Erasmus University of Rotterdam

Erasmus School of Economics

Msc Economics & Business

Master Specialisation in Policy Economics

Urban to Rural Migration as a Poverty Alleviation Tool: A Study on the NREGA Job Guarantee Scheme in India

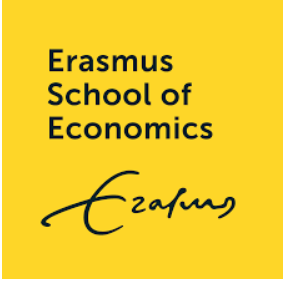
Author: Gaurav Peter Mathai

Student Number: 659514

Supervisor: Dr. Anne Boring

Second Assessor: Dr. Sacha Kapoor

Date: 17th July 2023



The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, the Erasmus School of Economics or the Erasmus University of Rotterdam.

Preface and Acknowledgements

I would like to thank my supervisor, Dr. Anne Boring, for guiding me through the process and reciprocating the same enthusiasm I had for the chosen research topic. I am grateful for her taking the time off to provide me with valuable insights and feedback.

I would like to thank my family and friends for supporting me through this journey.

Abstract

Migration is often considered as one of the greatest poverty alleviation measures undertaken by individuals, with a great amount of research done on migration out of less-developed rural areas to urban areas. The National Rural Employment Guarantee Act (NREGA), rolled out in three phases across India from 2006 to 2008, is one of the largest job guarantee schemes in the world, with guaranteed employment for the rural population. This paper uses the National Sample Survey from Rounds 55 and 64 to explore urban to rural migration using a difference-in-difference estimation strategy, given the possibility of guaranteed employment through the NREGA program; districts from the first two phases are the treatment group and the districts from the third phase are the control group. There is no significant urban to rural migration. However amongst the unemployed, 83.16% of their urban to rural migration is attributed to the NREGA program. There is a statistically significant increase in employment in the public works for urban to rural migrants. Additionally, values of a state, proxied by HDI (Human Development Index) is used to examine whether states with better values have a better implementation of the NREGA program; a unit increase in HDI of a state results in an increase in employment in the public works by 1.88 percentage points, statistically significant at the 1% level. Empirical evidence from this paper suggests the potential that urban to rural migration poses in poverty alleviation.

Table of Contents

1.0 Introduction

Haushofer and Shapiro (2016), and Muralidharan and Prakesh (2017) have noted that in research on socially beneficial outcomes, such as schooling enrolment and poverty alleviation, cash transfer programs are typically the focus. However, it is important to also consider the impact of labour and employment policies on these outcomes. Employment policies can have direct or spillover effects on social outcomes, such as poverty reduction and improving household outcomes (Srivastava & Srivastava, 2010). Therefore, it is crucial to study the impact of such policies, like the NREGA (National Rural Employment Guarantee Act) program, on social outcomes in India, specifically how migration resulting from the NREGA program could be an effective tool for alleviating those in poverty.

The NREGA program grants all households in rural India each 100 days of work per year at a state-level minimum wage, with a third of the jobs reserved for females (Ministry of Rural Development Government of India, 2005). The work from NREGA is short-term manual work that pays the worker a wage equivalent from the private sector or higher than what the market offers. The scale of NREGA as the largest employment programme in India allows for it to be used to study the effects of employment provisions on labour market outcomes.

The current literature on the NREGA program focuses heavily on female labour market outcomes and the educational outcomes of the children from those who find employment under the NREGA program. However, there has been a lack of empirical research on potential migration to rural areas. It is also important to note that employment in the public works is not constant throughout the year, with most work concentrated during the dry season before the monsoon rains (Imbert & Papp, 2015), as evidenced through Imbert and Papp's (2015) difference-in-difference estimations using weighted district controls.

Azam (2011) found that the NREGA policy in India has a positive impact on female employment outcomes, including a 2.4% increase in female labour force participation and an 8.3% increase in real wages for women workers, using a difference-in-difference estimation. Furthermore, the NREGA policy resulted in a 4.5 percentage point increase in women workers' wages relative to men. However, Imbert and Papp (2015) found that the wage impact of the policy between women and men decreases upon the addition of district controls, indicating the presence of other factors contributing to the relative wage increase between women and men workers, reduces the positive impact the policy had on female workers' wages.

In rural areas, the NREGA policy has a positive impact on female wage outcomes, as it offers attractive employment opportunities that provide better wages than existing ones and drives up current wages for female workers. Ghosh (2014) suggests that the NREGA programme pulls up the reservation wage rates for women in the whole labour market. Additionally, Dasgupta and Sudarshan (2011) demonstrated that the NREGA policy's uptake is negatively correlated with the gender wage gap within the agricultural sector using minimum wage data across states in India.

However, the effectiveness of the NREGA policy varies at the regional level and is constrained by localized context. In states such as Bihar and Uttar Pradesh, where there are strong social norms against women's participation in the labour market, women's registration for the NREGA program is often denied, resulting in a lower uptake rate of the policy (Pankaj & Tankha, 2010). On the other hand, regions like Kangra and Rajasthan, where social impediments to women's participation in the labour force are lacking, there are higher uptake rates of the NREGA policy due to economic necessity and women's desire to earn independently (Pankaj & Tankha, 2010), as evidenced through their field surveys.

Regarding children's educational attainment, Afridi et al. (2012) found that the increased employment of mothers under the NREGA policy led to a 50% rise in school attendance, with an increase of 3.6 hours per day spent in school, by manipulating spatial and temporal variation in the intensity of the implementation of the NREGA program. The positive effects of an increase in educational attainment are only restricted to younger children, as evidenced by Islam and Sivasankaran (2015), through their difference-in-difference estimations.

The NREGA program was progressively rolled out across India for 200 of the poorest districts in India from February 2006 for the first phase, another 130 districts from April 2007 for the second phase and the remaining rural districts from April 2008 for the final phase.

Previous studies have exploited the staggered implementation of the program with some exploiting the district and state codes to differentiate between regions that implemented the NREGA program before others. Following a similar methodology using a difference-in-difference estimation and additional data based on migration surveys, I estimate urban to rural migration given the guaranteed employment in rural areas, to determine whether the NREGA program did increase migration to rural areas and then examine the spillover effects on employment and educational outcomes on these urban to rural migrants. Additionally, I aim to understand and reconcile the heterogeneity between states in India, that led to a difference in

the efficacy of the NREGA program between states using HDI (Human Development Index) as a proxy for the values of a state and how progressive they are. HDI is a standardised composite tool assessing various factors that determine the quality of human development in a region. A high HDI can be owed to thorough implementation of various poverty alleviating programs; regions have a higher HDI because they implement programs better as they perceive greater benefits from these programs.

This study aims to answer the following question:

Did the NREGA program induce urban to rural migration in India with any spillover effects on migrants, and what role does a state's values play in effecting the NREGA program?

To explore the research question of this study, the following hypotheses need to be tested, to gain a holistic understanding of the research question:

1. The NREGA program will incentivise the least educated in urban areas to migrate to rural areas, given that they have less opportunities than those with a higher level of education.
2. Females that migrate could be migrating because of marriage, to follow their husbands, with employment opportunities being a secondary factor influencing their decision.
3. Employment opportunities are of greater importance for males who choose to migrate.
4. Migrants who relocate to rural areas because of the NREGA program will gain employment in either the public works or in the private sector.
5. A sudden increase in the rural population due to the NREGA program could put strains on the rural educational institutions, possibly lowering the educational attainment of the children of urban to rural migrants in relation to the existing rural population.
6. States with a higher HDI, are reflective of more progressive values. Better outcomes from the NREGA program will be expected in states with a higher HDI, such as employment in the public works. A state with higher HDI is an indication for potential migrants that the state has a better implementation of the policy.

I use the nationally representative Employment and Unemployment survey conducted by the National Sample Survey Organization (NSSO), for rounds 55 and 64. Round 55 consists of census data from July 1999 to June 2000 and round 64 consists of census data from July 2007 to June 2008. This data is supplemented with state-level HDI data from Global Data Labs.

The hypotheses are tested using a simple difference-in-difference empirical strategy, exploiting the staggered implementation of the NREGA program, with the main outcomes of

interest being urban to rural migration, employment in the public works, employment in the private sector and educational attainment. The difference-in-difference strategy allows for a differential comparison between “treated” districts and “control” districts.

I find no statistically significant results for urban to rural migration caused by the NREGA program. However, amongst the unemployed, I find that 83.16% of their urban to rural migration can be credited to the NREGA program. Employment in the public works increased by 0.456 percentage points for urban to rural migrants, significant at the 5% level. Females appear to be migrating to rural areas because of marriage, rather than the NREGA program. The results on education are not statistically different from zero after migrating because of the NREGA program. The study highlights that states with a unit higher HDI, have an increase in employment in the public works by 1.88 percentage points because of the NREGA program, reflecting better implementation of the program in states with higher HDI.

The rest of the paper is as follows. Section 2 examines related literature to provide a concrete basis for the hypotheses being tested. Section 3 describes the data used and the construction of variables, controls and outcomes. Section 4 discusses the methodology used to test the hypotheses. Section 5 describes and analyses the results following the methodology. Section 6 provides robustness checks to establish causality for the results found. Section 7 provides limitations, policy implications and recommendations for future research.

2.0 Literature Review

2.1 Migration

Oftentimes, migration is referred to as a poverty alleviation tool, specifically rural to urban migration. It is well established that stress migration is a result of looking for better opportunities for those that belong to a lower socio-economic status and are looking to alleviate themselves from poverty. It is often one of the most effective tools for poverty alleviation, especially if stuck in a geographically induced poverty trap; often remote and rural areas that lack development because of their geographical location.

According to Imbert and Papp (2020), for rural people to not migrate out to urban areas, their utility benefit associated with migration must be less than the NREGA pay they receive in a village, despite it being less than thirty five percent of the wage an individual can receive outside of the village. This implies that people look at other factors such as quality of life and job security when taking a decision to migrate out to urban areas. Additionally, costs of migration as estimated by Imbert and Papp (2020), are not substantial.

The flow-costs of migration to urban areas is substantial given higher living costs in urban areas (Imbert & Papp, 2020). Young (2013) mentions that the wage gaps between rural and urban areas (wage differentials between origin and destination) arise from the fact that jobs in urban areas are more skill-intensive and will attract more skilled workers. In the context of this study, the hypothesis that the least skilled proxied by being the least educated, will migrate from urban to rural areas, is highly plausible, if they have a greater chance of securing employment. The most impoverished in urban areas may not factor in the quality-of-life decision as greatly as an opportunity to seek employment.

According to Bhagat and Keshri (2013), individuals with the highest propensity to migrate in India are those with education below the primary level. Bhagat and Keshri (2013) focus on temporary labour migration as a means of survival for the rural poor in India. However, Bhagat and Keshri (2013) do no focus on the NREGA program, and the program's potential in alleviating the rural poor. Their findings support the hypothesis that the least educated have the highest propensity to migrate in search of better opportunities.

Hypothesis 1: The NREGA program will incentivise the least educated in urban areas to migrate to rural areas, given that they have less opportunities than those with a higher level of education.

Hypothesis 3: Employment opportunities are of greater importance for males who choose to migrate.

The authors note that temporary migration is seven times larger than that of permanent migration within India, with the demographic of the former being dominated by members from low caste communities with very little educational attainment (Bhagat & Keshri, 2013). However, given the nature of the NREGA program, which is unexamined by Bhagat and Keshri (2013), it is difficult to rule out migration from urban to rural areas amongst those from disadvantaged groups in urban areas; albeit I too am unable to determine whether the urban to rural migration results in section 5 of this paper are temporary or permanent.

Adewale (2005) notes that urban to rural migration in Nigeria is largely driven by individuals who are married, with 71% of the respondents to the survey reported being married. This could be suggestive that the hypothesis that there will be heterogeneity in migration between sexes, specifically for women, given the cultural context in India, could be possibly true. This is further corroborated by Fulford (2013) who finds that two thirds of Indian women who have migrated, have done so for marriage.

Hypothesis 2: Women that migrate could be migrating because of marriage, to follow their husbands, with employment opportunities being a secondary factor influencing their decision.

2.2 NREGA and Employment Effects

Imbert and Papp (2015) find an increase in employment in the public works in the dry season because of the NREGA program; they estimate an increase in public employment at 1.17 percentage points. They also note that there is an increase of 4.7 log points in deflated daily earnings, implying that the NREGA program increased wages for unskilled labour (Imbert & Papp, 2015). This corroborates the hypothesis that the least educated, presumed to have low or less skills, have an incentive to migrate to rural areas for an increase in wages.

Imbert and Papp (2015) assume competitive labour markets and find a decrease in private sector work, especially in the dry season. They note that the offset in private sector work in the dry season is reconciled by an increase in employment in public works (Imbert & Papp, 2015). This implies a crowding out effect; a surplus labour demand and a shortage of labour supply in the private sector in rural areas in the dry season. Within the context of the study, this opens the possibility of urban to rural migration to fill in the labour supply shortage in the private sector when the native rural population take up the guaranteed employment by the NREGA program.

Azam (2011) finds increased female labour force participation and increase in real wages for females because of the NREGA program, with an increase of 2.4% in labour force participation and increase of 8.4% in real wages. Ghosh (2014) corroborates this by suggesting that NREGA increases the reservation wages for females; the lowest wage at which an individual is willing to work. This study supports these findings by finding an increased participation by females in the public works in comparison to the private sector, where they are susceptible to greater discrimination.

Hypothesis 4: Migrants who relocate to rural areas because of the NREGA program will gain employment in either the public works or in the private sector.

2.3 NREGA and Education Outcomes

Afridi et al. (2012), found an increase of school attendance of roughly fifty percent arising from a mother gaining employment through NREGA, with the effect being larger for the lowest two wealth quartiles. Islam and Sivasankaran (2015) find an increase of 0.184 days in school attendance in the previous week for children aged between six to nine years old, but a decrease in 0.194 days in school attendance in the previous week for children aged between fifteen to seventeen. This could be a negative spillover effect created by changes in the local rural economy, with children foregoing education for work in the private sector. Given an increase in the uptake of jobs in the public works sector, a surplus of jobs in the private sector emanates consequently.

McKenzie and Rappoport (2010) find significant negative effects on education from migration between Mexico and the USA, with living in a migrant household reducing the probability of completing high school by 13% for males and 14% for females. Given the heterogeneity in culture, language, and development between certain states in India, an urban to rural migrant from a different state is susceptible to reduced educational attainment as a result. The findings by McKenzie and Rappoport (2010) lend some credence to the hypothesis that education outcomes for migrants may reduce because of additional strains placed on rural education institutions because of an increase in the rural population.

Hypothesis 5: A sudden increase in the rural population due to the NREGA program could put strains on the rural educational institutions, possibly lowering the educational attainment of the children of urban to rural migrants in relation to the existing rural population.

2.4 Values and Poverty Alleviation

According to Ravallion and Datt (2002), pro-poor growth was greater in states with initial higher literacy, farm productivity and rural living standards. They find that low education and poor health inhibits people from participating in non-farm work which is crucial to pro-poor growth. Pro-poor growth refers to equitable economic development, focusing on economic growth that alleviates poverty. The findings by Ravallion and Datt (2002) lend credence to the hypothesis, that states with higher HDI will have better implementation of the NREGA and hence more socially benefit outcomes. The HDI is a composite measurement considering a multi-dimensional approach in assessing human development, including life expectancy, education, and income.

It is important to note that despite using HDI to measure the progressive values of states, the NREGA program is susceptible to corruption, and a lack of accountability by the central government (Drèze, 2019). NREGA has the ability to enhance economic security, empowering women and promote social equity, but solid implementation, roll-out and political will is required for the act to succeed (Drèze, 2019).

Hypothesis 6: States with a higher HDI, are reflective of more progressive values. Better outcomes from the NREGA program will be expected in states with a higher HDI, such as employment in the public works. A state with higher HDI is an indication for potential migrants as higher perceived benefits from a better implementation of the policy.

3.0 Data

3.1 Description of Data

The primary source of data for this study comes from the nationally representative Employment and Unemployment Survey conducted by the National Sample Survey Organisation (NSSO). This study uses data from Round 55 (July 1999 to June 2000) and Round 64 (July 2007 to June 2008) to estimate migratory trends and other employment effects on specifically urban to rural migrants in India.

Data selected from Round 55 includes blocks 4 and 5, which primarily consists of migration data matched to demographics and the usual principal activity of each member of the household respectively. Blocks 4 and 5 of Round 55 have been merged to create a dataset representative of migration and employment activities of each member of the household. Data from round 64 includes blocks 4 and 6, consisting of the usual principal activity of each member of the household and migration data matched to demographics respectively. In a similar fashion, blocks 4 and 6 of Round 64 survey data have also been merged. The two datasets are then appended and merged with an additional dataset relating each district to which phase of the NREGA roll-out they were a part of. Data on state-level HDI from the Global Data Lab is additionally merged by state to the primary dataset from the years of 2000 and 2008, corresponding to the Rounds 55 and 64.

Data from the NSSO is stratified on a district level and has attached sampling weights to adjust for the oversampling of certain households. The sampling weights indicate the inverse of the probability of a household being selected into the survey, thus ensuring that the sample is representative of the population. The data used for this study is aggregated on the district level, and by merging the additional dataset relating each district to their respective phase of the NREGA roll-out, the scope of the study is narrowed down to districts that were a part of the NREGA program. Certain districts have both urban and rural areas, which leads to the estimates of treatment effects being underestimated because of a negative selection bias, as the survey does not distinguish regions on a sub-district level. The National Sample Survey does not differentiate between urban and rural areas within districts, thereby underestimating treatment effects as the differential impacts caused by heterogeneity between sub-district regions are not fully accounted for.

Exploiting the staggered implementation of the NREGA program, a simple difference-in-difference approach is employed, using Sub-Rounds 1 to 2 of Round 64 (July 2007 to

December 2007) as the pre-treatment period and Sub-Round 3 (January 2008 to March 2008) of Round 64 as the post-treatment period. NREGA was rolled out in three phases. In February 2006, the program was introduced to 200 districts, as part of the first phase. This was then extended to 130 districts in April 2007, as part of the second phase. Finally, by April 2008, the program was expanded to the rest of rural India, hereafter referred to as the third phase. Districts that were selected to be a part of the first two phases of the program will be the treated group and districts from the third phase will be the control group.

The pre-treatment period employed for this study is past the actual roll-out and implementation dates of the NREGA program. The first two phases were introduced in February 2006 and April 2007. However, in the context of this study, a reasonable assumption would be that migration is a difficult decision to make and one that takes time; individuals who have to make the decision to migrate need to base it off whether the program works and if people do get guaranteed employment in the public works. Allowing for a certain lag, allows for this decision-making process and for the actual process of migration to occur. This is evidenced by Treyz et al. (1993) who say that migration patterns are affected by economic factors that have taken place in the past; an individual's decision to migrate may be based on economic conditions and opportunities from the past than in the present. Another reason for the selected pre-program period is the lack of availability of migration surveys, with migration surveys only being conducted in Rounds 55 and Round 64 in the recent past, as part of the Employment and Unemployment Survey conducted by the NSSO.

Data from Round 55 is used to test for pre-program migratory and employment trends by conducting parallel trends checks for various outcomes to establish that the NREGA program has a causal effect on migratory and employment patterns.

3.2 Construction of Outcomes and Variables

The treatment group is denoted by the binary variable NREGA being one, if the district was part of either the first or second phase of the roll-out of the NREGA program, and is zero if the district is from the third phase. The pre-treatment period has a value of zero for the variable Post, indicating Sub-Rounds 1 to 2 (July 2007 to December 2007), and Post is equal to one for observations from Sub-Round 3 of Round 64 (January 2008 to March 2008).

For the primary outcome variable of urban to rural migration, the answer to the question of last location of usual principal activity is used to construct the binary variable urban_rural if an individual previously lived at an urban location, from block 4 of Round 55 and block 6 of

Round 64. The answer to the question is extrapolated to create the following variables, urban_rural_samedistrict, urban_rural_samestate and urban_rural_diffstate indicating urban to rural migration within the same district, same state and from a different state respectively.

Outcome variables for employment in public works, any employment in the private sector and attendance at an education are publicworks, pvtwork and educ. These variables use the answers to the survey from block 6 of Round 55 and block 4 of Round 64, based on an individual's usual principal activity.

State and district level controls have been created based on the unique state and district codes to control for unobserved heterogeneity arising from unique features of each state and district. Controls for sex and age are based on answers to questions from the survey for both blocks in both the rounds. Education levels have additionally been aggregated across both the rounds ranging from the values of 0 to 14, with an increase in the value signifying an increase in the level of education. Controls for primary level and secondary level education too have been created based on the former. Controls for marital status have been aggregated across both rounds using data from blocks 6 and 4 of Rounds 55 and 64 respectively, ranging from the values of 1 to 4. Additionally binary variables for being currently married and for ever being married are included, with the latter including being currently married and divorced/separated and widowed.

HDI is additionally used as a proxy for how progressive the values of a state are. Using state-level HDI data from the Global Data Lab for the years 2000 and 2008, controls for HDI have been created. Using HDI data from 2008 the following controls for low HDI, medium HDI, high HDI and very high HDI are characterised as below 0.550, 0.550 to 0.699, 0.700 to 0.799 and 0.800 and above respectively. No state in India in either 2000 or 2008 had very high HDI.

3.3 Descriptive Statistics

Table 1: Descriptive Statistics for Round 64

VARIABLES	(1) Observations	(2) Mean	(3) Standard Deviation
Employment in the Public Works	292,489	0.00134	0.0366
Employment in the Private Sector	292,492	0.249	0.432
Unemployed	292,492	0. 0136	0.116
Urban to Rural Migrants (Overall)	292,489	0.0272	0.163
Urban to Rural Migrants (Same District)	292,489	0.0120	0.109
Urban to Rural Migrants (Same State)	292,489	0.0100	0.0997
Urban to Rural Migrants (Different State)	292,489	0.005	0.0714
Illiterate	292,492	0.3988	0.490
Primary Education	292,492	0.146	0.353
Secondary Education	292,492	0.0717	0.258
Age	292,489	27.38	19.42
Currently Married	292,492	0.467	0.499
Female	292,492	0.492	0.500

Notes: This Table presents the number of observations, means and standard deviations of the described variables above. Data is from the NSSO from Round 64 (July 2007 to June 2008).

The summary statistics above is based on data from July 2007 to June 2008, Round 64. 0.13% of the sample from Round 64 was employed in the public works, and nearly 25% was employed in the private sector. 1.36% of the sample was unemployed in the period of the Round 64 survey. During the same period, about 2.7% of the respondents were urban to rural migrants, with 1.2% migrating from urban to rural areas within the same district, 1% migrating within the same state, and 0.5% migrating from a different state. About 49% of all respondents in the Round 64 survey were females.

The average respondent had an education level slightly below the primary level, but literate through informal education, through schemes such as TLC (Total Literacy Campaign). Nearly 40% of the respondents were illiterate. About 15% of the sample had primary level education, and roughly 7% had secondary level education. The average age of a respondent was 27 years old. The average respondent was either married or had been married. Nearly 47% of the sample was married at the time of the survey.

4.0 Methodology

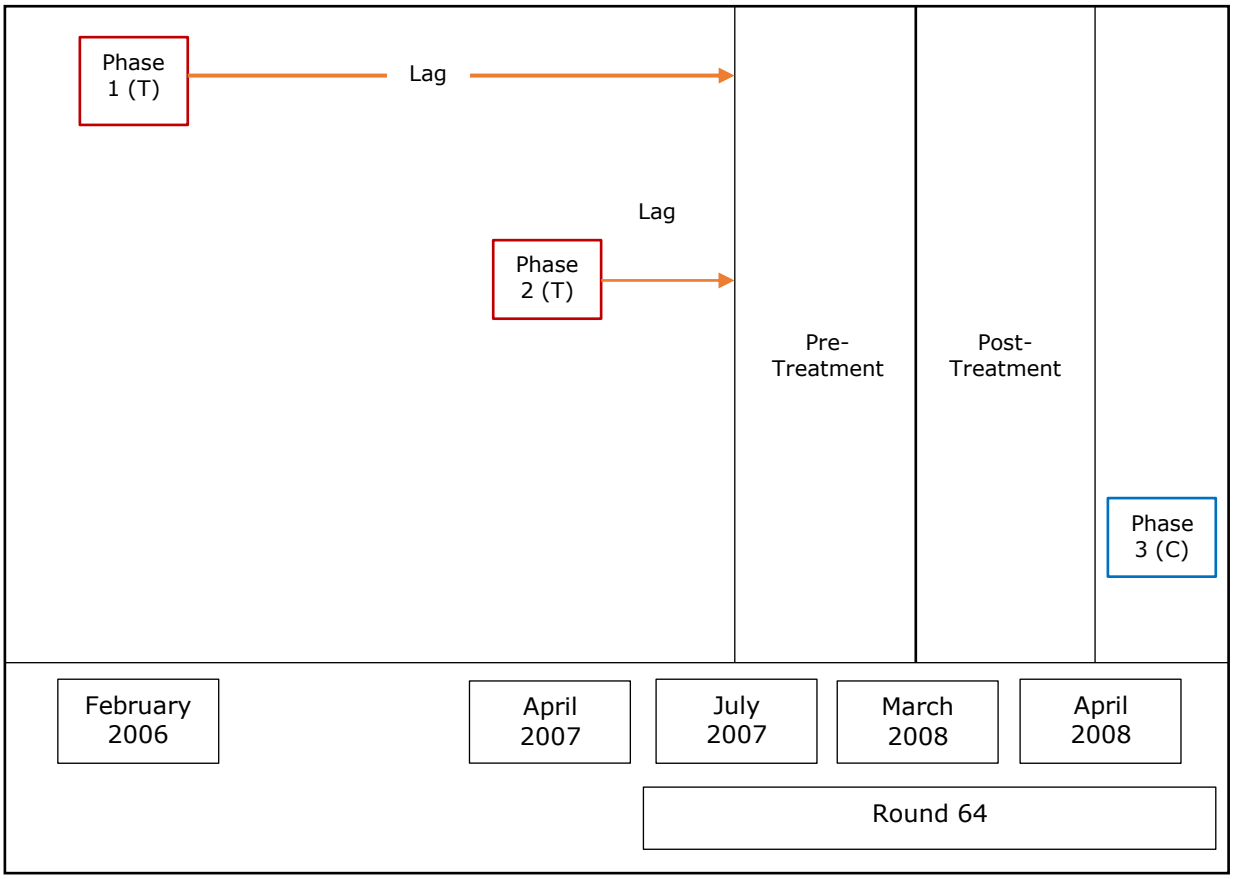
4.1 Urban to Rural Migration

Exploiting the staggered implementation of the NREGA program, a simple difference-in-difference estimation is used to compare changes in urban to rural migration between districts belonging to the first two phases and the districts in the third phase of the NREGA program, under the assumption that the decision to migrate is lagged, in line with Treyz et al. (1993).

The implementation of the NREGA program is exploited, allowing for lags from the implementation from the first phase and second phase till the implementation of the third phase, with the pre-program period defined from July 2007 to December 2007 and the post-program period defined as January 2008 to March 2008, leaving districts from phase 3 untreated through the time period selected, illustrated below in Figure 1.

A simple comparison between the two sets of districts will be naïve and will lead to biased estimates. This is because districts from the first two phases were chosen because they have lower agricultural wages, lower agricultural output per worker, and a higher proportion of scheduled castes (SC) and scheduled tribes (ST) in comparison to the later districts. These could be correlated with labour market outcomes, social outcomes, and migratory outcomes. Hence, a difference-in-difference approach is preferred to allow the migratory patterns to be only influenced by the NREGA program and to avoid biased estimates.

Figure 1: Roll-Out and Treatment Period



The regression equation to be estimated for migration from urban to rural areas is

$$UrbantoRuralMigration_{idt} = \beta_0 + \beta_1 \cdot NREGA_d + \beta_2 \cdot Post_t + \beta_3 \cdot NREGA_d * Post_t + \lambda \cdot X_{id} + \varepsilon_{it} \quad \textbf{(1)}$$

where the outcome variable captures urban to rural migration (urban_rural, urban_rural_samedistrict, urban_rural_samestate and urban_rural_diffstate) for an individual/household i in district d in period t. $NREGA_d$ is a binary variable equal to one for districts from the first two phases. $Post_t$ is a binary variable equal to one in the post period (January 2008 to March 2008, Sub-Round 3 of Round 64). β_0 is the mean outcome of the districts from the third phase, from the pre-period treatment of Sub-Rounds 1 to 2 (July 2007 to December 2007). X_{id} is a set of time-invariant controls that include sex, age, education level, state and district controls.

β_1 is the difference in urban to rural outcomes between the districts from the first two phases and the districts from the third phase at baseline, likely to be positive; urban to rural migration is expected to be greater in treated district from the first two phases than the control districts

because of the NREGA program intervention . β_2 is the difference in urban to rural migration in the pre and post periods for the third phase districts, and it is likely to be positive; urban to rural migration will be greater in the post-treatment period accounting for the lag, than in the period from when the program was implemented in phase 1 and 2 districts to the chosen post-treatment period. β_3 is the most important coefficient to establish a prediction, which is the difference of the difference between urban to rural migration before and after the implementation of NREGA in the first two phases' districts and the difference between outcomes before and after the implementation of NREGA in the third phase districts.

4.2 Heterogeneity Analysis

I conduct a heterogeneity analysis to examine gender differences and whether marriage can explain women's decisions to migrate. The following equation is a modified version of equation (1), still employing a difference-in-difference approach. The binary variables of Female_i and $\text{Currently Married}_i$, equalling one if an individual is a female and is married at the time of the survey in Round 64 respectively, are added to equation (1). These variables are interacted with each other and the former two variables of Post_t and NREGA_d, resulting in triple interactions and quadruple interactions.

$$Outcome_{idt} = \delta_0 + \delta_1 Female_i + \delta_2 NREGA_d + \delta_3 Post_t + \delta_4 CurrentlyMarried_i + \delta_5 Post_t * NREGA_d + \delta_6 Female_i * NREGA_d + \delta_7 Female_i * Post_t + \delta_8 \cdot Female_i * CurrentlyMarried_i + \delta_9 NREGA_d * CurrentlyMarried_i + \delta_{10} Post_t * CurrentlyMarried_i + \delta_{11} Female_i * CurrentlyMarried_i * Post_t + \delta_{12} Female_i * NREGA_d * CurrentlyMarried_i + \delta_{13} Female_i * NREGA_d * Post_t + \delta_{14} CurrentlyMarried_i * NREGA_d * Post_t + \delta_{15} Female_i * CurrentlyMarried_i * NREGA_d * Post_t + \eta \cdot X_{id} + \nu_{idt} \quad \textbf{(2)}$$

The two main coefficients of interest for the heterogeneity analysis are δ_8 and δ_{15}. δ_8 captures the effect on the outcome employment in the public works or urban to rural migration by females who are married, based on the interaction between the two variables. This would allow for a better understanding on the motivation of migration for females. The practice of arranged marriages is preponderant, with the wife expected to move into the husband's house in India. It could be highly plausible that females from urban areas get married to men from rural areas. δ_{15} captures the effects on employment in the public works or urban to rural migration by females who are married in the treated districts, in the post-treatment period of Sub-Round 3 of Round 64. This allows for a deeper look into whether females who are married

are also motivated by the NREGA program for their choice of migration from urban to rural areas in the post-treatment period.

4.3 Employment and Educational Outcomes

Following the same functional form as the former two equations, a difference-in-difference approach is used to estimate employment and education outcomes as a result of the NREGA program on urban to rural migrants. A simple comparison between the difference in employment and education trends between the districts from first two phases and the districts from the third phase will be naïve. Hence a difference-in-difference estimation is used, supplemented with interactions of urban to rural migration variables, to isolate the treatment effects on employment and education for urban to rural migrants because of the NREGA program.

The variable of UrbantoRuralMigrate (urban_rural) is added to equation (1) and then used to create individual interactions with the variables NREGA and Post and a triple interaction of all three variables, on various outcomes such as employment in the public works, any employment in the private sector and education.

$$Outcome_{idt} = \gamma_0 + \gamma_1 \cdot UrbantoRuralMigrate_{idt} + \gamma_2 \cdot NREGA_d + \gamma_3 \cdot Post_t + \gamma_4 \cdot NREGA_d * Post_t + \gamma_5 \cdot UrbantoRuralMigrate_{idt} * NREGA_d + + \gamma_6 \cdot UrbantoRuralMigrate_{idt} * Post_t + \gamma_7 \cdot UrbantoRuralMigrate_{idt} * NREGA_d * Post_t + \tau \cdot X_{id} + \omega_{idt} \quad \textbf{(3)}$$

γ_1 captures employment and education effects on those that migrated from urban to rural areas because of the NREGA program. γ_5 captures the differential effects of NREGA on education and employment between urban to rural migrants and non-migrants. γ_7 measures the difference in the impact of NREGA on employment and education between urban to rural migrants during the post period of implementation (Sub-Round 3 of Round 64: January 2008 to March 2008) and non-migrants. The remainder of the variables and the coefficients follow the same interpretation as in the sub-sections above, and the same set of time-invariant controls are used as well.

For the three difference-in-difference estimations to be valid, the parallel trends assumption must hold, and the assumption of a lag from actual NREGA implementation to the chosen post-treatment period must be reasonable. The former would imply that the trends of urban to rural migration and employment in the public works would have been the same in the absence of the NREGA program. The latter is tested by using a simple OLS using districts from

phase 3 to ensure that the results do not differ from the main estimations. The two are tested for in section 6 of the paper.

5.0 Results

5.1 Does Urban to Rural Migration Occur?

Table 2: Urban to Rural Migration (Break-down by District, State & Intra-State Level)

VARIABLES	(1) urban_rural	(2) urban_rural_samedistrict	(3) urban_rural_samestate	(4) urban_rural_diffstate
NREGA	-0.0110***	-0.00452***	-0.00497***	-0.00151***
	(0.000890)	(0.000590)	(0.000546)	(0.000402)
Post	0.00205	0.00190**	0.000640	-0.000491
	(0.00131)	(0.000887)	(0.000815)	(0.000555)
NREGA*Post	-0.00245	-0.00245**	-0.000208	0.000204
	(0.00156)	(0.00105)	(0.000965)	(0.000682)
Constant	0.0331***	0.0142***	0.0126***	0.00633***
	(0.000740)	(0.000489)	(0.000461)	(0.000328)
Observations	220,020	220,020	220,020	220,020
R-squared	0.001	0.001	0.001	0.000

*Notes: This Table presents results for urban to rural migration using data from Round 64. Column 1 presents total results of urban to rural migration within India. Column 2 presents results of urban to rural migration within the same district. Column 3 presents results of urban to rural migration within the same state. Column 4 presents results of urban to rural migration from different states. The results are estimated using equation 1, with NREGA*Post being the interaction term and coefficient of interest. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

The results presented above are estimated using equation (1), with no control variables. Overall urban to rural migration appears to have decreased by 0.245 percentage points, but not at a statistically significant level because of the NREGA program. This lack of urban to rural migration appears to be largely driven by no migration within districts, with within district urban to rural migration decreasing by 0.245 percentage points because of the NREGA program in the post-treatment period, statistically significant at the 5% level. There appears to be a statistically insignificant coefficient on the interaction term between NREGA and Post on migration within a state, while a positive yet statistically insignificant coefficient for migration from a different state.

These results contrast with a logical mechanism that should drive urban to rural migration within the same district; the greater ease of movement within an individual's district in comparison to another district within the same state or to another state, because of the less distance required to move from an urban to rural area, or the perceived lack of regional and cultural differences, given the heterogeneity between regions in India. The concomitant question then is who drives this migration or rather lack of migration, is it people with fewer opportunities in urban areas like those with less education?

Table 3: Urban to Rural Migration by Primary Education and No Literacy

VARIABLES	(1) urban_rural	(1) urban_rural
NREGA	-0.0111***	-0.0128***
	(0.000972)	(0.00123)
Post	0.00222	0.000877
	(0.00144)	(0.00175)
Primary Education	-0.00350*	
	(0.00199)	
NREGA*Post	-0.00290*	
	(0.00171)	
Primary Education*NREGA	0.000120	
	(0.00241)	
Primary Education*Post	-0.00111	
	(0.00350)	
Primary Education*NREGA*Post	0.00320	
	(0.00421)	
Illiterate		-0.0168***
		(0.00143)
NREGA*Post		-0.00118
		(0.00214)
Illiterate*NREGA		0.00699***
		(0.00173)
Illiterate*Post		0.00290
		(0.00257)
Illiterate*NREGA*Post		-0.00337
		(0.00306)
Constant	0.0336***	0.0392***
	(0.000809)	(0.00100)
Observations	220,020	220,020
R-squared	0.001	0.003

*Notes: This table presents urban to rural migration in the whole country using data from Round 64. Column 1 has no controls and uses interactions as stated above. The results are estimated using a modified equation 3. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Estimating a modified equation (3) with a focus on primary level education in column 1, we see a positive effect that is statistically insignificant on the coefficient of the triple interaction term between primary education, NREGA and Post in comparison to illiterates and higher education levels. The coefficient on the interaction between NREGA and Post, remains negative and statistically insignificant. It appears that individuals with a primary level of education do not appear to have a higher propensity to move to rural areas because of the NREGA program. This contrasts with the hypothesis that we would expect individuals with less education to be moving to seek employment, under the assumption that they do not have better opportunities in urban areas as compared to rural areas.

To test further the hypothesis whether the least educated still migrate from urban to rural areas, equation (3) is modified with a focus on those with no literacy, in column 2. The coefficient on the interaction term between Illiterate, NREGA and Post in column 2, is negative

and statistically insignificant in comparison to primary and higher education levels. There is a decrease in urban to rural migration by 0.337 percentage points for those with no literacy as a result of the NREGA program. The coefficient on the interaction term between NREGA and Post, is statistically insignificant using the regression estimation in column 2. Once again, the results do not corroborate the hypothesis that the least educated; a proxy for fewer opportunities in the urban areas migrate to rural areas for better opportunities arising from the NREGA program.

It is important to test the difference in urban to rural migration by sex and try to understand whether the NREGA program is driving this urban to rural migration or another prominent factor, such as marriage. It could be very well likely that a female's decision to migrate because of marriage may also be influenced by the possibility of gaining employment through the NREGA program in rural areas.

5.2 Heterogeneity Analysis

5.2.1 Heterogeneity by Sex

Table 4: Urban to Rural Migration Heterogeneity by Sex and Marriage

VARIABLES	(1) urban_rural	(2) urban_rural	(3) publicworks
Female	0.00214	0.00245	-0.000514**
	(0.00158)	(0.00157)	(0.000241)
NREGA	-0.00589***	-0.00424***	0.000693**
	(0.00126)	(0.00126)	(0.000315)
Post	-0.00166	-0.00135	0.000991**
	(0.00177)	(0.00177)	(0.000503)
Currently Married	0.0143***	0.00176	0.00165***
	(0.00188)	(0.00211)	(0.000475)
NREGA*Post	0.000681	0.000216	-0.000368
	(0.00213)	(0.00212)	(0.000684)
Female*NREGA	-0.00322*	-0.00265	-0.000537
	(0.00188)	(0.00187)	(0.000361)
Female*Post	0.000103	-0.000683	-0.000921*
	(0.00268)	(0.00266)	(0.000554)
Female*Currently Married	0.0274***	0.0353***	-0.00145***
	(0.00300)	(0.00299)	(0.000518)
Currently Married *NREGA	-0.00310	-0.00129	0.000144
	(0.00228)	(0.00227)	(0.000667)
Currently Married *Post	0.00921***	0.00839**	-0.00106
	(0.00340)	(0.00338)	(0.000889)
Female*Currently Married*Post	-0.00263	-0.00183	0.000853
	(0.00533)	(0.00530)	(0.000955)
Female*Currently Married*NREGA	-0.00791**	-0.00833**	0.000390
	(0.00361)	(0.00359)	(0.000745)
Female*NREGA*Post	0.00116	0.00192	0.000663
	(0.00319)	(0.00317)	(0.000778)

Currently Married*NREGA*Post	-0.01000**	-0.00889**	-0.00101
	(0.00405)	(0.00403)	(0.00116)
Female* Currently Married*NREGA*Post	0.00444	0.00360	0.00106
	(0.00637)	(0.00634)	(0.00132)
Marital_Status		0.0154***	
		(0.00122)	
Education Level		0.00426***	
		(0.000139)	
Age		-0.000548***	
		(9.05e-05)	
Age^2		7.27e-06***	
		(1.13e-06)	
district_control		0.000128***	
		(2.22e-05)	
state_control		0.000595***	
		(4.52e-05)	
Unemployed		0.0317***	
		(0.00447)	
Constant	0.0186***	-0.0305***	0.000723***
	(0.00105)	(0.00181)	(0.000209)
Observations	220,020	219,585	220,020
R-squared	0.010	0.021	0.001

*Notes: This table presents urban to rural migration and employment in the public works in the whole country using data from Round 64. Column 1 presents urban to rural migration as outcome and has no controls and uses interactions as stated above. Column 2 presents urban to rural migration as outcome, and adds on controls for age, the square of age overall education level, unemployed, state and district controls. Column 3 follows the same as column 1 but with the outcome as employment in the public works. The results are estimated using equation 2. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

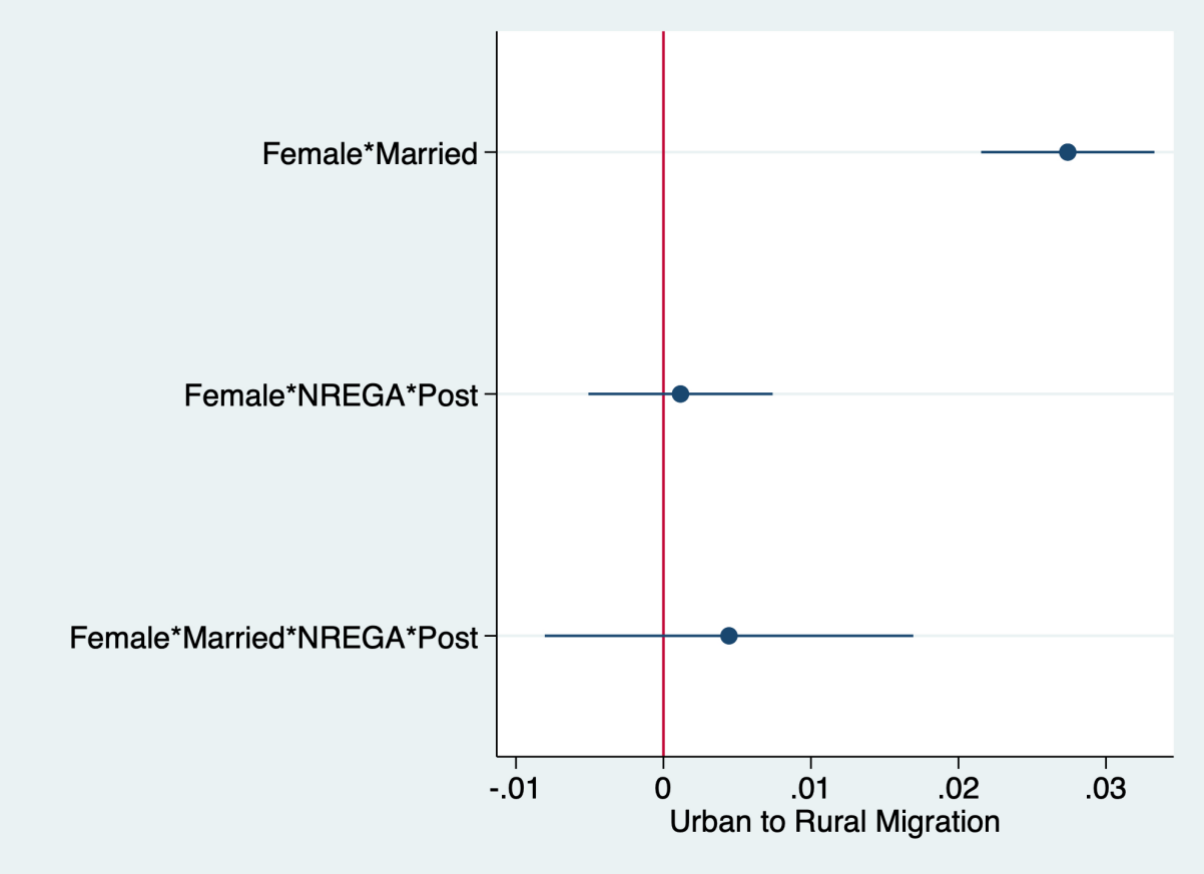
Estimating equation (2) with urban to rural migration set as the outcome results in a positive yet statistically insignificant result on the quadruple interaction term between female, currently married, NREGA and Post. This implies that females who were married, were unaffected by the NREGA program in the post-treatment period; their decision to migrate was unaffected by the potential benefits that NREGA could offer them, a third reservation of all jobs in the public works. This is further corroborated by the coefficient on the triple interaction term between female, Post and NREGA, which is positive and statistically insignificant from zero. This implies that women were unaffected by the NREGA program in the post period.

A possible driving mechanism behind this result, is the fact that their decision to migrate from urban to rural areas is driven by marriage. This is evidenced by the interaction term between female and currently married, statistically significant at the 1% level in column 1, implying an increase in urban to rural migration by 2.74 percentage points. Another possibility is that women do not see any benefit of the NREGA program. A rather perverse implication is that the values in certain regions hinders females from enrolling in the NREGA program. This is supported by the findings of Pankaj and Tankha (2010) in states such as Bihar and Uttar

Pradesh where there are strong social norms against females participating in the labour force. This is supported by the coefficient on the interaction term between female and currently married in column 3, where a decrease of 0.45 percentage points in employment in public works is seen for women who were married, statistically significant at the 1% level.

The above described treatment effects have been illustrated graphically in Figure 2 to show that urban to rural migration for women is largely driven by marriage, with NREGA having little to no effect on women's decisions to migrate from an urban to rural area.

Figure 2: Heterogeneity by Sex and Marriage



Notes: This figure represents the results from Table 5, estimated using equation (2) and data from Round 64. Only the coefficients of interest described are presented graphically.

It appears that women are slightly likely to migrate more than men, as seen on the coefficient of female, being positive but not statistically significant, in Table 4.

To gain a better understanding, further heterogeneity analysis is required. A rather straightforward analysis would be to test if there is heterogeneity in urban to rural migration by employment status, with unemployed people from urban areas migrating to rural areas in search of employment.

5.2.2 Heterogeneity by Employment Status

Table 5: Heterogeneity by Employment Status

VARIABLES	(1) urban_rural	(2) urban_rural
NREGA	-0.0111***	-0.0109***
	(0.000888)	(0.000885)
Post	0.00261**	0.00257**
	(0.00132)	(0.00131)
Unemployed	0.0370***	0.0364***
	(0.00918)	(0.00915)
NREGA*Post	-0.00328**	-0.00327**
	(0.00157)	(0.00156)
Unemployed*NREGA	0.00645	0.00734
	(0.0116)	(0.0116)
Unemployed*Post	-0.0428***	-0.0419***
	(0.0127)	(0.0127)
Unemployed*NREGA*Post	0.0568***	0.0561***
	(0.0180)	(0.0180)
Female		0.0121***
		(0.000682)
Age		0.00184***
		(5.58e-05)
Age^2		-2.02e-05***
		(8.52e-07)
district_control		0.000119***
		(2.22e-05)
state_control		0.000630***
		(4.54e-05)
Constant	0.0326***	-0.0135***
	(0.000740)	(0.00122)
Observations	220,020	220,020
R-squared	0.002	0.011

*Notes: This table presents urban to rural migration in the whole country using data from Round 64. Column 1 has no controls and uses interactions as stated above. Column 2 adds on controls for age, the square of age and overall education level. The results are estimated using a modified equation 3. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

By estimating a modified version of equation (3) with a focus on employment status, specifically if an individual is unemployed or not, we see a positive and statistically significant coefficient on the triple interaction term between being unemployed, NREGA and Post in both columns. In column 1 and 2, the coefficient is statistically significant at the 1% level, with an increase in urban to rural migration by 5.68 percentage points and 5.61 percentage points in columns 1 and 2 respectively. Being unemployed in the treated districts increases urban to rural migration in the post-treatment period. It appears that those unemployed in urban areas see merit in the NREGA program and its ability to give them employment, hence their choice of migration, as a poverty-allevation tool.

Table 6: Statistics for the Unemployed in Round 64

VARIABLES	(1) Observations	(2) Mean	(3) Standard Deviation
Urban to Rural Migrants	3,982	0.0683	0.252
Illiterate	3,982	0.0455	0.208
Primary Education	3,982	0.102	0.303
Secondary Education	3,982	0.190	0.392
Age	3,982	23.08	6.314
Currently Married	3,982	0.186	0.389
Female	3,982	0.278	0.448

Notes: This Table presents the number of observations, means and standard deviations of the described variables above. Data is from the NSSO from Round 64 (July 2007 to June 2008) only for those unemployed.

In order to understand the demographics of the unemployed, descriptive statistics of the unemployed from Round 64 are provided above. Roughly 6.83% of the unemployed are urban to rural migrants. Using the coefficient in column 1 from Table 5 of the triple interaction term, [(0.0568/0.0683)*100] = 83.16%. 83.16% of the urban to rural migration amongst the unemployed can be attributed to the NREGA program in the post-treatment period.

The average unemployed person had an education level equal to or higher than higher secondary education. 10% of unemployed people had primary education and about 4.6% were illiterate. Roughly 19% of the unemployed have secondary education. The average age was 23 years old, and roughly 18.6% were married. 27.8% of the sample of unemployed were female.

The average unemployed person appears to be more educated and younger than the average person from the population in Round 64. It is then surprising as this is in clear contrast to the hypothesis that the least educated have the fewest opportunities in urban areas, in line with Young (2013).

5.3 Employment Outcomes for Urban to Rural Migrants

5.3.1 Employment in the Public Works for Urban to Rural Migrants

Table 7: Employment in the Public Works for Urban to Rural Migrants

VARIABLES	(1) publicworks	(2) publicworks
Urban to Rural Migrants	-0.000386	-0.000341
	(0.000532)	(0.000532)
NREGA	0.000597***	0.000619***
	(0.000182)	(0.000183)
Post	0.000279	0.000283
	(0.000241)	(0.000241)
NREGA*Post	-0.000360	-0.000364
	(0.000329)	(0.000329)
Urban to Rural Migrants *NREGA	-0.000603	-0.000640
	(0.000748)	(0.000750)
Urban to Rural Migrants * Post	-0.000796	-0.000826
	(0.000570)	(0.000571)
Urban to Rural Migrants *NREGA *Post	0.00456**	0.00464**
	(0.00224)	(0.00224)
Female		0.00131***
		(0.000151)
Age		1.55e-05***
		(2.71e-06)
state_control		6.63e-06
		(7.10e-06)
Constant	0.000902***	-0.000312*
	(0.000126)	(0.000176)
Observations	220,020	220,020
R-squared	0.000	0.000

*Notes: This table presents the results for employment in the public works for urban to rural migrants in the whole country using data from Round 64. The interaction between Urban to Rural Migrants, NREGA and Post is the main coefficient of interest. The results are estimated using equation 3. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

One of the main objectives of the NREGA program is to provide guaranteed employment of 100 days per year in the public works, with additional objectives of reducing gender gaps in opportunities and employment. Table 7 uses equation (3) to estimate the effects on employment in the public works for urban to rural migrants. In both columns, the coefficient on the triple interaction term between NREGA, Post and Urban to Rural Migration, is positive and statistically significant at the 5% level. This implies that urban to rural migrants that have moved, seeking employment in rural areas through the NREGA program received an increase in employment in the public works. With no controls, urban to rural migrants had an increase in employment in the public works by 0.456 percentage points, and with controls, an increase of 0.464 percentage points, both significant at the 5% level.

In Table A1, in the appendix, the coefficient on migrate3_triple has a coefficient of

0.00880, statistically significant at the 10% level. This implies that urban to rural migrants from the same state have an increase in their employment in the public works by 0.880 percentage points because of the NREGA program. A possible driving mechanism as to why urban to rural migrants from the same state would face an increase in their employment in the public works could be elucidated by the heterogeneity between states in India, with differences in language, culture and societal demographics sometimes being very distinct from one state to the other, thereby making mobility and access to jobs within the same state easier than in a different state.

The coefficient on female in column 2 is positive and statistically significant at the 1% level, implying that being a female increases employment in the public works by 0.131 percentage points. This is in line with the fact that more females migrate to rural areas than males. This result stands even with state level controls, accounting for heterogeneity between states.

However, as mentioned by Imbert and Papp (2015), there could have been a difference in the roll-out and implementation of the program based on the values and political will in each state, thereby affecting the effect of employment in the public works for females depending on the region they have migrated to .

5.3.2 Employment in the Private Sector for Urban to Rural Migrants

One of the hypotheses mentioned is that urban to rural migrants will see an increase in either the public works or the private sector.

Using equation 3, effects on any employment in the private sector have been estimated in Table 8. There is no statistically significant change in an urban to rural migrant's employment in the private sector, if at all, a slight increase in their employment in the private sector in comparison to their employment in the private sector prior to their migration and the implementation of the NREGA program, as seen by the coefficient on the triple interaction term between NREGA, Post and Urban to Rural Migration. Across various specifications in Table A2, in the appendix, there appears to be ambiguity in the direction of employment in the private sector for urban to rural migrants across specifications including interactions with urban to rural migration within the same district, within the same state and from a different state as well.

Table 8: Any Employment in the Private Sector for Urban to Rural Migrants

VARIABLES	(1) Pvtwork	(2) pvtwork
Urban to Rural Migrants	0.0334***	-0.00100
	(0.0102)	(0.0100)
NREGA	0.0240***	0.0308***
	(0.00232)	(0.00223)
Post	-0.000674	0.000233
	(0.00310)	(0.00297)
NREGA*Post	-0.00114	-0.00223
	(0.00403)	(0.00386)
Urban to Rural Migrants *NREGA	0.0129	0.0141
	(0.0147)	(0.0144)
Urban to Rural Migrants * Post	0.0126	0.00839
	(0.0176)	(0.0170)
Urban to Rural Migrants *NREGA *Post	0.000870	0.0148
	(0.0255)	(0.0247)
Female		0.124***
		(0.00176)
Age		0.00441***
		(4.30e-05)
state_control		0.00715***
		(0.000103)
Constant	0.235***	-0.0715***
	(0.00178)	(0.00247)
Observations	220,020	220,020
R-squared	0.001	0.084

*Notes: This table presents the results for any employment in the private sector for urban to rural migrants in the whole country using data from Round 64. The interaction between Urban to Rural Migrants, NREGA and Post is the main coefficient of interest. The results are estimated using equation 3. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

It is once again interesting to note that the coefficient on female in column 2 of Table 8 is positive and statistically significant at the 1% level. Once again, it is evident that being a female increases any employment within the private sector by 12.4 percentage points in relation to being a male.

It appears that women are not discriminated in both the private and public sector as a result of the NREGA program. This is in contrast to evidence by Breitkreuz et al. (2017), and Narayanan and Das (2014). However, these results may be too naïve to conclude and would require analysis on solely the effect of the program on the intended rural female population.

5.3.3 Employment Effects for Females in Rural Areas

To test whether females benefited from the program, equation (3) is estimated with the outcomes being employment in the public works and any employment in the private sector with interaction terms using the binary variable of female instead of urban to rural migration.

Table 9: The Effect of NREGA on Female Employment Outcomes

VARIABLES	(1) publicworks	(2) pvtwork
NREGA	0.000587***	0.0238***
	(0.000178)	(0.00227)
Post	0.000431	0.0115***
	(0.000407)	(0.00438)
Female	-0.00134***	-0.118***
	(0.000182)	(0.00223)
NREGA*Post	-0.000642	-0.0181***
	(0.000502)	(0.00540)
Female*Post	-0.000363	-0.0234***
	(0.000433)	(0.00539)
Female*NREGA*Post	0.000795	0.0336***
	(0.000536)	(0.00642)
Constant	0.00155***	0.294***
	(0.000176)	(0.00217)
Observations	220,020	220,020
R-squared	0.000	0.020

*Notes: This table presents employment outcomes in the public works and private sector using data from Round 64 using a modified equation (3). Columns 1 and 2 include controls for female and double and triple interactions with NREGA and Post. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

In columns 1 and 2, in relation to males for NREGA districts in the post-treatment period, females have an increase in public works employment and in any employment in the private sector. There is an increase in employment in the public works by 0.08 percentage points, but not at a statistically significant level. The increase in employment in the private sector is 3.36 percentage points, statistically significant at the 1% level. This provides evidence to some efficacy of the NREGA program on its intended rural population, specifically for women. Women appear to not just be discriminated less in government mandated employment opportunities than opportunities in the private sector, contrasting the findings of Breitkreuz et al. (2017), and Narayanan and Das (2014). This could be driven by the special provisions made to females, such as a reservation of a third of the jobs for females, equal wages between sexes and special provisions for mothers and pregnant women, having some direct spillover effects into the private sector as well.

5.4 Educational Outcomes for Urban to Rural Migrants

Equation (3) is used to estimate the effects on educational attainment by urban to rural migrants because of the NREGA program. The outcome variable for education has been created based on the responses for an individual's usual principal activity, as attendance at an educational institution. The estimation uses the same set of time-invariant controls, including sex, age, and state-level controls.

The results of the estimation are presented below in Table 10.

Table 10: Educational Attainment for Urban to Rural Migrants

VARIABLES	(1) educ	(2) educ
Urban to Rural Migrants	-0.133***	-0.0540***
	(0.00755)	(0.00674)
NREGA	-0.0118***	-0.0243***
	(0.00233)	(0.00206)
Post	0.00410	0.00240
	(0.00318)	(0.00279)
NREGA*Post	0.00399	0.00406
	(0.00407)	(0.00358)
Urban to Rural Migrants *NREGA	-0.0208**	-0.0143
	(0.00994)	(0.00902)
Urban to Rural Migrants * Post	-0.0177	-0.00917
	(0.0125)	(0.0111)
Urban to Rural Migrants *NREGA *Post	0.0188	-0.00130
	(0.0171)	(0.0154)
Female		0.0396***
		(0.00161)
Age		-0.0101***
		(3.88e-05)
state_control		-0.000783***
		(9.09e-05)
Constant	0.251***	0.527***
	(0.00182)	(0.00281)
Observations	220,020	220,020
R-squared	0.003	0.218

*Notes: This table presents the results for educational attainment for urban to rural migrants in the whole country using data from Round 64. The interaction between Urban to Rural Migrants, NREGA and Post is the main coefficient of interest. The results are estimated using equation 3. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

The coefficient on the triple interaction term is not statistically significant but is a positive coefficient in column 1 and negative in column 2, implying no change or a very little increase or decrease in the education of urban to rural migrants because of the NREGA program. Across various specifications in Table A3, in the appendix, there appears to be ambiguity in the direction of educational attainment for urban to rural migrants, including in interactions with urban to rural migration within the same district, within the same state and from a different state as well. The coefficient on female in column 2 from Table 10 illustrates that being a female increases attendance at an educational institution by 3.96 percentage points as compared to a male, statistically significant at the 1% level. An additional year increase in age, reduces attendance at an educational institution by 1.01 percentage points, statistically significant at the 1% level. This is in line with the findings by Islam and Sivasankaran (2015); increases in educational attainment are restricted to children between the ages of six to nine, and there is a

decrease in educational attainment for children aged 15 to 17, who instead choose to work, with the opportunity cost of education weighing less than that of work.

5.5 HDI and the implementation of the NREGA Program

HDI is a standard measurement that follows a standardised calculation which allows for it to be a consistent tool to compare states; a composite measurement that captures a state's development holistically. This enables HDI to be a good proxy for the values each state is representative of. The underlying hypothesis is that states with a higher HDI had a better roll-out and implementation of the NREGA program; states with a higher HDI have greater perceived benefits of the NREGA program, and will therefore see better results from better implementation.

Using a modified version of equation (3) and employing state-level HDI from 2008, urban to rural migration is estimated below in Table 11.

Table 11: Urban to Rural Migration by HDI

	(1)	(2)
VARIABLES	urban_rural	urban_rural
NREGA	0.0442***	0.0498***
	(0.00994)	(0.00989)
Post	0.00236	-0.000607
	(0.0145)	(0.0144)
HDI	0.192***	0.135***
	(0.0143)	(0.0142)
NREGA*Post	-0.0111	-0.00787
	(0.0175)	(0.0174)
HDI*NREGA	-0.0892***	-0.0975***
	(0.0177)	(0.0176)
HDI*Post	-0.000821	0.00401
	(0.0253)	(0.0252)
HDI*NREGA*Post	0.0160	0.0105
	(0.0311)	(0.0310)
Female		0.0161***
		(0.000704)
Age		0.00125***
		(5.86e-05)
Age^2		-1.09e-05***
		(9.22e-07)
Education Level		0.00324***
		(0.000128)
Constant	-0.0788***	-0.0934***
	(0.00814)	(0.00808)
Observations	220,020	219,719
R-squared	0.004	0.015

*Notes: This table presents urban to rural migration in the whole country using data from Round 64, supplemented with state-level HDI from 2008. Column 1 has no controls and uses interactions as stated above. Column 2 adds on controls for age, the square of age and overall education level. The results are estimated using a modified equation 3.Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Estimating a modified equation (3) using state-level HDI from 2008 results in a positive coefficient yet statistically insignificant coefficient on the triple interaction term between HDI, NREGA and Post. A positive coefficient in columns 1 and 2 of Table 11 imply that a one unit increase in HDI in the treated district in the post-period, results in an increase of urban to rural migration by 1.6 percentage points and 1.05 percentage points because of the NREGA program. However, these are not statistically significant. However, the coefficient on HDI is positive and statistically significant at the 1% level, implying that a higher HDI correlates positively with urban to rural migration. To further test the hypothesis that the roll-out and implementation of the NREGA program was better in states with higher HDI, equation (3) is used as it was in Table 11, with the outcome set to employment in the public works. The results are presented in Table 12.

Table 12: Employment in the Public Works by HDI

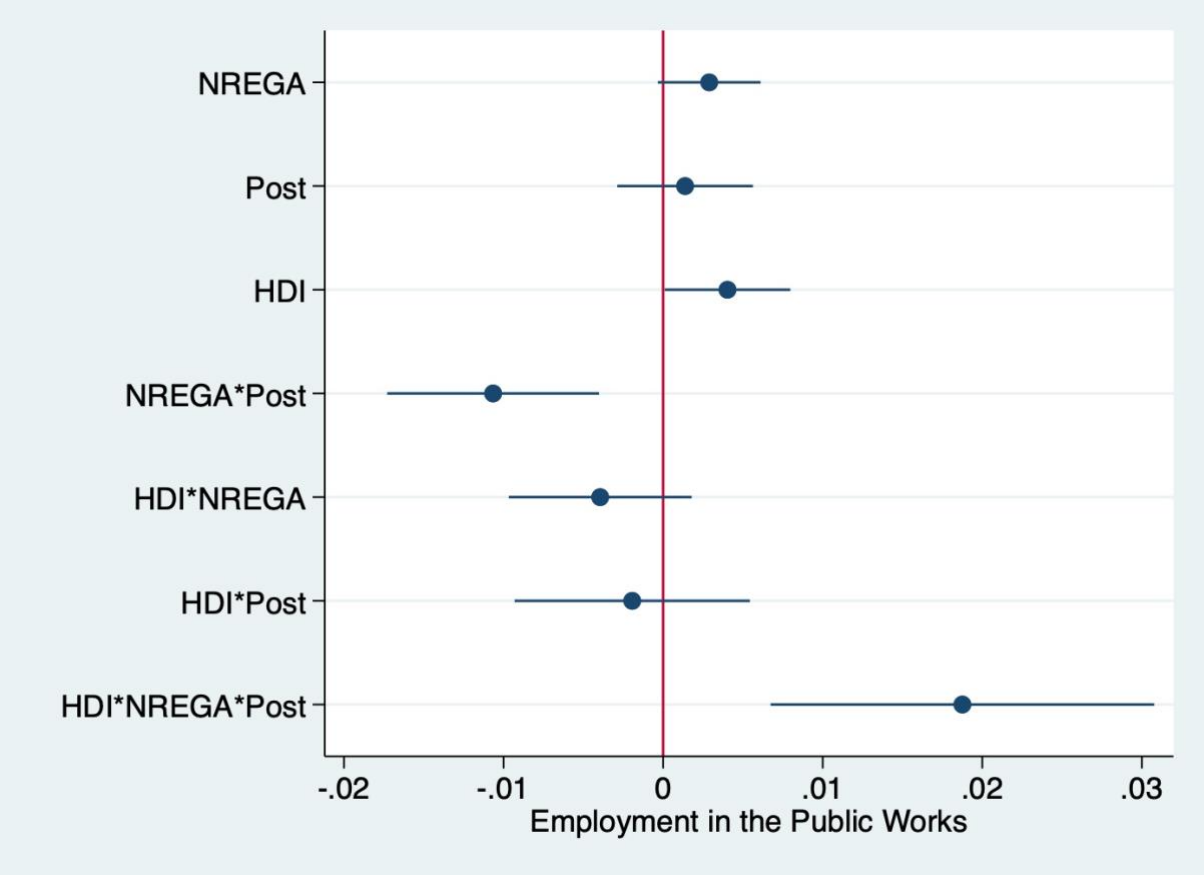
VARIABLES	(1) publicworks	(2) publicworks
NREGA	0.00289*	0.00286*
	(0.00164)	(0.00165)
Post	0.00138	0.00116
	(0.00217)	(0.00217)
HDI	0.00403**	0.00475**
	(0.00201)	(0.00200)
NREGA*Post	-0.0107***	-0.0105***
	(0.00339)	(0.00339)
HDI*NREGA	-0.00394	-0.00399
	(0.00293)	(0.00293)
HDI*Post	-0.00194	-0.00157
	(0.00376)	(0.00377)
HDI*NREGA*Post	0.0188***	0.0184***
	(0.00613)	(0.00613)
Female		-0.00152***
		(0.000159)
Age		0.000174***
		(1.40e-05)
Age^2		-2.41e-06***
		(2.07e-07)
Education Level		-0.000110***
		(2.43e-05)
Constant	-0.00146	-0.00257**
	(0.00114)	(0.00114)
Observations	220,020	219,719
R-squared	0.000	0.001

*Notes: This table presents employment in the public works in the whole country using data from Round 64, supplemented with state-level HDI from 2008. Column 1 has no controls and uses interactions as stated above. Column 2 adds on controls for age, the square of age and overall education level. The results are estimated using a modified equation 3. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

The coefficient on the triple interaction term between HDI, NREGA and Post is positive and statistically significant at the 1% level in both columns, when estimating the effects on

employment in the public works. In column 1, with no controls, the coefficient on the triple interaction term is statistically significant at the 1% level; a unit increase in HDI increases employment in the public works by 1.88 percentage points in the treated districts in the post-treatment period. The same holds true in column two, with a 1.84 percentage point increase in employment in the public works, statistically significant at the 1% level.

Figure 3: Effect of HDI on Urban to Rural Migration



Notes: This figure graphically presents the results from Table 12, using a modified version of equation (3), using data from Round 64.

This corroborates the hypothesis that, states with higher values perceive greater benefits from the NREGA program. This results in them having a better roll-out and implementation as evidenced by the positive and statistically significant coefficients on high HDI states for urban to rural migration and employment in the public works. A positive coefficient on high HDI states is also indicative of the trust people in high HDI states have in the efficacy of the NREGA program. The results described above have been illustrated in graphically in Figure 3. Specifically, it can be noted that the triple interaction between HDI, NREGA and Post yields a highly significant and positive effect on employment in the public works, one of the key goals of the NREGA program.

6.0 Robustness Checks

The primary means of estimating various outcomes throughout this paper has been through a difference-in-difference framework. In order for a difference-in-difference estimation to hold true and provide a causal result, the parallel trends assumption must hold. This implies that in the absence of the NREGA program, the average change in outcomes would have been the same between the treatment and control groups, districts from the first two phases and districts from the third phase. The trend in outcomes in the control and treatment groups would have been the same in the absence of the NREGA program.

The second assumption required for the estimates to be valid is that the lag between phase 1 and phase 2 districts till the pre-treatment and post-treatment periods is valid. This is tested for by using only phase 3 districts, with a simple OLS regression. If the results are similar to that of Table 2, then the assumption is valid.

If the parallel trends assumption holds and the validity of the assumption of the lag holds, then differences in the trends between the treatment and control groups can be causally attributed to the NREGA program.

Data from Round 55 is used to test for these parallel trends prior to the implementation of the program. Due to limitations in the availability of the migration surveys conducted by the NSSO as part of the Employment and Unemployment surveys, Round 55 is the only viable option to test for pre-program migratory and employment effects.

In order to test for the parallel trends, a modified version of equation (1) is estimated, as seen below:

$$Y_{idt} = \alpha_0 + \alpha_1 \cdot NREGA_d + \alpha_2 \cdot Post55_t + \alpha_3 \cdot NREGA_d * Post55_t + \epsilon_{it} \quad \textbf{(4)}$$

where the outcome variable Y_{idt} captures urban to rural migration as a whole country, employment in the public works and employment as a casual wage labourer in the private sector for an individual/household i in district d in period t. $NREGA_d$ is a binary variable equal to one for districts from the first two phases (treatment group) and equal to zero for districts from the third phase (control group). $Post55_t$ is a binary variable equal to zero for the pseudo-pre period (April 1999 to December 1999, Sub-Rounds 1 and 2 of Round 55), and $Post55_t$ is equal to one for the pseudo-post period (January 2000 to June 2000, Sub-Rounds 3 and 4 of Round 55). The coefficient of the interaction term between $NREGA_d$ and $Post55_t$, α_3, is the coefficient of

interest. For the parallel trends assumption to hold true, the coefficient of interest, α_3, must not be statistically different from zero.

Mathematically, the following is the parallel trends assumption, given equation (4):

$$E(Y_{idt}|NREGA_d = 0, Post55_t = 1) - E(Y_{idt}|NREGA_d = 0, Post55_t = 0) = E(Y_{0idt}|NREGA_d = 1, Post55_t = 1) - E(Y_{idt}|NREGA_d = 1, Post55_t = 0) \quad \textbf{(5)}$$

Equation (4) implies that the change in outcomes between the treatment and control groups would be the same in the absence of the NREGA program, the parallel trends assumption.

6.1 Check for Lag Assumption

Using data from Round 64, and limiting the sample to only phase 3 districts, the following equation is estimated to check for the validity of the assumption of the lag between the implementation periods of phase 1 and 2 districts, and the treatment period:

$$Urbanto\ RuralMigrate_{it} = \upsilon_0 + \upsilon_1 \cdot Post\_3_t + \epsilon_{it} \quad \textbf{(6)}$$

Post_3$_t$ is a binary variable equal to one, if the period is sub-round 4, April 2008 to June 2008, coinciding with the rollout of the NREGA program in April 2008 to phase 3 districts. The results of estimating equation (6) are presented in the Table 13.

Table 13: Urban to Rural Migration for Phase 3 Districts

VARIABLES	(1) urban_rural	(2) urban_rural_samedistrict	(3) urban_rural_samestate	(4) urban_rural_diffstate
Post_3	-0.000110	0.000379	0.000775	-0.00126**
	(0.00123)	(0.000834)	(0.000782)	(0.000492)
Constant	0.0338***	0.0149***	0.0127***	0.00617***
	(0.000612)	(0.000410)	(0.000380)	(0.000266)
Observations	115,518	115,518	115,518	115,518
R-squared	0.000	0.000	0.000	0.000

*Notes: This Table only uses Districts from Phase 3 to test for urban to rural migration, immediately after the implementation of the NREGA program in Phase 3 Districts using data from Round 64. It follows a simple OLS estimation, with Post = 1 if Sub Round is 4. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

The coefficients on Post_3 are ambiguous across various specifications and are not statistically different from zero, apart from the coefficient for estimating urban to rural migration from a different state, which is negative and statistically significant at the 5% level.

The paper mainly uses overall urban to rural migration, and the results from Table 13 corroborate the validity of the assumption of the lag, as the coefficient in column 1 of Table 13 is negative and not statistically different from zero, as is the case with the coefficient on the

interaction term between NREGA and Post in column 1 of Table 2. Hence the assumption of a lag is reasonable, and the results presented in the paper are valid.

6.2 Balancing Checks

The districts considered to be as part of the treatment group are from phases 1 and 2, and the control group is districts from phase 3. As mentioned earlier, the implementation of the NREGA program was not random, with the poorest districts consisting of the most socially disadvantaged groups receiving the program before it was rolled out to the rest of rural India in Phase 3. Using data from Round 55, some statistics of districts from the three different phases are provided to see how different these districts were prior to the implementation of the program. Table 14 provides summary statistics of some variables of interest between the districts from the three phases.

Table 14: Balancing Check between Districts from different Phases

Illiterate			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.4118961	0.4135697	0.3941505
Observations	54,606	35,815	64,587
Primary Education			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.1438303	0.1433198	0.1541642
Observations	54,606	35,815	64,587
Secondary Education			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.0716222	0.0765043	0.0795052
Observations	54,606	35,815	64,587
Married			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.4431747	0.4589697	0.4576153
Observations	54,606	35,815	64,587
Urban to Rural Migrants			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.0303263	0.0292056	0.0346045
Observations	54,606	35,815	64,587
Public Works			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.0012087	0.0014798	0.0010683
Observations	54,606	35,815	64,587
Private Work			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.2499542	0.2474661	0.2364563
Observations	54,606	35,815	64,587

Female			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	0.4826942	0.4890409	0.496555
Observations	54,606	35,815	64,587
Age			
Phase	**Phase 1**	**Phase 2**	**Phase 3**
Mean	25.44240284	25.40494564	25.2707932
Observations	54,606	35,815	64,587

Notes: This Table presents the means and observations of various statistics between districts from the 3 phases of NREGA roll-out using data from Round 55.

The districts from the first two phases are very similar across the various variables, however there are appear to be a few differences in the control group; districts from the third phase. It appears that on average there are fewer illiterate people in districts from the third phase in relation to districts from the first two phases, and districts from the third phase have a slightly higher percentage of people who have completed primary education in relation districts from the first two phases.

In comparison to districts from the first two phases, there are more urban to rural migrants in districts from the third phase. It is also interesting to note that fewer people were employed in the public works in the districts from the third phase, in relation to the districts from the first two phases.

Even though the treatment (districts from the first two phases) and the control groups (districts from the third phase) are slightly different, a difference-in-difference estimation would be valid as long as the pre-program trends between the treatment and control groups are the same, implying that the only difference in trends in the post-program period is a result of the NREGA program.

6.3 Parallel Trends Check for Urban to Rural Migration

Using equation (4), pre-NREGA migratory trends are tested, specifically on the outcome of urban to rural migration within the whole country, and presented below in Table 15.

The coefficient on the interaction term between NREGA and Post_55 is positive but not statistically significant from zero. This implies that between the first two sub-rounds and the last two-sub rounds of Round 55, there was no difference in the average outcomes of urban to rural migration between districts from the first two phases of the NREGA program (treatment group) and districts from the third phase of the NREGA program (control group).

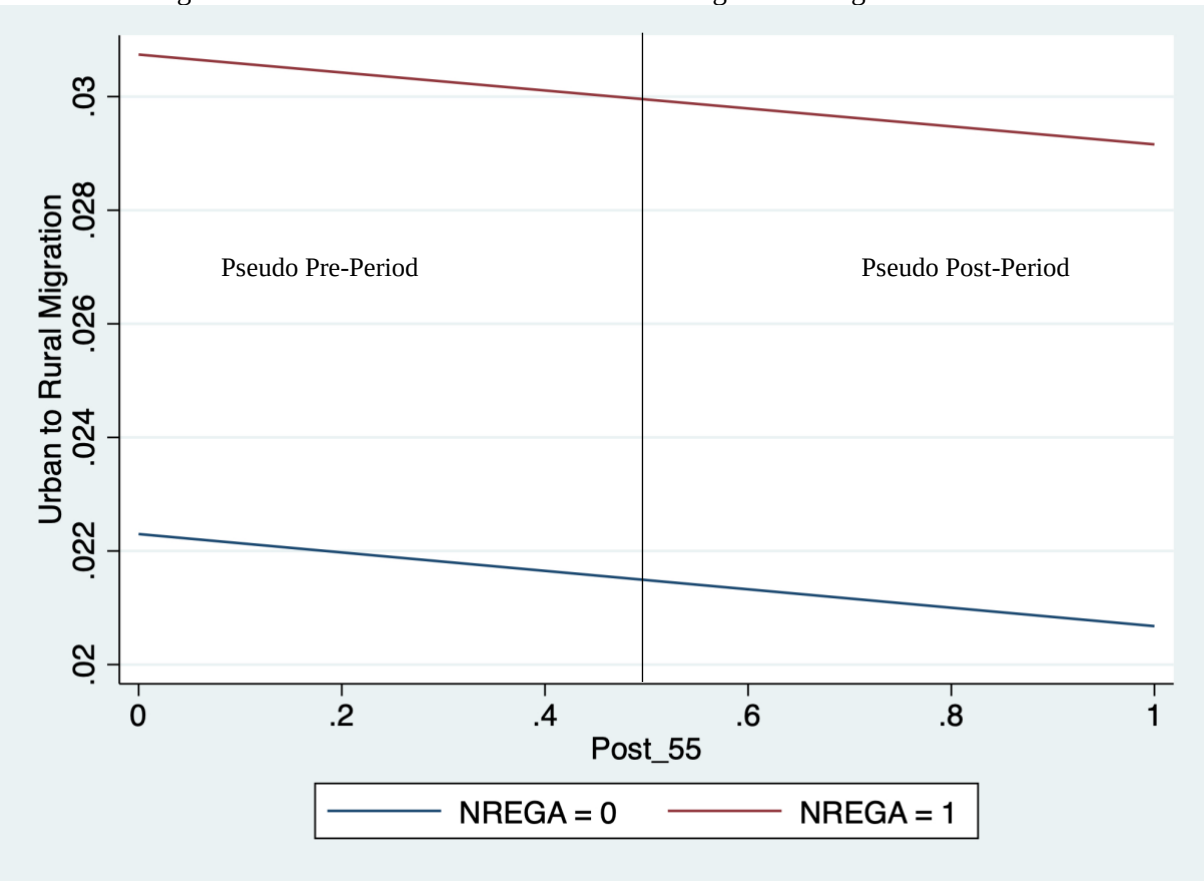
Table 15: Parallel Trends for Urban to Rural Migration

VARIABLES	(1) urban_rural
NREGA	0.00844***
	(0.000929)
Post_55	-0.00162***
	(0.000503)
NREGA*Post_55	3.96e-05
	(0.00125)
Constant	0.0223***
	(0.000377)
Observations	427,235
R-squared	0.001

*Notes: This table presents urban to rural migration within the whole country using data from Round 55. NREGA*Post_55 is the interaction term and coefficient of interest. The results are estimated using equation 4 with no controls. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

The following trends are illustrated graphically in Figure 4, to illustrate how the trends in urban to rural migration between the treatment and control group remain parallel through the entirety of Round 55, from July 1999 to June 2000.

Figure 4: Parallel Trends for Urban to Rural Migration using Round 55 Data



Notes: This figure illustrates the parallel trends for urban to rural migration estimated using equation (3) with the results presented in Table 11. The table uses data from Round 55. The blue line is the control group indicated by NREGA = 0, and the red line is the treatment group indicated by NREGA = 1. The x-axis is the time variable Post_55, indicating the evolution of trends in urban to rural migration through the period of July 1999 to June 2000, Round 55.

6.4 Parallel Trends Check for Employment in the Public Works

Once again, using equation (4), pre-NREGA trends for employment in the public works is tested. This is a key check for parallel trends, given that the NREGA program explicitly aims to provide guaranteed employment for 100 days a year in the public works to those from rural districts. It is important to ensure that the NREGA program has a causal effect on variables tested in Section 5, and hence parallel trends need to be present.

The results for the parallel trends check for employment in the public works is presented below in Table 16.

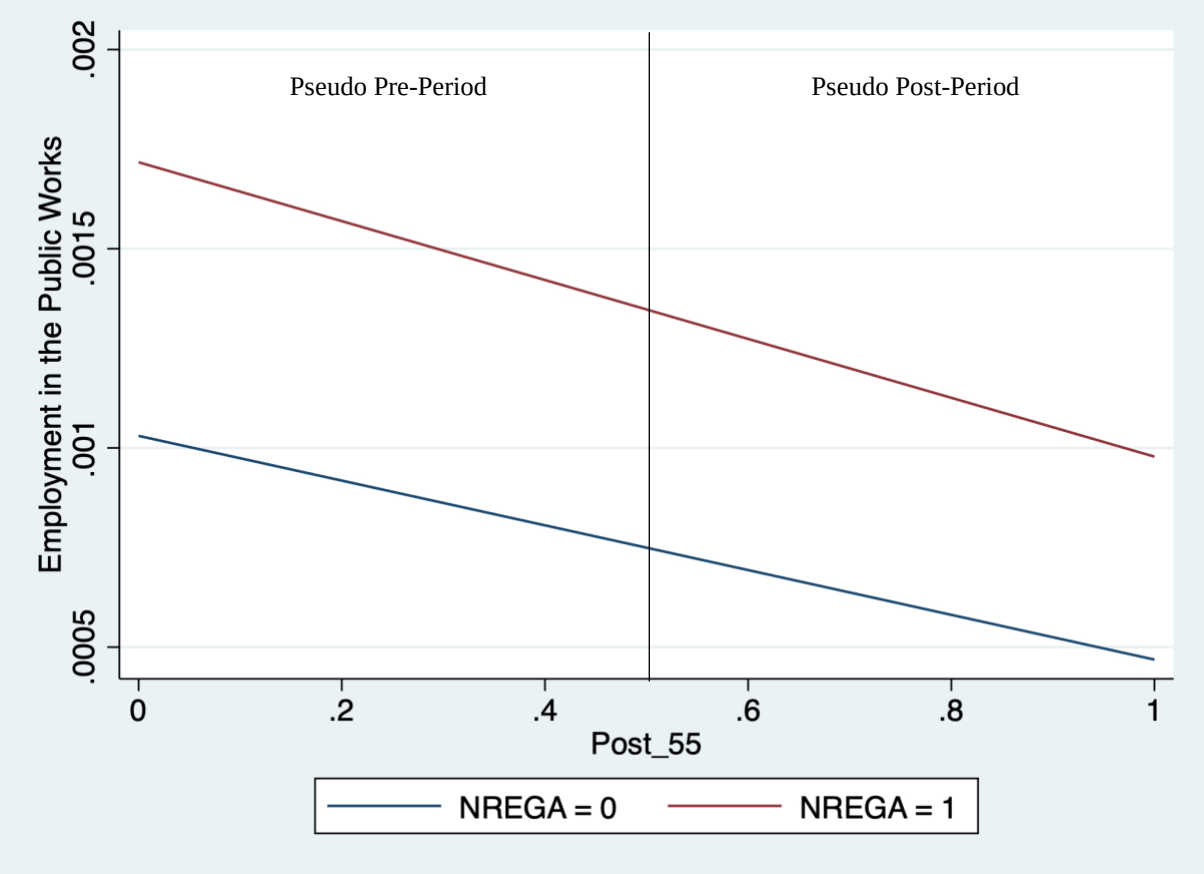
The coefficient on the interaction term between NREGA and Post_55 is negative but not statistically different from zero. This implies that between the first two sub-rounds and the last two-sub rounds of Round 55, there was no difference in the average outcomes of employment in the public works between districts from the first two phases of the NREGA program (treatment group) and districts from the third phase of the NREGA program (control group). The following trends are illustrated graphically in Figure 5, to illustrate how the trends in employment in the public works between the treatment and control group remain parallel through the entirety of Round 55, from July 1999 to June 2000.

Table 16: Parallel Trends Check for Employment in the Public Works

VARIABLES	(1) publicworks
NREGA	0.000687***
	(0.000219)
Post_55	-0.000562***
	(9.63e-05)
NREGA*Post_55	-0.000177
	(0.000266)
Constant	0.00103***
	(8.19e-05)
Observations	427,235
R-squared	0.000

*Notes: This table presents employment in the public works using data from Round 55. NREGA*Post_55 is the interaction term and coefficient of interest. The results are estimated using equation 4 with no controls. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Figure 5: Parallel Trends for Employment in the Public Works using Round 55 Data



Notes: This figure illustrates the parallel trends for employment in the public works estimated using equation (3) with the results presented in Table 12. The table uses data from Round 55. The blue line is the control group indicated by NREGA = 0, and the red line is the treatment group indicated by NREGA = 1. The x-axis is the time variable Post_55, indicating the evolution of employment in the public works through the period of July 1999 to June 2000, Round 55.

6.5 Parallel Trends for Employment in the Private Sector as a Casual Wage Labourer

The NREGA program as stated earlier provides 100 days of guaranteed employment as a casual labourer in the public works. It is concomitant to assume that NREGA would have a causal effect on the direct opposite, employment in the private sector as a casual wage labourer. In order to test this, a binary outcome variable casualpvt has been created based on responses to the survey question of an individual's usual principal activity. This is then estimated using equation (4), with the results presented in Table 17. The coefficient on the interaction term between NREGA and Post_55 is positive but not statistically significant from zero. This implies that between the first two sub-rounds and the last two-sub rounds of Round 55, there was no difference in the average outcomes of employment as a casual wage labourer in the private sector between districts from the first two phases of the NREGA program (treatment group) and districts from the third phase of the NREGA program (control group).

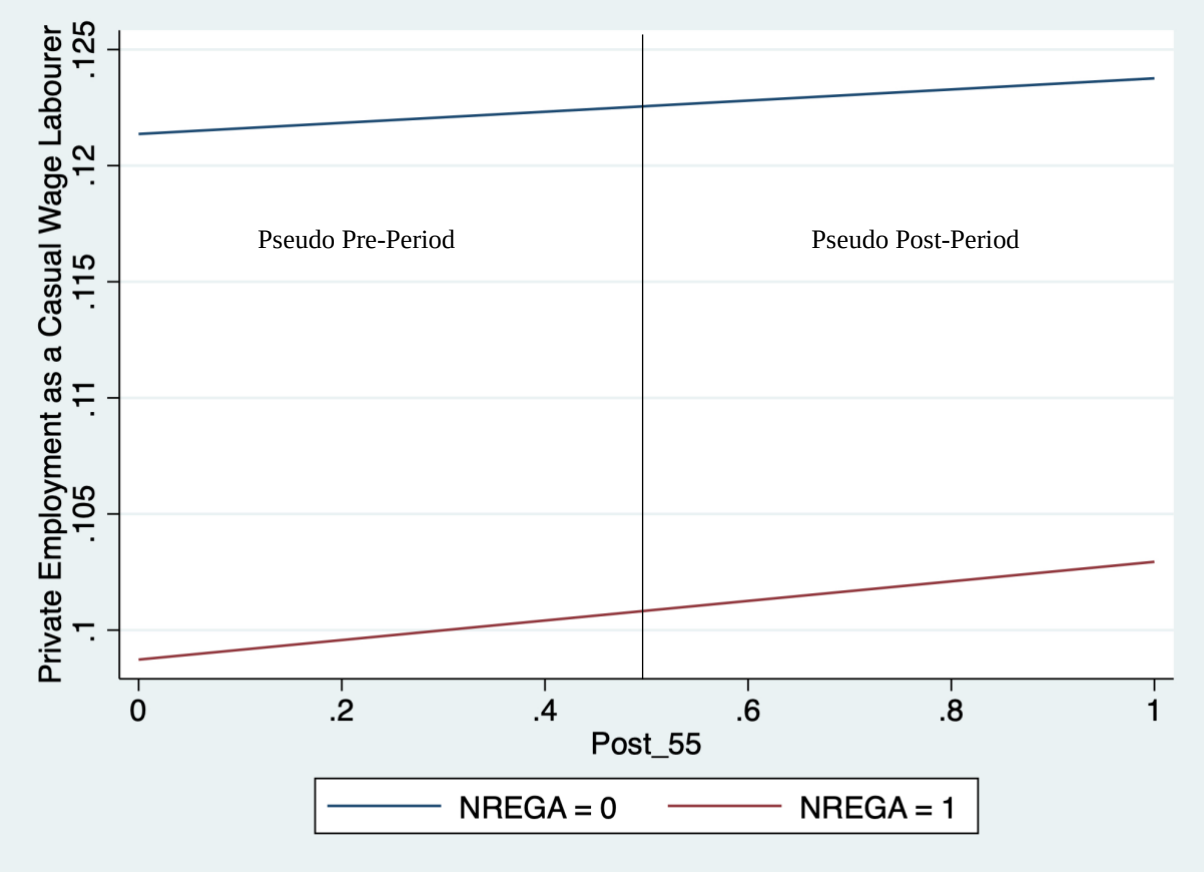
Table 17: Parallel Trends Check for Private Employment as a Casual Wage Labourer

VARIABLES	(1) casualpvt
NREGA	-0.0226***
	(0.00169)
Post_55	0.00240**
	(0.00113)
NREGA*Post_55	0.00181
	(0.00231)
Constant	0.121***
	(0.000834)
Observations	427,235
R-squared	0.001

*Notes: This table presents employment in the private sector as a casual wage labourer using data from Round 55. NREGA*Post_55 is the interaction term and coefficient of interest. The results are estimated using equation 4 with no controls. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

The following trends are illustrated graphically in Figure 6, to illustrate how the trends in employment in the private sector as a casual wage labourer between the treatment and control group remain parallel through the entirety of Round 55, from July 1999 to June 2000.

Figure 6: Parallel Trends for Private Employment as a Casual Wage Labourer using Round 55 Data



Notes: This figure illustrates the parallel trends for private employment as a casual wage labourer estimated using equation (3) with the results presented in Table 13. The table uses data from Round 55. The blue line is the control group indicated by NREGA = 0, and the red line is the treatment group indicated by NREGA = 1. The x-axis is the time variable Post_55, indicating the evolution of employment in the public works through the period of July 1999 to June 2000, Round 55.

7.0 Conclusion

7.1 Limitations

The migration surveys as part of the Employment and Unemployment Surveys conducted by the NSSO is highly irregular. The past two migration surveys conducted were in rounds 55 and 64. This causes an issue of the pre-program and post-program period not matching too well with the migration surveys. This issue has been overcome by assuming a lag between the implementation of the NREGA program in the first two phases and the actual decision and act of migration, a fair assumption given the decision-making process, in line with Treyz et al. (1993).

It can be assumed that the results estimated in this study are local average treatment effects (LATE), given the pre-planned staggered implementation of the program, cultural context of India and the fact that the NREGA program is focused only on the rural population of India. Additionally, there is no migration data post July 2008, hindering analysis on long-term migratory effects and effects of the NREGA program on various other social outcomes.

Given the lack of migration data post-July 2008, it is hard to distinguish between temporary and permanent migration. Bhagat and Keshri (2013) establish temporary migration trends in their paper, and the urban to rural migratory effects estimated in this paper could be temporary stress-migration as a means of survival. It is hard to distinguish between temporary and permanent migration; most people may be migrating for temporary work; or they could be migrating to rural areas to live but continue work in urban areas.

The issue of not having migration data past Round 64, also inhibits this study from truly understanding employment effects on urban to rural migrants. It is fair to assume that it takes time for people to realise employment in the public works given that they have just migrated. This study is unable to analyse whether employment and educational attainment is fully realised or not for these migrants. If people migrate, it takes time to see effects on employment and education despite the lags created for this study.

The final limitation is that the assumption of a lag between roll-out of NREGA to certain districts and migration from urban areas could be violated relatively easily. If migration is a means of survival, then no lag should exist, and the decision to migrate would be a rather quick one. In that case, the results presented in this study, given that the treatment groups are from districts of the first two phases of the NREGA program, may be an under-estimation of the true urban to rural migratory effects because of the NREGA program.

7.2 Suggestions for Future Research

Migration is a decision that is hard to tease out through just quantitative data, especially with the datasets used for this study. More qualitative data elicited through qualitative surveys, to understand preferences, is required to understand the decisions and motivations behind an individual's decision to migrate.

A more appropriate research setting would be to focus in on certain rural areas within a certain proximity to urban areas. By focusing on out- and in-migration of these rural areas, a more detailed analysis of the type of migration can be understood. By focusing on rural areas within a certain proximity to urban areas, migration is more likely. Controls for infrastructure such as roads and availability of public transportation can be further added to the study.

It is worth noting that the NREGA program provides employment in the public works; infrastructure in rural areas should improve as a result of that. A study on whether the infrastructure has really improved, and whether this improvement has led to an increase in social and human capital outcomes would be worth exploring as well.

7.3 Policy Implications

This study highlights that urban to rural migration does take place in certain localised contexts, whether temporary, permanent or as a response to gaining a better life or just surviving. It is worth noting that the NREGA program has no explicit features to facilitate or promote migration. The government can facilitate this migration and leverage it as a poverty alleviation tool. Given India's rapid increase in population, and increasing population density in urban areas, facilitating migration and simultaneously alleviating the poorest urban dwellers from poverty is worthwhile. It could be achieved by directly incentivising the poorest in urban areas to migrate to rural areas, and provide them with guaranteed employment through the NREGA program. This incentivising can be done by just educating the poor about the NREGA program or providing additional benefits that would motivate them to migrate.

Additionally, the study highlights that having a higher HDI improves the implementation of policies and programs, possibly even incentivising state level governments to have their own initiatives to improve their population's welfare. The central government of India should place a higher priority on equitable development, with a focus on building up infrastructure in states with low HDI. The improved infrastructure would increase human capital outcomes, health outcomes and education outcomes, thereby increasing HDI. As opined in this study, states with higher HDI appear to have greater trust placed in their governments, which is crucial to

effectively implementing and rolling out programs that result in a high take-up rate. This would be crucial in the long-run to effectively allow poverty-alleviation programs to work, and make for a more equitable, pro-poor growth.

References

Adewale, J. G. (2005). Socio-economic factors associated with urban-rural migration in Nigeria: A case study of oyo state, Nigeria. *Journal of Human Ecology*, *17*(1), 13–16. https://doi.org/10.1080/09709274.2005.11905752

Afridi, F., Mukhopadhyay, A., & Sahoo, S. (2012). Female labour force participation and child education in India: The effect of the National Rural Employment Guarantee Scheme. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2085180

Azam, M. (2011). The impact of Indian Job Guarantee Scheme on labor market outcomes: Evidence from a natural experiment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1941959

Breitkreuz, R., Stanton, C., Brady, N., Pattison-Williams, J., King, E. D., Mishra, C., & Swallow, B. (2017). The mahatma gandhi national rural employment guarantee scheme: A policy solution to rural poverty in India? *Development Policy Review*, *35*(3), 397–417. https://doi.org/10.1111/dpr.12220

Drèze, J. (2019). *Sense and solidarity: Jholawala economics for everyone*. Oxford University Press.

Fulford, S. (2013). The Puzzle of Marriage Migration in India. *Boston College Working Papers in Economics*, *820*.

Ghosh, J. (2014). Can employment schemes work? the case of the Rural Employment Guarantee in India. *Contributions to Economic Theory, Policy, Development and Finance*, 145–171. https://doi.org/10.1057/9781137450968_7

Global Data Lab. (n.d.). *Subnational HDI (v7.0)*. Global Data Lab. https://globaldatalab.org/shdi/download/?levels=1%2B4&years=2008%2B2000&interpolation=0&extrapolation=0

Haushofer, J., & Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the Poor: ExperimentalEvidence from Kenya*. *The Quarterly Journal of Economics*, *131*(4), 1973–2042. https://doi.org/10.1093/qje/qjw025

Imbert, C., & Papp, J. (2015). Labor market effects of social programs: Evidence from India's employment guarantee. *American Economic Journal: Applied Economics*, *7*(2), 233–263. https://doi.org/10.1257/app.20130401

Imbert, C., & Papp, J. (2020). Costs and benefits of rural-urban migration: Evidence from India. *Journal of Development Economics*, *146*, 102473. https://doi.org/10.1016/j.jdeveco.2020.102473

Islam, M., & Sivasankaran, A. (2015). How does child labor respond to changes in adult work opportunities? Evidence from NREGA. *Harvard University, Cambridge, MA.*

Keshri, K., & Bhagat, R. B. (2013). Socioeconomic determinants of temporary labour migration in India. *Asian Population Studies*, *9*(2), 175–195. https://doi.org/10.1080/17441730.2013.797294

McKenzie, D., & Rapoport, H. (2010). Can migration reduce educational attainment? evidence from Mexico. Journal of Population Economics, 24(4), 1331–1358. https://doi.org/10.1007/s00148-010-0316-x

Narayanan, S., & Das, U. (2014). Women Participation and Rationing in the Employment Guarantee Scheme. Economic and Political Weekly, 49(46), 46–53. http://www.jstor.org/stable/24480958

National Sample Survey Office. (2000). India - Employment and Unemployment Survey, July 1999 - June 2000, NSS 55th Round [Data file]. Retrieved from http://www.microdata.gov.in/NADA43/index.php/catalog/90/datafile/F5

National Sample Survey Office. (2008). India - Employment and Unemployment Survey, July 2007 - June 2008, NSS 64th Round [Data file]. Retrieved from http://www.microdata.gov.in/NADA43/index.php/catalog/117

Pankaj, A., & Tankha, R. (2010). Empowerment Effects of the NREGS on Women Workers: A Study in Four States. Economic and Political Weekly, 45(30), 45–55. http://www.jstor.org/stable/20764337

Ravallion, M., & Datt, G. (2002). Why has economic growth been more pro-poor in some states of India than others? *Journal of Development Economics*, *68*(2), 381–400. https://doi.org/10.1016/s0304-3878(02)00018-4

Treyz, G. I., Rickman, D. S., Hunt, G. L., & Greenwood, M. J. (1993). The Dynamics of U.S. Internal Migration. The Review of Economics and Statistics, 75(2), 209–214. https://doi.org/10.2307/2109425

Young, A. (2013). Inequality, the urban-rural gap, and migration*. The Quarterly Journal of Economics, 128(4), 1727–1785. https://doi.org/10.1093/qje/qjt025

Appendix

Table A1: Employment in the Public Works for Urban to Rural Migrants (Multiple Specifications)

VARIABLES	(1) publicworks	(2) publicworks	(3) publicworks	(4) publicworks	(5) publicworks
urban_rural	-0.000386	-0.000341			
	(0.000532)	(0.000532)			
NREGA	0.000597***	0.000619***	0.000595***	0.000588***	0.000589***
	(0.000182)	(0.000183)	(0.000179)	(0.000180)	(0.000179)
Post	0.000279	0.000283	0.000274	0.000254	0.000251
	(0.000241)	(0.000241)	(0.000237)	(0.000237)	(0.000235)
NREGA_Post	-0.000360	-0.000364	-0.000284	-0.000321	-0.000250
	(0.000329)	(0.000329)	(0.000326)	(0.000325)	(0.000325)
migrate1_NREGA	-0.000603	-0.000640			
	(0.000748)	(0.000750)			
migrate1_post	-0.000796	-0.000826			
	(0.000570)	(0.000571)			
migrate1_triple	0.00456**	0.00464**			
	(0.00224)	(0.00224)			
sex		0.00131***			
		(0.000151)			
Age		1.55e-05***			
		(2.71e-06)			
state_control		6.63e-06			
		(7.10e-06)			
urban_rural_samedistrict			0.000320		
			(0.00121)		
migrate2_NREGA			-0.000635		
			(0.00168)		
migrate2_post			-0.00148		
			(0.00123)		
migrate2_triple			0.00282		
			(0.00302)		
urban_rural_samestate				-0.000901***	
				(0.000125)	
migrate3_NREGA				-0.000588***	
				(0.000180)	
migrate3_post				-0.000254	
				(0.000237)	
migrate3_triple				0.00880*	
				(0.00488)	
urban_rural_diffstate					-0.000895***
					(0.000124)
migrate4_NREGA					-0.000589***
					(0.000179)
migrate4_post					-0.000251
					(0.000235)
migrate4_triple					0.000250
					(0.000325)
Constant	0.000902***	-0.000312*	0.000885***	0.000901***	0.000895***
	(0.000126)	(0.000176)	(0.000124)	(0.000125)	(0.000124)

Observations	220,020	220,020	220,020	220,020	220,020
R-squared	0.000	0.000	0.000	0.000	0.000

*Notes: This table presents the results for employment in the public works for urban to rural migrants in the whole country using data from Round 64. migrate1_NREGA is the interaction between urban_rural and NREGA. migrate1_post is the interaction between urban_rural and post. migrate1_triple is a triple interaction term between NREGA, Post and urban_rural, and is the main coefficient of interest. Coefficients that start with migrate2 follow the same as the previous interactions but using urban to rural migration within the same district. Coefficients that start with migrat3 follow the same as the previous interactions but using urban to rural migration within the same state. Coefficients that start with migrate4 follow the same as the previous interactions but using urban to rural migration between different states. The results are estimated using equation 2. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Table A2: Any Employment in the Private Sector for Urban to Rural Migrants (Multiple Specifications)

VARIABLES	(1) pvtwork	(2) pvtwork	(3) pvtwork	(4) pvtwork	(5) pvtwork
urban_rural	0.0334***	-0.00100			
	(0.0102)	(0.0100)			
NREGA	0.0240***	0.0308***	0.0238***	0.0243***	0.0238***
	(0.00232)	(0.00223)	(0.00230)	(0.00230)	(0.00230)
Post	-0.000674	0.000233	-0.000466	-0.000276	-0.000246
	(0.00310)	(0.00297)	(0.00307)	(0.00307)	(0.00306)
NREGA_Post	-0.00114	-0.00223	-0.00109	-0.00158	-0.00122
	(0.00403)	(0.00386)	(0.00400)	(0.00400)	(0.00399)
migrate1_NREGA	0.0129	0.0141			
	(0.0147)	(0.0144)			
migrate1_post	0.0126	0.00839			
	(0.0176)	(0.0170)			
migrate1_triple	0.000870	0.0148			
	(0.0255)	(0.0247)			
sex		0.124***			
		(0.00176)			
Age		0.00441***			
		(4.30e-05)			
state_control		0.00715***			
		(0.000103)			
urban_rural_samedistrict			0.0138		
			(0.0151)		
migrate2_NREGA			0.0205		
			(0.0217)		
migrate2_post			0.0172		
			(0.0256)		
migrate2_triple			-0.0126		
			(0.0377)		
urban_rural_samestate				0.0515***	
				(0.0168)	
migrate3_NREGA				-0.0136	
				(0.0243)	
migrate3_post				0.00602	
				(0.0288)	
migrate3_triple				0.0318	
				(0.0421)	
urban_rural_diffstate					0.0374
					(0.0232)
migrate4_NREGA					0.0428
					(0.0327)
migrate4_post					0.0172
					(0.0420)
migrate4_triple					-0.0287
					(0.0583)
Constant	0.235***	-0.0715***	0.236***	0.235***	0.236***
	(0.00178)	(0.00247)	(0.00177)	(0.00177)	(0.00176)
Observations	220,020	220,020	220,020	220,020	220,020

R-squared	0.001	0.084	0.001	0.001	0.001

*Notes: This table presents the results for any employment in the private sector for urban to rural migrants in the whole country using data from Round 64. migrate1_NREGA is the interaction between urban_rural and NREGA. migrate1_post is the interaction between urban_rural and post. migrate1_triple is a triple interaction term between NREGA, Post and urban_rural, and is the main coefficient of interest. Coefficients that start with migrate2 follow the same as the previous interactions but using urban to rural migration within the same district. Coefficients that start with migrat3 follow the same as the previous interactions but using urban to rural migration within the same state. Coefficients that start with migrate4 follow the same as the previous interactions but using urban to rural migration between different states. The results are estimated using equation 2. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Table A3: Educational Attainment for Urban to Rural Migrants (Multiple Specifications)

VARIABLES	(1) educ	(2) educ	(3) educ	(4) educ	(5) educ
urban_rural	-0.133***	-0.0540***			
	(0.00755)	(0.00674)			
NREGA	-0.0118***	-0.0243***	-0.0111***	-0.0113***	-0.0110***
	(0.00233)	(0.00206)	(0.00230)	(0.00230)	(0.00229)
Post	0.00410	0.00240	0.00382	0.00351	0.00314
	(0.00318)	(0.00279)	(0.00314)	(0.00313)	(0.00312)
NREGA_Post	0.00399	0.00406	0.00410	0.00476	0.00511
	(0.00407)	(0.00358)	(0.00402)	(0.00401)	(0.00400)
migrate1_NREGA	-0.0208**	-0.0143			
	(0.00994)	(0.00902)			
migrate1_post	-0.0177	-0.00917			
	(0.0125)	(0.0111)			
migrate1_triple	0.0188	-0.00130			
	(0.0171)	(0.0154)			
sex		0.0396***			
		(0.00161)			
Age		-0.0101***			
		(3.88e-05)			
state_control		-0.000783***			
		(9.09e-05)			
urban_rural_samedistrict			-0.125***		
			(0.0115)		
migrate2_NREGA			-0.0314**		
			(0.0149)		
migrate2_post			-0.0237		
			(0.0184)		
migrate2_triple			0.0426*		
			(0.0259)		
urban_rural_samestate				-0.135***	
				(0.0118)	
migrate3_NREGA				-0.0216	
				(0.0158)	
migrate3_post				-0.0171	
				(0.0195)	
migrate3_triple				0.0137	
				(0.0267)	
urban_rural_diffstate					-0.131***
					(0.0168)
migrate4_NREGA					-0.00217
					(0.0223)
migrate4_post					-0.00102
					(0.0301)
migrate4_triple					-0.0249
					(0.0390)
Constant	0.251***	0.527***	0.248***	0.248***	0.247***
	(0.00182)	(0.00281)	(0.00180)	(0.00180)	(0.00179)
Observations	220,020	220,020	220,020	220,020	220,020
R-squared	0.003	0.218	0.001	0.001	0.001

*Notes: This table presents the results educational attainment for urban to rural migrants in the whole country using data from Round 64. migrate1_NREGA is the interaction between urban_rural and NREGA. migrate1_post is the interaction between urban_rural and post. migrate1_triple is a triple interaction term between NREGA, Post and urban_rural, and is the main coefficient of interest. Coefficients that start with migrate2 follow the same as the previous interactions but using urban to rural migration within the same district. Coefficients that start with migrat3 follow the same as the previous interactions but using urban to rural migration within the same state. Coefficients that start with migrate4 follow the same as the previous interactions but using urban to rural migration between different states. The results are estimated using equation 2. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*