	2.094 (1.514)	0.0140 (0.0187)	4.008 (3.700)
Rajasthan x Summer	6.358 (0.854)	-5.947 (1.315)	-0.0548 (0.0199)	-5.889 (2.439)
Village PairSeason FE	Yes	Yes	Yes	Yes
ControlsSeason FE	Yes	Yes	Yes	Yes
Observations	5,123	5,138	5,142	1,281
Other States Outside Summer	.62	10.7	.2	53.95
Other States in Summer	6.23	24.29	0.39	62.34

Notes: The unit of observation is an adult in a given season. The sample is restricted to observations for which there was overlap in a propensity score for being in Rajasthan estimated using the controls in Table 1. Column 1 presents results from a regression of days spent working on the NREGA during a particular season on a set of explanatory variables. The outcome is missing for 14 observations. In Column 2 the outcome is the number of days spent away for work. It is missing for 3 observations. In Columns 3 the outcome is a dummy variable for adults who migrated during the season. In Columns 4 the outcome is the number of days spent away for work for adults who migrated. It is missing for adults who did not migrate and for 3 adults who did migrate. Rajasthan is a dummy variable for whether the adult lives within a village in Rajasthan. Summer is a dummy variable for the summer months (mid-March to mid-July) Other States Outside Summer is the mean of the outcome variable for the monsoon and winter season in Madhya Pradesh and Gujarat villages. Other States in Summer is the outcome mean for the summer months in Madhya Pradesh and Gujarat villages. Controls are the variables shown in Table 1 Standard errors are clustered at the village-level.

Table A.6: Effect of the NREGA on seasonal migration excluding Gujarat villages

	NREGA Days	Migration Days	Migrated?	Migration Days if Migrated
	(1)	(2)	(3)	(4)
Rajasthan	-0.378 (0.143)	-0.457 (1.144)	-0.00330 (0.0145)	-0.872 (3.792)
Rajasthan x Summer	7.554 (1.108)	-8.382 (1.279)	-0.0903 (0.0164)	-8.129 (3.286)
Village PairSeason FE	Yes	Yes	Yes	Yes
ControlsSeason FE	Yes	Yes	Yes	Yes
Observations	4,714	4,725	4,728	1,105
Other States Outside Summer	.85	8.80	.17	50.29
Other States in Summer	8.45	25.91	.4	63.66

Notes: The unit of observation is an adult in a given season. The sample is restricted to pairs of villages in Madhya Pradesh and Rajasthan only. Column 1 presents results from a regression of days spent working on the NREGA during a particular season on a set of explanatory variables. The outcome is missing for 14 observations. In Column 2 the outcome is the number of days spent away for work. It is missing for 3 observations. In Columns 3 the outcome is a dummy variable for adults who migrated during the season. In Columns 4 the outcome is the number of days spent away for work for adults who migrated. It is missing for adults who did not migrate and for 3 adults who did migrate. Rajasthan is a dummy variable for whether the adult lives within a village in Rajasthan. Summer is a dummy variable for the summer months (mid-March to mid-July) Other States Outside Summer is the mean of the outcome variable for the monsoon and winter season in Madhya Pradesh and Gujarat villages. Other States in Summer is the outcome mean for the summer months in Madhya Pradesh and Gujarat villages. Controls are the variables shown in Table 1 Standard errors are clustered at the village-level.

Table A.7: Effect of the NREGA on seasonal migration (All Adults)

	NREGA Days	Migration Days	Migrated?	Migration Days if Migrated
	(1)	(2)	(3)	(4)
Rajasthan	-0.297 (0.130)	-1.141 (1.270)	-0.0132 (0.0157)	-1.087 (3.218)
Rajasthan x Summer	6.969 (0.817)	-6.620 (1.119)	-0.0494 (0.0145)	-5.415 (2.486)
Village PairSeason FE	Yes	Yes	Yes	Yes
ControlsSeason FE	Yes	Yes	Yes	Yes
Observations	8,144	8,126	8,166	2,179
Other States Outside Summer	0.57	14.71	.25	61.04
Other States in Summer	5.93	25.98	.41	65.25

Notes: The unit of observation is an adult in a given season. The sample includes all adults, including those who could not be interviewed in person, and whose information was given by the household head. Column 1 presents results from a regression of days spent working on the NREGA during a particular season on a set of explanatory variables. The outcome is missing for 22 observations. In Column 2 the outcome is the number of days spent away for work. It is missing for 40 observations. In Columns 3 the outcome is a dummy variable for adults who migrated during the season. In Columns 4 the outcome is the number of days spent away for work for adults who migrated. It is missing for adults who did not migrate and 40 who did. Rajasthan is a dummy variable for whether the adult lives within a village in Rajasthan. Summer is a dummy variable for the summer months (mid-March to mid-July) Other States Outside Summer is the mean of the outcome variable for the monsoon and winter season in Madhya Pradesh and Gujarat villages. Other States in Summer is the outcome mean for the summer months in Madhya Pradesh and Gujarat villages. Standard errors are clustered at the village-level.

Table A.8: Cross-state comparison of permanent migration in the last five years

	Any Migrant		Number of Migrants	
	(1)	(2)	(3)	(4)
Rajasthan	0.0293 (0.0263)	0.102 (0.0309)	0.0880 (0.133)	0.235 (0.145)
Village Pair Fixed Effect	Yes	Yes	Yes	Yes
Household and Village Controls	No	Yes	No	Yes
Observations	702	702	702	702
Mean in Control	.39	.39	1.23	1.23

Notes: The unit of observation is a household. Results in Panel B are based on pairs of villages in Madhya Pradesh and Rajasthan only. In Column 1 and 2 the dependent variable is a dummy which equals one if any member of the household left within the past five years. In Column 3 and 4 it is the number of household members who left within the past five years. Controls include village and household controls presented in Table 1. Mean in Control is the average outcome in non-Rajasthan villages. Standard errors are clustered at the village level. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

Table A.9: Predictions of NREGA and migration earnings for migrants in the summer 2009

	Daily NREGA earnings		Daily Migration Earnings	
Female	-1.581	(1.319)	-12.49	(3.555)
Primary education	-2.035	(2.961)	7.671	(6.160)
Education beyond primary	-0.547	(3.627)	2.366	(6.059)
Age 30 to 45	-1.490	(2.324)	7.067	(4.348)
Age higher than 45	-3.777	(2.546)	-13.91	(4.665)
Married	0.759	(2.819)	0.206	(4.103)
Speaks Gujarati	-1.971	(4.438)	4.709	(5.804)
Speaks Hindi	1.389	(2.654)	-5.874	(4.037)
Number of adults	-0.371	(0.550)	0.492	(0.975)
Number of Children (below 12)	-0.00103	(0.733)	-0.364	(1.103)
Hindu	-4.296	(3.167)	3.838	(5.954)
Scheduled Tribes	-1.240	(4.569)	-6.692	(6.632)
HH has kuccha roof	0.220	(1.994)	-3.143	(4.183)
HH has dirt floor	-2.542	(5.093)	7.383	(5.997)
HH has cell phone	-0.915	(2.838)	4.094	(4.721)
HH has electricity	-0.201	(2.320)	-7.218	(4.855)
HH main income source is farming	-0.0245	(2.226)	1.264	(4.316)
Total Population	0.00490	(0.00340)	-0.00645	(0.00557)
% Population Literate	-27.19	(13.89)	-32.11	(22.51)
Education Facility	-8.568	(4.479)	-8.021	(12.30)
Drinking water facility?	5.131	(4.993)	22.37	(16.56)
Medical facility?	0.656	(3.978)	-7.225	(6.701)
Approach with paved road?	-5.982	(4.535)	6.023	(7.717)
Bus Service?	-5.592	(2.828)	2.319	(6.858)
Post and telecommunication facility?	-0.888	(5.399)	-6.156	(8.155)
Distance to Town (km)	-0.0597	(0.0699)	0.00547	(0.180)
Observations	816		569	
R-Square	0.062		0.103	
Mean observed earnings	64.41		100.81	
Mean predicted earnings	64.1		95.8	

Notes: The unit of observation is an adult. In column 1 the sample is restricted to the 816 respondents who worked for NREGA in the summer 2009, and for which information on NREGA earnings is available (20 respondents who did NREGA work have missing earnings). In column 2 the sample is restricted to the 569 respondents who migrated in the summer 2009 and for which information on migration earnings is available (200 respondents migrated and have missing earnings). The dependent variable are "Daily NREGA Earnings", which is the ratio between NREGA earnings in the summer 2009 and the total number of days worked on the program during that season and "Daily Migration Earnings", which is the ration between total migration earnings in the summer 2009 and the total number of days away for work during that season. "Mean observed earnings" is the mean of the dependent variable for the regression sample, and "Mean predicted earnings" is the mean of the linear prediction based on regression estimates for the whole sample. Standard errors are clustered at the village-level and reported in parentheses.

Table A.10: Structural Estimates: Alternative Utility Function

	Flow Cost of Migration		Fixed Cost of Migration		Home Productivity
	Mean	SD	Mean	SD	Constant
	(1)	(2)	(3)	(4)	(5)
Estimate	63.0	20.6	280	1063	407
Average Estimate	59.1	19.8	324	1280	434
Standard Error	(7)	(3.9)	(128.8)	(314.5)	(63.8)
5% Conf. Interval	[50.7;66.4]	[12.4;26.2]	[95;571]	[363;1807]	[347;539]

Note: This table presents estimates from the method of simulated moments described in section 4.3. Average estimates, standard errors, and confidence intervals (2.5th and 97.5th percentiles) are computed using 100 bootstrapped samples composed of independent village draws.

Table A.11: Model Fit and Counterfactuals: Alternative Utility Function

		Targeted	Actual	Fitted	100 Days	No NREGA
		(1)	(2)	(3)	(4)	(5)
Non-Rajasthan villages						
(i)	Migration rate	Yes	38.8%	39.9%	35.0%	29.9%
(ii)	Average migration duration	Yes	61.3	61.3	63.3	62.4
(iii)	Standard deviation of duration	Yes	32.4	28.6	30.2	32.9
Rajasthan vs other states						
(iv)	Difference in migration rate	Yes	-0.084	-0.069	-	-
(v)	Difference in migration duration	Yes	-4.92	2.43	-	-
All Migrants						
(vi)	% did or wanted NREGA work	No	89.3%	50.4%	47.0%	37.0%

Note: This table presents six empirical moments. (i) is the migration rate in Madhya Pradesh or Gujarat villages. (ii) is the average migration duration and (iii) is the standard deviation of the migration duration in these villages. (iv) is the difference in migration rate and (v) is the difference in average duration between Rajasthan villages and the other villages. (vi) is the proportion of migrants who work or want to work for NREGA. Column 1 indicates which moments were targeted in the estimation. Column 2 displays the moments observed in the data. Column 3 displays the simulated moments based on the main estimates presented in Table 5. Column 4 presents counterfactual moments if all workers who wanted work were given up to the legal entitlement of 100 days per household. Column 5 presents the counterfactual moments if no NREGA work was provided.

Table A.12: Expenditure items in NSS Employment Unemployment Survey and construction of urban deflators

Serial Number	Expenditures Item	In deflator with no durable or housing	In deflator with food only
(1)	(2)	(3)	(4)
1	cereals & cereal products	x	x
2	pulses & pulse products	x	x
3	milk	x	x
4	milk products	x	x
5	edible oil and vanaspati	x	x
6	vegetables	x	x
7	fruits & nuts	x	x
8	egg, fish & meat	x	x
9	sugar	x	x
10	salt & spices	x	x
11	other food items	x	x
12	pan, tobacco & intoxicants	x	x
13	fuel & light	x	-
14	entertainment	x	-
15	personal care and effects	x	-
16	toilet articles	x	-
17	sundry articles	x	-
18	consumer services excluding conveyance	x	-
19	conveyance	x	-
20	rent/ house rent	-	-
21	consumer taxes and cesses	x	-
22	medical expenses	x	-
24	medical (institutional)	-	-
25	tuition fees & other fees	-	-
26	school books & other educational articles	-	-
27	clothing and bedding	-	-
28	footwear	-	-
29	furniture and fixtures	-	-
30	crockery & utensils	-	-
31	cooking and household appliances	-	-
32	goods for recreation	-	-
33	jewellery & ornaments	-	-
34	personal transport equipment	-	-
35	therapeutic appliances	-	-
36	other personal goods	-	-
37	repair and maintenance (of residential buildings)	-	-

Notes: This table presents expenditure categories in the "Household Consumer Expenditures" (Section 9) of the 2009-10 NSS Employment Unemployment survey. Column 3 indicates which expenditure items are included in the deflator used in Column 8 of Table 4, excludes durable and housing expenditures. Column 4 indicates which expenditure items are included in the deflator used in Column 7 of Table 4, which only includes food expenditures.