The Impact of Water Access on Short-Term Migration in Rural India

Esha D. Zaveri[a]
Douglas H. Wrenn[b]
Karen Fisher-Vanden[c]

December 24, 2019

[a] Water Economist, the World Bank, Washington, DC, ezaveri@worldbank.org
[b] Corresponding Author: Agricultural Economics, Sociology, and Education, Pennsylvania State University, 112A Armsby Building, University Park, PA 16802, dhw121@psu.edu
[c] Agricultural Economics, Sociology, and Education, Pennsylvania State University, 113 Armsby Building, University Park, PA 16802
We are grateful to David Abler, S. Chandrasekhar, and participants at the Population Association of America 2015 annual meeting for helpful comments. We thank S. Chandrasekhar for sharing data and A. Gupta and G. Sharma at the Ministry of Water Resources for data assistance. This work was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program, Integrated Assessment Research Program, Grant No. DE-SC0005171

Abstract: Migration is an important risk-reduction strategy for households in developing countries. In this paper, we examine the impact of rainfall variability and irrigation availability on short-term migration decisions in India. Our results show that both rainfall shocks and the availability of irrigation impact the decisions of households to dispatch migrants. For irrigation, we find that migration responds to costs and that deep fossil-water wells, which provide a constant source of water, eliminate any benefit of short-term migration. This suggests that regions with access to more secure and stable sources of water are less likely to rely on migration as an income-smoothing mechanism, at least in the short run. Whether this holds in the long run will depend on the continued stability and availability of irrigation water.

Keywords: Migration, Groundwater, Irrigation, Climate, Indian Monsoon

JEL Codes: D74, I32, O13, O15, Q15, Q16

1. Introduction

Scientists expect that increased climate variability will amplify future stresses on the hydrologic cycle (World Bank 2016). The Indian subcontinent and its monsoon climate are particularly vulnerable to these changes. In recent decades, the monsoon circulation has weakened and precipitation has declined (Singh *et al*. 2014). In addition, studies have found evidence of both historical and future increases in rainfall variability (Menon *et al*. 2013; Singh *et al*. 2014). This increased variability represents one of the most significant sources of risk facing Indian households as it directly impacts agricultural output and household income. Indeed, evidence from as far back as the 1800s has demonstrated a significant correlation between rainfall shortages and harvest failures (Burgess and Donaldson 2010; Roy 2016). Thus, rural households have resorted to a variety of strategies to address the risks associated with increased rainfall variability.[1]

In this paper, we study the relationship between two important adaptation mechanisms used by rural households in India: irrigation and short-term migration. Irrigation is one of the most important adaptation methods used by farmers in response to risks associated with rainfall variability with groundwater providing the most reliable source of water (Fishman 2012; Taraz 2018). In India, approximately 90 million households utilise some form of groundwater during critical periods of crop growth (Sekhri 2011). However, as the water levels in many aquifers have dropped it has become increasingly difficult to access groundwater without improvements in irrigation technology. Of the different types of groundwater infrastructure used, deep tube wells (i.e., fossil-water wells) provide the most reliable source of water. Thus, uncertain access to groundwater resources based on deep tube wells has become an increasingly binding

[1] Lanjouw and Shariff (2004) show that rural households develop many methods to reduce the variance of household income in response to risk, especially risks related to low agricultural productivity.

constraint on agricultural livelihoods in many parts of India and in turn has contributed to water shortages and uncertain incomes in rural areas.

In addition to irrigation, and partially because of it, many households have resorted to additional coping mechanisms. One of the most important is short-term migration. While short-term migration plays a role in reducing income uncertainty on its own, our interest in this paper is in how short-term migration interacts with the different sources and dimensions of groundwater irrigation. As far as we are aware, this is the first paper to integrate irrigation access with short-term migration decisions to demonstrate how multiple adaptation strategies interact in the context of increasing rainfall variability, a fundamental environmental change impacting the Indian subcontinent.

To facilitate our analysis, we combine multiple datasets related to migration, irrigation, and groundwater availability. Our main data source for migration comes from the National Sample Survey Organization's (NSS) 2007-08 nationally representative survey, which is considered the most comprehensive data source on short-term migration in India providing short-term migration histories for 79,000 rural households. We use the NSS data because other existing datasets only capture permanent migration, which, for reasons other than marriage, is rare and very low in India (Munshi and Rosenzweig 2016).[2] Short-term migration from Indian villages, on the other hand, is very common and a large portion of this migration is for short periods of time.[3]

[2] Munshi and Rosenzweig (2016) develop a theory of caste networks to explain low levels of permanent migration for males.

[3] Short-term migration is well documented in the literature on developing countries (Breman 1996; Banerjee and Duflo 2007; Deshingkar and Farrington 2009; Badiani and Safir 2010; Keshri and Bhagat 2012; Coffey *et al*. 2014; Bryan *et al*. 2014; Morten 2019).

The results from our empirical model suggest that short-term migration decisions respond to past rainfall variability and to the agricultural opportunity costs associated with irrigation. Precisely, our results demonstrate that access to secure water resources determine the relative benefits of short-term migration with more reliable sources reducing the need for short-term migration. Specifically, tube wells, which provide a consistent source of water, allow individuals to specialize in agricultural related activities and help small landholders to profitably farm even in times of water scarcity, which in turn reduces the benefit of short-term migration. We confirm these results using plausibility exogenous variation in the geological characteristics of groundwater aquifers, which have influenced the development of groundwater infrastructure and related technology.

This paper contributes to a growing body of literature that studies adaptive development (Agrawal and Lemos 2015; Castels-Quintana *et al*. 2018; Lemos *et al.* 2013) and environmental migration (Millock 2015). While economic models have been developed to study migration, only recently have economists started to estimate the interactions between environmental change and migration. In developing countries, the effects of gradual changes in the environment have been shown to impact migration through the wage channel (Marchiori *et al*. 2012) and the agricultural productivity channel (Feng *et al*. 2010; Chen and Muelle, 2018). Global studies that focus on international migration have also found that agricultural incomes play an important role in influencing the climate-migration relationship in low and middle-income countries (Beonnier *et al.* 2018; Cattaneo and Peri 2016; Cai *et al.* 2016), and that access to irrigation has the potential to modulate this relationship (Beonnier *et al.* 2018).

In India, studies that have examined the impacts of environmental change on mobility have focused on permanent migration based on state-level migration data from the census data

(Viswanathan and Kumar 2012; Dallmann and Millock 2013) or household data from smaller geographical areas within a state (Fishman *et al*. 2015). Dallmann and Millock (2013) find that inter-state migration rises marginally in response to an increase in drought frequency, and Viswanathan and Kumar (2015) find some evidence for inter-state migration in response to weather changes via the channel of falling wheat and rice yields. In a recent study closest to ours, Fishman *et al*. (2015) empirically study permanent male migration in response to groundwater depletion. Their evidence is based on a targeted geographic area in two districts in the north-western state of Gujarat that have already experienced very large declines in groundwater levels.[4] They find that in water scare villages there are higher rates of permanent male migration to urban areas from households that belong to a relatively richer landholding class. This paper complements this literature and studies the influence of groundwater irrigation on short-term migration, a form of migration that is economically more important for small and marginal farmers. Moreover, our analysis spans the entire country thus providing evidence for the relationship between two key adaptation responses that are valid over a larger geographic area.

Migration is a complex phenomenon and having access to secure and stable irrigation is one of the many factors contributing to short-term rural migration, especially in a country that is facing rapid economic and social change. While our work highlights important correlations in this area, we do not attempt to suggest that agricultural opportunity costs are the only drivers for short-term migration. However, given the risk posed by groundwater depletion and rainfall variability to rural India it is necessary that we understand the impact of irrigation availability and access on decision-making. Moreover, by examining whether groundwater irrigation has an economically significant impact on short-term migration decisions we can highlight important

[4] The study uses deep-lying geological features that are hydrologically responsible for an increase in the fall of water tables to distinguish between water scarce and water abundant villages.

interactions between two different adaptation responses. Such evidence will be critical when designing and evaluating climate change adaptation policies going forward.

The rest of the paper is organized as follows. Section 2 describes the structure of the data used in the empirical model; Section 3 presents our empirical strategy; and Section 4 presents our main results and a series of robustness checks. Section 5 concludes the paper.

2. Data

2.1. Migration

Our empirical analysis integrates individual-level migration data with district-level data on weather and irrigation.[5] In general, there is a paucity of reliable migration data in India, and this is particularly true of short-term migration which occupies an important share of migration in the country but lacks proper documentation in official statistics. To remedy this deficiency, we use household-level data from the National Sample Survey Organization's national survey of migration conducted from July 2007 through June 2008 (the NSS 64th round). The NSSO uses a recall-based interview method, respondents are asked questions about things that happened to them over the previous year, and the survey is conducted in four rounds: (1) July 2007-September 2007; (2) October 2007-December 2007; (3) January 2008-March 2008; and (4) April 2008-June 2008, which we exploit when attaching our weather variables to the migration data in different years. The 64th round of the NSS is one of the richest sources of household migration history, and it is the first dataset to capture short-term migration for the entire country. Since we

[5] Districts are administrative units within states and the administrative level at which detailed weather and irrigation data is available. The average district is about 5000 sq. km and contains about 2M people.

are interested in the impacts of rural water infrastructure on short-term migration, this paper primarily focuses on rural India where its occurrence is widespread.

According to the survey, around 12.58 million rural residents are short-term migrants compared to one million urban residents.[6] This translates into short-term migrants making up about 1.69% of the rural population, 1.96% of prime-age adults, and 80% of all rural out-migrants in India. (Figure A1 in the Appendix shows the spatial distribution of short-term migrants in India.) The survey also identifies if the movement is within the same district, to another district in the same state, or to another state. Most movements are to a different district within the same state or to other states rather than within the same district. Given that our weather and irrigation variables are defined at the district level, we define migration as district-to-district movements to capture the impact of movements away from source regions.

2.2. Weather data

We use observed temperature and precipitation data which we acquired from a relatively new gauge-based observationally-gridded daily dataset – Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) (Yasutomi *et al*. 2011; Yatagai *et al.* 2012) – which was compiled by the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of Japan, Meteorological Agency (MRI/JMA). The precipitation and temperature data in the APHRODITE database are available at a spatial resolution of 0.25° x 0.25°. We re-scale these gridded weather

[6] Despite the large numbers, it is likely that these NSS estimates underestimate the flows. Independent surveys have found higher short-term migration rates that those reported in the NSS in certain regions (Coffey *et al*. 2014). The reasons for this discrepancy are the way households are defined and the time frame used to define a short-term migrant. As per the NSS survey, a household comprises a group of people who live together and share a common kitchen, excluding guests, visitors and all those who stay away from the household for more than 6 months. However, some who stay away for more than 6 months and reside at home for the rest of the months are still technically part of the household and could potentially be termed short-term migrants (Chandrasekhar *et al*.., 2014).

data to the district level by taking an area-weighted average of grid values in each district using GIS maps corresponding to 2001 district boundaries.

2.3. Irrigation data

The NSS 2007-08 survey provides detailed information on migrants but lacks irrigation and agricultural information for households. The most comprehensive data available for irrigation in India is the Minor Irrigation Census (MIC) conducted by the Ministry of Water Resources on a quinquennial basis. We make use of these data from the 2006-07 round, which most closely aligns with the timing of the NSSO migration survey.[7] The MIC accounts for the entire population of groundwater structures and surface water schemes in India. The data includes information about area irrigated by different sources (surface water vs. wells) as well as sown area, cultivated area, average water table depth, and other important features. We match district-level variables from the MIC to the districts included in the NSSO survey to get irrigated area by resource type for each district.

2.4. Aquifer coverage

Finally, we digitised multiple historical maps drawn from the Geological Survey and National Atlas of India. The 1969 Geohydrology Map from the Geological Survey is used to measure district area shares covering the unconsolidated aquifer as illustrated in Figure 1. Further, we combine this with the 1982 Water Resources plates from the National Atlas of India, which contains hydrological maps of the presence of three categories of aquifer within these unconsolidated formations: thickest (aquifer greater than 150 meters), fairly-thick (aquifer thickness between 100 and 150 meters), and thick (aquifer thickness up to 100 meters). We code

[7] The other three rounds in this series are: 1986-87, 1993-94, and 2000-01.

each of these categories as a binary variable (Figure 1). The thickness of the aquifer reflects groundwater abundance and is determined prehistorically. The earliest Indian aquifer formations date back to 3,500 million years ago, with the youngest aquifers dating back to the Pleistocene age (Jain *et al*. 2007). Therefore, it does not measure the water table or annual water depth within the aquifer but captures a long-term geological potential (Jain *et al*. 2007). The purpose of the process was to create a map of the location of prehistoric water aquifers in India to use as a source of plausibly exogenous variation in groundwater in a robustness check model.

3. Empirical model

Migration has long been thought to play a key role in the efficient allocation of labor. In the seminal Harris and Todaro (1970) model, the benefits of farming prevent rural migration in the hopes of better urban wages. Although the Harris and Todaro (1970) model was focused on long term and permanent migration, the short-term migrants we study also appear to trade off an agricultural opportunity cost with the expected benefits of migration. In our case, seasonal migration offers a diversified stream of income as explained by the Stark (1978) model of risk sharing. Here, migration splits households and diversifies income streams across household members so that livelihoods are divided across seasons within a given year.

In an area where irrigation is important, the opportunity cost of staying when there is no access to irrigation is presumably high. In other words, on-farm labor and irrigation can act as complements for mean consumption. With more capital assets in irrigation and a lower percentage of assets in other forms of capital, such as tractors that are good substitutes for labor, more labor is required on the farm to help with production activities and so less labor migrates out. On the other hand, since on-farm labor and irrigation are both substitutes in reducing income variance the propensity to migrate can rise with more irrigation. The following model tests these hypotheses empirically and finds that the complementarity effect dominates.

To explore the relationship between short-term migration, rainfall variability, and irrigation, we estimate the following binary choice model:

$$STM_{ihd} = \beta_1 R_d + \sum_{j=1} \beta_{2j} W_{dj} + \delta_1 Z_d + \delta_2 H_{hd} + \delta_3 I_{ihd} + \eta_t + \mu_s + \epsilon_{ihd} \quad (1)$$

In equation (1), STM_{ihd} is a latent binary variable that takes a value of one if an individual, i, living in household h and district $d,$ is a short-term migrant. Recall that the NSSO uses a recall survey approach, so during the time of the interview individuals are asked if they spent 1 to 6 months away from their village for work over the previous year. Based on this question, and our own interpretation of what constitutes migration, we define a person as a short-term migrate if they answered "yes" to the survey question and their migration destination was outside of their home district.

While the 64th NSS is the best data available data on short-term migration, it is still limited in its temporal specificity. While we know whether an individual was a short-term migrant over during year preceding the interview, we do not know the specific time of departure during that year and we do not know when the person was interviewed. The best we can say, based on the survey design, is whether a person was interviewed in one of the four three-mouth sub-rounds (July-September 2007, October-December 2007, January-March 2008, or April-June 2008) and whether they migrated over the year prior to being interviewed. We exploit this fact below to associate different weather and irrigation variables with each person based on which year, 2007 or 2008, they were interviewed. While imperfect, this allows for identification based on both spatial and temporal variation.

The variable R_d is a district-level measure of annual rainfall variability. To produce this variable, we compute the absolute deviation of rainfall, in each year, from the 30-year average, in meters, within each district. In the model, we use a one-year lag of this variable given the recall nature of the survey question – respondents are asked to recall what happened over the previous year which suggests that this recollection, in turn, would have been based on what happened over the year preceding the decision. We also exploit the sub-round nature of the data and use 2006 weather shocks for the 2007 sub-rounds and the 2007 weather shocks for the 2008 sub-rounds. Ideally, we would be able to attach monthly survey responses to monthly weather data but given data limitations this is not possible.

The variable W_{dj} is our measure of irrigation coverage where the d subscript is for the district and the j subscript if for the type of irrigation. We define irrigated area in one of the three ways. First, we specify the share of cultivated area in each district that is irrigated regardless of what type of irrigation is used. Second, we break out the percentage of cultivated area in each district into the portion irrigated using groundwater and the portion irrigated using surface water. And finally, we further break out groundwater irrigated area into the percentage of cultivated area that is irrigated by different types of well technology: deep tube wells; shallow tube wells; and dug wells. We chose to focus on the share of cultivated area, as opposed to the number of wells, to assure that we capture actual utilization along with access. A limitation of using our key irrigation variables at the district level is that they do not capture household heterogeneity within a district, an issue we address in the robustness checks. At the same time, since our key variables are measured at the district level concerns about endogeneity bias are reduced.

In addition to our variables of interest, we control for a number of district, household, and individual effects. At the district level, Z_d, we include average annual growing degree days (GDD), average monsoon rainfall over the preceding 30-year period, population, and GDP growth.[8] At the household level, H_{hd}, we include an indicator landownership (households with land at or below 0.4 hectares, which is the median level of land ownership in the data), the natural log of monthly per capita expenditure (MPCE), and household size. MPCE is calculated using the total average value of goods and services a household consumes per month and is often used as a proxy for household income (National Sample Survey Organization, 2010).[9] The amount of land owned represents household assets or wealth which may reduce a household's risk aversion (Kurosaki and Fafchamps, 2002). It also raises the productivity of own farming. Therefore, individuals that are small landholders are more likely to engage in off-farm work through temporary migration. At the individual level, I_{ihd}, we include the level of education of the household members (illiterate, primary school, middle secondary, or higher secondary and above), their employment status (if their primary activity at the time of the survey was casual agricultural or non-agricultural laborer, cultivator, business owner or salaried worker), whether they considered themselves not in the labor force, social group[10] (scheduled caste (SC), scheduled tribe (ST), other backward classes (OBC), and others), religion (Hindu, Muslim, Christian, and others), sex, age, and marital status.

---

[8] To calculate population growth by district between 1991 and 2001 censuses, we use Kumar and Somanathan (2009) who provide population weights that allow for the construction of population totals using boundaries of the 1991 or 2001 census as the base as well as "Districts of India" (www.statoids.com/yin.html) that documents changes in district boundaries since 1982. Annual compounded GDP growth rate between 1999 and 2004 are also constructed.

[9] As it is difficult to collect reliable income data, the National Sample Survey Organization collects data on consumption expenditure in its surveys

[10] Traditionally, caste hierarchy was linked to individuals' occupations: upper castes were landowners, middle-ranked castes the farmers and artisans, and the lowest-ranked (scheduled) castes the laborers who performed menial tasks (Anderson, 2011). While this type of employment rigidity has decreased, the salience of belonging to a caste remains.

Recognizing that seasonality is critical to short-term migration decisions, we include quarter fixed effects (η_t) to account for the different interview-periods spanning July-September, October-December, January-March, and April-June, which are in turn associated with different timings of departure from the source-district. Since most Indian states have existed for more than fifty years, we also include state fixed effects, μ_s, to account for state-level characteristics and policies that could affect the economic conditions that govern the patterns of migration.

Given the village-level sampling design of the NSS, standard errors are clustered at the village level. We also present, as a robustness check, results with standard errors clustered at the district level. Additionally, survey weights are incorporated in all models. Summary statistics for all weather, irrigation, and aquifer variables are shown in Table 1; statistics for district, household, and individual control variables are shown in Table A1.

For our primary analysis, we employ a linear probability model. Linear models are preferable since nonlinear approaches require unrealistically strong model assumptions, especially on the behavior of the error term in the stipulated underlying structural model (Angrist 2001; Wooldridge 2002).[11] Several papers have demonstrated the advantages of a linear probability model over nonlinear models (Mullahy 1990; Klaassen and Magnus 2001; Horrace and Oaxaca 2006), especially if the main purpose is to estimate the marginal response, in terms of a percentage point change in average probability, for small change in some independent variable. We present estimates from the linear probably model in the main paper, and those from a nonlinear probit as a robustness check.

[11] Linear probably models are also advantageous as they allow for easy inclusion of fine-scaled fixed effects and the identification of marginal effects, which is ultimately the parameter of interest in most economic settings.

4. Results

4.1. Main Results

We begin with the main results in Table 2 on the impact of rainfall variability and irrigation availability on short-term migration. The column headings denote the three methods used to specify irrigation coverage. All results are produced using NSS sample weights with the standard errors clustered at the village level. Robustness results with standard errors clustered at the district level are shown in Table A2 of the Appendix.

Beginning with precipitation, we see that positive rainfall shocks are associated with reduced short-term migration. From column (3), we observe that a 1-cm increase in rainfall is associated with a -0.000168 (-0.0168%) percentage point drop in the probability of migration. Given the baseline migration probably of 1.96% in Table 1, this represents a 0.82% change in the probability of short-term migration.[12] For a 1-meter rainfall shock, we get a 0.0168 (1.68%) percentage point reduction in the probability of migration, which represents an 82% drop over baseline migration. While it is unlikely that the change in probabilities would be linear between a 1-cm and 1-meter shock, these values at least give an idea of the range of changes that may occur because of average shocks and extreme shocks.

Turning to irrigation, in column (1) we see that irrigated area has a negative impact on the likelihood that a household dispatches a migrant. While this may be true, it is more likely that this outcome is driven by our aggregation of the different irrigation technologies. Recall that for the model in column (1) we defined irrigation as the percentage of cultivated area that is

[12] To get our estimate for a 1-cm shock to rainfall, we multiplied the parameter estimate of -0.0168 from column (3) by 0.01 to convert the rainfall shock variable from meters to centimeters. The choice of one centimeter was based on the average shock size of -0.01 in Table 1.

irrigated, irrespective of the source. Thus, we are not able to differentiate the effects of different irrigation types – surface vs. ground water and deep vs. shallow tube wells.

In column (2), we divide irrigated area into the share from well water and the share from surface water. From these results, we see that well-water irrigation produces a negative and significant effect and surface water a positive and significant effect. Specifically, we find that one percentage point increase in the area irrigated with wells reduces the probability of migration by 0.02%, and one percentage point increase in the area irrigated with surface water increases the probability of migration by 0.02%. This reduction in short-term migration, for a percentage point increase in well area, represents around a 1% change, relative to baseline, for short-term migration, which is like the reduction produced for 1-cm increase in rainfall.

These results, for the short-term migration response to a change in well-irrigated area, are intuitive given that well water, especially water from deep tube wells, reduces the uncertainty associated with water availability in farming and thus reduces the need to send out short-term migrants to smooth income. Groundwater irrigation is often colloquially called "irrigation on demand" as it enhances average farm productivity and stabilizes output during periods of low rainfall removing the need to engage in income diversification strategies through temporary migration. Several studies have found that it accrues larger benefits to rural economies than surface irrigation (Sekhri 2014; Sekhri 2013).

In column (3) of Table 2, we test this hypothesis – that different well technologies produce different behavioral responses – by breaking out well irrigation into the share performed with deep tube wells, shallow tube wells, and dug wells. In this model, we find that all well coefficients are negative or insignificant and that the surface water coefficient is positive and significant. In addition, we see that the coefficient on deep tube-well irrigation is largest in absolute value and larger than the coefficient for the shallow tube wells. As stated above, this result reflects fundamental differences between the three technologies. Dug wells, which remain a common source of irrigation in India, are the shallowest wells and rely on use of suction pumps that run on diesel or electricity (Dubash 2002; Jain *et al.* 2007). As water levels fall, however, it becomes increasingly difficult to use dug wells as suction limits the height from which water can be drawn – typically 8-10 meters (Dubash 2002; Sekhri 2011). Conversely, tube wells and submersible force pumps can lift water from greater depths providing 3-15 times as much water as dug wells (Jain *et al.* 2007). Thus, tube wells allow groundwater to be pumped from even greater depths compared to dug wells and can provide longer-term access to groundwater. For instance, shallow tube wells provide approximately 2-3 times the water available in comparison to dug wells, while deep tube wells can provide 15 times the water and the highest level of certainty (Jain *et al.* 2007).

Overall, the results in Table 2 suggest that short-term migration decisions respond to agricultural opportunity costs, especially those associated with rainfall variability and different types of irrigation technology. To provide some additional support for this conclusion, in Table 3 we re-estimate the model in column (2) of Table 2 with interaction terms added between the rainfall shock and the ground and surface water variables. The hypothesis, based on the discussion above, is that increased rainfall (positive rainfall shocks) dampen the positive migration effect associated with increased surface water irrigation. Specifically, rainfall increases make surface water irrigation more reliable and thus reduce the need to migrate to smooth income. The impact for ground water is ambiguous but given that groundwater relies less on rainfall it is likely that it will not my have an interactive effect.

The results in Table 3 support this hypothesis. As before, the direct effect of a positive rainfall shock is to reduce the propensity to migrate, and the direct effect of ground and surface water is to reduce and increase short-term migration, respectively. However, the interactive effect between surface water irrigation and rainfall is negative. This negative coefficient suggests that while surface irrigation increases the overall need to migrate to smooth income, a positive rainfall shock, which increases surface water availability, dampens this need as it reduces the risk associated with not having enough water. The interaction effect between ground water and rainfall is positive, but statistically insignificant.

4.2. Robustness Checks

In this section, we present results from a series of models designed to assess the robustness of the findings in the previous section. While our main results make sense and align closely with the behavioral response that we would expect given a change in rainfall variability and/or irrigation availability, the outcomes cannot be taken as causal given our lack of a clear identification strategy. To address this deficiency, we provide additional support by testing some of the key assumptions of the model. All models are presented with standard errors clustered at the village level and using the NSS sampling weights unless stated otherwise.

We begin by assessing whether the choice of the linear probability model, and its linear form, are impacting our results. In Table 4, we present results (marginal effects) from a nonlinear probit model using the same data and variables as Table 2. The results are average marginal effects with standard errors calculated using the Delta method. While the results in Table 4 change some, compared to Table 2, they are qualitatively similar suggesting that the linearity assumption in the LPM is not impacting our findings, at least not in terms of the marginal effects.

In our second model, we apply an alternative method in specifying our rainfall shock. In Table 2, we used an absolute difference approach. Here, we calculate the z-score for each rainfall observation. Specifically, we use the same 30 years of rainfall data, at the district level, as we did in Table 2, but calculate the z-score for each observation in 2006 and 2007 based on these data. The results from this process are shown in Table 5; results clustered at the district level are given in Table A3 of the Appendix.

The results in Table 5 are very similar to those in Table 2. The rainfall coefficients, while different as a result of using a different variable, have the same sign and significance as those in Table 2. In addition, the results for irrigation are identical to Table 2. Specifically, we find, in column (3), that a one standard deviation increase in rainfall leads to a 0.003 (0.3%) percentage point drop in the probably of migration. To compare this result to the result in column (3) of Table 2, we note that a one standard deviation increase in the rainfall is equal to 0.22 meters of rain. If we multiple the value from Table 2 (0.0161) by 0.22 we get 0.0035 (0.35%), which is very similar to the results in column (3) of Table 5. Thus, based on these results it does not appear that the way we are specifying our rainfall shock in Table 2 is significantly impacting our findings.

For our third robustness check, we assess the extent to which using district-level weather and irrigation data with individual migration information may impact our conclusions. Since our variables of interest, irrigation and rainfall, are measured at the district level, it is possible that the actual weather shocks and irrigation coverage faced by individuals, within district, may differ from the district-level proxies we include in the model, i.e., heterogeneity may exist in how individuals experience irrigation and weather outcomes within district. To the extent that this is occurring, our estimates will be biased and inconsistent. To address this, we estimate two different models. First, we estimate a district-level model where the outcome and the regressors are on the same spatial scale. Then, we estimate an individual-level model where we parametrically control for district-level heterogeneity. Each model represents a different way of accounting for heterogeneity bias.

For the first approach, we compress all individual data to the district level using NSS sample weights and estimate a fractional probit model. For this model, the outcome variable is measured as the share of individuals aged 15 to 65 within in each district level that are listed as short-term migrants.[13] As in Table 2, we include rainfall shocks and estimate three separate models for each of our irrigation measures. Given the nonlinear nature of the fractional probit model, we present marginal effects for each of the variables.

The results from the fractional probit models are shown in Table 6. From these models, we see that the results, compared to Table 2, are almost identical for the irrigation variables and are very similar for the weather shocks. For example, from column (3) we find that a 1-cm increase in rainfall leads to a 0.000109 (0.0109%) percentage point drop in the share of people migrating. This drop represents a 0.56% change away from the baseline for the outcome in the fractional probit model and is very similar to the 0.86% drop from column (3) in Table 2.

For the second heterogeneity model, we estimate a random effects panel data model with random terms specified at the district level. This model allows us to explicitly test if intra-district heterogeneity is impacting our results.[14] The results are shown in Table 7. Comparing these results with those in Tables 2, we find qualitatively similar outcomes. The coefficients for

[13] We also estimated a negative binomial model where the outcome was the count of migrants in each district, and the right-hand side of the model included an exposure term to control for the number of possible migrants in each district. The marginal effects from that model are very similar to those in the fractional probit model.
[14] We also estimated a random effects model with random effects specified at the household-level (Table A4 of the Appendix). This model controls for intra-household heterogeneity in how individuals experience weather shocks and irrigation technology. The results are like the intra-district model.

irrigation are almost identical, and the coefficients for the weather shocks, while smaller, are statistically significant and generally agree with the columns in Table 2. Based on the results in Tables 6 and 7, it does not appear that the aggregation of our whether and irrigation data is meaningfully impacting our findings.

For our final robustness check, we address the concern that our irrigation variables may be biased by the endogeneity of historical irrigation infrastructure investment – i.e., we attempt to test to what extent the variation in the irrigation infrastructure that we exploit in our model may be endogenous to migration decisions or determined by factors that jointly affect irrigation investment and short-term migration our results will be biased. To alleviate this concern, we exploit prehistoric spatial variation in the thickness of underground aquifers using geo-referenced hydrological maps.

Prior work has highlighted that regions with access to greater groundwater endowments, as proxied for by the thickness of aquifers, saw large increases in cultivated area under high-yield varieties during the Green Revolution (Rud 2012; Sekhri 2014). In our context, we are interested in whether the properties of the aquifers in each district have led to increased usage tube and dug-well technology. To confirm this, we collect district-level data on irrigation shares from the national agricultural censuses for the years 1970-2005 and relate the change in the usage of different technologies of the properties of the aquifers in each district using the following model:

$$Y_{dt} = \alpha_0 + \alpha_1 Thickest_d + \alpha_2 Fairly - Thick_d + \lambda_t \tag{2}$$

$$+ \sum_t \alpha_{1t}(Thickest_d * T_t) + \sum_t \alpha_{2t}(Fairly - Thick_d * T_t) + \epsilon_{dt}$$

In this model, the outcome variable, Y_{dt}, represents either tube-well or dug-well irrigated area in each district, in each year, $Thickest$ is an indicator that equals one if a district has access to the thickest aquifers, and $Fairly\text{-}Thick$ is an indicator that equals one if a district has access to fairly-thick aquifers with the excluded group the thick aquifer category. λ_t are year fixed effects and T_t are year dummies interacted with aquifer thickness. The standard errors are clustered at the district level. The coefficients of interest are α_{1t} and α_{2t}, where positive and significant values indicate that districts with greater access to an aquifer category irrigate more via tube wells or dug wells. Increasing values of α_{1t} and α_{2t} imply that over time the factors responsible for successful high-yield crop adoption also triggered a divergence in the type of irrigation used.

To demonstrate that aquifer properties do indeed impact irrigation outcomes, we estimate the model in equation (2) for both tube and dug wells and plot the yearly coefficients for the districts with the thickest aquifers. The results from this process are shown in Figures 2 and 3. From these figures, we see districts with the greatest groundwater endowments saw the highest levels of tube-well irrigation (Figure 2), but a fall in dug well irrigation (Figure 3). These results demonstrate that the variation in groundwater resources clearly influenced the subsequent development of irrigation technologies in India and thus serve as a source of exogenous variation for tube-well and dug-well irrigation investment.

Having validated the use of aquifer properties as a proxy for the development of irrigation infrastructure, we re-estimate the model in equation (1) replacing our irrigation variables (columns 1-3 in Table 2) with one of two methods for accounting for the type of aquifer underlying each district. For the first method, we use a simple measure of the share of land area in each district covered by an unconsolidated aquifer. For the second method, we break out the unconsolidated shares into the percentages covered by different types (thicknesses) of aquifers.

The results from our aquifers models are shown in Table 8. Column (1) uses just the unconsolidated vs. consolidated designation, and column (2) breaks aquifers into different thicknesses. All models use NSS sample weights with standard errors clustered at the household level. The results in Table 8 look very similar to those in Table 2. For the rain-shock variables, the coefficients are almost identical. For the aquifer variables, while their coefficients are different from the coefficients on the irrigation variables in Table 2, the pattern is the same. The thickest aquifers have the highest coefficient values, in absolute terms, which is what we would expect given that these aquifers most likely proxy for where deep tube wells are drilled. Thus, as we saw in Table 2, access to reliable sources of water reduces the likelihood of a short-term migration falls in districts that overlay the thickest aquifer, and where area covering the unconsolidated aquifer is the greatest, mirroring the results for tube well irrigation.

4.3. Heterogeneity Analysis

The previous two sections presented our main findings and the robustness of those results for the full sample. In this section, we expand on our main results (Table 2) and demonstrate how they vary across individuals based on the land ownership and occupation. Using the same sample and model as in column (3) of Table 2, we divide the sample into two groups based on the landownership status of the household. Specifically, we use a cutoff of 0.4 hectares, which represents the median level of landownership in the full sample, and define those households at or below this value as Small Landowners (SLH) and those above it as Large Landowners

(LLH).[15] Then, using those two sub-samples, we create an indicator for whether a person is classified as one of five types of workers – casual non-agricultural laborer (nonagcausal), salaried worker, business worker, casual agricultural worker (agcausal), or cultivator – and interact it with our weather shock variable. The purchase of this exercise is to: (1) look at how rainfall shocks impact households based on landownership and (2) determine if individuals involved in agriculture are differentially impacted by rainfall shocks relative those that are not.

The results from our heterogeneity models are shown in Table 9. Column (1) shows SLH results and column (2) shows LLH results. From these results, we first observe that SLHs are more impacted by rainfall shocks. Specifically, a 1-cm shock to rainfall leads to a 0.017% decrease in short-term migration for SLHs vs. a 0.014% drop for LLHs. This result likely stems from the fact the land ownership signifies greater wealth and thus results in a greater ability to hedge income risks. Turning to the interactive effect of rainfall shocks with different occupations, we see that agcausal with and without land and land-owning business owners have an additional effect, in terms of short-term migration, relative to the base group nonagcasual. For agcasual in SLHs, the additional effect on short-term migration, of a 1-cm increase in rainfall, is 0.015%, and for agcasual in LLH the additional effect is 0.064%. Both results suggest, as expected, that for those households engaged casual agriculture the impact of a positive rainfall shock is more negative in terms of short-term migration – i.e., under conditions of greater rainfall it is less necessary to immigrate short term to smooth income. The fact that the effect is larger for LLHs also seems plausible given their direct access to land to farm that becomes relatively more valuable in times of excess rainfall. The explanation for the additional effect for land-owning

[15] The data does not actually show specific values for land ownership but has ten different categories. We use the categories to roughly approximate the median ownership rate in the sample.

businesspeople is less clear, but it is possible that there is an inactive effect between increased rainfall, better agriculture outcomes, and business success.

The results in Table 9, while still not causal, do provide a more nuanced picture of how rainfall shocks impact short-term migration and indicate that is agricultural households that are most impacted by these shocks. Specifically, they show that while rainfall shocks impact the tendency for all rural households to migrate to some extent, the impact is greater for households that depend on casual agriculture as a source of income and greater still for households with more land farm.

5. Discussion and conclusions

In this paper, we highlight the relationship between two distinct adaptation responses: groundwater irrigation and short-term migration and show that groundwater availability and access does have an economically significant impact on rural, short-term labor mobility.
We find that access to tube-well irrigation allows individuals to specialize in agricultural related activities thereby reducing the likelihood of short-term migration. Irrigation can, therefore, serve as an alternative to short-term migration as a risk mitigation strategy so long as it is available and utilized. A simple quantification exercise suggests that at least 10% (or approximately 1.2 million people) of the baseline number of short-term migrants in rural areas would not move temporarily if groundwater irrigation coverage increased by 10 percentage points.

From a policy perceptive, shutting down access to groundwater in response to growing depletion will have significant effects on temporary labor mobility and likely household welfare, especially if the reduction in availability happens very quickly. On the flip side, however, if water is not managed properly it may be that in the long run its availability will also disappear quickly. In the short run, low productivity or negative rural sector shocks driven by less groundwater irrigation or rainfall shocks can spill over to the rest of the economy. In the long-run (for any given technology frontier) productivity growth in agriculture will be limited by fixed factors of production (for example, through the depletion of groundwater). In such cases where the fixed factor of production such as water limits production, preference should be given to speeding the transition out of agriculture to the non-agricultural sector. Thus, any policy must balance these trade-offs. This would suggest that the optimal strategy ought to be one of: (i) improving farm productivity and value added in the short run since it is often the case that farms provide livelihoods where there is no social security, (ii) while promoting resilience and sustainability in the long run where fixed factor supplies of a resource (water) could limit growth and (iii) technology change to ease the binding constraint.

In India, although the census and national surveys have focused primarily on long term migration, short-term migration is widely practiced, and comprises more than five times the proportion of households that migrate for longer durations (Colmer 2015). This paper shows that irrigation plays a critical role as a measure of the household's opportunity cost of sending a short-term migrant. Since these types of migrants are negatively selected on education and measures of economic status, the types of jobs available to them in receiving areas are limited. Multiple studies have shown that many of the seasonal migrants get absorbed in the construction industry (Chandrasekhar et al. 2019). The conditions of workers in this industry are meant to be regulating by a variety of acts[16](Chandrasekhar *et al*. 2019), but these are not streamlined across states and the workers often do not accrue the benefits promised by these regulations. India, as of now, does not have in place a comprehensive policy framework to tackle to the flow of short-term migrants (Chandrasekhar *et al.* 2019)

[16] Building and Other Construction Workers (Regulation of Employment and Conditions of Services) Act, 1996 and Building and Other Construction Workers Welfare Cess Act, 1996

There is an important caveat to the interpretation of our results. Due to data limitations and the cross-sectional nature of our analysis, we only observe short-term migration over one survey round. It is possible that over time new forms of migration might emerge. For instance, if irrigation enables farm households to improve their agricultural productivity and household income and break the income constraint, then it could encourage male members to switch from temporary migration to more permanent remittance-based migration over time. Irrigation access, therefore, might reduce certain types of migration, but over time, also encourage other types of migration. Despite this caveat, we believe the paper makes an important contribution towards understanding the interaction between different adaptation responses, a critical exercise when evaluating climate change adaptation policies going forward.

The data used in the paper offer only a snapshot of the seasonal migration phenomenon in the country. More data is needed to understand the consequences of such type of migration on income and if it can offset the income effects of losing irrigation access. More work is also needed to understand the types of adjustment costs and spatial frictions involved in the migration process, and the resulting reallocation of labor across space and economic sectors that may arise. For instance, the economic consequences of a local negative productivity shock such as a lack of access to irrigation water can be even greater if people are unable to move away. This has important implications for broader issues of economic development and structural transformation. Researchers have argued that urban economic growth has brought significant gains to the rural as well as urban poor (Datt and Ravallion 2009). Short-term migration could be one channel by which some of the benefits of urban growth could be brought to rural villages before villages completely transition out of agriculture (Coffey *et al.* 2014). Without knowing the exact destination of short-term migration flows, it is difficult to analyze the contribution of migration flows to the urbanization process. At the same time, patterns of vulnerability can persist in destination regions continuing exposure to environmental risks (Singh and Basu 2019). Future research could shed light on these long-term dynamics and the consequent spillovers on human capital, and welfare.

References

Agrawal, A. and Lemos, M. C. (2015). Adaptive development. Nature Climate Change, 5(3),185-187.

Anderson, S. (2011). Caste as an Impediment to Trade. *American Economic Journal: Applied Economics* 3, 239-263.

Angrist, J. D. (2001). Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice. *Journal of Business & Economic Statistics* 19, 2–16

Badiani, R. and Safir A. (2010). Coping with aggregate shocks: Temporary migration and other labor responses to climatic shocks in rural India. World Bank Mimeo.

Banerjee, A. V. and Duflo, E. (2007). The economic lives of the poor. *The Journal of Economic Perspectives*, 21, 141-168.

Benonnier, T., Millock, K. and Taraz, V. (2019). Climate change, migration, and irrigation. *Technical report, HAL Archives ouvertes.*

Breman, J. (1996). *Footloose Labour: Working in India's Informal Economy.* Cambridge University Press, Cambridge, UK.

Bryan, G., Chowdhury, S. and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica, 82*(5), 1671-1748.

Burgess, R. and Donaldson, D. (2010). Can Openness Mitigate the Effects of Weather Shocks? Evidence from India's Famine Era. *Am. Econ. Rev.* 100(2), 449-53.

Cai, R., Feng, S., Oppenheimer, M. and Pytlikova, M. (2016). Climate variability and international migration: The importance of the agricultural linkage. *Journal of Environmental Economics & Management* 79, 135–151.

Cattaneo, C. and Peri, G. (2016). The migration response to increasing temperatures. *Journal of Development Economics* 122, 127–146.

Chandrasekhar, S., Das, M. and Sharma, A. (2014). Short-term migration and consumption expenditure of households in rural India. Indira Gandhi Institute of Development Research Working Paper.

Chen, J., and Mueller, V. (2018). Coastal climate change, soil salinity and human migration in Bangladesh. *Nature Climate Change* 8(11), 981-985.

Coffey, D., Papp, J. and Spears, D. (2014). Short-Term Labor Migration from Rural North India: Evidence from New Survey Data. *Population Research and Policy Review*, 1-20.

Colmer, J. (2015). Urbanization, growth, and development: Evidence from India. *Mimeo.*

Dallmann, I. and Millock, K. (2013). Climate Variability and Internal Migration: A Test on Indian Inter-State Migration. Paris School of Economics Working Paper.

Datt, G. and Ravallion, M. (2009). Has India's Economic Growth Become More Pro-Poor in the Wake of Economic Reforms? *The World Bank Development Research Group*, Policy Research Working Paper 5103.

Deshingkar, P., Farrington, J. (2009). A Framework for Understanding Circular Migration. In: Deshingkar P, Farrington J (eds) *Circular Migration and Multilocational Livelihoods in Rural India*, pp.1-36. Oxford University Press, New Delhi, India.

Dubash, N. K. (2002). *Tubewell capitalism: groundwater development and agrarian change in Gujarat*. Oxford University Press, New Delhi, India

Feng, S., Krueger, A. B. and Oppenheimer, M. (2010). Linkages among climate change, crop yields and Mexico-US cross-border migration. *Proceedings of the National Academy of Sciences* 107, 14257-14262.

Fishman, R. M. (2012). Climate change, rainfall variability, and adaptation through irrigation: Evidence from Indian agriculture. Working Paper

Fishman, R., Jain, M. and Kishore, A. (2015). When Water Runs Out: Scarcity, Adaptation and Migration in Gujarat. International Growth Center Working Paper.

Harris, J. and Todaro, M. P. (1970). Migration, unemployment and development: A two sector analysis. *American Economic Review* 60 (1), 126-142.

Horrace, W. C. and Oaxaca, R.L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economic Letters* 90(3), 321-327.

Jain, S., Agarwal, P. and Singh, V. (2007). *Hydrology and Water Resources of India*. Springer Science & Business Media, Dordrecht, The Netherlands.

Keshri, K. and Bhagat, R. (2012). Temporary and seasonal migration: Regional pattern, characteristics and associated factors. *Economic and Political Weekly* 47, 81-88.

Klaassen, F. and Magnus, J. (2001). Are Points in Tennis Independent and Identically Distributed? Evidence from a Dynamic Binary Panel Data Model. *Journal of the American Statistical Association* 96(454), 500- 509

Kurosaki, T. and Fafchamps, M. (2002). Insurance Market Efficiency and Crop Choices in Pakistan. *Journal of Development Economics* 67(2), 419–453.

Lanjouw, P. and Shariff, A. (2004). Rural Non-Farm Employment in India: Access, Incomes and Poverty Impact. *Economic and Political Weekly* 39(40), 4429-4446.

Lemos, M. C., Agrawal, A., Eakin, H., Nelson, D. R., Engle, N. L. and Johns, O. (2013). Building adaptive capacity to climate change in less developed countries. In: Climate science for serving society, pp 437–457. Springer, Dordrecht, The Netherlands.

Marchiori, L., Maystadt, J.-F. and Schumacher, I. (2012). The impact of weather anomalies on migration in sub-Saharan Africa. *Journal of Environmental Economics and Management* 63, 355-374.

Menon, A., Levermann, A., Schewe, J., Lehmann, J. and Frieler, K. (2013). Consistent increase in Indian monsoon rainfall and its variability across CMIP-5 models. *Earth Sys. Dynam.* 4(2), 287-300.

Morten, M. (2019). Temporary migration and endogenous risk sharing in village India. *Journal of Political Economy* 127(1), 1-46.

Mullahy, J. (1990). Weighted least squares estimation of the linear probability model, revisited. *Economics Letters* 32(1), 35-41.

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

National Sample Survey Organization (2010). Employment and Unemployment situation in India 2007-08 NSS 64th Round July 2007 to June 2008. Technical report, Ministry of Statistics and Programme Implementation, Government of India.

Roy, T. (2016). Were Indian Famines 'Natural' Or 'Manmade'? *LSE Working Paper No. 243*

Rud, J. P. (2012). Electricity provision and industrial development: Evidence from India. *Journal of Development Economics* 97(2), 352-367

Sekhri, S. (2011). Public provision and protection of natural resources: Groundwater irrigation in rural India. *American Economic Journal: Applied Economics* 3(4), 29-55.

Sekhri, S. (2013). Missing Water: Agricultural Stress and Adaptation Strategies in response to Groundwater Depletion among Farmers in India. Working Paper

Sekhri, S. (2014). Wells, Water, and Welfare: The Impact of Access to Groundwater on Rural Poverty and Conflict. *Am Econ J Appl Econ* 6(3), 76-102.

Singh, D., Tsiang, M., Rajaratnam, B. and Diffenbaugh, N.S. (2014). Observed changes in extreme wet and dry spells during the South Asian summer monsoon season. *Nature Climate Change* 4(6), 456-461.

Singh, C. and Basu, R. (2019). Moving in and out of vulnerability: interrogating migration as an adaptation strategy along a rural-urban continuum in India. *The Geographical Journal* https://doi.org/10.1111/geoj.12328

Stark, O. (1978). Economic-demographic interactions in agricultural development: The case of rural to urban migration. In: U.N.F. for Population Activities (ed) *Population and agricultural development: Selected relationships and possible planning uses*, Volume 6. Food and Agricultural Organization of the United Nations.

Taraz, V. (2018). Can farmers adapt to higher temperatures? Evidence from India. *World Development* 112, 205–219.

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

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, MA.

World Bank (2016). *High and Dry: Climate Change, Water, and the Economy*. World Bank, Washington D.C.

Figures and Tables

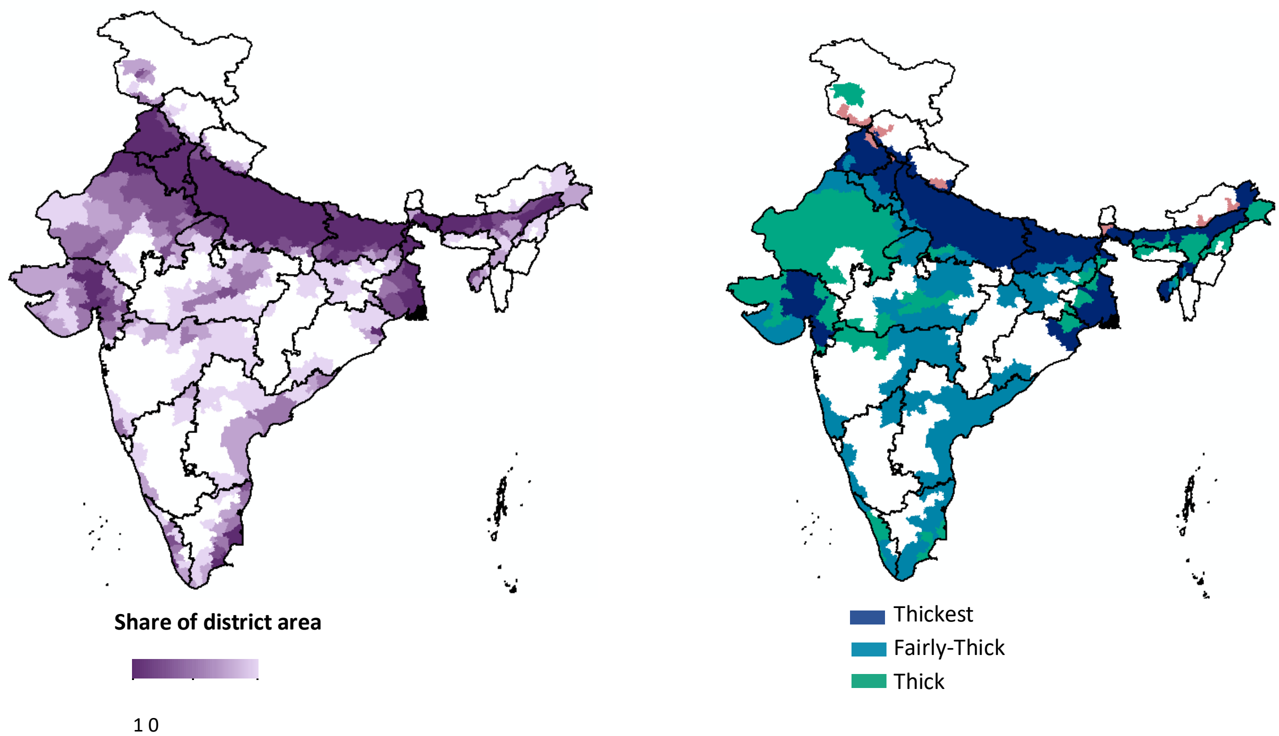

Figure 1 Aquifer Coverage

Notes: The map on the left shows the share of district area overlaying an unconsolidated aquifer for each district using 2001 census boundaries. The map on the right classifies the unconsolidated aquifer into 3 categories of thickness: Thickest (> 150 m), Fairly-Thick (100 - 150m) and Thick (≤100 m). State boundaries are in black. Data is from the 1969 Geohydrology Map of India and the 1982 Water Resources Plates in the National Atlas of India.

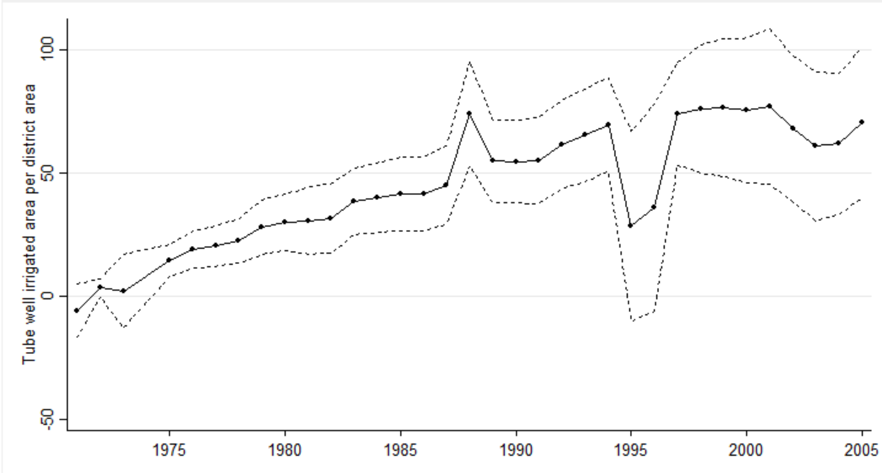

Figure 2 Differential trends in tube well irrigated area by aquifer capacity

Notes: Figure plots coefficients from equation (2) that capture year-varying effects of groundwater coverage related to the Thickest aquifer on tube well irrigated area. Area is measured in 1000 hectares. Regressions are weighted by district area. The dashed lines represent 95% confidence intervals. Data is from Directorate of Economics and Statistics, Ministry of Agriculture, Geological Survey and National Atlas of India.

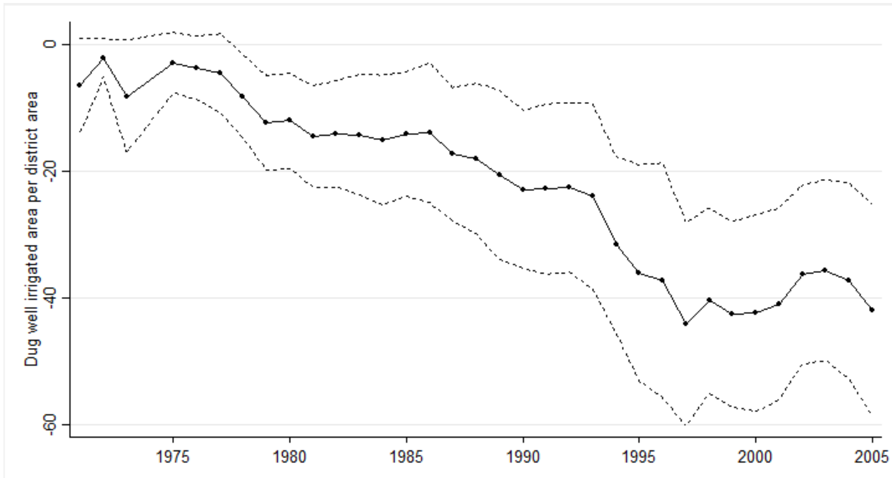

Figure 3 Differential trends in dug well irrigated area by aquifer capacity
Notes: Figure plots coefficients from equation (2) that capture year-varying effects of groundwater coverage related to the Thickest aquifer on dug well irrigated area. Area is measured in 1000 hectares. Regressions are weighted by district area. The dashed lines represent 95% confidence intervals. Data is from Directorate of Economics and Statistics, Ministry of Agriculture, Geological Survey and National Atlas of India.

Table 1 Summary statistics

Variables	Mean	St. Dev.	Min.	Max.	Level of Variable	Data Source
Migration						
Short-term migrants	1.96	13.85	0	1	Individual	NSS 2007-08
Weather						
MonsoonPrecip (30-year mean in meters)	6.44	0.66	0.22	7.72	District	APHRODITE
GDD (30-year mean in 1000s of degree days)	1.04	0.54	0.14	5.09	District	APHRODITE
RainfallShock (Deviation from 30-year mean)	-0.01	0.22	-1.01	1.23	District	APHRODITE
RainfallShock (z-score)	-0.11	1.14	-3.17	3.16	District	APHRODITE
Irrigation						
IrrigatedArea/CultivatedArea (%)	35.77	28.12	0	100	District	Minor Irrigation Census 2006-07
SurfaceIrrigatedArea/CultivatedArea (%)	6.25	10.75	0	100	District	
WellIrrigatedArea/CultivatedArea (%)	29.95	29.36	0	100	District	
DeepTubewellIrrigatedArea/CultivatedArea (%)	4.95	10.04	0	100	District	
ShallowTubewellIrrigatedArea/CultivatedArea (%)	19.66	28.31	0	100	District	
DugwellIrrigatedArea/CultivatedArea (%)	6.23	11.84	0	100	District	
Depth to groundwater level (m)	6.98	6.18	0.98	71.72	District	Groundwater Board 2005-09
Aquifer						
UnconsolitdatedAquifer (%)	46.39	43.24	0	100	District	
ThickestAquifer (> 150 meters)	36.05	48.02	0	100	District	
FairlyThickAquifer (100-150 meters)	12.84	33.46	0	100	District	
ThickAquifer (< 100 meters)	24.24	42.85	0	100	District	

Note: This table presents summary statistics for our outcome variable (migration) and our weather, irrigation, and aquifer variables. Summary statistics for the individual and household variables included in all models are shown in Table A1 in the Appendix. All statistics are generated using NSS sample weights.

Table 2 Short-term migration response to rainfall and irrigation

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0038*	0.0034*	0.0039*
	(0.002)	(0.002)	(0.002)
MonsoonPrecip (30-year mean in meters)	-0.0008	-0.0015	-0.0013
	(0.001)	(0.001)	(0.001)
RainfallShock (Deviation from 30-year mean)	-0.0162***	-0.0164***	-0.0168***
	(0.005)	(0.005)	(0.005)
IrrigatedArea/CultivatedArea (%)	-0.0001*		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002*	0.0002*
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0003***
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001***
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			-0.0000
			(0.000)
Constant	0.0696***	0.0732***	0.0673***
	(0.026)	(0.026)	(0.026)
Observations	213379	213379	213197
R-sq	0.0301	0.0306	0.0307

Note: This table presents our main results. The short-term migration sample is composed of all individuals aged 15 to 65 interviewed from July 2007 through June 2008. The dependent variable is a binary indicator equal to 1 if an individual has spent one to six months away from home during the last year. All regressions include sub-round and state fixed effects and district, household, and individual controls. The rainfall shock used in these models is measured at the district level and represents the deviation in average annual rainfall from the 30-year within district mean in the year before the sample was collected. For rounds 1 and 2 of the NSS this is the deviation in rainfall for the year 2006, and for rounds 3 and 4 of the NSS this is the deviation in the year 2007. The columns represent our three methods for defining district-level irrigation. Standard errors are clustered at the village level. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table 3 Short-term migration response to interaction between rainfall and irrigation

GDD (30-year mean in 1000s of degree days)	0.0033*
	(0.002)
MonsoonPrecip (30-year mean in meters)	-0.0019
	(0.001)
RainfallShock (Deviation from 30-year mean)	-0.0144**
	(0.007)
WellIrrigatedArea/CultivatedArea (%)	-0.0002***
	(0.000)
WellIrrigatedArea-x-RainfallShock	0.0001
	(0.000)
SurfaceIrrigatedArea/CultivatedArea (%)	0.0002**
	(0.000)
SurfaceIrrigatedArea-x-RainfallShock	-0.0008*
	(0.001)
Constant	0.0725***
	(0.026)
Observations	213379
R-sq	0.0355

Note: This table presents results with an interaction between irrigation technology and rainfall. The results are based on column (2) in Table 2 with well and surface water irrigation interacted with the rainfall shock. NSS sample weights are used, and the standard errors are clustered at the household level.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table 4 Short-term migration response to rainfall and irrigation (binary probit model)

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0068**	0.0061*	0.0071**
	(0.003)	(0.003)	(0.003)
MonsoonPrecip (30-year mean in meters)	-0.0024	-0.0035*	-0.0033*
	(0.002)	(0.002)	(0.002)
RainfallShock (Deviation from 30-year mean)	-0.0178***	-0.0177***	-0.0188***
	(0.004)	(0.004)	(0.004)
IrrigatedArea/CultivatedArea (%)	-0.0001		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002**	0.0002**
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0004***
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001***
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			0.0000
			(0.000)
Observations	213246	213246	213064
Pseudo R-sq	0.1884	0.1904	0.1913

Note: This table presents results for the same models as Table 2 estimated using a binary probit model. The results are average marginal effects with the standard errors clustered at the village level. Standard errors for the marginal effects are generated using the Delta method. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table 5 Short-term migration response to rainfall and irrigation (z-score for rainfall shocks)

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0034*	0.0030	0.0035*
	(0.002)	(0.002)	(0.002)
MonsoonPrecip (30-year mean in meters)	-0.0011	-0.0018	-0.0016
	(0.001)	(0.001)	(0.001)
RainfallShock (z-score)	-0.0027***	-0.0028***	-0.0030***
	(0.001)	(0.001)	(0.001)
IrrigatedArea/CultivatedArea (%)	-0.0001		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002*	0.0002*
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0003***
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001**
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			0.0000
			(0.000)
Constant	0.0726***	0.0761***	0.0700***
	(0.026)	(0.026)	(0.026)
Observations	213379	213379	213197
R-sq	0.0300	0.0304	0.0306

Note: The models in this table are the same as in Table 2, but with the rainfall shock variable replaced with a z-score variable representing the within district standard deviation of rainfall from the 30-year within district mean. The within district z-scores used are based on the year before the sample was collected. For rounds 1 and 2 of the NSS this is the deviation in rainfall for year 2006, and the rounds 3 and 4 of the NSS this is the deviation in year 2007. Standard errors are clustered at the village level. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table 6 District-level migration response to rainfall and irrigation (fractional probit model)

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0037**	0.0031*	0.0035*
	(0.002)	(0.002)	(0.002)
MonsoonPrecip (30-year mean in meters)	0.0015	0.0012	0.0012
	(0.002)	(0.002)	(0.002)
RainfallShock (Deviation from 30-year mean)	-0.0110***	-0.0105***	-0.0109***
	(0.004)	(0.004)	(0.004)
IrrigatedArea/CultivatedArea (%)	-0.0001**		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002***	0.0002**
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0003***
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001***
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			-0.0000
			(0.000)
Observations	2047	2047	2043
Pseudo R-sq	0.0723	0.0743	0.0746

Note: The results in this table are marginal effects from a fractional probit model estimated on short-term migration data aggregated to district level. The outcome in each model is the share of individuals in each district during each of the four sub rounds of the NSS that were listed as short-term migrants. All individual and household variables are aggregated up to the district level using sampling weights from the NSS. All weather and irrigation variables are measured at the district level. Each model includes state fixed effects and the natural log of population in each district. Standard errors are robust.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table 7 Short-term migration response to rainfall and irrigation (random effect model)

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0035	0.0031	0.0035
	(0.002)	(0.002)	(0.002)
MonsoonPrecip (30-year mean in meters)	-0.0013	-0.0020	-0.0018
	(0.002)	(0.002)	(0.002)
RainfallShock (Deviation from 30-year mean)	-0.0061***	-0.0061***	-0.0062***
	(0.002)	(0.002)	(0.002)
IrrigatedArea/CultivatedArea (%)	-0.0001**		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002**	0.0002**
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0003***
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001***
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			-0.0000
			(0.000)
Constant	0.0653***	0.0637***	0.0188
	(0.013)	(0.013)	(0.013)
Observations	213379	213379	213379
R-sq-within	0.0231	0.0231	0.0232
R-sq-between	0.146	0.151	0.151
R-sq-overall	0.0297	0.0302	0.0303

Note: This table presents results from a random effects panel data model with random effects specified at the district level. All regressions include sub-round and state fixed effects and district, household, and individual controls. The rainfall shock used in these models is measured at the district level and represents the deviation in average annual rainfall from the 30-year within district mean in the year before the sample was collected. For rounds 1 and 2 of the NSS this is the deviation in rainfall for year 2006, and the rounds 3 and 4 of the NSS this is the deviation in year 2007. Standard errors are clustered at the village level. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table 8 Short-term migration in response to aquifer properties

	(1)	(2)
GDD (30-year mean in 1000s of degree days)	0.0062***	0.0058***
	(0.002)	(0.002)
MonsoonPrecip (30-year mean in meters)	0.0005	-0.0000
	(0.001)	(0.001)
RainfallShock (Deviation from 30-year mean)	-0.0164***	-0.0146***
	(0.005)	(0.005)
UnconsolitdatedAquifer (%)	-0.0002***	
	(0.000)	
ThickestAquifer (> 150 meters)		-0.0243***
		(0.007)
FairlyThickAquifer (100-150 meters)		-0.0154**
		(0.008)
ThickAquifer (< 100 meters)		-0.0093
		(0.006)
Constant	0.0541**	0.0576**
	(0.026)	(0.026)
Observations	234057	234057
R-sq	0.0307	0.0309

Note: The results in this table are the same as Table 2, but with the irrigation variables replaced with variables describing the share of different types of aquifers underlying each district. Model (1) uses a variable for the share of each district covered by an unconsolidated aquifer. Model (2) uses three variables describing the share of different types of aquifers, by depth, covering each district where the left-out share is for the shallowest aquifer. The depths are described in the table. Standard errors are clustered at the village level. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table 9 Heterogeneity analysis

	(1)	(2)
GDD	0.0053***	0.0023
	(0.002)	(0.003)
MonsoonPrecip	-0.0022	0.0006
	(0.001)	(0.002)
RainfallShock	-0.0173***	-0.0135**
	(0.006)	(0.006)
RainfallShock-x-Salary	0.0013	-0.0197
	(0.007)	(0.015)
RainfallShock-x-Business	0.0048	-0.0286*
	(0.008)	(0.017)
RainfallShock-x-AgCasual	-0.0152*	-0.0643*
	(0.009)	(0.035)
RainfallShock-x-AgCultivator	0.0109	0.0031
	(0.008)	(0.006)
Salary	0.0025	0.0239***
	(0.005)	(0.005)
Business	-0.0027	0.0063**
	(0.003)	(0.003)
AgCasual	0.0147***	0.0405***
	(0.003)	(0.009)
AgCultivator	0.0112***	0.0059***
	(0.003)	(0.001)
SurfaceIrrigatedArea (%)	0.0001	0.0004*
	(0.000)	(0.000)
DeepTubewellIrrigatedArea (%)	-0.0002**	-0.0004***
	(0.000)	(0.000)
ShallowTubewellIrrigatedArea (%)	-0.0001***	-0.0001**
	(0.000)	(0.000)
DugwellIrrigatedArea (%)	0.0000	-0.0000
	(0.000)	(0.000)
Constant	0.0420	0.0877***
	(0.031)	(0.032)
Observations	124242	88955
R-sq	0.0297	0.0268

Note: This table presents results from a heterogeneity analysis of the impact of rainfall shocks on different types of people based on employment and land ownership. The results are based on the same model used in column (3) of Table 2. We first divide the sample into two groups based on land the landownership status of the household where the cutoff is the median amount of land owned by each household (0.4 hectares). Second, we create an indicator for whether a person is classified as one of five types of workers – nonagcasual, salary, business, ag casual, or cultivator. We interact this variable with our weather shock variable. Column (1) shows results non landowners and column (2) shows results for landowners. All regressions include sub-round and state fixed effects and district, household, and individual controls. Standard errors are clustered at the village level. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Appendix

A. Motivational Figures and Tables

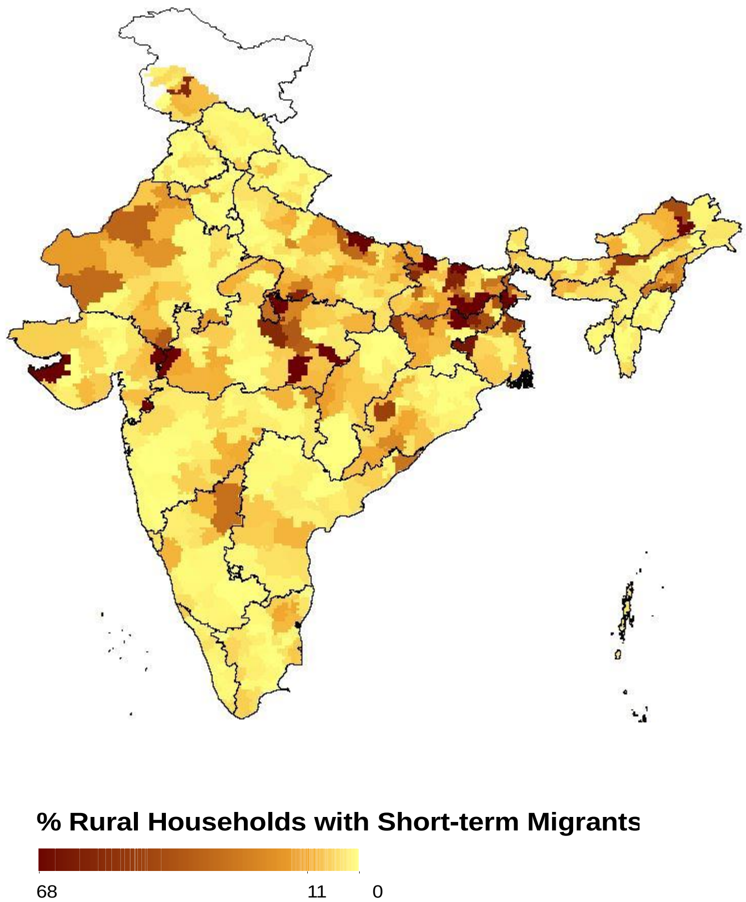

Figure A1 Short-term migration

Notes: The map shows the percentage of rural households with short-term migrants for each district. State boundaries are shown in black. The data used for this map is from the NSS 2007-08.

Table A1 Summary statistics for individual and household variables

Variable	Mean	St. Dev.	Min.	Max.	Level of Variable	Data Source
District Controls						
Population growth 1991-2001	22.89	10.15	-2.80	95.16	District	Census 1991, 2001
Avg. Annual GDP growth 1999-2004	4.65	2.62	-9.72	13.17	District	Planning Commission
Households and Individual controls						
Household size	5.88	2.73	1	30	Household	
Log monthly per capita expenditure (mpce)	8.11	0.54	3.99	12.18	Household	
Female	0.49	0.50	0	1	Individual	
Age	34.94	14.16	15	65	Individual	
Married	0.47	0.50	0	1	Individual	
Land Holdings						
Non-Land Household	0.56	0.42	0	1	Household	
Land Household	0.44	0.40	0	1	Household	
Religion						
Hindu	0.78	0.41	0	1	Individual	
Muslim	0.11	0.31	0	1	Individual	
Christian	0.07	0.25	0	1	Individual	
Others	0.04	0.20	0	1	Individual	
Social Caste						NSS 2007-08
Schedules Caste/ Schedulted Tribe (SC/ST)	0.36	0.48	0	1	Individual	
Other Backward Class (OBC)	0.40	0.49	0	1	Individual	
Others	0.23	0.42	0	1	Individual	
Education levels						
Illiterate	0.39	0.49	0	1	Individual	
Primary	0.34	0.47	0	1	Individual	
Middle and Secondary	0.21	0.40	0	1	Individual	
Higher Secondary and above	0.06	0.24	0	1	Individual	
Employment Status Before Migrating						
Non-Agricultural Casual	0.41	0.12	0	1	Individual	
Salaried Work	0.06	0.34	0	1	Individual	
Business	0.06	0.30	0	1	Individual	
Agricultural Casual	0.13	0.42	0	1	Individual	
Cultivator	0.34	0.19	0	1	Individual	

Note: This table presents summary statistics for the individual and household variables included in all models. All statistics are generated using weighted observations based NSS sampling weights.

Table A2 Short-term migration response to irrigation (standard errors clustered at the district level)

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0038	0.0034	0.0039
	(0.003)	(0.003)	(0.003)
MonsoonPrecip (30-year mean in meters)	-0.0008	-0.0015	-0.0013
	(0.002)	(0.002)	(0.002)
RainfallShock (Deviation from 30-year mean)	-0.0162**	-0.0164**	-0.0168**
	(0.007)	(0.007)	(0.007)
IrrigatedArea/CultivatedArea (%)	-0.0001		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002	0.0002
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0003**
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001*
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			-0.0000
			(0.000)
Constant	0.0696**	0.0732**	0.0673**
	(0.030)	(0.030)	(0.030)
Observations	213379	213379	213197
R-sq	0.0301	0.0306	0.0307

Note: This table presents the same modeling results as Table 2 in the paper with the standard errors clustered at the district level. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table A3 Short-term migration response to irrigation (z-score for rainfall shocks and standard errors clustered at the district level)

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0034	0.0030	0.0035
	(0.003)	(0.003)	(0.003)
MonsoonPrecip (30-year mean in meters)	-0.0011	-0.0018	-0.0016
	(0.002)	(0.002)	(0.002)
RainfallShock (z-score)	-0.0027**	-0.0028**	-0.0030**
	(0.001)	(0.001)	(0.001)
IrrigatedArea/CultivatedArea (%)	-0.0001		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002	0.0002
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0003*
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001*
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			0.0000
			(0.000)
Constant	0.0726**	0.0761**	0.0700**
	(0.030)	(0.031)	(0.030)
Observations	213379	213379	213197
R-sq	0.0300	0.0304	0.0306

Note: This table presents the same modeling results as Table 3 in the paper with the standard errors clustered at the district level. All models use weighted observations based NSS sampling weights.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level

Table A4 Short-term migration response to rainfall and irrigation (random effects model with household-level random terms)

	(1)	(2)	(3)
GDD (30-year mean in 1000s of degree days)	0.0035	0.0031	0.0035
	(0.002)	(0.002)	(0.002)
MonsoonPrecip (30-year mean in meters)	-0.0013	-0.0020	-0.0018
	(0.002)	(0.002)	(0.002)
RainfallShock (Deviation from 30-year mean)	-0.0061***	-0.0061***	-0.0062***
	(0.002)	(0.002)	(0.002)
IrrigatedArea/CultivatedArea (%)	-0.0001**		
	(0.000)		
WellIrrigatedArea/CultivatedArea (%)		-0.0002***	
		(0.000)	
SurfaceIrrigatedArea/CultivatedArea (%)		0.0002**	0.0002**
		(0.000)	(0.000)
DeepTubewellIrrigatedArea/CultivatedArea (%)			-0.0003***
			(0.000)
ShallowTubewellIrrigatedArea/CultivatedArea (%)			-0.0001***
			(0.000)
DugwellIrrigatedArea/CultivatedArea (%)			-0.0000
			(0.000)
Constant	0.0653***	0.0637***	-0.0041
	(0.013)	(0.013)	(0.011)
Observations	213379	213379	213064
R-sq-within	0.0231	0.0231	0.0332
R-sq-between	0.146	0.151	0.0695
R-sq-overall	0.0297	0.0302	0.0349

Note: The results in this table are the same as for the random effects model in Table 7, but with the random effects terms specified at the household level.

* Significant at 10% level; ** Significant at 5% level; *** Significant at 1% level