Where have all the men gone?

Temporary Migration & Rural Public Works - A Discontinuity Approach

Kasturi Kandalam*

Last Updated: November 8, 2025

Abstract

This paper examines the impact of the National Rural Employment Guarantee Scheme (NREGS) on temporary migration in India, using a fuzzy regression discontinuity design that exploits the phased rollout of the program. Drawing on nationally representative data from the NSS Employment–Unemployment Survey (2007–08), the study finds that the introduction of NREGS significantly increased the seasonal migration of young males, while the effect on young females was negligible. The robustness of the findings is confirmed across different model specifications, bandwidths, and by using both sharp and fuzzy designs. The results suggest that NREGS enabled male migration by either substituting female labor into public works or by reducing the risks associated with migration through the availability of fallback employment. These findings highlight the complex and sometimes unintended consequences of rural workfare programs on household labor allocation and migration strategies.

Keywords: Temporary Migration, Public Works, Employment, India

*kasturiveemohan@gmail.com

1 Introduction

In India, it is observed that the majority of internal migration tends to be seasonal[1] (Rajan, 2020), (Haberfeld et al., 1999), (Deshingkar and Farrington, 2009), & (Keshri and Bhagat, 2012). Approximately 86% of migrants are males and 73.3% of them move to urban areas for work or in search of work.[2] Most young males migrate "looking for good pay and experience of city life" (Deshingkar and Start, 2003, p 20). Temporary migration is argued to be a "permanent part of households' long-term economic strategies" (Coffey et al., 2015, p 3). However, development policies and welfare programs often aim to reduce labour migration (de Haan and Rogaly, 2002). The approach put forth in the "*Report of the working group on rural poverty alleviation programmes*" is to identify backward districts and set in place public works that reduce distress migration due to unavailability of work (Planning Commission, 2001, p 63).

This paper estimates the effect of rural public works such as the National Rural Employment Guarantee Act (NREGS) on the temporary migration of young males. NREGS is a rural workfare programme that provides gainful employment to households from rural areas for 100 days during the agricultural off-season. We leverage the variation in NREGS initial allocation across rural districts of India to estimate the causal effect of NREGS on short-term migration.

NREGS was intended to be allocated based on the 'backwardness rank' drawn from an index considering various parameters measuring distress within a district. NREGS allocation in actuality did not adhere to the recommended rule. This non-adherence allows us to employ a fuzzy regression discontinuity design (FRDD).

We use the National Sample Survey Office (2008) - Employment Unemployment Survey (NSS-EUS) 64^{th} round to pursue FRDD. The 64^{th} round of NSS-EUS extensively captures migration patterns of individuals who are moving temporarily or permanently. We find that NREGS increases migration of males aged less than 35 years by 11.2 percentage points. The result for young females is small and insignificant. The results remain positive with varying bin lengths, turning the FRDD sharp and using a normalised index instead of the rank.

We attribute the increase in temporary migration because of NREGS to two prospective mechanisms - an increase in females' uptake of NREGS is allowing males to migrate for work, or job search (Tumbe, 2015). The existence of NREGS is also allowing the migrants to take up more migration because there is a fallback at origin (NREGS). In other words, NREGS insures migrants against risky migration. According to Morten (2019), the ability to migrate provides households with a way to insure themselves against risks, which could alter their motivations for participating in informal risk-sharing. Additionally, Bryan et al. (2014) found introducing an additional treatment that offers migration insurance leads households

[1] We use the words temporary and seasonal interchangeably.

[2] Based on authors' calculations.

to migrate.

The rest of the paper proceeds as follows: Section 2 describes a survey of relevant literature on migration and NREGS. Section 3 explicates the context of NREGS roll-out and implementation and the dataset being used. Section 4 explains the reduced form approach to estimate the impact of the NREGS on short-term migration. Section 5 presents the results and section 6 concludes.

2 A survey of literature on Migration and NREGS

In this section, we conduct a comprehensive review of existing studies on temporary migration and the National Rural Employment Guarantee Scheme (NREGS). The aim is to explore the complexities of temporary migration and how NREGS impacts it. This will help us identify areas that require further exploration.

2.1 Migration

The first formal understanding and articulation of seasonal and circular labour migration were done in the 1970s (Fagerlid and Tisdel, 2020). In the Indian context, Rao (1994) defined temporary migration to be "characteristically short term, repetitive or cyclical in nature, and adjusted to the annual agricultural cycle". This notion contested the linear model of migration and its relation with urban expansion.

Seminal work by Harris and Todaro (1970) analyses, in a two-sector model[3], the reason behind unemployment in the urban areas. The Harris-Todaro model suggests that individuals will choose to migrate if the anticipated benefits from migration to urban areas surpass the benefits of remaining in rural areas. But, in the case of temporary migration, the costs are low and fixed based on pre-existing networks (Munshi and Rosenzweig, 2016) and 'time-varying opportunity costs' (Coffey et al., 2015). The migrants we are focussing on frequently move from rural areas for work or in search of work.

Our paper is influenced by the theoretical framework of Stark and Bloom (1985), who interpret labour migration from rural regions as a strategy to alleviate the risks inherent in agricultural livelihoods. We approach temporary migration as a way for households to diversify their income streams seasonally. Coffey et al. (2015) note temporary migration as a repeated annual strategy for households.

Temporary migration, which is documented to be "seven times larger than permanent migration" (Keshri and Bhagat, 2013, p 190), is predominantly a rural phenomenon characterized by migration from rural to urban areas. Factors such as low economic, educational, and social status significantly drive temporary labour migration, as opposed to permanent labour migration (Keshri and Bhagat, 2013).

[3]The two sectors comprise of an urban, modern sector and a rural, traditional sector.

Temporary migrant's experiences and outcomes at their destination are heavily dependent on social identities such as gender and caste. From fieldwork in Mahbubnagar, Andhra Pradesh, Korra (2011) illuminates the need to strengthen options at the origin to curtail migration to Hyderabad. Most migrants were looking for opportunities "to pay off debts and to get their daughters married" (Korra, 2011, p 70). Short-term migration of upper-caste individuals to Hyderabad is typically for white-collar or blue-collar jobs, while lower-caste migration is generally associated with agricultural or construction work. Temporary migration is also gendered. Mazumdar et al. (2013) highlight the lack of gender sensitivity in data collection processes, noting that migration has not diversified opportunities for women. The gendered nature of labour migration and the failure to collect detailed data on female mobility results in an inability to recognize the gender bias in the occupations held by migrants.

Understanding of temporary migration remains inadequate if it is only considered as an action, a rational agent takes to optimise their livelihoods with only economic motives in mind (Lal, 1989). Drawing from fieldwork in Tapu village of Jharkhand, Shah (2006) explicates seasonal casual labour migration as a "dynamic sociopolitical process". She argues that the migrants choice to work in brutal conditions of brick kilns is not just with an economic motive but as a "temporary escape to gain independence from parents, and live out prohibited amorous relationships" (Shah, 2006, p 12). The drudgery most Economics literature portrays about temporary migration is rightly estimated but barely scratches the surface of migrants' experiences both at origin and destination.

2.2 NREGS & Migration

Numerous researchers have conducted studies to understand the impact of NREGS on a variety of economic and social outcomes. Notable ones adjacent to this work are Zimmermann (2012) uses the initial phase-wise allocation of NREGS and our paper replicates their regression discontinuity framework, Coffey et al. (2015) conducted a primary survey on migration in North-West India and depict short-term migration as a livelihood diversification strategy, Morten (2019) structurally estimated migration decisions of households. In both Imbert and Papp (2020a) and Imbert and Papp (2020b), states are classified based on NREGS performance and migration decisions are understood. National Rural Employment Guarantee Scheme (NREGS) was explicitly set up to arrest migration (Planning Commission, 2001). The line of reasoning from Solinski (2012), Deshingkar and Start (2003) & Dodd et al. (2018) is that internal labour migration burdens the public infrastructure and overwhelms the job market in urban areas leading to sub-optimal outcomes for both the native population and the migrants.

Imbert and Papp (2020b) causally estimate the effect of NREGS on rural-to-urban migration and subsequently urban labour markets.[4] They mainly find that the existence of NREGS actively decreased migration compared to districts where NREGS was not implemented. They unpack the relation between the existence of NREGS on rural wages, migration, and urban wages. An interesting aspect of their

[4]The main research question and the data (NSS 64^{th} round) used are the same as ours but they estimate using a difference-in-difference approach.

work is that it sheds light on urban spillovers - the urban areas might have attracted migrants from both districts where NREGS was active and those where it was not.

They categorise rural districts as *early districts* that were selected to receive the program and compare them in *star states*[5] versus others. Their main finding is that NREGS increased rural wages and reduced short-term migration to urban areas. This is in line with Morten (2019) who jointly determines migration and risk sharing. Temporary migration is modelled as a form of self-insurance for households and implies that the existence of NREGS in the model reduces both risk-sharing and migration by increasing rural incomes. The reduced form estimation of this implication is what we are interested in. It is established that migration enables people to diversify their livelihood opportunities. We hypothesise that NREGS could curtail migration (because work is available at the source) or spur it (men move for longer periods of time and women join NREGS).

Hagen-Zanker and Leon-Himmelstine (2013) conduct a cross-country analysis that looks at the varying impacts of employment guarantee programmes on migratory flows. They conclude that the effects are usually mixed. There is a considerable amount of heterogeneity in the way NREGS affects migration in different parts of India. In Thoradipattu, Vellimalai, Kalrayan Hills Block and Neelamangalam Panchayats in Villupuram district of Tamil Nadu Jacob (2013) finds no impact. In village panchayats of Anchetty, Thaggatti, and Madakkal in the Krishnagiri district of Tamil Nadu, Dodd et al. (2018) find mixed effects. Across 30 panchayats in Andhra Pradesh, National Federation of Indian Women (2008) find that NREGS decreases migration. In Dokur village of Mahbubnagar in Andhra Pradesh, Singh (2013) find NREGS increased migration.

The impact of the National Rural Employment Guarantee Scheme (NREGS) exhibits variability due to several factors such as low pay or payment delays, age and gender dynamics - older males and females tend to opt for NREGS, while the younger population migrates for work and NREGS increases rural wages by providing competitive wages that deter seasonal migration.

An underscoring link that emerges from the literature survey is that rural households perceive NREGS both as a complement and/or a substitute for temporary migration. This perception hinges on structural reasons (such as agricultural productivity by season) and personal factors (such as marriage or death in the household) (Dodd et al., 2018) & (Solinski, 2012). Our research contribution lies in demonstrating that seasonal migration patterns are significantly influenced by shifts in employment opportunities and to shed light on the nuances of seasonal migration and public works.

[5] Star States as coined by Imbert and Papp (2015) are states where the fraction of working days allotted for NREGS exceeds 1%. The underlying implication of the high provision of employment guarantee is that the states have better means to run the programme, hence the *star*.

3 Background & Data

In this subsection, we present the context of the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), discuss its relevance for the identification strategy and explore the data used for empirical analysis.

3.1 Identification of Backward Districts

On 23^{rd} August 2005, the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA)[6] was passed. It was implemented in February 2006 under the UPA government. MGNREGA draws from the spirit of the Right to Information Act (2005), ensuring transparency in operations and easy access to administrative records. This guaranteed scrutiny since the nascent stages of NREGS. The act forms the legal basis for the National Rural Employment Guarantee Scheme (NREGS). In rural India, the scheme aims to provide willing adults of a household, an opportunity to self-select into 100 days of unskilled manual work for a state-specific minimum wage.

In practise, the gram panchayat provides a 'job card' to all the adult members of a rural household willing to provide unskilled manual labour. An eligible individual willing to work applies in writing and work must be provided within 15 days and at a distance less than or equal to 5 km of the village (Narayanan, 2020). During the agricultural lean season, MGNREGA targeted the creation of land assets, water conservation, rural connectivity and flood control among other activities (Planning Commission, 2009).

NREGS was implemented in districts in a phased manner (Planning Commission, 7 08, p 352). The allotment of NREGS was based on the district-level backwardness rank created by the task force set up by the Planning Commission headed by Rohini Nayyar (Nayyar Committee henceforth) that published "*Identification of Backward Districts for Wage Employment and Self-Employment Programmes*".[7] The backwardness rank was created using data from 55th round of the National Sample Survey Organisation - Employment & Unemployment Survey (National Sample Survey Office, 2000). The report investigates 447 districts in 17 states.[8] Measures on the following various parameters were calculated and divided into two indices.[9]

[6]On 2nd October 2009, the nomenclature of the act was amended from National Rural Employment Guarantee Act (NREGA) to Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA). Both are the same. I will use MGNREGA when referring to the act and NREGS when referring to the scheme that was formulated from the act

[7]The Nayyar Committee report was published as a part of the 2003 Planning Commission report.

[8]447 districts in 17 states out of 482 districts in 28 states were covered. The excluded 35 districts were urban areas/state capitals.

[9]The reason for creating two indices is the unavailability of data. To measure the incidence of poverty, 50^{th} round of NSS-EUS was used instead of the 55^{th} round. Hence, the first index was used to categorise and recommend districts for wage employment and self-employment programmes whereas the second index with five parameters was used to establish robustness (Planning Commission, 2003, p 7, line 14).

$$1^{st} \text{ index} \begin{cases} \text{Agricultural wages} \\ \text{Output per agricultural worker} \\ \text{SC/ST population of the district} \end{cases}$$

$$2^{nd} \text{ index} \begin{cases} \text{Per hectare agricultural productivity} \\ \text{Drought-proneness and Desert-proneness} \\ \text{Rural connectivity} \\ \text{Incidence of poverty} \\ \text{Unemployment rate} \end{cases}$$

For each district, two indices were created to ensure accurate targeting. Notably, 111 districts spanning 14 states are common to both indices. This implies that even when different parameters were used to ascertain backwardness, 111 districts commonly exhibit backwardness in these 14 states. Totalling the values from each of the parameters resulted in a composite index value. A district is considered more backward if it has a lower composite index value, and consequently, it receives a lower rank. For instance, according to the first index, the district that is awarded the first rank is *Dangs*, located in Gujarat. It has a composite index value of 0.078. This district is characterized by low wages, low agricultural productivity, and a high SC/ST population.

The report concluded that based on the backwardness rank created by the first index, the first 150 districts, "which form a core of under-developed areas" (Planning Commission, 2003, p 14) must receive wage employment programmes. Furthermore, it was decided that self-employment programmes must be taken up in all the districts.

3.2 MGNREGA

The formulation of employment guarantee programmes started in the 1960s. First attempts to identify backward districts were also made during the same time. The Government of India (1961)'s *Third five year plan* states that by the end of the third year plan, the "backlog" of unemployment is likely to increase from 9 million to 12 million. The plan proposes "organizing a rural works program aimed at providing work for an average of about 100 days in the year" (Government of India, 1961, p 392, line 18). Narayanan (2020), explains that an Employment Guarantee Scheme (EGS) was initiated in response to a serious drought in Maharashtra. This offered inspiration for MGNREGA. Social protection programmes in India included workfare programmes such as the Food for Work Programme, the Jawahar Gram Samridhi Yojana (JGSY), the Employment Assurance Scheme (EAS) and Sampoorna Grameen Rojgar Yojana (SGRY) but without the guarantee. This suggests that concurrent workfare programmes were in operation alongside NREGS, indicating that we might be underestimating the impact by only taking NREGS into account.

The distinguishing feature of MNREGA, in contrast to other employment guarantees, is the thorough investigation it has been subjected to and the large scale at which it has been carried out. A commendable review of literature on NREGS's implementation across India was undertaken by Sukhtankar (2017). He proposes a "third law of NREGS" which states "*every result has an equal and opposite result*" while mentioning the complications of investigating NREGS. The heterogeneity in implementation implied that numerous studies indicate varying effects of NREGS on various outcomes (Sukhtankar, 2017, p 232). The state-level and district-level variations in the implementation of NREGS are very well captured by Witsoe (2014) who in his ethnographic study noted,

> "There are, in fact, three distinct NREGAs. The first is the NREGA enacted through legislation, the vision of which is operationalized through a centrally maintained documentary system. The second is the NREGA practiced by a vast bureaucracy under control of state governments, whose main task is the production of documentation within the broad parameters of the centrally maintained architecture. Since this documentation is compiled into data and reports, this is the NREGA most visible to academics. And lastly, there is the NREGA as practiced in villages."

3.3 NREGS Roll-Out and Implementation

NREGS, in phase I was launched in 200 districts across the country. One interpretation of the NREGS implementation is that it was integrated into districts where central-government-sponsored programmes like Rashtriya Sam Vikas Yojana (RSVY), Sampoorna Grameen Rozgar Yojana (SGRY)[10] and the National Food for Work Programme (NFFWP) were already in effect (Planning Commission, 7 08, p 350) & (Klonner and Oldiges, 2022). The backbone for the selection of districts for RSVY, NFFWP, and subsequently NREGS was the Task Force report that identified backward districts.

Out of 200 districts, a subset of the sum of 147 districts where RSVY was functional and 150 districts where NFFWP was operating were chosen to receive NREGS. Another interpretation of NREGS implementation proposed by Klonner and Oldiges (2022) is the 'strict rule' that allotted NREGS to 55 extremist districts with Maoist conflict and 'soft rule' that allotted NREGS to 140 districts based on within-state backwardness. Klonner and Oldiges (2022) note that 195 (out of 200 districts) had a rationale behind the NREGS allotment, and the other 5 districts were allotted at the "discretion of the Planning Commission's chairman"(Klonner and Oldiges, 2022, p 42).

In March 2007, Government of India (2007a) released a press note stating 130 additional districts were allotted NREGS in phase II. The press release states the criteria for choosing the districts - 50 districts were decided from the report on agricultural indebtedness (Government of India, 2007b), 15 districts from areas with a high incidence of farmer suicides, 5 districts that were partitioned from phase 1 districts, 34 districts based on the Planning Commission (2003) task force report, 22 districts for equitable distribution across states and 4 as 'Learning Impact Districts'.[11] In phase III, from 1^{st} April 2008, the government

[10] The aim for SGRY was to be completely subsumed in NREGS by 2008-09 (Planning Commission, 7 08, p 350)

[11] Klonner and Oldiges (2022) notes that these 4 districts were chosen at the discretion of the Planning Commission

has decided to cover all of rural India. Table 1 shows the frequency of phase-wise allocation of NREGS across districts.[12]

Table 1: National Rural Employment Guarantee Scheme Timeline

Time Period	No. of Districts	Phase
Apr, 2006	200	I
Apr, 2007	130	II
Apr, 2008	295	III

3.3.1 Timeline of Estimation

From table 1 we can see that NREGS's phase I was implemented in 2006, phase II in 2007 and phase III in 2008. NSS EUS 64$^{\text{th}}$ round was administered from June 2007 to July 2008. This implies that if we compare phase I with phase II, we would be comparing having NREGS for one year with recently being allotted NREGS. We choose our base specification to be a comparison between phase II and phase III which compares districts where NREGS exists versus districts where NREGS does not exist. Here, the treatment and control are uncorrupted.

3.4 Dataset construction

We leverage the initial phase-wise implementation of NREGS at the district level. Our *independent variable* is created by merging data from Planning Commission (2003)'s task force report on *Identification of Backward Districts for Wage Employment and Self-Employment Programmes* with phase-wise allocation in districts.[13] This allowed us to match the district backwardness rank with the phase in which the district received NREGS.

For our *dependent variable*, the measures of temporary migration are from the Employment and Unemployment Survey (EUS), 64th round[14] (July 2007 - June 2008) carried out by the National Sample Survey Organisation (NSSO). This nationally representative survey contains particulars on migration history at individual and household levels. Information is collected on migrant households, out-migrants, seasonal or circular migrants, and long-term migrants. The survey instrument aimed to track migrants both at source and destination. Information on the number of spells and the industry they are engaged with during the longest spell are present. This survey covered over 125,000 households and is the richest

chairman.

[12] This is depicted using the political map of India in Appendix **??** - map 12

[13] Link to the phase-wise implementation of NREGS in districts or here

[14] NSS-EUS 64th round is thin but because of its large sample size it is comparable to a thick round (Chadha and Nandwani, 2018). It is evident from National Sample Survey Office (2008) that it conducts thick rounds every five years with a large sample size and data being representative at the NSS district level. Thin rounds usually have smaller sample sizes, with the data being representative only at the state level and are conducted in the years between two successive thick rounds.

source of data on all kinds of migrants to date.

All the statistics and estimates in this paper using the NSS data are adjusted using the sampling weights provided with the data. NSSO collects data in four quarters during a year, this implies that the data is representative of the outcomes throughout the year.[15]

NSS-EUS 64$^{\text{th}}$ round defines a migrant as "a household member whose last usual place of residence is different from the present place of enumeration". Our focus is on short-term migrants who move to seek employment. Schedule 10.2 (questionnaire) asks "**whether stayed away from village/town for 1 month or more but less than 6 months during last 365 days for employment or in search of employment**" (National Sample Survey Office, 2008). We consider anyone as a temporary migrant who answers *yes* to this question. Map 1 shows the district-level percentage of temporary migrants.

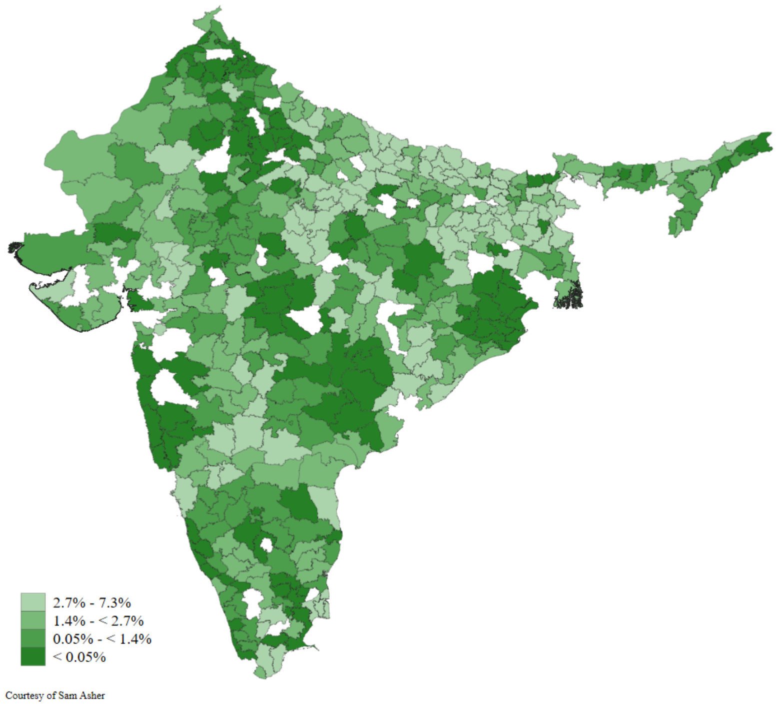

Figure 1: Map of Temporary migration rates in Rural India*Note: White spaces indicate no data.*

The final data was created by merging the districts with backwardness rank from the task force report with the phase-wise allocation of NREGS and then with the NSS 64th round.[16] Data consists of 17 states and 447 districts ranked in the task force report.

The unit of analysis is at the individual level. The sample excludes urban areas and individuals aged

[15] Imbert and Papp (2020b) explain that it is not strictly representative at the district*quarter level as the quarterly samples are not randomly drawn.

[16] The districts are merged based on the names given in Appendix I, a list of FOD sub-regions from the NSS 64th round documentation

below 18 years or above 60 years[17]. Around 3% of the sample reported as temporary migrants who moved for employment or in search of employment. This national-level statistic hides a lot of heterogeneity. We can understand further about temporary migrants and their characteristics from table 2 in comparison with the total sample. We note a dominance of males in short-term migration for work (86%). 55% of short-term migrants are aged between 18-30 years. The percentage of illiterates among short-term migrants is similar to the sample (40%). 85% of those who migrate for a short term own less than a hectare of land. The percentage of Hindus in the sample is similar to the temporary migrant sample (85%). In the sample, 11% belong to Scheduled Tribes whereas 21% belong to the Scheduled Castes, 45% to the OBC group and 23% to others. In the sample of short-term migrants, 20% belong to the Scheduled Tribes, 24% belong to the Scheduled Castes, 40% to the OBC group and 17% to others.

NSS 64$^{\text{th}}$ round records the number of spells[18], the destination and the industry they primarily engaged with during the longest spell. The destination during the longest spell is categorised as - the same district (rural or urban), other districts in the same state (rural or urban), another state (rural or urban), and another country. From table 2, we can see that 69% of short-term migrants go to urban areas 43% migrate to pursue jobs in construction and 24% migrate to pursue jobs in the agricultural sector.

Informed by existing literature on temporary migration and data showing the gendered nature of said migration, we aim to causally estimate the effect of NREGS's existence on young male temporary migrants who are aged less than 35 years as 72% of the sample is comprised of them.

4 Research Design & Empirical Strategy

"La nature ne fait jamais des sauts"

(Leibniz, 1765, p 469)

Translation: "nature does not make jumps"

4.1 Program effect on migration

Estimating migratory flows is challenging for several reasons. Firstly, much has been noted about the bluntness of survey instruments when collecting information on migratory behaviours. There are different kinds of migration yet there exists a huge dearth of data focusing on all kinds. Secondly, most data that captures migratory behaviours only focuses on long-term migrants. This means that their residence is captured at the destination. Migration data providing both source and destination is rare. Thirdly, migration is a complex and nuanced experience that impacts many lives. The data does not capture crucial information about the livelihoods of migrants as noted by Tumbe (2012)[19], Tumbe (2015) and Jha

[17] Most research on migration considered age groups between 18-60 years.

[18] A spell is defined as staying away from a village or town for 15 days or more.

[19] Tumbe (2012)'s *'India Migration Bibliography'* is a rich resource of literature, data sources and bibliometrics on different kinds of migration in India.

Table 2: Descriptive Statistics from NSS 64^{th} Round

	All		**Temporary Migrants**	
	Mean	SD	Mean	SD
Male	0.50	0.50	0.86	0.35
Female	0.50	0.50	0.14	0.35
Age Categories:				
18-25 years	0.27	0.44	0.38	0.48
26-35 years	0.26	0.44	0.34	0.47
36-45 years	0.21	0.41	0.19	0.39
46-55 years	0.17	0.37	0.07	0.25
56-60 years	0.06	0.25	0.01	0.12
Educational Attainment:				
Illiterate	0.42	0.49	0.41	0.49
Primary or below	0.25	0.43	0.32	0.47
Upper Primary or above	0.33	0.47	0.27	0.44
Household Size	5.49	2.55	5.88	2.62
Land Owning:				
$<$ 0.005 hect	0.12	0.32	0.11	0.32
$\geq$ 0.005 to $<$ 1 hect	0.65	0.48	0.74	0.44
$\geq$ 1 hect	0.24	0.43	0.15	0.36
Religion:				
Hindu	0.15	0.35	0.15	0.36
Non-Hindu	0.85	0.35	0.85	0.36
Social Groups:				
ST	0.11	0.31	0.20	0.40
SC	0.21	0.41	0.24	0.43
OBC	0.45	0.50	0.40	0.49
Others	0.23	0.42	0.16	0.36
Migrant Spell			2.25	1.60
Destination during longest spell:				
Same district: rural			0.09	0.28
Same district: urban			0.08	0.27
Same state but another district: rural			0.14	0.34
Same state but another district: urban			0.23	0.42
Another state: rural			0.09	0.28
Another state: urban			0.38	0.49
UPA status at destination:				
Worked in agriculture			0.24	0.43
Worked in manufacturing and mining			0.16	0.37
Worked in construction			0.43	0.50
Worked in other sector (including services)			0.16	0.37
Observations	1,54,117		10,890	

Notes: The sample is composed of rural adults aged 18-60 years of 17 states.

Each statistic is computed using sampling weights.

Source: NSS Employment-Unemployment Survey 64^{th} round

(2023). On the other hand, estimating the impact of NREGS also presents difficulties. The substantial costs to set up the programme and its aim for poverty alleviation imply that policymakers would not allow for a random allocation of NREGS in districts. The constraints are binding but research on this still exists.

Imbert and Papp (2015) estimate the causal effect of NREGS on short-term migration using a Difference-in-Differences strategy. They use National Sample Survey Office (2000)'s 55^{th} round & National Sample Survey Office (2008)'s 64^{th} round that includes migration information. Imbert and Papp (2015) using National Sample Survey Office (2000)'s 55^{th} round defined a short-term migrant as someone who answers no to "whether an individual stayed in the village for the last six months or more" and yes to "whether an individual spent two to six months away from the village for work within the past year" excluding people who have moved into the household over the last six months (who changed their "usual place of residence"). This definition of short-term migration is restrictive and does not match with the short-term migration question with National Sample Survey Office (2008)'s 64^{th} round. This specific selection of demographics only represents individuals who moved to the village less than 6 months ago, and who did not change their place of residence and are looking for work. This does not capture individuals who move for work more than once and this demographic is not representative of the short-term migrant sample. Hence, we use a fuzzy regression discontinuity design.

We obtain causal identification drawing from the NREGS's setup in subsection 3, this subsection elaborates on the identification strategy, its assumptions and empirical specifications. The aim is to compare changes in migration flows of young male migrants from rural areas with and without the program. Regression discontinuity design (RDD) is a form of causal inference that allows a researcher to investigate natural experiments based on cutoffs that occur because of a policy or a rule. These rules or policies are established based on considerations of fairness, ethical implications, practicality, or a combination thereof. This allows some units to be in the treatment group and some in the control group. The defining feature of RDD is the probability of getting treated changed abruptly at the known threshold. When there is perfect compliance with treatment assignment, it is called a Sharp regression discontinuity. When there is imperfect compliance with treatment assignment, it is called a Fuzzy regression discontinuity (FRDD).

We use the Fuzzy Regression Discontinuity Design (FRDD) due to observed discrepancies in NREGS implementation. Specifically, there are instances where districts, despite their backwardness rank, did not receive NREGS as they should have, and some districts were recipients of NREGS contrary to their rank. The expected treatment assignment and the receipt of the program don't completely coincide. This is the fuzziness in the regression discontinuity design.

We specifically draw methodological inputs from Zimmermann (2012). Her paper proposes an algorithm to predict NREGS, this is based on the backwardness index rank and is used as a forcing variable. She proposes that there is a state-wise cutoff where each state includes at least one district that has NREGS in phase II irrespective of the ranking. Then districts are sorted proportional to the prevalence of poverty and then based on their backwardness rank. Zimmermann (2012) & Klonner and Oldiges (2022) base their causal analysis on the Planning Commission reports and so do we.

We develop an algorithm that uses actual NREGS implementation data, combined with the backwardness index, to determine the number of districts per state that should have received NREGS. Using the number of districts that received NREGS in each phase, we rank the districts supposed to receive NREGS by their backwardness rank and match them with the districts that actually received NREGS. Table 3 demonstrates the algorithm's effectiveness in predicting which districts should receive NREGS in each phase. Two significant observations about the algorithm are that its predictions closely align with the findings of Zimmermann (2012)[20], and it exhibits a considerable heterogeneity in its prediction of NREGS allocation among districts for each state.

State	**No. of Districts**	**Actual Phase1**	**Predicted Phase 1**	**Actual Phase2**	**Predicted Phase 2**
Andhra Pradesh	21	13	10	6	5
Assam	23	7	7	6	3
Bihar	36	22	16	14	14
Chhattisgarh	15	11	9	3	3
Gujarat	20	6	4	3	3
Haryana	18	2	0	1	1
Jharkhand	20	18	17	2	2
Karnataka	26	5	4	6	1
Kerala	10	2	1	2	2
Madhya Pradesh	42	18	13	10	9
Maharashtra	30	12	11	6	4
Orissa	30	19	17	5	5
Punjab	15	1	1	2	2
Rajasthan	31	6	5	6	2
Tamil Nadu	26	6	4	4	4
Uttar Pradesh	64	22	19	17	14
West Bengal	17	10	8	7	7
Total	447	180	146	100	81
Success Percentage			81.11%		81%

Table 3: Predicted Success of NREGS Phase-wise Implementation in 17 Indian States

Based on the state-specific backwardness rank, we normalise the prediction of NREGS receipt for easy comparison. In each state, for example, the district that received NREGS first in Phase II is marked as zero, the rest of the districts that are eligible for Phase II are marked with a positive rank and districts that are eligible for Phase I are given a negative rank. It is important to note that the districts that last received NREGS are not hinting at a time period but are just less backward based on the 3-parameter index.

The three components that make an FRDD are - a score, a cutoff, and a treatment. In this case, the

[20]Refer to table 13 from Zimmermann (2012) in the appendix for comparison

score or the running variable is the state-specific rank variable determining treatment (NREGS). The cutoff is set at zero and the treatment is being allotted NREGS. The score affects the probability of allotment of NREGS discontinuously.

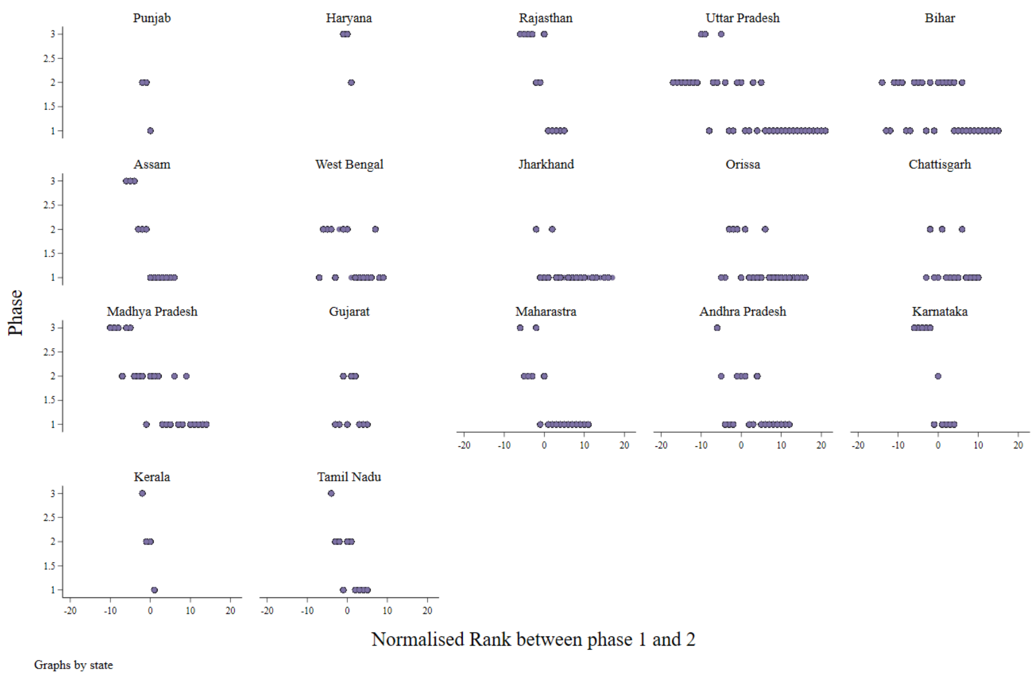

Figure 2: Normalised Rank by state between phase 1 and 2

Figures 2 and 3 display the state-specific normalized rank cutoff on the x-axis and the actual count of observations for each phase on the y-axis. In all 17 states, at least one district received NREGS in phase 2. Both table 3 and figures 2 and 3 convey a consistent narrative - the actual receipt of the program does not entirely align with the intended assignment.

At the heart of FRDD is the problem of imperfect compliance. The validity of the FRDD analysis lies in the running (threshold) variable that is un-manipulable by those who are being benefitted from the treatment. The identification strategy assumes the integrity of the running variable. According to Lee and Lemieux (2010), the integrity of the running variable can be established statistically or institutionally. Institutionally, we establish the timeline for the creation of the backwardness rank and allotment of NREGS. The normalized rank variable is created based on the backwardness Index that was published in 2003. The data that Nayyar Committee used for the task force report is from the 1990s which comprises NSS 50^{th} (1993-1994) and 55^{th} (1999-2000) rounds. We propose that there was no strategic misreporting of information in these NSS rounds foreseeing NREGS implementation in 2006. Hence, institutionally, there exists no systematic manipulation of the running variable. Statistically, this can be established using McCrary density tests. We fail to reject the null hypothesis that discontinuity is zero at the threshold. The density plot 14 is in the Appendix ??.

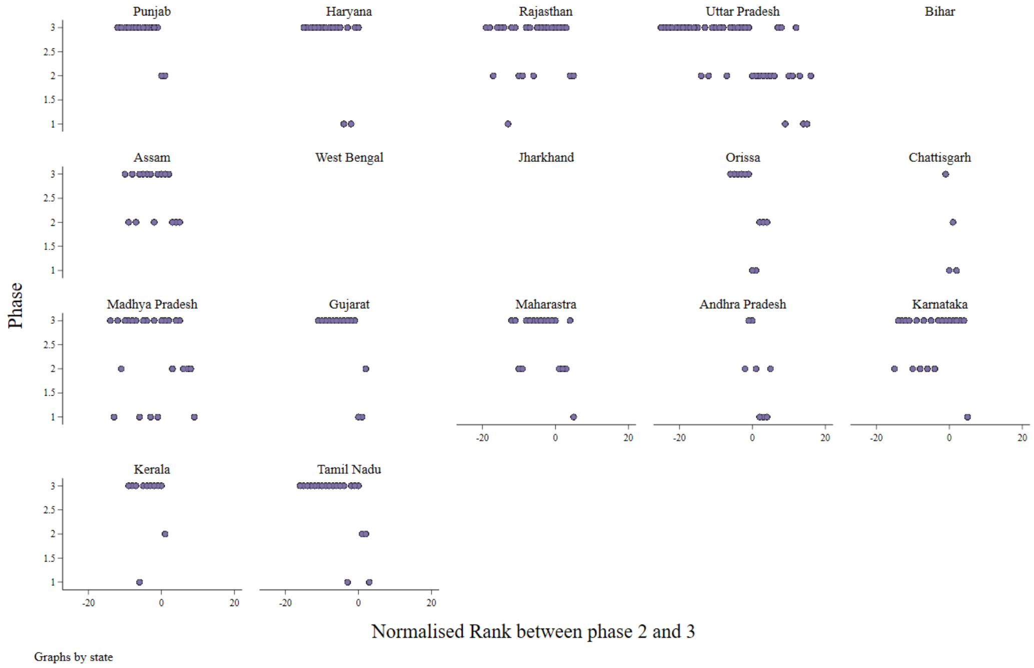

Figure 3: Normalised Rank by state between phase 2 and 3

The scatter plot 4 presents the composite index across the backwardness ranks of 447 districts in 17 states. The plot shows no signs of manipulation, and it's evident that the composite index increases in line with the parameters derived from the data. This indicates that the running variable (the composite index) is smooth, continuous, and increases monotonically. This implies that the influence of all other factors, apart from the receipt of NREGS, remains constant on either side of the discontinuity threshold. In other words, all other factors that could potentially affect migration change smoothly in relation to the normalized rank. This assumption is crucial for the validity of the regression discontinuity design used in this analysis. We also plot figure 15 to show the observations at each state-specific rank for Phases 2 and 3 (in the appendix **??**). It reveals that all 17 states have ample observations (short-term migrants) closer to the cutoff and the distribution is not skewed.

4.2 Anatomy of RDD

Fuzzy regression discontinuity is employed when the cutoff doesn't perfectly determine treatment but creates a likelihood of receiving the treatment. We are trying to estimate the effect of NREGS phase-wise implementation on short-term migration in Phase II versus Phase III.

The idea behind a classic regression discontinuity design is to compare the units on either side of the cutoff. The assumption is without the cutoff, the units would have similar fits because they are of similar characteristics. It is safe to assume that farther from the cutoff the characteristics of the districts would be different. The causal effect would be the jump in the relationship between the outcome variable and

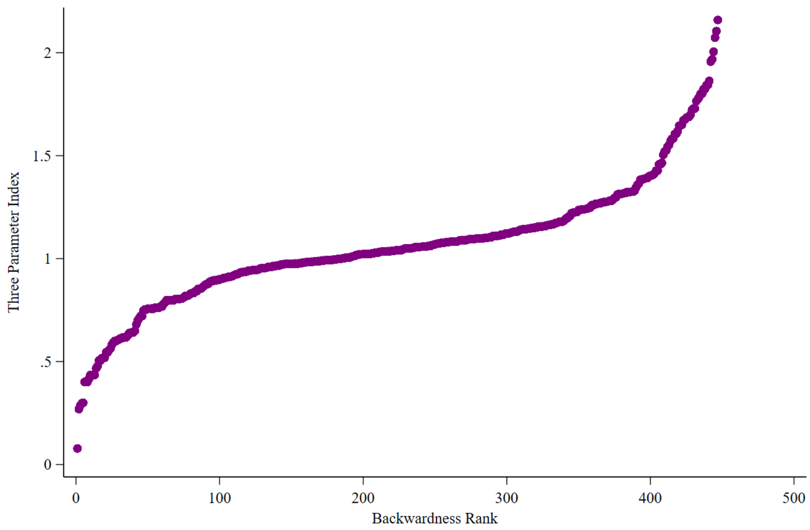

Figure 4: General Distribution of Index over Ranks

the running variable divided by the jump in the relationship between the treatment variable and the running variable. We base our analysis on the backwardness rank (our running variable) - this implies that the causality will be established for 'intention to treat'. This means the causal estimates are for the intended treatment assignment - NREGS, based on the backwardness index. This approach preserves the random assignment of the original groups, which is crucial for ensuring that the groups are comparable while taking into account the deviation from the rule.

To solve the problem of non-compliance, an Instrumental Variables Approach (IV) is used. We use predicted treatment status based on the proposed algorithm, indicating which side of the cut-off a district falls on to get NREGS as an instrument. This instrument only affects temporary migration through NREGS's existence. The exclusion restriction holds because there is no other programme that uses phase-wise allocation similar to NREGS. Here, which side of the cutoff the district falls on affects short-term migration only through an increase in the probability of being allotted NREGS. A fuzzy regression discontinuity design can be viewed as an instrumental variable approach used near a threshold. The advantage of RDD over other causal mechanisms is its 'visual transparency'. We make two important decisions - on bin length and polynomial to ensure the transparency translates.

4.2.1 Regarding optimal Bin length

We create plots by selecting a bin length on either side of the cutoff. A bandwidth (or bin length) - h is chosen to avoid having treated and untreated observations together. After selecting a bin length, the average value of the dependent variable is calculated for each bin and graphed against the mid-points of the bins, then we calculate the mean values of the outcome within bins.[21] By comparing outcomes on

[21] We select the bandwidth using *rdplot* command that is a companion command of *rdrobust* and *rdbwselect*. The number of bins are calculated to compute local sample means so that the integrated mean square error (IMSE) of the is minimized at a given threshold.

either side of a cutoff, we can infer whether the treatment has the expected effect. This makes RDD a potent tool for causal inference.

A legitimate concern is if we have fixed h but our data is sparsely populated, then it might lead to biased estimates. Calonico et al. (2015), studies different kinds of regression discontinuity (RDD) plots and proposes an alternative binning method - quantile spaced (QS) binning from the usual evenly spaced (ES) binning of the data. The former method forces all the bins to have approximately the same number of observations. In our own explorations of RDD plots, We try varying bin lengths and deduce that a lesser bin length does not fit the points around the cutoff and overestimates the jump. As we increase the bin length, the slope rises underestimating the jump.

According to the methods proposed by Calonico et al. (2015), we choose an optimal bandwidth of 10 on the right of the threshold and 11 on the left of the threshold. and for robustness, estimate the same reduced form with $h/2$ and $2h$ to check if the direction of the coefficients is changing. We note that the contention is between bias and variance. If we pick a larger bandwidth, the estimates are more precise but bias also increases. We choose to include household and individual-level controls to improve the efficiency of the estimates.

4.2.2 Regarding choice of polynomial and kernels

We fit a polynomial to capture the behaviour of underlying conditional expectation functions (Calonico et al., 2015, p 1754). We select a local linear polynomial (of degree 1) because it captures, by design, any potential discontinuities at a given cutoff and provides a "disciplined representation of the overall variability of the data" (Calonico et al., 2015, p 1755). Polynomials are sensitive to the points near the cutoff. Using a global polynomial leads to the problem of over-fitting. We fit a local linear polynomial in our reduced-form specification as it is least sensitive to the changes near the cutoff.

We choose uniform (or rectangular kernels) for our reduced form because (Lee and Lemieux, 2010, p 319) writes, "choice of kernel typically has little impact in practice". One can think of a regression discontinuity as a specific form of weighted local linear regression. Usage of kernels - uniform, triangular or epanechnikov determines how much weight to give observations near the threshold. Uniform kernel weights give the same weight to all the observations. Triangular kernels weigh distant observations less but in a linear fashion. Epanechnikov weighs far away observations less in the form of a curve.[22] We note that a change in kernel weights from uniform to triangular does not change the coefficients significantly.

There is a general agreement in the discipline about the importance of regression discontinuity as a method, but its formal properties are not consolidated and remain unknown (Cattaneo et al., 2019) and (Lee and Lemieux, 2010). This could be one of the reasons, the choice of bin length, and fit of the polynomial makes the application quite tedious and subjective. Hence, the plausibility of the researcher's

[22] This is graphically depicted here.

design is questioned because it can be easily misrepresented.

4.2.3 Regarding other choices

We believe that the variance of the error term varies at the district level because treatment varies by district. Hence we cluster our standard errors at this level. This implies that our reduced form specification allows for within-district correlation. We employ state-level fixed effects to control for time-invariant factors such as the implementation of public works, rainfall...etc.

4.3 Empirical Specification

Under the assumption of continuity of all other migrant and district characteristics other than NREGS treatment at the treatment threshold, the fuzzy RD estimator calculates the local average treatment effect (LATE) of the district receiving NREGS if the district falls below the threshold. Following the recommendations of Lee and Lemieux (2010) and Imbens and Lemieux (2008), our primary specification uses local linear regression within a given bandwidth of the treatment threshold, and controls for the running variable on either side of the threshold. We use the following two-stage instrumental variables specification:

$$nregs_{ds} = \alpha_1 + \phi Z_{ds} + \beta I_{ids} + +\delta_1 nrank_{ds} + \delta_2 nrank_{ds} * Z_{ds} + \eta_s + \epsilon_{ids} \quad (1)$$

$$Y_{ids} = \alpha_2 + \lambda_{2SLS}\widehat{nregs}_{ds} + \gamma_1 I_{ids} + \gamma_2 nrank_{ds} + \eta_s + \epsilon_{ids} \quad (2)$$

In specifications 1 and 2, the subscripts refer to an observation in each individual i, in district d and state s. Equation 1 represents the first stage regression. Here, $nregs_{ds}$ is the binary treatment variable that takes 1 if the district gets NREGS and 0 otherwise. Z_{ds} is the binary instrument that determines which side of the cut-off a district falls on to get NREGS. $nrank_{ds}$ is a district's rank based on the state-specific normalized rank. I_{ids} is a vector of individual-level controls such as age group, education levels, religion, gender, marital status, land ownership, and household composition. η_s are state-fixed effects.

Equation 2 represents the second stage regression. Our outcome variable of interest is Y_{ids} equals 1 if the individual is a male temporary migrant under 35 years of age. Conversely, it takes the value 0 if the individual is a male under 35 years old but not a temporary migrant. $\widehat{nregs}_{ds}$ is the coefficient of the first stage regression. The coefficient λ_{2SLS} is the causal effect in this two-stage least-squares regression. I_{dt} and η_d are the same as equation 1.

5 Results

5.1 Main Results

Here we present all results using the preferred specifications 1 and 2. The base specification on which we build is the effect of NREGS's existence of male temporary migrants (<35 yrs). The main result of this paper - table 4 suggests that the existence of NREGS (phase 2 versus phase 3) causes a statistically significant (at 95% CI) 11.2 percentage point increase in the migration of young males for work or in search of work. The fuzzy regression discontinuity plot 5 reflects the same. Each point on the plot represents a bin, on the x-axis is the normalised rank and on the y-axis is the average computed within each bin. The interaction term from our specification allows linear fits with different slopes on either side of the threshold.

Table 4: Main Results -Dependent Variable: Male Temporary Migrants (<35 yrs)

	OLS	FRDD
	(1)	(2)
NREGS (in phase 2)	0.0223***	0.112**
	(0.00791)	(0.0471)
Observations	18019	10478
Controls	Yes	Yes
State Fixed Effects	Yes	Yes

Standard Errors are clustered at the district level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Independent variable takes 1 if NREGS is implemented (phase 2) and 0 if NREGS is not yet implemented (phase 3).

IV was implemented with an instrument that determines which side of the cut-off a district falls on to get NREGS. The specification was implemented with a local linear polynomial with cutoff zero and uniformly distributed kernel weights

The drop in observations from a simple OLS to an FRDD specification is because of the choice of bin length. In table 5, when the same specification was run for the effect of recently being allotted NREGS (phase 2) and having NREGS for a year (phase 1), we see no statistical significance yet the direction of the coefficient is positive in both the OLS and FRDD columns. We note a 1.7 percentage point increase in temporary migration of males in districts where NREGS was operational for one year. The magnitude of the coefficients shows economic insignificance. This is unsurprising because the time difference between availing NREGS in phase 1 versus 2 is 1 year. We understand that the introduction of NREGS would have taken longer to impact a frequent phenomenon such as migration which is affected by many personal and structural factors. This is reflected in the figure 6.

Employing the same specification for young female short-term migrants does not result in a statistically significant change in their migrant flows because of NREGS's existence in the district. In districts where NREGS exists, we note a 1.6 percentage point increase in the temporary migration of young females. The magnitude of the coefficient hints at economic insignificance. This is reflected in the figure 7.

When the same specification is administered for the whole short-term migrant population for all age

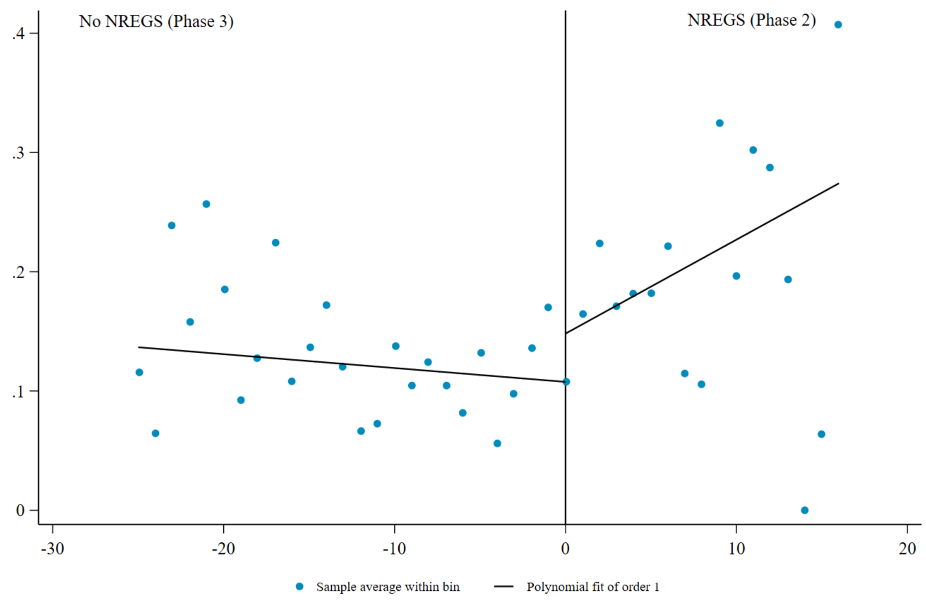

Figure 5: Reduced form: Effect of having NREGS on male short-term migrants aged 18-34 yrs with a linear polynomial

Table 5: Effect of NREGS (phase 1 vs phase 2) on Male Temporary Migrants (<35 yrs)

	OLS	FRDD
	(1)	(2)
NREGS for 1 yr (in phase 1)	0.0127	0.0176
	(0.0111)	(0.0474)
Observations	20437	16104
Controls	Yes	Yes
State Fixed Effects	Yes	Yes

Standard Errors are clustered at the district level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Independent variable takes 1 if NREGS is implemented for the last 1 year and 0 if NREGS is implemented recently.

IV was implemented with an instrument that determines which side of the cut-off a district falls on to get NREGS. The specification was implemented with a local linear polynomial with cutoff zero and uniformly distributed kernel weights

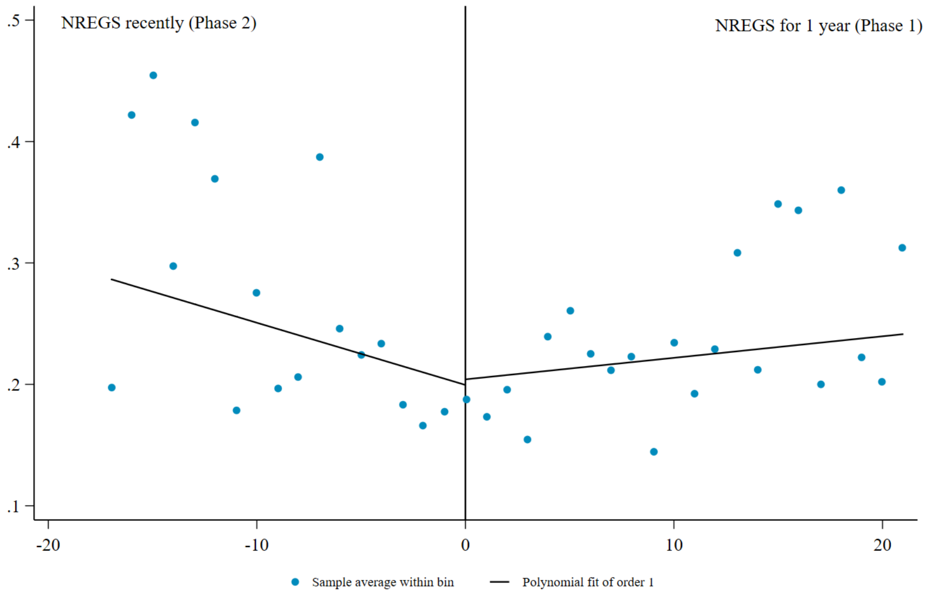

Figure 6: Reduced form: Effect of having NREGS recently vs for one year on male short-term migrants aged 18-34 yrs with a linear polynomial

Table 6: Effect of having NREGS on Female Temporary Migrants (<35 yrs)

	OLS	FRDD
	(1)	(2)
NREGS (in phase 2)	0.000459	0.0164
	(0.00167)	(0.00990)
Observations	18468	6056
Controls	Yes	Yes
State Fixed Effects	Yes	Yes

Standard Errors are clustered at the district level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Independent variable takes 1 if NREGS is implemented (phase 2) and 0 if NREGS is not yet implemented (phase 3).

IV was implemented with an instrument that determines which side of the cut-off a district falls on to get NREGS. The specification was implemented with a local linear polynomial with cutoff zero and uniformly distributed kernel weights.

Controls at individual-level are marital status, household size, land owning, education, social group, and labor force participation.

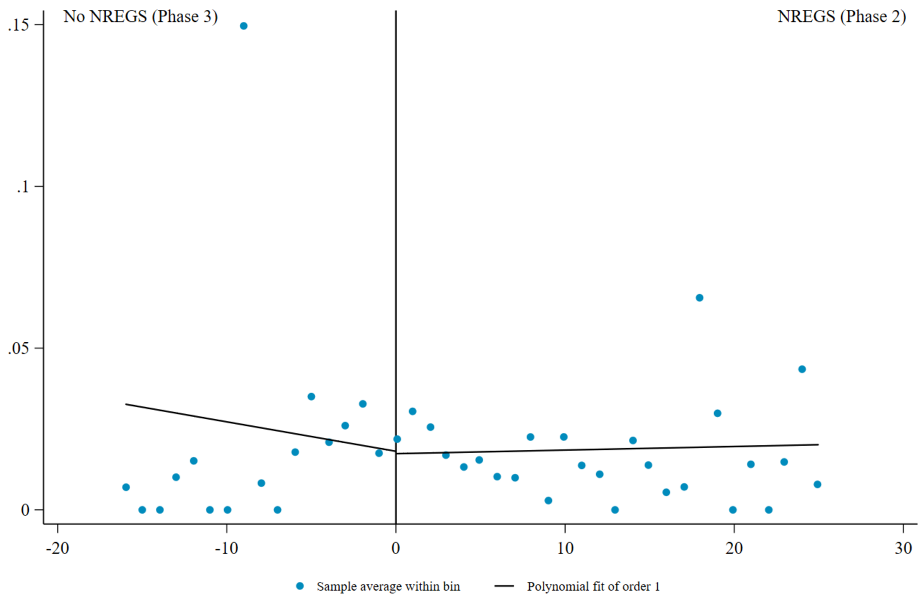

Figure 7: Reduced form: Effect of having NREGS on young female temporary migrants with a linear polynomial

groups we find a positive causal effect. In table 7, we note that the OLS coefficient represents a very small change in temporary migration when NREGS exists in a district. The FRDD specification shows a statistically significant (at 99% CI) 4.5 percentage point increase in short-term migration in districts where NREGS exists. This is reflected in the figure 8.

Table 7: Dependent Variable: All Temporary Migrants (<35 yrs)

	OLS	FRDD
	(1)	(2)
NREGS (in phase 2)	0.00828**	0.0459***
	(0.00325)	(0.0158)
Observations	88564	46862
Controls	Yes	Yes
State Fixed Effects	Yes	Yes

Standard Errors are clustered at the district level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Independent variable takes 1 if NREGS is implemented (phase 2) and 0 if NREGS is not yet implemented (phase 3).

IV was implemented with an instrument that determines which side of the cut-off a district falls on to get NREGS. The specification was implemented with a local linear polynomial with cutoff zero and uniformly distributed kernel weights

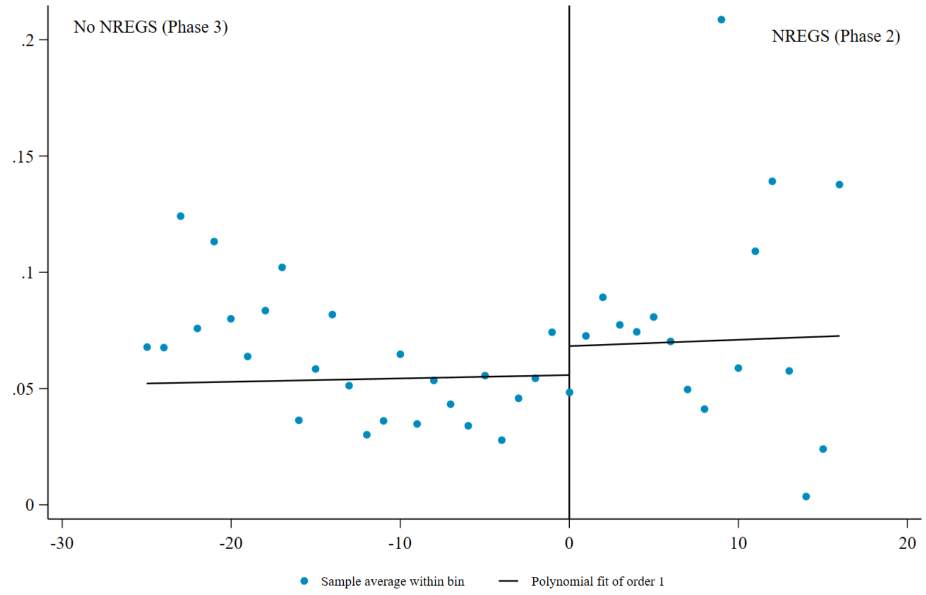

Figure 8: Reduced form: Effect of having NREGS on all short-term migrants with a linear polynomial

To check the sensitivity of the results, we employ the same model with varying bin lengths as suggested by Calonico et al. (2015). In table 8, we see that the direction of the coefficient remains unchanged with varying bin lengths and the magnitude is large enough indicating economic significance.

Table 8: Varying bandwidthsDependent Variable: Male Temporary Migrants (<35 yrs)

	h/2=5.5	h=11	2h=22
	(1)	(2)	(3)
NREGS	0.0912	0.112**	0.120**
(in phase 2)	(0.0678)	(0.0471)	(0.0590)
Observations	6373	10478	13518
Controls	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes

Standard Errors are clustered at the district level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: h denotes bin length or bandwidth

Independent variable takes 1 if NREGS is implemented (phase 2) and 0 if NREGS is not yet implemented (phase 3).

IV was implemented with an instrument that determines which side of the cut-off a district falls on to get NREGS. The specification was implemented with a local linear polynomial with cutoff zero and uniformly distributed kernel weights.

Controls at individual-level are marital status, household size, land owning, education, social group, and labor force participation.

As a second robustness check, we remove the fuzziness (19% noted in table 3) by matching the districts that the algorithm predicts to have received NREGS in phase 2 with the districts that actually received NREGS in phase 2. This omission of fuzziness turns the design into a sharp regression discontinuity. We note that the coefficient is statistically insignificant but the direction is positive. This is reflected in the figure 9.

Table 9: Sharp RDDDependent Variable: Male Temporary Migrants (<35 yrs)

	OLS	RDD
	(1)	(2)
NREGS (in phase 2)	0.0451***	0.0149
	(0.0129)	(0.0254)
Observations	12818	13396
Controls	Yes	Yes
State Fixed Effects	Yes	Yes

Standard Errors are clustered at the district level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Independent variable takes 1 if NREGS is implemented (phase 2) and 0 if NREGS is not yet implemented (phase 3).

The specification was implemented with a local linear polynomial with cutoff zero and uniformly distributed kernel weights.

Controls at individual-level are marital status, household size, land owning, education, social group, and labor force

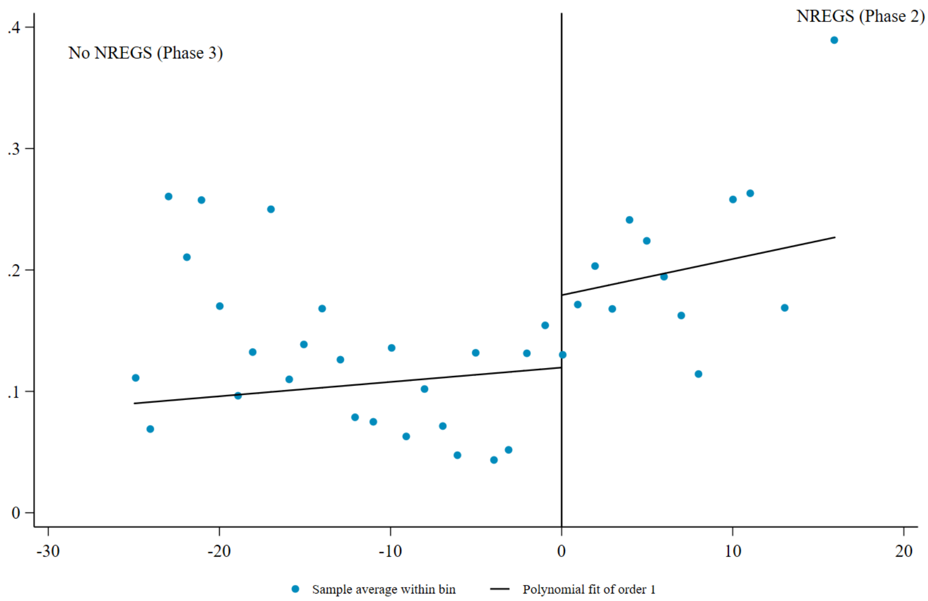

Figure 9: Sharp RDD | Effect of having NREGS on young male short-term migrants with a linear polynomial

As a third robustness check, we normalise the three-parameter index as our running variable. We see a 9.1 percentage point increase in young male temporary migrants in districts where NREGS exists. We see this represented in the figure 10. The plot is difficult to interpret when compared to other regression discontinuity plots because there are districts that share the same index value and the optimal binlength that was calculated by Calonico et al. (2015) was less than the other specifications we ran.

Table 10: Using Three Parameter IndexDependent Variable: Male Temporary Migrants (<35 yrs)

	FRDD
	(1)
NREGS (in phase 2)	0.0912
	(0.0678)
Observations	6373
Controls	Yes
State Fixed Effects	Yes

Standard Errors are clustered at the district level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Independent variable takes 1 if NREGS is implemented (phase 2) and 0 if NREGS is not yet implemented (phase 3).

IV was implemented with an instrument that determines which side of the cut-off a district falls on to get NREGS. The specification was implemented with a local linear polynomial with cutoff zero and uniformly distributed kernel weights.

Controls at individual-level are marital status, household size, land owning, education, social group, and labor force

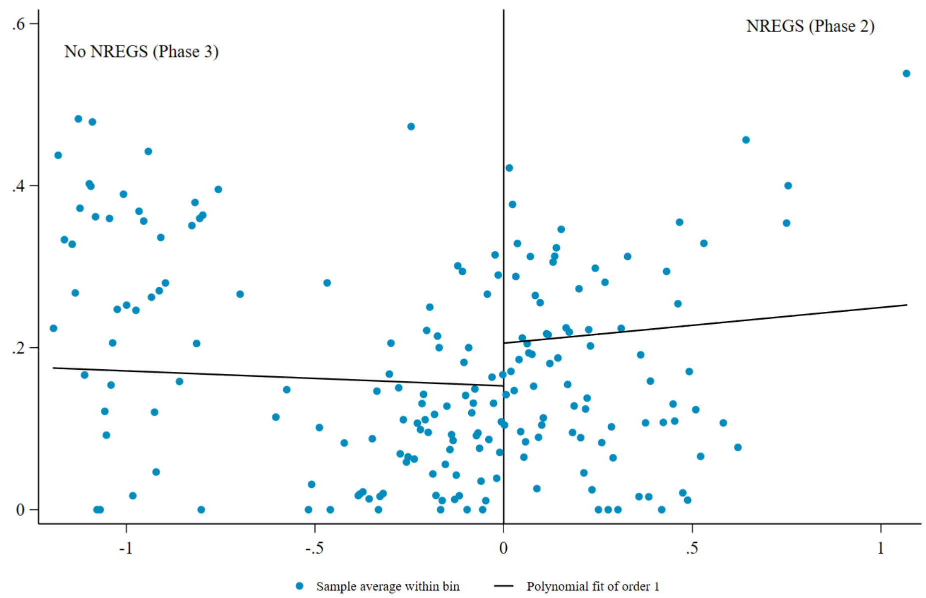

Figure 10: Normalised Index | Effect of having NREGS on young male short-term migrants with a linear polynomial

5.2 Discussion

From the results, it is clear that we note a positive causal effect of NREGS on young male temporary migrants. The effect on young female temporary migrants and the effect of being allotted NREGS in phase 1 versus phase 2 is positive but not statistically significant. We note that the direction of the effect is retained when we implement a sharp RDD and use the normalised index value instead of the rank.

The effect seems dramatic compared to research of Morten (2019), Imbert and Papp (2020b) and Coffey et al. (2015) which is adjacent to ours. Much of the research that is done on the effect of NREGS on temporary migration finds a negative effect. They conclude, that NREGS fulfils its aim of reducing migration. However, the scope of the research is limited based on the geographical unit they chose - Morten (2019) draws data from 6 ICRISAT villages over time. She finds that NREGS reduces migration by 17 percentage points. Imbert and Papp (2020b) using the same data as ours, classify states into star and non-star and compare migrant flows in 'early districts' - based on NREGS implementation. They find that NREGS increased rural wages and reduced short-term migration to urban areas. Their focus in this paper and in Imbert and Papp (2015) was to explicate the effect of NREGS on urban areas, they find that NREGS affects urban wages through migration.

Coffey et al. (2015) uses data from a novel survey conducted during the agricultural lean season (summer) in a historically high migration region overlapping the borders of Rajasthan, Gujarat, and Madhya Pradesh. They find that migration choice varies by season, especially in places where the costs of migration are low. They find that there is an increase in the probability of migration for adults aged between 15 to 30 years. However, as age increases the number of trips falls. They also find that migration is 'negatively selective' based on the education and economic status of the migrants and their households.

Singh (2013), according to focus group discussions conducted notes that seasonal migration is very high in Dokur village of Mahbubnagar district. Most migration took place in search of better employment opportunities. She writes,

> "Wages under NREGA are Rs.120 whereas if they go to work on construction site in Hyderabad they get Rs.200 for female workers and Rs.300 for male workers. It was found that NREGA only provides seasonal employment and therefore for rest of the season people migrate. Thus it can be said that NREGA is not helping to reduce migration in this village. It was found that only female and old males work under NREGA and young workers migrate to earn good wages." (Singh, 2013, p 25, line 13)

Drawing from the literature, two prospective mechanisms emerge - NREGS is bearing the cost of job search among young males. As the agricultural workforce at the origin is increasingly feminized, men are migrating for work or in search of work. Our result is corroborated by the "missing men" puzzle Tumbe (2015) puts forth. Our second prospective mechanism is insurance - having NREGS has increased livelihood diversification strategies by increasing rural wages. Now, young men are moving because the fallback option exists (which is NREGS). Having NREGS made migration a less risky proposition (Morten, 2019) & (Bryan et al., 2014)

6 Conclusion

Migration is a complex and nuanced phenomenon. Our investigation yielded this conclusion that aligns with prior research on migration. We are confident that these findings will continue to be corroborated by future studies. There is abundant work stating the conclusions on migration depend on the variation in NREGS's implementation that a researcher is investigating.

This research can be taken forward in at least two ways. Firstly, a fuller analysis would examine the role of time when the survey was administered to make the results robust. Including other factors that affect the NREGS's implementation would result in finer evidence of the impact of NREGS's relative existence. Secondly, examining how district and state borders influence the migratory flows would unearth further nuances on migrant spells.

All current studies on migration highlight the constraints posed by data limitations in estimating migratory movements. Foster and Rosenzweig (2007) explains that these limitations arise due to the practical challenges in locating migrants, the complexities migration introduces when defining sampling frames and households, the tendency of survey respondents to forget or ignore short trips for work, and most questionnaires often do not inquire about these short trips. There could be many unobservables that we cannot account for, because of data restrictions, the complexity of our outcome variable, or the discretion at which some districts were allotted NREGS.

Our estimates admittedly do not capture every dimension of temporary migration in young males but offer results that explain the complex nature of migration. Understanding of migration moved further away from being a concept that 'pushes' people out of poverty or 'pulls' them towards better opportunities. Our attempt to understand the effect of NREGS on temporary migration is not a novel one, but this exploration does illuminate pathways for further investigation, underscoring the scarcity of available data. This attempt, if anything, captures the nuanced and complex nature of migration and reinstates the need for data on migrants and their remittances at both their origin and destination to say anything conclusive.

7 Appendix

7.1 Background and Data

7.1.1 NREGS Roll-Out and Implementation

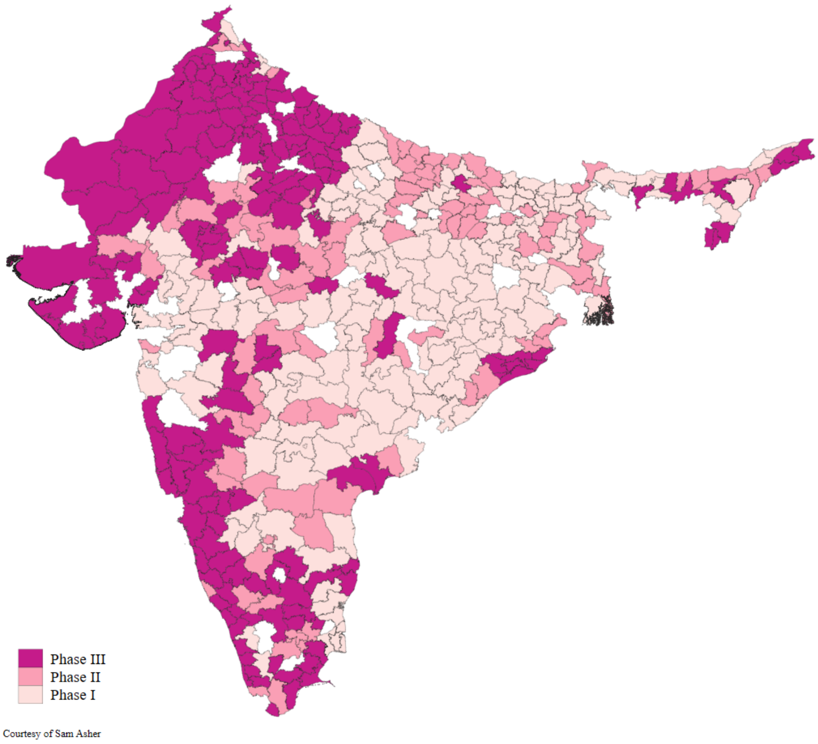

Figure 11: Phase-wise allocation of NREGS in rural districts

Figure 12: *

Note: White spaces indicate no data.

Back to subsection 3.4

7.2 Methodology and Specifications

Table 1: Predictive Success of Algorithm for Major Indian States

		actual NREGS		prediction success rate	
	N	Phase 1	Phase 2	Phase 1	Phase 2
Andhra Pradesh	21	13	6	0.90	0.75
Assam	23	7	6	0.91	0.75
Bihar	36	22	14	0.81	1.00
Chhattisgarh	15	11	3	0.73	1.00
Gujarat	20	6	3	0.80	0.93
Haryana	18	2	1	0.72	0.94
Jharkhand	20	18	2	0.85	1.00
Karnataka	26	5	6	0.88	0.52
Kerala	10	2	2	0.77	1.00
Madhya Pradesh	42	18	10	0.76	0.88
Maharashtra	30	12	6	0.93	0.56
Orissa	30	19	5	0.73	0.91
Punjab	15	1	2	1.00	0.93
Rajasthan	31	6	6	0.90	0.72
Tamil Nadu	26	6	4	0.88	0.95
Uttar Pradesh	64	22	17	0.88	0.79
West Bengal	17	10	7	0.76	1.00
Total	447	180	100	0.84	0.82

Note: Table includes all districts with non-missing development index value for 17 major Indian states (the only missing districts in these states are urban districts according to the Planning Commission report definition from 2003 and therefore include either the state capital or an urban agglomeration of at least one million people). Column 1 provides the number of non-missing index districts in each state. Columns 2 and 3 give the actual number of treatment districts per state in a given phase of NREGS rollout. Columns 4 and 5 give the success rate of the algorithm in predicting a district's treatment status (NREGS or no NREGS) in a given phase according to the two-step algorithm explained in the text. The number of districts treated in Phase 3 is the difference between the number of districts in a state (N) and the sum of the districts in a state actually treated in Phase 1 and Phase 2.

Figure 13: Predictive Success of Algorithmfor Major Indian States Zimmermann (2012)

Back to table 3

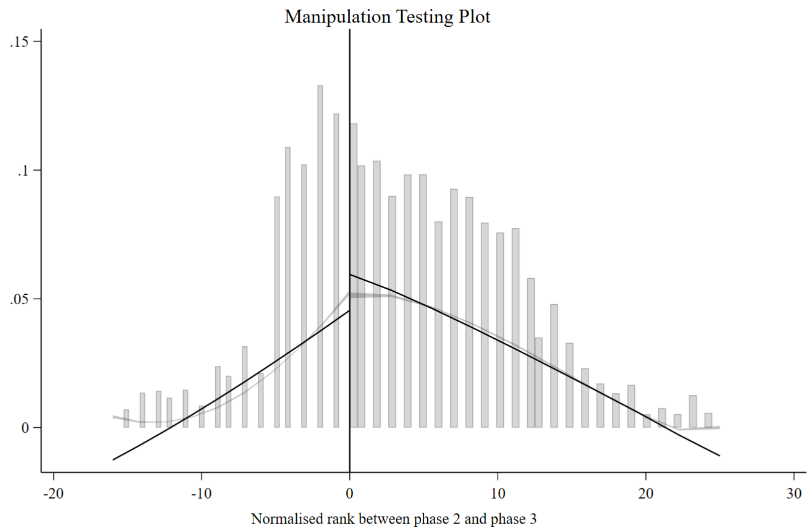

Figure 14: Mccrary Density test for Normalised Rank between phase 2 and 3

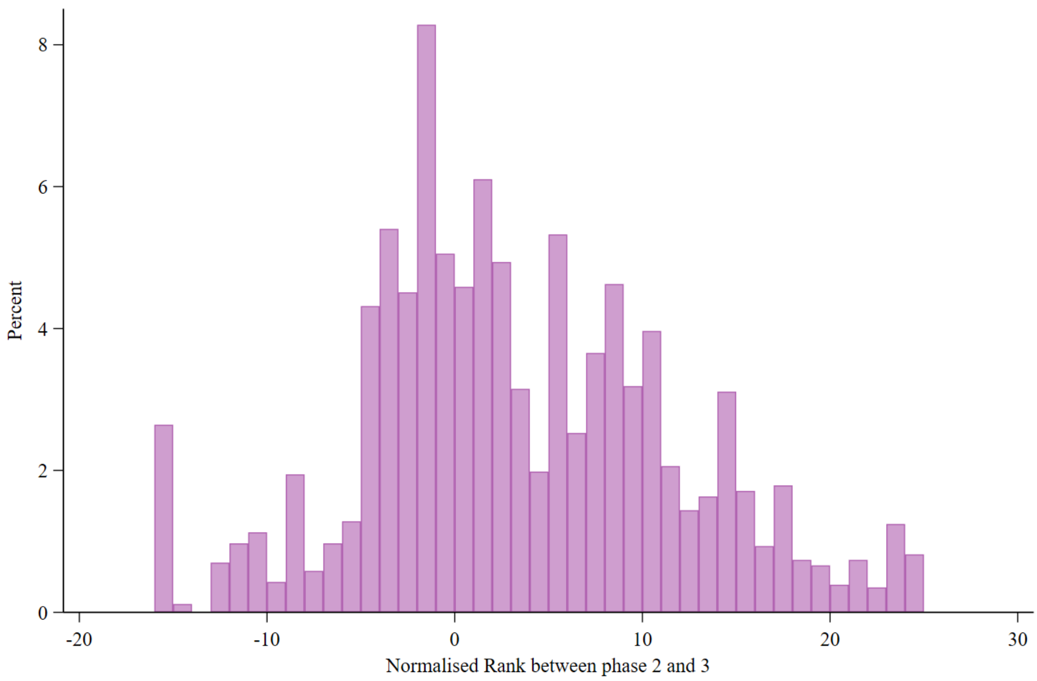

Figure 15: Distribution per state-specific rankbetween Phase 2 and 3

Back to graph 3

References

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

Calonico, Cattaneo, and Titiunik (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.

Cattaneo, Idrobo, and Titiunik (2019). *A Practical Introduction to Regression Discontinuity Designs: Foundations*.

Chadha, N. and Nandwani, B. (2018). Ethnic fragmentation, public good provision and inequality in india, 1988–2012. *Oxford Development Studies*, 46(3):363–377.

Coffey, Papp, and Spears (2015). Short-term labor migration from rural north india: Evidence from new survey data. *Population Research and Policy Review*, 34(3):361–380.

de Haan, A. and Rogaly, B. (2002). Introduction: Migrant workers and their role in rural change. *Journal of Development Studies*, 38(5):1–14.

Deshingkar, P. and Farrington, J. (2009). A framework of understanding circular migration. *Circular Migration and Multinational Livelihood Strategies in Rural India*, pages 1–36.

Deshingkar, P. and Start, D. (2003). *Seasonal migration for livelihoods in India: Coping, accumulation and exclusion*, volume 111. Overseas Development Institute London.

Dodd, W., Wyngaarden, S., Humphries, S., Patel, K., Majowicz, S., Little, M., and Dewey, C. (2018). The Relationship Between MGNREGA and Internal Labour Migration in Tamil Nadu, India. *The European Journal of Development Research*, 30(2):178–194.

Fagerlid, C. and Tisdel, M. A. (2020). *Introduction: Literary Anthropology, Migration, and Belonging*, pages 1–17. Springer International Publishing, Cham.

Foster, A. and Rosenzweig, M. (2007). Chapter 47 economic development and the decline of agricultural employment. volume 4 of *Handbook of Development Economics*, pages 3051–3083. Elsevier.

Government of India, G. o. I. (1961). India's third five year plan.

Government of India, G. o. I. (2007a). Press release - 130 additional districts included under nrega.

Government of India, G. o. I. (2007b). Report of the expert group on agricultural indebtedness.

Haberfeld, Y., Menaria, R., Sahoo, B. B., and Vyas, R. (1999). Seasonal migration of rural labor in india. *Population Research and Policy Review*, 18:471–487.

Hagen-Zanker, J. and Leon-Himmelstine, C. (2013). What do we know about the impact of social protection programmes on the decision to migrate? *Migration and Development*, 2:117.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, 60(1):126–142.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635. The regression discontinuity design: Theory and applications.

Imbert and Papp (2015). Labor market effects of social programs: Evidence from india's employment guarantee. *American Economic Journal: Applied Economics*, pages 233–263.

Imbert and Papp (2020a). Costs and benefits of rural-urban migration: Evidence from India. *Journal of*

Development Economics, 146:102473.

Imbert and Papp (2020b). Short-term migration, rural public works, and urban labor markets: Evidence from India. *Journal of the European Economic Association*, 18(2):927–963.

Jacob, N. (2013). The Impact of NREGA on Rural-Urban Migration: Field survey of Villupuram District, Tamil Nadu. Working papers, eSocialSciences.

Jha, A. (2023). What does mis tell us about problems faced by migrants? https://www.hindustantimes.com/india-news/what-does-mis-tell-us-about-problems-faced-by-migrants-101678711884749.html. [Accessed 06-04-2024].

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

Keshri, K. and Bhagat, R. (2013). Socioeconomic determinants of temporary migration in india: A regional analysis. *Asian Population Studies*, 9:175–195.

Klonner, S. and Oldiges, C. (2022). The welfare effects of India's rural employment guarantee. *Journal of Development Economics*, 157:102848.

Korra, V. (2011). Labour migration in mahabubnagar: Nature and characteristics. *Economic and Political Weekly*, 46(2):67–70.

Lal (1989). The hindu equilibrium. vol. ii: Aspects of indian labour. *Oxford: Clarendon Press*.

Lee, D. S. and Lemieux, T. (2010). Regression Discontinuity Designs in Economics -. *American Economic Association*.

Leibniz, G. (1765). Nouveaux essais sur l'entendement humain. page 469.

Mazumdar, Neetha, and Agnihotri (2013). Migration and gender in india. *Economic and Political Weekly*, 48(10):54–64.

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

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(01):46–98.

Narayanan, S. (2020). The Continuing Relevance of MGNREGA. https://www.theindiaforum.in/article/continuing-relevance-mgnrega#. [Accessed 31-03-2024].

National Federation of Indian Women, M. (2008). Mahatma gandhi national rural employment guarantee scheme (mgnregs): A study of socio-economic empowerment of women in andhra pradesh. *Ministry of Rural Development*.

National Sample Survey Office, N. (2000). National Sample Survey 1999-2000 (55th round) - Ministry of Statistics and Programme Implementation (MOSPI).

National Sample Survey Office, N. (2008). National Sample Survey 2007-2008 (64th round) - Schedule 10.2 - Employment, Unemployment and Migration Particulars, Ministry of Statistics and Programme Implementation (MOSPI).

Planning Commission, G. o. I. (2001). Report of the working group on rural poverty alleviation programmes. Technical report, Planning Commission, Government of India.

Planning Commission, G. o. I. (2003). Report of the task force: Identification of districts for wage and self employment program. Technical report, Planning Commission, Government of India.

Planning Commission, G. o. I. (2007-08). Annual report. Technical report, Planning Commission, Government of India.

Planning Commission, G. o. I. (2009). Annual report. Technical report, Planning Commission, Government of India.

Rajan, S. I. (2020). Covid-19-led migrant crisis a critique of policies. *Economic and Political Weekly*, pages 13–16.

Rao, U. (1994). Palamoor labour: A study of migrant labour in mahabubnagar district. *Hyderabad: CD Deshmukh Impact Centre, Council for Social Development.*

Shah, A. (2006). The labour of love: Seasonal migration from jharkhand to the brick kilns of other states in india. *Contributions to Indian Sociology*, 40.

Singh, N. (2013). Impact of mgnrega on migration and asset creation. *Report submitted to International Crop Research Institute for the semi-Arid Tropics*, pages 1–36.

Solinski, T. (2012). Nrega and labour migration in india: Is village life what the 'rural' poor want? *The South Asianist Journal*, 1(1).

Stark, O. and Bloom, D. E. (1985). The new economics of labor migration. *The American Economic Review*, 75(2):173–178.

Sukhtankar, S. (2017). India's national rural employment guarantee scheme: What do we really know about the world's largest workfare program? In *Brookings-NCAER India Policy Forum*, volume 13, pages 231–286.

Tumbe, C. (2012). India migration bibliography. *Available at SSRN 2117805.*

Tumbe, C. (2015). Missing men, migration and labour markets: evidence from india. *The Indian Journal of Labour Economics*, 58(2):245–267.

Witsoe, J. (2014). The practice of development: An ethnographic examination of the national rural employment guarantee act in bihar. *Mimeo,Union College.*

Zimmermann, L. (2012). Labor market impacts of a large-scale public works program: evidence from the Indian Employment Guarantee Scheme.