The Spillover Effects of Public Works on Labor Allocation and Wages: Evidence from National Rural Employment Guarantee Scheme, India

Ashesh Prasann*

Job Market Paper

Abstract

Public works programs guaranteeing work at above-market wages are intended to provide security to the seasonally unemployed. Impact evaluations of India's National Rural Employment Guarantee Scheme (NREGS), the largest such public works program in the world, have estimated a variety of labor market effects, but this literature is agnostic about general equilibrium spillovers generated by the program, potentially biasing these impact estimates. This paper tests for spillovers from NREGS to neighboring labor markets by exploiting plausibly exogenous variation in wage differentials and exposure induced by the program's staggered rollout across contiguous districts. The results show that unassigned districts exposed to program neighbors experienced an 8.7% rise in casual wages relative to unexposed districts and these spillovers are higher among women. These findings demonstrate that the impact of NREGS on wages in the unskilled labor market is larger than previously estimated. They also provide empirical support for the theory that increased seasonal migration to contiguous program areas is the mechanism generating spatial spillovers in neighboring areas which did not receive NREGS over the study period.

Keywords: Spillovers, labor markets, seasonal migration, public works, safety nets, NREGS, India

JEL Classification: I38, J61, O15, O18, R23

I am grateful to comments from Chris Ahlin, Andrew Dillon, Songqing Jin, Leah Lakdawala, Maria Porter, Mywish Maredia, and seminar participants at Michigan State University, the Midwest Economic Association (MEA) meeting – Minneapolis, Midwest International Economic Development Conference (MWIEDC) – Madison, and the Agricultural & Applied Economic Association (AAEA) meetings – San Francisco.

*Michigan State University, Ph.D. candidate, prasann2@msu.edu.

1. Introduction

Public works programs guaranteeing work at a pre-determined wage are intended to provide security to the unemployed and underemployed in the short-term. These public works programs are an increasingly used feature of labor market policy for developing countries. Between 2007 and 2009, many countries started public works programs in response to the food, fuel and finance crises, with 24 being supported by the World Bank (Zimmerman 2013). Recent examples include flagship programs in Argentina, Ethiopia, and India, among others (Subbarao et al. 2012). Since long durations of guaranteed employment at close to the prevailing market wage are likely to put upward pressure on the local wage rate (Subbarao 2003), large-scale public works programs could widen regional wage differentials if implemented non-uniformly across proximate labor markets. In labor markets linked by migration, these wage differentials should equalize at equilibrium with corresponding changes in aggregate employment, thus leading to spillovers from areas assigned the intervention to areas which were unassigned. This type of spillover effect presents a challenge for the estimation of causal impacts through natural experiment approaches because it leads to a violation of SUTVA, i.e. Stable Unit Treatment Value Assumption (Imbens & Rubin 2009), and could result in biased estimates. Accounting for these spillovers is necessary for policy evaluation and is particularly important given the need for cost-benefit analyses of large-scale public works programs, which are increasingly being used as anti-poverty schemes which provide higher and more stable incomes for the poor during economic shocks. Policy decisions about continuing, expanding, or restricting public works programs depend on a precise estimation of their impacts, which could be understated in the presence of spillovers to non-participants (Murgai & Ravallion 2005).

1.1 National Rural Employment Guarantee Scheme

Enacted as law in September 2005, India's National Rural Employment Guarantee Scheme (NREGS), the largest public works program in the world, guarantees every rural household 100 days of employment at the state-specific minimum wage for agricultural laborers[1]. In 2010-11, NREGS provided 2.27 billion person-days of employment to 53 million households, and its budget accounted for 0.6% of India's GDP. The work offered by the program consists of short-term, unskilled, manual work like digging wells and tanks, construction and repair of embankments, planting of trees and building of roads, among others. The jobs provided are similar to private sector casual labor jobs and work provision is concentrated during the dry season, when private sector demand for unskilled labor in agriculture is low. There is also a strong gender component to the program's provisions. Firstly, it is mandated that men and women are paid equally and in cash[2]. Further, at least one third of the NREGS workforce in a village is required to be female and childcare facilities are to be provided at worksites when more than five children under the age of six are present[3]. Rolled out non-randomly in 200 of the most "backward"[4] Indian districts in February 2006, the act was gradually extended to 130 districts in April 2007 and to the rest of rural India in April 2008 (Table 1).

[Insert Table 1 here]

The impact of NREGS on labor market outcomes has been subjected to scrutiny by a recent set of papers. On employment outcomes, the program has been attributed with increasing labor force participation

[1] In 2009, the central government uncoupled NREGS wages from state-level statutory minimum wages and set an all-India uniform wage of 100 rupees per day, but adjusted for state-specific inflation (Dutta et al 2014). Since this change occurred after the study period in this paper and state-level variation in program wage still exists, I continue to work with the assumption that program wage is mandated at the state level for this analysis.

[2] In a 2008 survey of NREGS workers in six Hindi speaking states in North India, only 30% of female respondents reported earning cash income from sources other than the program in the past three months. The comparable figure for men was 55% (Khera & Nayak 2009).

[3] Section 5 and Schedule II of NREGA, 2005.

[4] The Planning Commission used district-level data on caste composition, agricultural productivity and agricultural wage rates from the mid-1990s to calculate a "poverty index" and ranking for 447 districts in 17 states (Planning Commission 2003). This index was then combined with state poverty headcounts, which are not publicly available, to allocate early phase districts to states, with each state receiving at least one district in Phase I. Comparing the list of 199 Phase I districts with only the poverty index ranking, it is clear that higher ranked districts of richer states received the program because of the imperative to introduce the program in at least one district of each state, thus leading to a wider geographic spread than warranted by only the index.

(Azam 2012) and crowding out private sector work (Imbert and Papp 2015, Zimmerman 2013). In terms of wage outcomes, higher private sector (Imbert and Papp 2015), and unskilled labor wages (Azam 2012, Berg et al. 2012) have been estimated; Zimmerman (2013) has also found evidence supporting the program's success in serving as a safety net, with take-up increasing after bad rainfall shocks and men moving out of private casual sector into more risky family employment. In terms of migration impacts, the slowing down of short-term migration from rural to urban areas has been estimated (Imbert and Papp 2014, Bhatia & Ranjan 2009, Jacob 2008). The wider literature on NREGS has demonstrated the program's ripple effects – increased use of labor-saving agricultural technology (Bhargava 2014), reduction in school enrollments among older children (Shah & Steinberg 2015), increase in child labor (Shah & Steinberg 2015, Islam & Sivasankaran 2014), positive impact on grade progression and test scores (Mani et al. 2014), and a rise in consumption, nutritional intake and asset accumulation (Liu & Deininger 2010).

The growing empirical literature on NREGS has frequently exploited its staggered rollout[5] by employing quasi-experimental methods like difference-in-difference (Azam 2012, Berg et al. 2012, Imbert and Papp 2015, Liu and Deininger 2010) and regression discontinuity (Bhargava 2014, Zimmerman 2013) to identify its impacts on labor market outcomes. However, these studies are agnostic about the sign and magnitude of spillovers from program to non-program districts, reflecting the general paucity of studies that test for 'between' (Bayliss & Ham 2015) spillovers, a problem which exists in the experimental literature as well.[6] Multiple randomized experiment based studies have estimated spillovers from treated peers to ineligible individuals or households *within* treated units; examples include deworming externalities (Miguel & Kremer 2004), cash transfers effects of PROGRESA on ineligible households (Angelucci & De Maro 2015, Angelucci & De Giorgi 2009), information spillovers (Oster & Thornton

[5] The timeline of NREGA's three-phase rollout was 199 districts in Phase I (Feb 2006), 128 districts in Phase II (April 2007) and the remaining 261 districts in Phase III (April 2008).

[6] In the theoretical literature, Fields & Raghunathan (2014) have modeled the effect of NREGS on inter-temporal productivity spillovers as part of a two-period seasonal agriculture market.

2012, Miller & Mobarak 2013), and general equilibrium effects of rainfall insurance (Mobarak & Rosenzweig 2013). The corresponding estimations of between spillovers in experimental designs are limited to cross-school deworming externalities (Miguel & Kremer 2004) and cross-village spillovers of PROGRESA on school participation (Bobba & Gignoux 2014).

Conceptually, both within and between spillovers lead to interference and violation of the SUTVA assumption, which requires that the treatment status of any unit does not affect the potential outcomes of other units, and is necessary for both experimental and quasi-experimental approaches to estimating causal impacts. In the NREGS context, quasi-experimental approaches to estimating the program's causal impact assume that labor market outcomes of unassigned districts are the appropriate counterfactual for the assigned districts in absence of the public works program, as a group ("parallel paths assumption" for DID estimators) or locally (for RDD estimators). This assumption is invalidated if changes in outcomes of unassigned districts outcomes occurred precisely because of exposure to assigned districts, leading to biased causal impacts (Duflo et al. 2007), incorrect cost-benefit calculations, and potentially flawed policy decisions about continuing or discontinuing the program. In this analysis, I test for the presence of between spillovers by exploiting plausibly exogenous variation in exposure to NREGS within the set of unassigned districts. Conditional on non-assignment, the comparison of wage and time allocation outcomes in exposed and unexposed districts enables us to compute spillover effects that I argue are driven by changes in labor flows between assigned and unassigned areas. One of the major contributions of this paper is that it is the first known estimation of labor market spillovers from the world's largest public works program.

It is well-established that wage differentials across labor markets linked by migration should lead to equalization of wages in a competitive equilibrium. In fact, capital flows, migration, and goods trade are each, by themselves, sufficient for equalization of wages (Robertson 2000). Given that the study period for this analysis is relatively short and that the demand for NREGS is seasonal in the agricultural labor market, it is unlikely that capital flows across neighboring districts in the form of lumpy, long-term investments in land and equipment would be the driving mechanism for equalization. The case for goods trade is even weaker because of the similarity in relative endowments and the mix of goods produced across neighboring districts. It must be highlighted here that in a long-run general equilibrium setting, *nominal* wage differentials could persist even in absence of capital flows and goods trade, if shocks to demand or supply of labor in a local market are fully capitalized in the price of land (Rosen 1979, Roback 1982), thus equalizing *real* wages spatially. Given the seasonality of NREGS and the relatively short study period of this analysis, this long-run result from the Rosen-Roback framework, the "workhorse of spatial equilibrium analysis" (Glaeser 2001), has limited applicability to this study. Nevertheless, it does necessitate the use of inflation-adjusted wages for the purposes of empirical analysis.

In principle, the enforcement of a mandated minimum wage in excess of market wage in assigned districts and its absence in neighboring unassigned districts would have created wage differentials across interlinked labor markets. Using NSS data from 2004-05, Murgai and Ravallion (2005) showed that 75% of all casual laborers in India, the group targeted by NREGS, worked for less than the state-specific minimum wage, making the program attractive to them[7]. Even in 2009-10, when the program had been rolled out, two-thirds of agricultural labor days were paid less than the minimum wage for agricultural unskilled labor (Dutta et al. 2014). Given this differential, two kinds of migration-linked spillovers could affect residents of non-program districts. First, if the NREGS raised private sector wages in program districts, residents of non-program districts could seasonally migrate or commute to program districts if

[7] In India, the legal enforcement of The Minimum Wages Act 1948 is shared between the central and state governments. Yet, enforcement is weak due to the profusion of minimum wages across states and sectors, poor human resource capacity, and low availability of funds in the state labor departments. There are 45 central government labor regulations (on which states can make further amendments) and in addition, hundreds of state laws (see Debroy (2005) and Anant et al. (2006)). Consequently, there is a gap between the number of available officers and their demand in enforcing these regulations, thus reducing the effectiveness of and pressure on the existing staff (Soundarajan 2013).

destination wages are higher, *even without participating in the program*[8], thus lowering aggregate labor supply and raising wages in their home districts[9]. Second, if NREGS lowered out-migration or commutes from program districts, non-program destinations would experience a reduction in labor supply. Together, these amplifying effects could raise wages in the non-program districts, thus resulting in wage spillovers for districts considered to be the control group in previous evaluations.

Previous estimations of NREGS's labor market impacts have justified abstracting away from migration spillovers by citing the low fraction of rural, inter-district migrants in the population, but this reasoning is flawed on two counts. First, the relevant statistic for gauging the importance of migration spillovers is the fraction of mobile workers – short-term migrants and inter-district commuters – relative to total number of unskilled casual workers, not the entire population. Given that NREGS is designed to target unskilled labor with its timing, wage and type of work offerings, the size of inter-district labor flows relative to the market for unskilled labor is important. Although the NSS definition of short-term migrant leads to an estimated 6.75%[10] of all casual workers falling under the category, this is likely an underestimate[11]. By other estimates, about 10% of agricultural laborers – could be seasonal migrants (Srivastava 2011)[12].

[8] Commuting to a program district in order to work in NREGS is ruled out in principle, by the rules of the program, which specify that only village residents are eligible for job cards and work allocation. This does not rule out long term in-migrants benefiting from the program but such migration is low and deterred by high costs.

[9] This is similar to the effect on emigration on labor market outcomes in source countries, a question not given great attention in the empirical literature, with a notable exception being Mishra (2007), which estimated that the outflow of Mexican workers to the US between 1970 and 2000 has increased the wage of an average Mexican worker by about 8%. In related work, Robertson (2000) finds that U.S. wage shocks are transmitted from border to interior cities in Mexico by way of labor migration from interior to the border.

[10] Author's calculation using NSS 64 data.

[11] NSS 64 asked individuals whether they had migrated for 1-6 months in the last 365 days. This is likely an underestimate because: i) in many cases, the seasonal/circular migration cycle is longer than six months, and ii) quite often, entire households and not individuals participate in seasonal migration (Srivastava 2011). In absolute terms, there were an estimated 15.2 million short-duration out-migrants, of whom 12.9 million (85.1 per cent) were male, and 13.9 million (71 per cent) were rural out-migrants. The overall out-migration rate was 1.33 - 1.72 for rural areas and 0.4 for urban areas.

[12] Their salience to labor markets affected by NREGS is also reflected by the fact that in 2007-08, 56.6% of all seasonal migrants reported working in construction (36.2%) and agriculture (20.4%), sectors most likely to compete with the government program for workers.

Further, inter-district commuters are unaccounted for in NSS, and there is no other nationally representative survey that records commuting data. Second, while the decline in rural-urban migration dominates the popular discourse and empirical literature with respect to NREGS' effect on migration in India, inter-state rural-urban migrants accounted for 36.4% of all short-duration migration in 2007-08 (Srivastava 2011). An almost equal amount of seasonal migration occurred across districts in the same state, which suggests that changes in cross-district wage differentials could have had a sizable effect on the labor markets in close proximity to NREGS districts.

This paper studies the labor market spillovers from assigned to unassigned districts during the study period (July 2005 to March 2008), by exploiting the plausibly exogenous variation in exposure to program neighbors. In spirit, this approach is similar to McKinnish (2005), which studied welfare migration in border counties of U.S. states with large cross-border benefit differentials for Aid to Families with Dependent Children (AFDC). The study found that border counties in the high-benefit state experienced higher program participation and expenditures relative to interior counties. It is also related to the estimation of spatial spillovers from natural resource booms to counties not experiencing booms (Allcott & Keniston 2014) and the effects of civil wars on economic outcomes in neighboring countries (Murdoch & Sandler 2002). My results show that, conditional on not receiving the program, the real wage for casual labor increased by 8.7% in exposed districts relative to unexposed districts, with women experiencing larger wage increases than men. When heterogeneity of exposure intensity within exposed districts is accounted for, I estimate that a 10% increase in exposure intensity leads to a 1.03% rise in casual wage and an increase in weekly labor force participation. Lastly, I present evidence that short-term, short-distance migration to contiguous program areas is the mechanism generating between spillovers in districts which did not receive NREGS over the study period.

The rest of this paper is organized as follows. Section 2 provides a basic theoretical framework for the analysis and develops predictions for empirical estimation. Section 3 describes the relationship between program rollout and measures of exposure used in this analysis. Section 4 describes the data, with estimation strategy outlined in section 5. Section 6 presents the main results while section 7 discusses robustness checks. Section 8 presents evidence of seasonal migration being the mechanism generating spillovers, and section 9 offers concluding remarks.

2. Model

This section presents a simple model motivating the optimization problem faced by an individual in an origin district with an outside option of seasonal migration or commute to a neighboring destination. Since the results can be extended to the case with many destinations without loss of generality, this model is presented to fix ideas. This individual has a utility function u(c_i, l_i) over consumption and leisure, with the function increasing and concave over both arguments. Her time endowment, T, is split between leisure l_i, work at origin (L_i^O), and work at destination (L_i^D)**.** The origin district wage is w^o, while both migrants and commuters earn w^d at destination district. Work outside the home district is associated with an additional variable cost v_i, which is heterogeneous across individuals. While transportation cost is an example of a variable cost for commuters, it is fixed for seasonal migrants. Meanwhile, additional rent at destination is an example of variable cost for seasonal migrants but not for commuters. This model abstracts from the distinction between the variable costs faced by seasonal migrants and commuters because it does not change the individual's problem. The marginal wage rate she faces depends on which wage rate is higher – origin wage (w^o) or the net destination wage (w^d – v_i). Individuals also have non-labor income y_i, which can be thought of as profits from agricultural production in this setting. It is worth noting that the production function f(D_i) only allows the use of labor input D_i, thus ruling out capital flows across labor markets by assumption. Individuals thus choose consumption and leisure to solve:

$$\max u(c_i, l_i) \quad \text{s.t.} \quad c_i + W_i l_i = y_i + W_i T \tag{1}$$

$$L_i^D + L_i^O + l_i = T \tag{2}$$

$$W_i = \max \{w^o, w^d - v_i\} \tag{3}$$

$$y_i = \Pi_i = f(D_i) - w^o D_i \tag{4}$$

Solving the first order conditions, the individual's standard demand functions for leisure and consumption are given by:

$$l^* = l^* (W_i, y_i + W_i T) \tag{5}$$

$$c^* = c^*(W_i, y_i + W_i T) \tag{6}$$

In this model, it is optimal for an individual to either work in the home district or outside depending on the marginal wage rate, but not both. For a given $\{w^o, w^d\}$ pair and wage differential $(w^o - (w^d - v_i))$ across districts, individuals with low variable costs $(w^d - v_i > w^o)$ work outside the home district, henceforth termed "leavers." On the other hand, individuals with high variable costs $(w^d - v_i < w^o)$ optimally allocate their labor to the home district, hereafter referred to as "stayers." The aggregate labor supply *within* home district is the sum of individual-level labor supply of S stayers and labor inflow from destination to home districts (L_o):

$$L_S^O = \sum_{i}^{S} L_i^{O*} (W_i, y_i + W_i T) + L_o \tag{7}$$

while the aggregate labor supply at the destination is the sum of individual-level labor supply of L leavers and labor supplied by residents of destination districts (L_d)*:*

$$L_S^D = \sum_{i}^{L} L_i^{D*} (W_i, y_i + W_i T) + L_d \tag{8}$$

with (S+L) representing the total population of the home district. The labor market equilibrium conditions before the introduction of NREGS for home and destination districts can be written as:

$$L_S^O = \sum_{j}^{N} D_j^* (w^o) \tag{9}$$

$$L_S^D = \sum_{i}^{K} D_i^* (w^d) \tag{10}$$

where home district wage w^o and destination wage w^d clear the respective labor markets.

2.1 Introduction of NREGS

In this setup, I assume that only destination districts can receive NREGS. If the destination district receives the program, the guaranteed employment at above-market wage[13] ($G(w^p)$) raises aggregate labor demand and destination wage (w^d).

$$L_S^O = \sum_j^N D_j^*(w^d) + G(w^p)$$

As w^d increases at destination, the inflow of labor to home districts, L_o, would decline because of the high program wage being offered by NREGS. The rising destination wage w^d simultaneously changes the labor allocation decision of home district residents. As the wage differential ($w^o - (w^d - v_i)$) widens, previously marginal stayers would choose to allocate their labor to the program neighbor and leave, unless v_i is too high for all stayers. The amplifying effects of reduced inflow and increased outflow of labor lead to an inward shift in aggregate labor supply of home districts (L_S^O), which continues until w^o rises sufficiently and equalizes the wage differential for remaining stayers (eq. 7). The rising home district wage could induce the remaining stayers to either supply more labor or buy more leisure at the individual-level, with this ambiguity being the result of offsetting income and substitution effects. It is noteworthy here that if the destination district does not receive NREGS, the labor market equilibrium characterized by equations (9) and (10) remains unchanged. In other words, w^o remains unchanged and the home district experiences no spillovers to wage and aggregate employment if it is not 'exposed' to a program neighbor.

2.2 Exposure Intensity and Non-Program Labor Markets

I can now extend the model outlined above to multiple destinations to allow greater variation in exposure, i.e. number of program neighbors for home districts. I also assume that for a given w^d, the *effective* out of

[13] This is supported by empirical evidence on the minimum wage guaranteed by the program exceeding the prevailing market wage for casual labor in fifteen of the eighteen states in India in 2007 – 08 (Azam 2012).

district wage w_E^d for home district residents is a monotonically increasing function of exposure intensity E, a measure of linkages between home and destination labor markets where the program is implemented.

$$w_E^d = g(w^d, E),\ g_E(w^d, E) \geq 0$$

It is axiomatic that as the choice set of program destinations expands for home districts, the highest w^d offered outside the district is non-decreasing. It is noteworthy that this assumption might not hold in practice if neighboring districts are subject to different policies designed to stop the movement of labor, i.e., ones that keep labor markets segmented, thus delinking the relationship between the number of program neighbors and the maximum market wage offered outside a non-program district. Since the introduction of NREGS was not conditioned on any mobility restrictions for non-participants and this model aims to capture intra-country mobility within India, which constitutionally guarantees the right to move and reside in any part of the country (Part III, Constitution of India), there is little evidence of market segmentation and thus minimal cost to making this assumption. Given that the new wage differential ($w_E^d - (w^o - v_i)$) is an increasing function of exposure intensity, the following predictions are generated from this model:

<u>Prediction 1:</u> For a positive shock to destination wage w^d, origin wage w^o in unassigned districts is increasing in exposure intensity E.

<u>Prediction 2:</u> For a positive shock to destination wage w^d, aggregate employment in unassigned districts is decreasing in exposure intensity E.

<u>Prediction 3:</u> For a positive shock to destination wage w^d, individual-level labor supply for residents of unassigned districts is ambiguously related to exposure intensity E.

It is evident from the discussion above that the problem is symmetric for program districts, where the introduction of NREGS would lead to a reduced outflow and increased inflow of labor. Aggregate employment and market wage would increase as is to be expected with a shock to labor demand at a guaranteed minimum wage above prevailing market wage. In practice though, there is evidence that the employment guarantee offered by NREGS is fuzzy, with extensive rationing of work and significant unmet demand for it in some states (Dutta et al. 2014). It is worth emphasizing here that labor market spillovers to unassigned districts, like private sector labor market impacts in assigned districts, will be decreasing in the level of rationing. Hence, the extent of spillovers generated by the program is fundamentally an empirical question.

3. Program and Exposure

This analysis is restricted to Phase III districts which received NREGS last, henceforth referred to as "late" districts (Phase I and II will be referred to as "early" districts). The study period is from July 2004 – March 2008, when late districts had not yet received the program. However, late districts were exposed to it from February 2006 onwards by virtue of proximity and labor market linkages to neighboring early districts. Since this analysis is conditioned on being a late district, the non-random rollout of the program is not inherently a threat to its internal validity, unless proximity to early phase districts is systematically correlated with unobservable individual and district-level characteristics, an issue which is addressed in the next two sections of the paper. Using data from two rounds of the nationally representative National Sample Survey's (NSS) conducted before and after introduction of NREGS (Table 2), the spillover impacts of NREGA on wages and labor allocations in late districts can be estimated. The maps shown in Figures 1 and 2 in the appendix visually depict the spatial distribution of districts across NREGS phases and the classification of early and late phase districts used in this analysis.

[Insert Table 2 here]

[Insert Figures 1 and 2 here]

3.1 Exposure as Binary Indicator

In order to capture exposure, the following binary measure is computed for each late district:

$$E_1 = \begin{cases} 1 \ if\ at\ least\ one\ contiguous\ neighbor\ is\ "Early" \\ 0\ otherwise \end{cases}$$

In other words, a late district is considered exposed if it shares a border with one or more early district neighbors and is classified as unexposed if surrounded by late district neighbors. It is noteworthy that in principle any two districts or, for that matter, any two points in space can be considered neighbors, depending on how neighborhood is defined. Given the absence of theory about the geographic scale of labor markets, empirical choices in the literature have largely been driven by data constraints and the objective of the analysis. While spatial data on a finer scale identifying smaller administrative units is available, the NSS data used to construct outcomes in this analysis is only identified up to the district-level, making finer measures of neighborhood redundant. Additionally, since the objective of this exercise is to capture spillover impacts driven by variation in proximity to NREGS, in otherwise similar labor markets, first-order contiguity is used as the criterion for neighborhood (see Murdoch & Sandler 2002, Robertson 2000 for similar criteria). It is possible that some distant labor markets are better linked by idiosyncratic transport or social networks than adjacent districts, but this study abstracts from those linkages because on average, first-order contiguity enables greater comparability of exposed and unexposed districts similar to and in close proximity to each other. The map in Figure 3 of the appendix highlights all districts that are classified as exposed using this measure. In the working sample of 215 late districts used for the main analysis drawn from India's largest states, 80% (173) of the districts are exposed and the remaining 20% (42) districts are categorized as unexposed using this measure.

[Insert Figure 3 here]

3.1 Exposure Intensity as Ratio

Since E_1 constrains exposure to being a binary variable and does not capture heterogeneity across exposed districts, the following measure of exposure intensity is also computed for late phase districts.

$$E_2 = \frac{Number\ of\ contiguous\ "Early"\ neighbors}{Total\ number\ of\ contiguous\ neighbors}$$

E_2 takes on fractional values ($E_2 \in [0, 1]$) and increases with every additional contiguous early district, assuming that the total number of contiguous neighbors does not change. This measure gives higher weight to districts with more early neighbors that the binary measure of exposure, but also penalizes large districts which might have more adjacent neighbors by virtue of size, by adjusting the ratio downward. It also enables this analysis to investigate the robustness of spillover effects to an alternative measure of exposure which captures intensity. In the working sample, while late districts have 5.5 districts on average, the mean number of early neighbors is 1.9, with a standard deviation of 1.6. Conditional on exposure to at least one early neighbor, the average exposure intensity is 0.38 and standard deviation is 0.28 (see Table 3).

[Insert Table 3 here]

4. Data

As outlined in section 2, this analysis utilizes two sources of data: i) two rounds of employment surveys – NSS 61 (July 2004 – June 2005) and NSS 64 (July 2007 – June 2008) and ii) spatial data on district boundaries based on the Indian Administrative Census (2001). The individual is the primary unit of analysis, and the sample is restricted to all adults aged 15 to 59[14] without tertiary education in the 215 late districts[15]. The NSS survey is comprised of four sub-rounds designed to coincide with *rabi* and *kharif*, the two growing seasons in Indian agriculture, as well as post-harvest quarters. The study period for this analysis is restricted to July 2004 – March 2008 by dropping the last sub-round from NSS 64, which ensures that labor market changes in late districts occurring after assignment in April 2008 do not contaminate the working sample. Since the survey is uniformly distributed across sub-rounds by design, this restriction does not systematically change the working sample. Together, seven sub-rounds of data drawn from NSS 61 and NSS 64.

This analysis utilizes the *current daily status* measure of NSS employment surveys to construct weekly time allocation for each individual in three mutually exclusive and exhaustive categories: work days (casual labor, salaried work, domestic work, public sector work and self-employment), unemployment, and non-labor force participation (NLFP)[16]. I also compute daily wages for individuals who worked as casual laborers since this segment of the labor force is most likely to be directly impacted by spillover from public sector casual labor offered by NREGS. Daily wages for individuals who worked as salaried laborers is also used as an outcome variable to validate this assumption and ensure that this analysis captures spillovers induced by NREGS and not unrelated trends correlated with exposure. Both wages have been inflation-adjusted using state-level, quarterly CPI for agricultural workers. Here, it must be

[14] National Sample Survey Organization (NSSO), the agency which carried out the NSS, defines individuals aged 15 to 59 as the "economically active population" and uses this sample to calculate employment and unemployment rates. This analysis adopts the same convention.

[15] The sample drops union territories, the conflict affected state of Jammu & Kashmir, and small, sparsely populated north-eastern states. Completely urban districts are also dropped from the sample.

[16] NSS 61 and 64 recorded the time disposition of respondents during the week preceding the interview, coding the intensity of their activities as 0.5 or 1 for each day. In this analysis, these activities are classified in one of the four categories and the intensity of that category is summed across the week to get weekly time allocations. There are two other measures of employment also available in the NSS: *usual status* (based on a recall of a year) and *current weekly status* (based on recall of previous week). However, since these variables capture only the principal activity of each individual during a given reference period, they cannot shed light on intensive margin changes in time allocation in response to NREGS so they are not used in this analysis.

emphasized that all time allocation and wage outcomes are observed for stayers of late districts as the surveys does not track leavers. Thus, it is possible that the wage and labor outcomes of individuals who migrate to early neighbors in response to program exposure are not observed in the post-NREGS period and this attrition is more likely to occur in districts with higher exposure intensity. In principle, it is also possible that variable costs associated with working outside the home district are systematically lower for less productive workers, thus mechanically raising average wages for stayers. If variable costs associated with seasonal migration, in particular transportation costs, are also systematically lower for exposed districts because of better road or rail networks, this would lead to higher average wages for exposed district stayers in the pre-exposure period.

In light of the above discussion, it is useful to compare wage and labor outcomes across unexposed (E_1 = 0) and exposed (E_1 = 1) late districts before the introduction of the program. Table 4 summarizes individual-level statistics for all the dependent variables used in the analysis and salaried days in 2004 – 05, the survey year preceding Phase I assignment. While the weekly allocation of work days, casual days, salaried days and non-labor force participation is similar across exposed and unexposed districts, weekly unemployment accounts for 0.9 more days for the average economically active adult in exposed districts. In terms of wages, salaried wage is similar across both categories but casual wage is lower in exposed districts. While the latter could be indicative of lower variable costs associated with migration, they could also reflect other district characteristics, thus necessitating the use of district fixed effects in my estimation strategy.

[Insert Table 4 here]

Further, although the pre-exposure differences in unemployment days and casual labor wage are significant across exposed and unexposed districts, they do not invalidate the difference-in-difference estimation strategy undertaken in this analysis, which instead relies on the identifying assumption of parallel trends in outcomes. Since this assumption cannot be directly tested with the two rounds of survey data used in my study sample, I graph the annual evolution of real casual wage through two additional pre-program rounds (1993 – 1994 and 1999 – 2000) along with the study period survey rounds (2004 – 2005 and 2007 – 2008) in Fig. 4 below. As mentioned earlier, the last quarter from 2007 – 2008 has been dropped to exclude the post- Phase III period when late districts were assigned the program. Visually, it is quite evident that the casual wage trends in exposed and unexposed late districts run parallel between 1999 and 2005 and then their gap narrows after introduction of the program in the first two phases (PI and PII in the figure). This increases my confidence that pre-existing difference in casual wage across exposed and unexposed districts documented in Table 3 is not necessarily an indicator of pre-existing non-parallel trends for this outcome.

[Insert Figure 4 here]

Next, it is instructive to compare observable characteristics across late districts to explore if the non-random rollout of the program necessarily implied systematic differences across exposed and unexposed districts. In order to do so, Table 5 reports averages for individual and household-level characteristics during the pre-exposure period using NSS survey data. It also reports averages for demographic variables using district-level data from the 2001 Census. It is clear from the Census statistics that unexposed districts are more rural, comprise of larger households, and have a higher ratio of women to men (overall and caste-differentiated) than exposed districts. However, they do not differ from exposed districts in terms of population size, caste distribution of population, and literacy (overall and gender-differentiated). Since the Census is decadal, I do not have information on these characteristics for the post-exposure period in my sample (July 2007 – March 2008) to explicitly control for them. If time-invariant though, these district-level characteristics are accounted for by including district fixed effects in my regression specification. On the other hand, the comparison of means for individual and household-level variables using NSS data has much higher power given the large size of my study sample. It shows that in the pre-exposure period, the economically active population of unexposed districts was younger, less likely to belong to Scheduled Tribes or Other Backward Castes, and more likely to be male and belonging to the other caste category. In terms of land possession, literacy and likelihood of belonging to Scheduled Castes, there is no statistical difference between unexposed and exposed districts. All these variables as well are controlled for in my regression specifications. It is also possible that other time-varying individual, household or district-level unobservables are correlated with exposure. I carry out robustness checks by way of an out of study sample placebo analysis and impose sample restrictions in order to investigate the effect, if any, of these unobservables.

[Insert Table 5 here]

5. Estimation

In my strategy, I compare changes in outcomes of exposed late districts to unexposed late districts in order to calculate NREGS induced spillovers from early districts. In other words, the spillover effects on log daily wages and weekly time allocations are estimated by exploiting the plausibly exogenous shock to variation in exposure across late districts due to introduction of the program in contiguous early neighbors. I estimate variations of the following specification:

$$\mathrm{Y}_{\mathrm{idt}} = \alpha + \beta_1 Post_t + \beta_2 Exposure_d + \beta_3 Post_t{*}Exposure_d + \beta_4\delta_t + \beta_5\mu_d + \boldsymbol{\Theta}\mathbf{X}_{\boldsymbol{idt}} + \varepsilon_{\mathrm{idt}} \qquad (1)$$

where i indexes individual, d indexes district, and t indexes quarter-year. In the first specification (1), a difference-in-difference approach is employed and $Exposure_d$ is the binary measure of exposure used, i.e. "treatment" is defined as having at least one contiguous early neighbor. $Post_t$ is an indicator variable which is 0 for all late districts in the first four quarters predating the program (July 2004 – June 2005) and is 1 in the post-exposure quarters (July 2007 – March 2008). The variable of interest is the interaction term, $Post_t$* $Exposure_d$, and the parameter β_3 estimates the impact of exposure on individual-level time allocation and wage variables. This impact of exposure is a "treatment effect" which captures labor market spillovers in exposed districts relative to unexposed districts, conditional on non-assignment. Inflation-adjusted log daily wages for casual labor, salaried labor and time allocations and serve as the dependent variables (Y_{idt}) in separate regressions and $\mathbf{X}_{idt}$ represents individual, household, and district-level controls. Specifically, individual-level factors like age, education, and gender which are correlated with seasonal migration[17], and household-level factors like caste grouping and land possessed are controlled for. Quarter-year fixed effects (δ_t) are included to control for seasonal and secular changes in labor market outcomes through the study period. Since the exposure variation being exploited in this estimation is at the district-level, district fixed effects (μ_d), which account for pre-existing, time-invariant district characteristics, are also included. The error term ε_{idt} captures individual-level heterogeneity in the variable costs associated with working outside the home district. To account for intra-district correlation of individual-level errors, standard errors are clustered at the district-level for all reported results. All regressions are implemented at the individual-district-quarter level of analysis and sampling weights provided in the surveys are used to weight these estimates.

While the coefficients for the specification in (1) are intuitive and relatively easy to interpret, they do not capture heterogeneity of exposure intensity across exposed districts. The second type of specifications (2) estimated in this analysis employs the ratio-based measure of exposure intensity, $ExposureIntensity_d$, which represents the fraction of contiguous neighbor districts that receive the program in the early periods. Using this measure of exposure, β_3 represents the linear, marginal effect of increased exposure intensity; in other words β_3 is the effect of being surrounded entirely by early phase neighbors, relative to having no early phase neighbors.

[17] "The socioeconomic profile of the short-duration/seasonal out-migrants is very different from the other migrants. These migrants are much more likely to be from socially deprived and poorer groups, have low levels of education, less land and more likely to be engaged in casual work. They are also more likely to be of prime working age (two-thirds are15-29 years old) and are predominantly male" (Srivastava 2011).

5. 1 Plausible Exogeneity of Exposure

Irrespective of measure, the main source of potential bias in employing this econometric strategy is if the introduction of NREGS into *neighboring* districts is non-random, conditional on quarter-year and district fixed effects and individual and household-level controls. If exposure is actually a proxy for some other unobservable, and is correlated with differential district-level trends in outcomes, then the estimates of spillover impacts on late districts will be biased. One way to test for this is to see whether estimating (1) on the pre-program data can generate similar "effects" of having early NREGS program neighbors. In order to carry out this check formally for all outcome variables, specifications (1) and (2) are implemented using a placebo sample consisting of two survey rounds (1993 – 94 and 1999 – 2000) conducted before the introduction of NREGS. These results are reported in section 7. A visual representation of this placebo test can be seen in Figures 5 and 6 of the appendix, which show the quarterly evolution of inflation-adjusted casual wage in the placebo and study sample respectively. While the line marker partitions actual pre and post-exposure quarters in the study sample, it separates quarters before and after "fake" exposure in the placebo sample. It is clear that while the gap between exposed and unexposed districts narrows in the last two of three post-exposure quarters for the study sample, it fluctuates in both directions in the placebo sample, thus suggesting that differential trends are related to actual exposure.

[Insert Figures 5 and 6 here]

There are two other threats to exposure being plausibly exogenous. Firstly, if late districts were able to manipulate their exposure to the program, then it could be the case that exposure is correlated with other unobserved variables that differentially affect outcome trends for exposed districts. Here, it must be emphasized that my analysis is restricted to late districts and is thus conditioned on receiving NREGS last. While late district administrators may have tried to manipulate early reception of the program, it is not clear how they would have influenced being exposed to early district neighbors, given that they were unsuccessful. Since the central government's selection of early districts was non-random with economically underdeveloped districts being selected first, the sample of late districts is richer than early districts[18], but that is not the comparison being made in my analysis. Unless exposure is systematically correlated with differential outcome trends *within* the sample of late districts, the non-random rollout of the program does not threaten the internal validity of this analysis. As discussed earlier, the balance between exposed and unexposed districts for five of the seven outcomes used in this paper and the absence of pre-existing non-parallel casual wage trends increases my confidence that the two groups of late districts do not experience differential outcome trends, even if they are systematically different from early districts as a group.

A second, less serious source of bias is that if late district residents correctly anticipated the program rollout and the identified early districts which would receive NREGS before their home district, their behavioral response could be to migrate to NREGS districts *before* the shock, thus resulting in diminished or no spillover effects on labor markets being estimated. In fact, a large scale migration of this sort would be a threat to any evaluation of the program's impact, not just the estimation of spillovers associated with it. Given that the assignment of early phase districts is imperfectly predicted even using the index made publicly available after the introduction of the program (Zimmerman 2013), it is quite improbable that individuals would have correctly anticipated which districts would receive the program early and migrate to them. Furthermore, recent work has shown that public knowledge about the program remained low in Bihar, one of the poorest states in India, even three years after its initial implementation (Ravallion et. al 2013). In this setting, it is highly unlikely that anticipatory migration by unskilled workers, which requires information about the program and its wage offerings in other districts before implementation, is

[18] Zimmerman (2013) has reconstructed this two-step algorithm using state poverty headcounts from the 2001 Census to imperfectly predict assignment for a RD design based impact evaluation of NREGS. Since poverty index and corresponding rank is missing disproportionately for late phase districts, this analysis cannot utilize these ranks as additional controls.

a relevant concern. In the unlikely case of this concern being valid, my estimates simply serve as a lower bound for spillover effects.

6. Main Results

6.1 Exposure and Spillovers

The results from estimation of specification (2) are presented in Table 6. The dependent variables in the first two columns are logs of inflation-adjusted daily casual and salaried labor wages, conditional on having earned a positive daily wage. The next three columns are weekly time allocations for unemployment, work, and non-labor force participation days, which are mutually exclusive and exhaustive of time endowment. I estimate that, on average, spillovers from NREGS resulted in inflation-adjusted daily casual wage increasing by 8.7% (significant at the 5% level) more in exposed districts, relative to the increase in unexposed districts. This result provides empirical support to the prediction of home wages rising faster with exposure in late phase districts. The absence of a similar effect on salaried wage supports my conjecture that the relevant market for measuring spillovers from NREGS is the casual labor market and that my results reflect the impact of exposure to the program. There is no statistically significant effect of exposure on weekly time allocation variables as well. Given the limitations of repeated cross-section data, which does not track leavers in both rounds, the absence of predicted impacts on employment can be attributed to the changing composition of the sample. In particular, since leavers are likely to be concentrated in exposed districts during the post-period, the estimated effects of exposure on aggregate employment are attenuated. Further, the small fraction of casual workers relative to the overall sample could contribute to small effects not being detected. Although the contexts are not strictly comparable, the basic result is similar to the Alcott & Keniston (2014) finding that earnings spillovers from natural resource booms were concentrated in nearby counties, relative to faraway counties.

[Insert Table 6 here]

6.2 Exposure Intensity and Spillovers

When specification (2) is implemented, I estimate that, on average, spillovers from NREGS resulted in inflation-adjusted daily casual wage increasing by 1.03% with every additional 10% increase in exposure intensity (significant at the 5% level). For example, a late district with ten neighbors would experience a 1.03% increase in casual wage with every additional neighbor receiving the program early. Comparing the magnitude of these spillovers to previous evaluations of the program, it is instructive that a 50% increase in exposure intensity is roughly equivalent to the 4.7% and 5.1% increases in casual wages in program districts estimated by Imbert and Papp (2015) and Berg et al.(2012), respectively. In other words, program assignment to half their neighbors has about the same impact on casual wages in late districts as receiving the program. In Table 7, the change in log daily salaried wage remains statistically insignificant, corroborating the result from Table 6 and increasing my confidence that the estimated impacts on casual wages are not a reflection of secular wage increases across all labor markets in exposed districts, relative to unexposed districts. Interestingly, the time allocation results differ across exposure and exposure intensity, with labor supply increasing with exposure intensity. The non-labor force participation of residents in completely exposed districts decreased by 0.21 days, relative to completely unexposed districts (significant at the 5% level). This 7.2% increase in labor force participation is accompanied by positive but statistically insignificant increases in work and unemployment days. Since the survey does not track leavers, the increase in labor force participation should be interpreted as the effect of exposure intensity on the average labor supply responses of stayers, and not on aggregate district-level labor supply. Given that the model's predicted effect on individual-level labor supply of stayers is ambiguous (as discussed in section 2), rising casual wages accompanied with increased labor force participation by stayers is indicative of an upward sloping labor supply curve in this wage range. Further, the estimation of this additional time allocation effect suggests that while the dichotomous exposure variable performs fairly well in terms of approximating exposure intensity impacts on casual wages, there are efficiency gains when heterogeneity of exposure is accounted for.

[Insert Table 7 here]

6.3 Spillovers by Gender

Given that seasonal migrants who leave for employment purposes are overwhelmingly male in India, it is of interest to evaluate whether spillovers due to exposure are gender-differentiated. In order to estimate these gender-differentiated impacts, indicators for men and women are interacted with the $\beta_3 Post_{jt}*Exposure_j$ term in specification (2) and the coefficients are reported in Table 8. Additionally, I report effects on time allocated to casual labor, a subset of workdays, in order to focus on the gender-specific labor responses of stayers in this particular market. I estimate that, on average, casual wage for women increased by 9.3% more in exposed districts relative to the increase in unexposed districts. On the other hand, the increase in casual wage for men was estimated to be 8.5%. If NREGS raised women's wages more than men's on account of a bigger differential between program and private sector wages for women, as the findings from Azam (2012) suggest, it follows that the wage differentials across program and non-program districts would also be higher for women. Higher casual wage spillovers for women in exposed districts despite seasonal migrants being predominantly male, thus also support the prior that given relatively low labor force participation, even small flows of female migrant labor could have had large impacts on home district casual labor markets. On the other hand, weekly male labor force participation increased by 0.19 days in exposed districts and almost all of it translated into an increase of time allocated to casual labor, i.e. 0.17 days. Weekly unemployment days for men also rose by 0.08 days even as changes in time allocations for women are not statistically significant. The presence of time allocation spillovers among men despite higher wage spillovers among women, suggests that the elasticity of labor supply is positive among male stayers while income and substitution effects induced by higher wages could be offsetting each other for women. Lastly, the simultaneous increases in male unemployment, labor supply and casual wage, signals the presence of search costs or wage rigidities which prevent the labor market from clearing in late districts.

[Insert Table 8 here]

7. Robustness Checks

7.1 Placebo Analysis

In order to ensure that the effects estimated for late districts are genuinely spillovers and not being driven by a correlation between unobserved variables and exposure, a placebo analysis is carried out, and the effects of fake exposure to early program districts are estimated. In other words, if the parallel paths assumption underlying the difference-in-difference estimation is violated, differential trends in outcomes unrelated to exposure would be reflected in the placebo sample as well. In particular, the impacts on casual wages, for which positive and significant effects have been estimated in the main sample, would be replicated even with fake exposure if this concern was valid. To carry out the placebo analysis, specifications (1) and (2) are implemented using two rounds of data – NSS 50 (July 1993 – June 1994) and NSS 55 (July 1999 – June 2000) – which preceded the introduction of NREGS. In order to make the sample comparable with the study sample, the last sub-round of NSS 55 is dropped from the analysis. Since no districts received the program during this period, *a priori*, exposure to contiguous neighbors which received the program more than a decade later should have no effect on casual wages. In the results reported in Table 9, I observe that changes in inflation-adjusted casual labor wage are statistically insignificant in exposed late districts, relative to unexposed late districts. Similarly, there is no statistically significant effect of exposure intensity on casual labor wage in Table 10. These results validate the interpretation that the study sample results are not being driven by a spurious correlation between exposure and labor market outcomes but are instead genuine spillover effects generated by exposure-induced seasonal migration. Additionally, Tables 8 and 9 show that time allocation changes in the placebo sample are not statistically significant like the main sample results, but since predictions from theory cannot be tested for these outcomes because of the reasons discussed in section 6, no conclusion can be drawn from this similarity between placebo and study sample results.

[Insert Tables 9 and 10 here]

7.2 Sample Restriction

<u>Low Coverage States:</u> Since early phase districts were concentrated in eastern and central Indian states, the late districts in these states are more likely to be exposed and have higher exposure intensity than late districts in western and southern India. It could be the case that the spillover effects estimated in the study sample are driven by "high coverage" states where late district individuals have greater choice in terms of migration to high wage destinations. To test this hypothesis, specifications (1) and (2) are estimated on a restricted sample consisting only of late districts in large, "low coverage" states, defined as having less than half of their districts receiving the program early. The results in Tables 11 and 12 follow the same qualitative pattern observed in study sample results, but with larger magnitudes of increases in casual wage due to exposure (13.5%) and exposure intensity (18.3%). The effects are significant at 1% and 10% levels of significance, respectively, despite the reduced sample size. Estimating larger impacts in "low coverage" states relative to the original sample suggests that there may be diminishing returns to exposure and exposure intensities in terms of casual wage spillovers. Changes in salaried wage and time allocations continue to remain insignificant.

[Insert Tables 11 and 12 here]

8. Seasonal Migration and Exposure Intensity

Although the spillover impacts on casual wages estimated in this analysis are motivated as resulting from changes in flows of seasonal migrants to and from late phase districts, the direct impact of exposure on seasonal migration could not be estimated because migration information was not collected in NSS 61 (the round immediately preceding NREGS). While this information was collected in NSS 55 (July 1999 – June 2000) and NSS 64 (July 2007 – June 2008), differences in the definition of short-term migrants across rounds prevent comparability[19] and thus, direct estimation of exposure's direct impact on short-term migration. However, some inferences can be drawn about the empirical relationship between exposure intensity and seasonal migration by exploiting the temporal dimension of NSS 64's survey design, which comprises four quarterly sub-rounds (July – Sep 2007, Oct – Dec 2007, Jan – Mar 2008, and April – June 2008), broadly coinciding with the agricultural cycle.

As a first step, it is useful to visually observe the relationship between seasonal (short-term) migration and exposure intensity, and further explore whether this pattern differs for short-distance seasonal migration. Figure 7 in the appendix graph two measures: i) the percentage of seasonal migrants in the population, and ii) the percentage of seasonal migrants in the population who moved to a destination outside the district, within the same state. Both measures are graphed across three exposure intensity terciles, before and after program assignment to late districts (April 2008). Firstly, Fig. 7 shows that the relationship between short-distance seasonal migration and exposure intensity is inverted U-shaped in both periods, supporting the higher wage spillovers estimated in low-coverage states relative to the entire sample. Secondly, it is evident that the responsiveness of seasonal migration to program assignment varies by distance and exposure intensity in the April – June 2008 quarter, which coincides with the pre-monsoon dry season. Seasonal migration to all destinations increases, suggesting that overall seasonal migration continues to serve as a coping mechanism in the dry season despite the introduction of NREGS, supporting the findings of significant unmet demand for the program (Dutta et al. 2014). On further disaggregation, we find that seasonal migration to districts within the same state declines, and this decline is *proportional* to exposure intensity. This result is an indicator that in the absence of program assignment, seasonal, short-distance migration was the relevant form of labor movement induced by exposure to early neighbors. In other words, seasonal short-distance migration to early neighbors served as a substitute for program assignment in late districts in the pre-assignment period.

[19] NSS 55 defines short-term migrants as individuals who stayed away from their usual place of residence for 2 – 6 months during the last year. NSS 64 changed this period to 1 – 6 months.

[Insert Figure 7 here]

9. Conclusion

The findings from this analysis demonstrate that exposure to NREGS produced significant general equilibrium spillovers in the form of higher real casual labor wage in districts where it was not rolled out during the study period. Comparing the magnitude of these spillovers to previous evaluations of the program, it is striking that the impact of exposure (8.7%) is higher than the 4.7% and 5.1% increases in casual wages in program districts estimated by Imbert and Papp (2015) and Berg et al.(2012), respectively. Given the heterogeneity in exposure intensity across exposed late districts, it is more instructive to note that program assignment to half their contiguous neighbors is roughly equivalent to program assignment itself, if previously estimated intent-to-treat effects on casual wages are taken at face value. These spillovers are also marginally higher among women than men and are stronger in large, low-coverage states, indicating diminishing returns to exposure for late districts.

Given the presence and magnitude of these spillovers, it is evident that the gains from public works programs are not completely accrued by the jurisdictions in which they are implemented. Since public works programs are often started at the level of sub-national governments -- the precursor to NREGS was the state of Maharashtra's Employment Guarantee Scheme, started in early 1970s (Murgai & Ravallion 2005) -- there is a strong incentive to free-ride on a neighboring state or district's program given its spillover benefits. Conversely, if a state chooses to discontinue a public works program, district-level labor markets in neighboring states will also experience the end of spillover benefits. In either case, policymakers need to know the relevant parameters estimated in this paper to assess how their jurisdiction is likely to be impacted by public works programs in adjoining jurisdictions. The political economy dimension of these findings is that in periods of economic distress when the stabilization benefits of public works programs are most needed, government revenues also decline, thus making them less likely to be enacted given the benefit to waiting for a neighboring jurisdiction to start the program.

References

Allcott, H., & Keniston, D. (2014). “Dutch Disease or Agglomeration? The Local Economic Effects Of Natural Resource Booms In Modern America.” (No. w20508). National Bureau of Economic Research.

Angelucci, M. and G. De Giorgi (2009). “Indirect Effects of an Aid Program: How do Cash Injections Affect Ineligibles' Consumption?” *American Economic Review* 99(1), 486-508, March.

Angelucci, M., & Di Maro, V. (2015). “Programme Evaluation and Spillover Effects.” *Journal of Development Effectiveness*, (ahead-of-print), 1-22.

Azam, M. (2012). “The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment.” IZA Discussion Paper.

Baylis, K., & Ham, A. (2015, May). “How Important is Spatial Correlation in Randomized Controlled Trials?” In *2015 AAEA & WAEA Joint Annual Meeting, July 26-28, San Francisco, California* (No. 205586). Agricultural and Applied Economics Association & Western Agricultural Economics Association.

Bobba, M., & Gignoux, J. (2014). “Neighborhood Effects and Take-Up of Transfers in Integrated Social Policies: Evidence from Progresa.”

Berg, Erlend, S. Bhattacharya, R. Durgam & M. Ramachandra. (2012). “Can Rural Public Works Affect Agricultural Wages? Evidence from India.” Centre for the Study of African Economies, Oxford University. Working Paper WPS/2012-05.

Bhargava, A. K. (2014). “The Impact of India's Rural Employment Guarantee on Demand for Agricultural Technology.”

Duflo, Glannerster and Kremer (2007). “Using Randomization in Development Economics Research: A Toolkit.”

Dutta, P., Murgai, R., Ravallion, M., & Van de Walle, D. (2014). *Right to Work?: Assessing India's Employment Guarantee Scheme in Bihar*. World Bank Publications.

Glaeser, E. L. (2001). “The Economics of Location-based Tax Incentives.” Harvard Institute of Economic Research, Discussion Paper Number 1932.

Imbens, G. W. and D. Rubin. (2009). *Causal Inference in Statistics, and in the Social and Biomedical Sciences*. New York: Cambridge University Press.

Imbert, C. and J. Papp. (2014, August). “Seasonal Migration and Rural Workforce Programs: Evidence from India's Employment Guarantee.”

Imbert, C. and J. Papp (2014). “Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee.” *American Economic Journal - Applied Economics* (forthcoming).

Islam, M. and Sivasankaran, A. (2014). “How does Child Labor Respond to Changes in Adult Work Opportunities? Evidence from NREGA.”

Jacob, N. (2008). “The Impact of NREGS on Rural-Urban Migration: Field Survey Of Villupuram District, Tamil Nadu.” CCS Working Paper No. 202.

Kremer, M. and E. Miguel (2007). “The Illusion of Sustainability”, *Quarterly Journal of Economics*, 2007, 122(3), 1007-1065.

Li, T., & Sekhari, S. (2013). “The Unintended Consequences of Employment Based Safety Net Programs.”

Liu, Y., & Deininger, K. (2010). “Poverty Impacts of India's National Rural Employment Guarantee Scheme: Evidence from Andhra Pradesh.” *Selected Paper prepared for presentation at the Agricultural & Applied Economics Association*, 25-27.

Lipscomb, M. and A. M. Mobarak. (2008). “Decentralization and Water Pollution Spillovers: Evidence from the Redrawing of County Boundaries in Brazil.”

Mani, S., Behrman, J. R., Galab, S. and P. Reddy. (2014). “Impact of the NREGS on Schooling and Intellectual Human Capital.”

McKinnish, T. (2005). “Importing the Poor Welfare Magnetism and Cross-Border Welfare Migration.” *Journal of Human Resources*, *40*(1), 57-76.

Miguel, E., and M. Kremer. (2004). “Worms: Identifying Impacts on Education and Health in the Presence Of Treatment Externalities.” *Econometrica*, *72*(1), 159-217.

Miller, G. and A. M. Mobarak (2013). “Learning about New Technologies Through Social Networks: Experimental Evidence on Non-Traditional Stoves in Bangladesh,” *Marketing Science* R&R.

Mishra, P. (2007). “Emigration and Wages in Source Countries: Evidence from Mexico.” *Journal of Development Economics* 82.1: 180-199.

Mobarak, A. M. and M. R. Rosenzweig. (2013). “Informal Risk Sharing, Index Insurance, and Risk Taking in Developing Countries.” *American Economic Review*, 103(3): 375-80.

Moretti, E. (2011). “Local Labor Markets.” *Handbook of Labor Economics*, *4*, 1237-1313.

Murdoch, J. C., and T. Sandler. (2002). “Economic Growth, Civil Wars, and Spatial Spillovers.” *Journal of Conflict Resolution*, *46*(1), 91-110.

Murgai, R., & Ravallion, M. (2005). “Employment Guarantee in Rural India: What Would it Cost and How Much Would it Reduce Poverty?” *Economic and Political Weekly*, 3450-3455.

Oster, Emily and Rebecca Thornton. (2012). “Determinants of Technology Adoption: Private Value and Peer Effects in Menstrual Cup Take-Up,” *Journal of the European Economic Association*, December.

Raghunathan, K. and G. Fields (2014). “For Better or For Worse? The Effects of an Employment Guarantee in a Seasonal Agricultural Market.” IZA

Ranjan, A. and K. Bhatia (2009). “Alternative to Migration.” *Frontline 16.*

Ravallion, M., Van de Walle, D. P., Dutta, P., & Murgai, R. (2013). “Testing Information Constraints on India's Largest Antipoverty Program.” *World Bank Policy Research Working Paper*, (6598).

Roback, J. (1982). “Wages, Rents and the Quality of Life.” *Journal of Political Economy* 90 (December), 1257–1278.

Robertson, R. (2000). “Wage Shocks and North American Labor-Market Integration.” *American Economic Review*, 742-764.

Rosen, S. (1979). “Wage-based Indexes of Urban Quality of Life.” In: Miezkowski, Peter N., Straszheim, Mahlon R. (Eds.), Current Issues in Urban Economics. Johns Hopkins University Press, Baltimore, MD, pp. 74–104.

Shah, M., & Steinberg, B. M. (2015). “Workfare and Human Capital Investment: Evidence from India*”* (No. w21543). National Bureau of Economic Research.

Soundararajan, V. (2013, March). “Minimum Wages and Enforcement in India: Inverted U-Shaped Employment Effects.” In *8th IZA/World Bank Conference on Employment and Development*.

Srivastava, R. (2011). “Internal Migration in India: An Overview of its Features, Trends and Policy Challenges.” *Social and Human Sciences Sector*, UNICEF.

Subbarao, K. (1999). “Public Works as an Anti-Poverty Program: An Overview of Cross-Country Experience.” *American Journal of Agricultural Economics*. 84(2), 678–683 (1999)

Subbarao, K. (2003). “Systemic Shocks and Social Protection: Role and Effectiveness of Public Works Programs.” Social Protection, World Bank.

Subbarao, K., Del Ninno, C., Andrews, C., & Rodríguez-Alas, C. (2012). “Public Works as a Safety Net: Design, Evidence, and Implementation.” World Bank Publications.

Zimmermann, L. (2013, October). Why Guarantee Employment? Evidence from a Large Indian Public-Works Program.

Appendix

Figure 1: Program Rollout

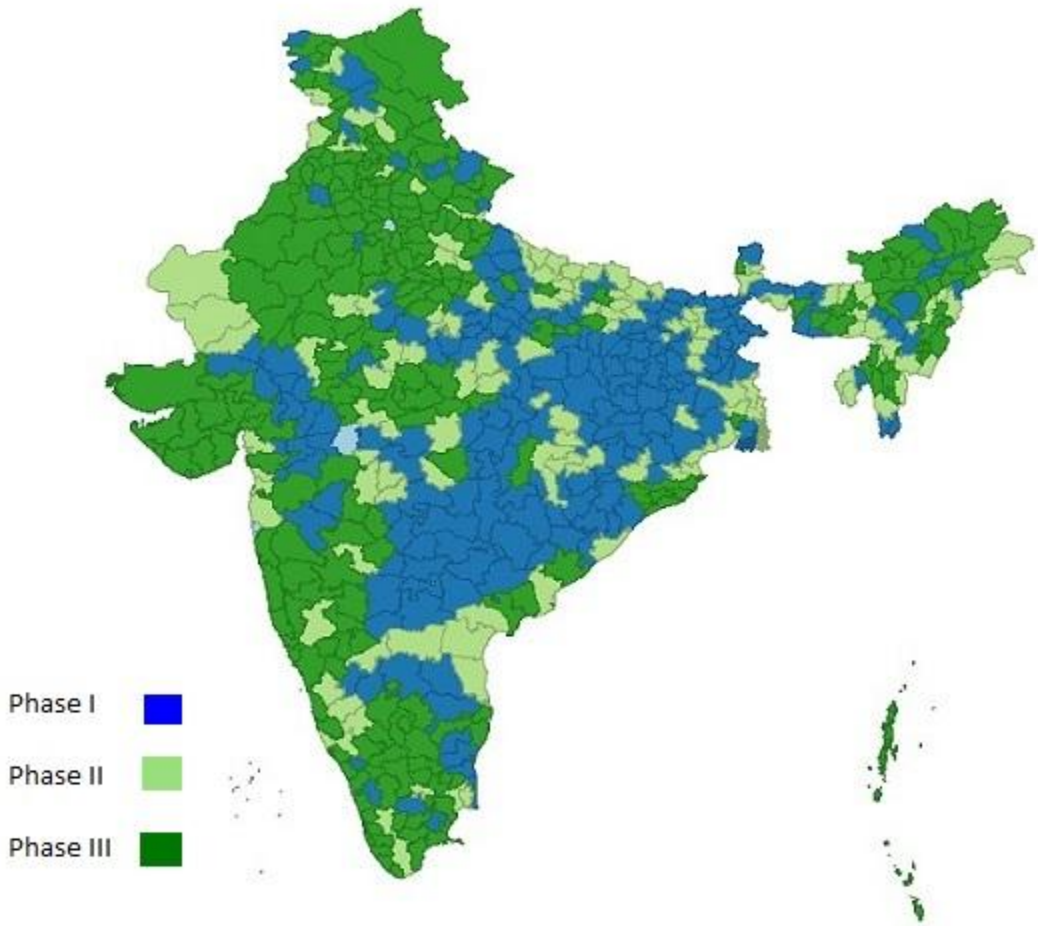


Figure 2: Early vs Late Districts

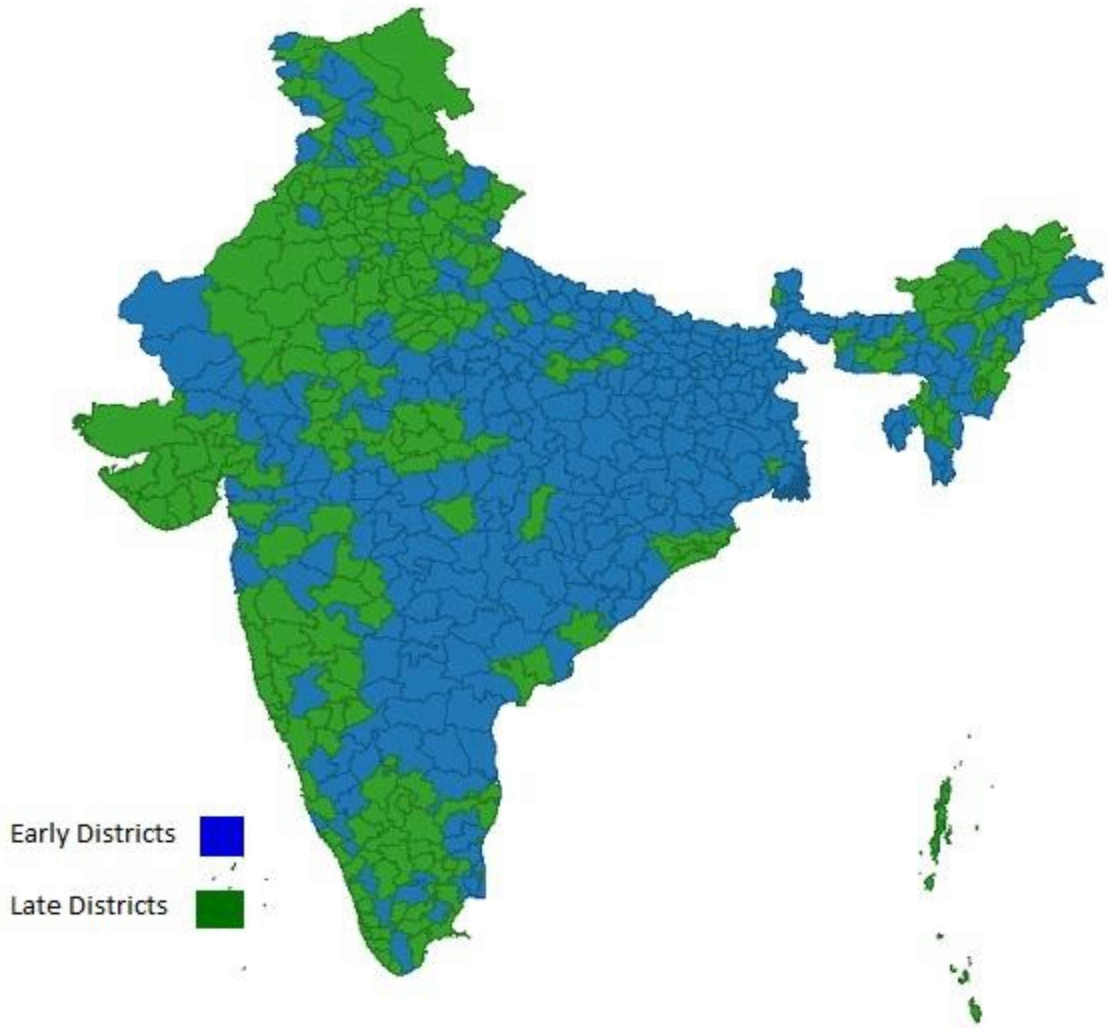
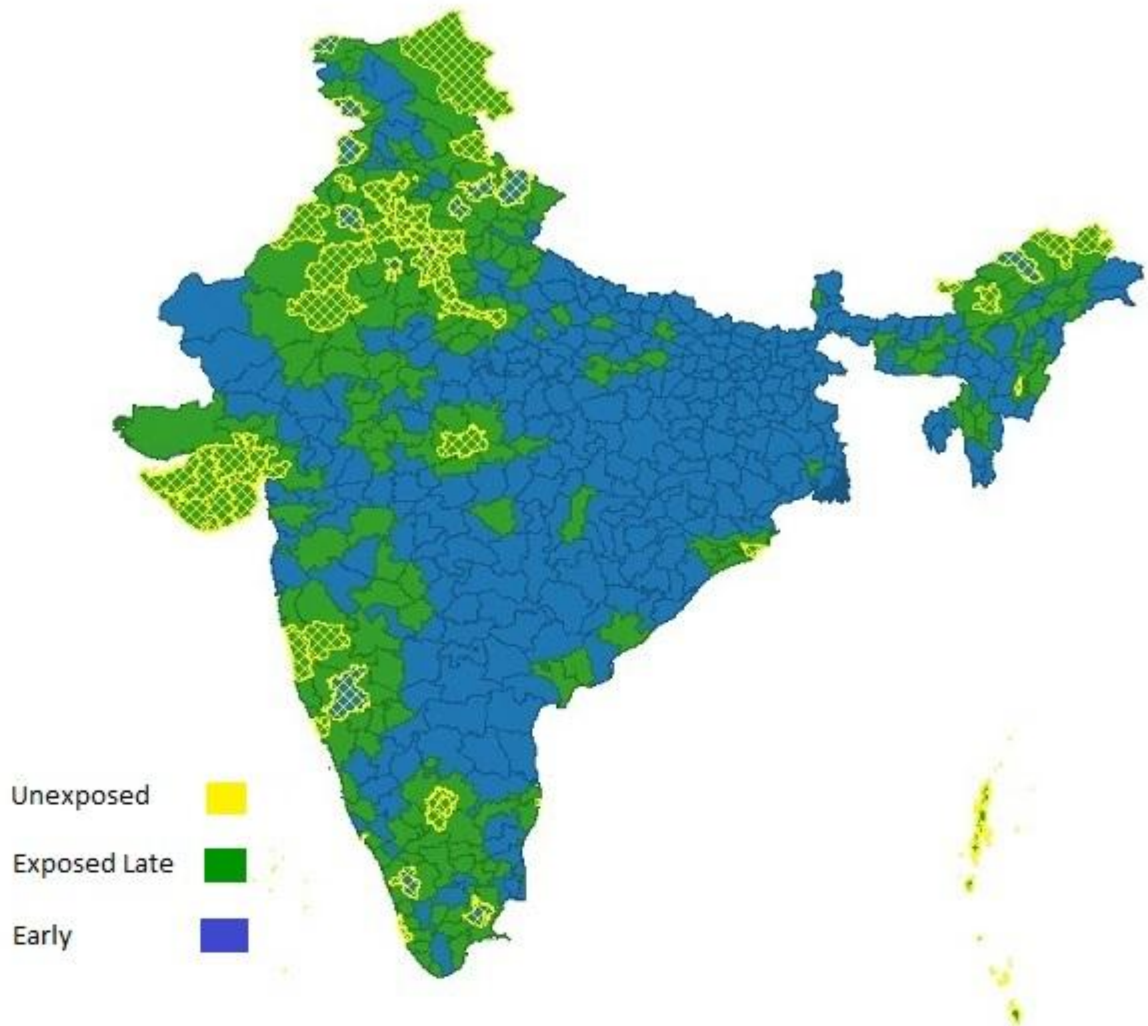
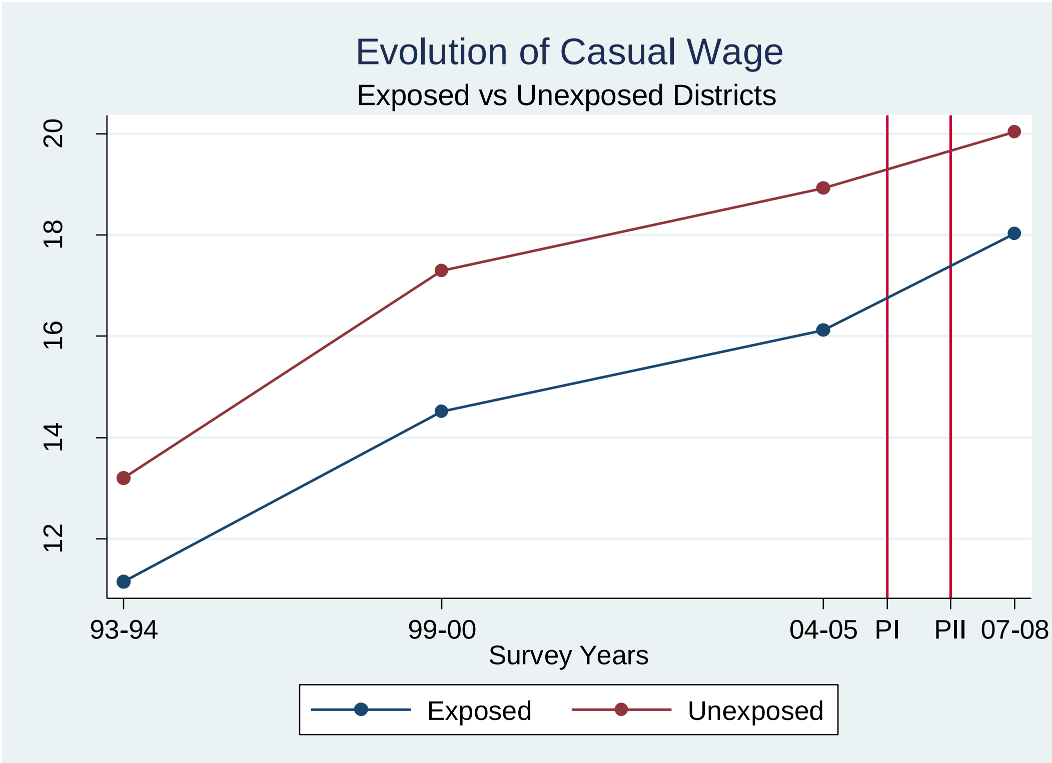


Figure 3: Exposure in Early and Late Phase Districts



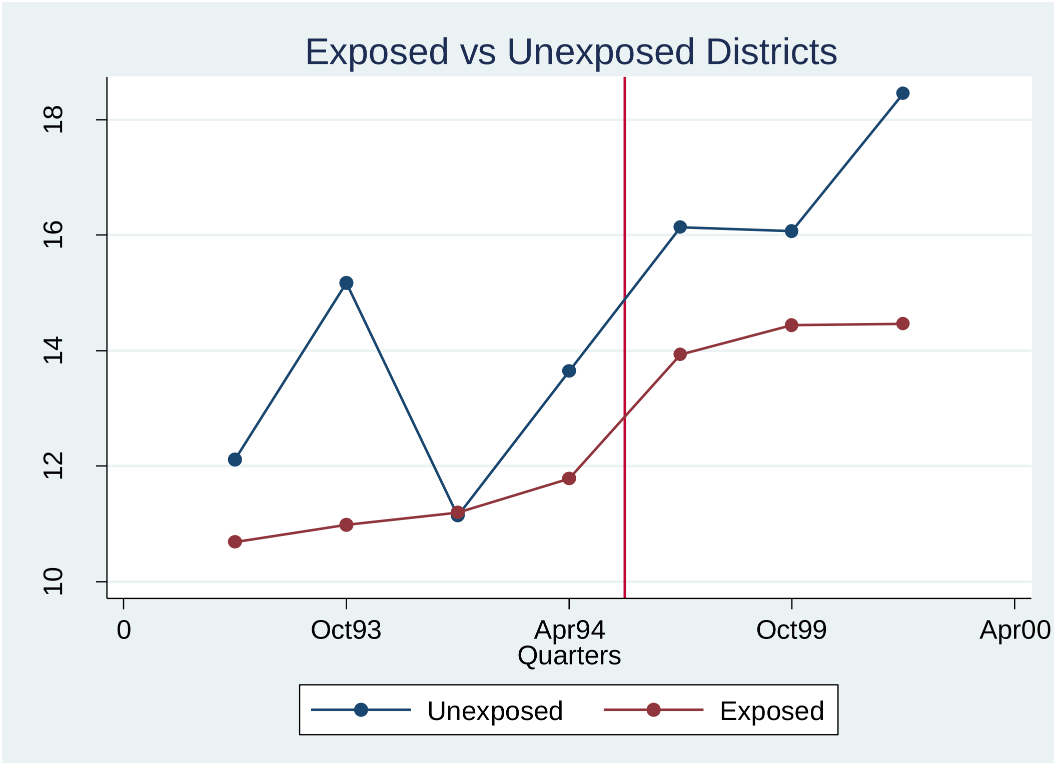
Note: The category of unexposed districts showed in this map includes Early districts, for the purpose of explication. The working sample used for analysis is restricted to Late districts.

Fig. 4: Annual Casual Wage Trends in Exposed and Unexposed Districts



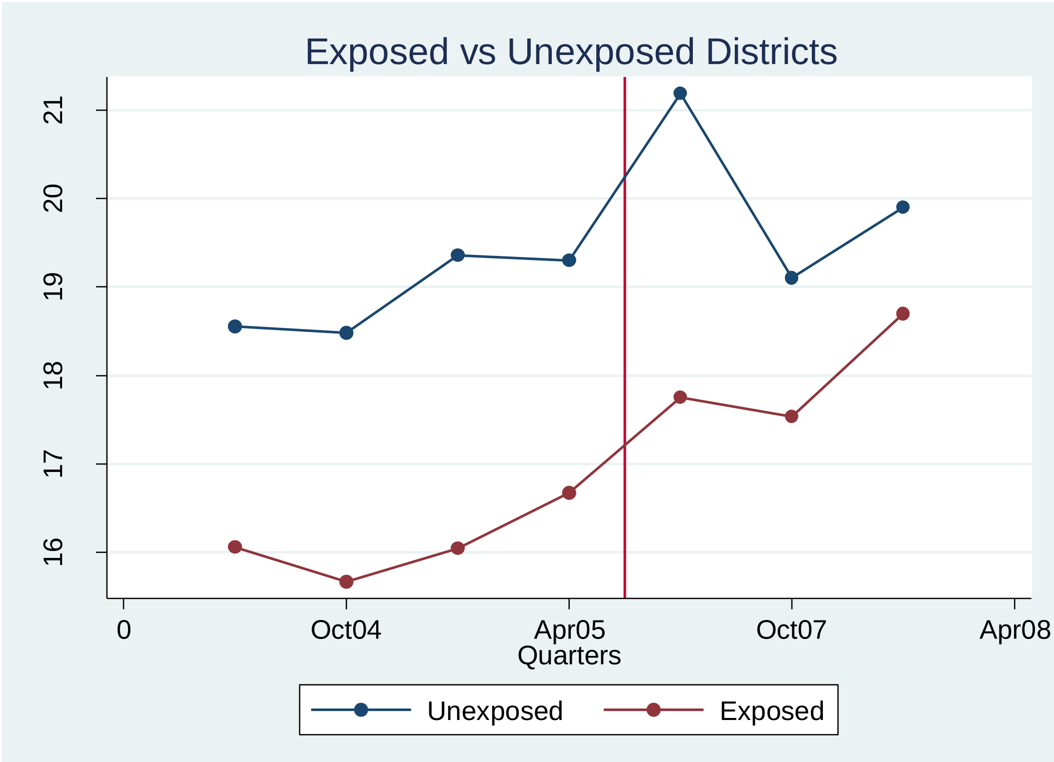
Note: Casual Wage is inflation-adjusted using the state-level Consumer Price Index for Agricultural Labor (CPI – AL) from the Indian Labour Bureau and reported in base year (1986-87) prices. The reference markers PI (Feb 2006) and PII (April 2007) indicate the first and second phase rollout dates of NREGS.

Figure 5: Quarterly Casual Wage and fake Exposure in Placebo Period



Note: Casual Wage is inflation-adjusted using the state-level Consumer Price Index for Agricultural Labor (CPI – AL) from the Indian Labour Bureau and reported in base year (1986-87) prices. The reference markers PI (Feb 2006) and PII (April 2007) indicate the first and second phase rollout dates of NREGS.

Figure 6: Quarterly Casual Wage and Exposure in Study Period



Note: Casual Wage is inflation-adjusted using the state-level Consumer Price Index for Agricultural Labor (CPI – AL) from the Indian Labour Bureau and reported in base year (1986-87) prices. The reference markers PI (Feb 2006) and PII (April 2007) indicate the first and second phase rollout dates of NREGS.

Figure 7: Seasonal Migration & Exposure Intensity in Late Districts

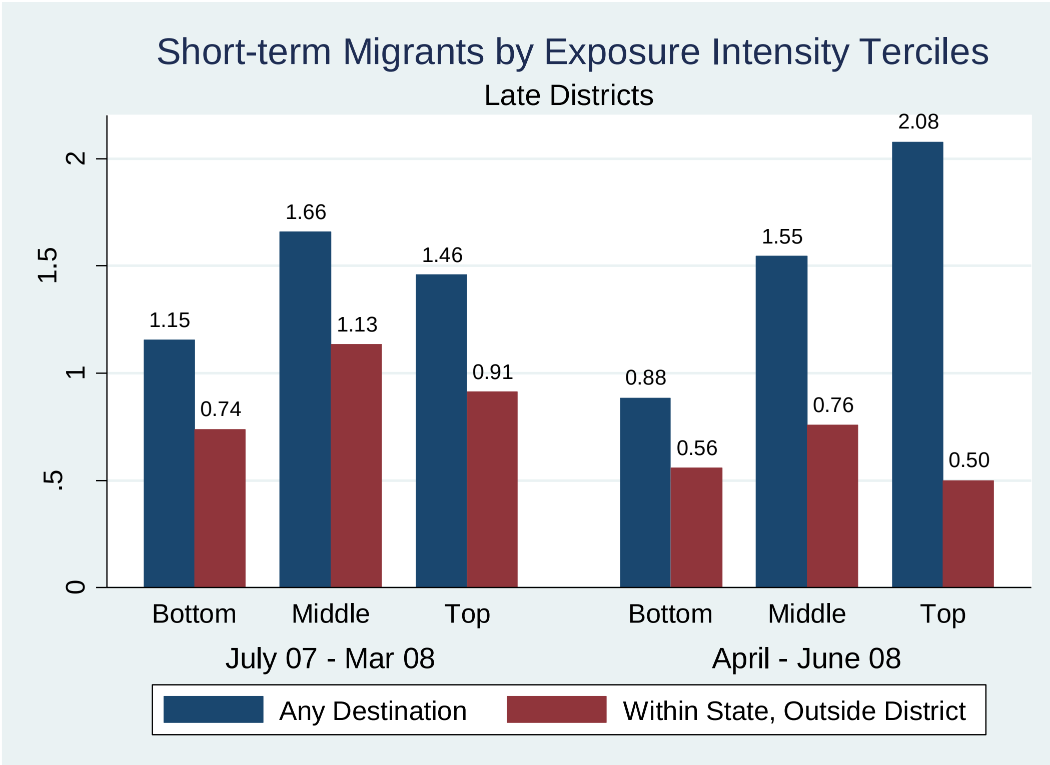


Table 1: NREGS Rollout

Phase	Phase I	Phase II	Phase III
Rollout date	Feb 2006	April 2007	April 2008
Number of Districts	200	130	261
Study Classification	Early	Early	Late

Table 2: NSS Survey Rounds & NREGS Rollout

Timing of Survey	Before Phase I	After Phase II	After Phase III
Round	NSS 61	NSS 64 (Sub-round 1–3)	NSS 64 (Sub-round 4)
Survey Year	July 2004 – June 2005	July 2007 – March 2008	April 2008 – June 2008
Number of Districts	NREGS = 0	NREGS = 330	NREGS = 588

Table 3: Exposure and Exposure Intensity– Late Districts

	Mean	S.D.
% Exposed	80.5	
% Unexposed	19.5	
Neighbors	5.5	1.6
Early Neighbors	1.9	1.6
Exposure Intensity	0.35	0.26
Exposure Intensity | Exposure	0.39	0.22

Table 4: Pre-Exposure Outcomes in Late Districts

	Unexposed	Exposed	
	Mean	Mean	p-value
Unemployment Days	0.28	0.37	0.049
NLFP Days	2.92	2.89	0.687
Work Days	3.80	3.74	0.512
Casual Days	0.87	0.93	0.501
Salaried Days	0.78	0.61	0.136
Casual Labor Wage (Real)	65.17	55.49	0.012
Salaried Labor Wage (Real)	110.74	108.63	0.678

Note: These estimates were computed using NSS 64 (July 2004 – June 2005). Casual Labor Wage and Salaried Labor Wage are reported in Rupees/Day (July 2004 prices)

Table 5: Pre-exposure Summary Statistics – Late Districts

Controls	Unexposed	Exposed	p-value	Source
Age	31.8	32.4	0.019	NSS 61
% Male	51.3	49.8	0.015	NSS 61
% Scheduled Caste (SC)	19.3	19.2	0.993	NSS 61
% Scheduled Tribe (ST)	1.3	4.9	0.000	NSS 61
% Other Backward Caste	40.7	45.9	0.071	NSS 61
% Others	38.7	29.9	0.004	NSS 61
% Literate	30.4	32.8	0.339	NSS 61
% Land Possessed < 1ha	72.4	77.1	0.130	NSS 61
Population	1,925,621	1,837,777	0.673	2001 Census
% Rural Population	64.4	71.3	0.026	2001 Census
Sex Ratio	906.7	939.9	0.007	2001 Census
% Literacy	59.8	59.3	0.783	2001 Census
% Female Literacy	59.4	58.7	0.746	2001 Census
% Male Literacy	80	79.7	0.860	2001 Census
% Scheduled Caste (SC)	2.5	4.2	0.236	2001 Census
% Scheduled Tribe (ST)	15.6	16.6	0.397	2001 Census
Sex Ratio (SC)	595.7	827.7	0.000	2001 Census
Sex Ratio (ST)	897.8	941.6	0.000	2001 Census
Household Size	5. 7	5.4	0.050	2001 Census
Number of Districts	42	173		

Note: The NSS 61 sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 2004 to June 2005, before the introduction of NREGS. The estimates for late phase district-level characteristics are taken from 2001 Census, which is representative of the entire district population. Sex ratio is reported as the number of women per 1000 men.

Table 6: Exposure induced Spillovers to Late Districts

	Log deflated casual wage	Log deflated salaried wage	Unemployment Days	NLFP Days	Work Days
Post*Exposure	0.087**	0.004	0.052	-0.051	-0.001
	(0.032)	(0.062)	(0.036)	(0.082)	(0.082)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	19,677	17,513	192,124	192,124	192,124

Note: Each column represents results from a separate regression. The sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 2004 to June 2005 and from July 2007 – March 2008. Unemployment days, non-labor force participation days and work days are calculated using time allocation responses for the week preceding interview. Log casual wage and log salaried wage is the log of earnings per day worked for people who report working in casual labor and salaried work respectively. Daily casual wage and salaried work wage are deflated using the monthly, state-level price index for agricultural laborers from the Indian Labour Bureau. Exposure is an indicator for whether district has at least one contiguous early-phase neighbor. Individual-level controls (age, age squared, gender, indicators for literacy, land possessed, and caste group), rural fraction of district population, year-quarter and district fixed effects are included in all regressions. All estimates are computed using sampling weights. Standard errors are clustered at district-level to control for intra-district correlation. ***Significant at 1% level, **Significant at 5% level, *Significant at 1% level.

Table 7: Exposure Intensity induced Spillovers to Late Districts

	Log deflated casual wage	Log deflated salaried wage	Unemployment Days	NLFP Days	Work Days
Post*Exposure	0.103**	-0.072	0.083	-0.212*	0.128
	(0.051)	(0.091)	(0.066)	(0.114)	(0.138)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	19,677	17,513	192,124	192,124	192,124

Note: Each column represents results from a separate regression. The sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 2004 to June 2005 and from July 2007 – March 2008. Unemployment days, non-labor force participation days and work days are calculated using time allocation responses for the week preceding interview. Log casual wage and log salaried wage is the log of earnings per day worked for people who report working in casual labor and salaried work respectively. Daily casual wage and salaried work wage are deflated using the monthly, state-level price index for agricultural laborers from the Indian Labour Bureau. Exposure is the ratio of the number of contiguous early-phase neighbor to number of contiguous neighbors. Individual-level controls (age, age squared, gender, indicators for literacy, land possessed, and caste group), rural fraction of district population, year-quarter and district fixed effects are included in all regressions. All estimates are computed using sampling weights. Standard errors are clustered at district-level to control for intra-district correlation. ***Significant at 1% level, **Significant at 5% level, *Significant at 1% level.

Table 8: Exposure induced Spillovers to Late Districts by Gender

	Log deflated casual wage	Unemployment Days	NLFP Days	Work Days	Casual Days
Post*Exposure*Women	0.093***	0.031	0.077	-0.108	0.044
	(0.034)	(0.035)	(0.096)	(0.098)	(0.063)
Post*Exposure*Men	0.085**	0.075*	-0.188**	0.113	0.169**
	(0.034)	(0.041)	(0.089)	(0.087)	(0.076)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	19,677	192,124	192,124	192,124	192,124

Note: Each column represents results from a separate regression. The sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 2004 to June 2005 and from July 2007 – March 2008. Unemployment days, non-labor force participation days and work days are calculated using time allocation responses for the week preceding interview. Log casual wage and log salaried wage is the log of earnings per day worked for people who report working in casual labor and salaried work respectively. Daily casual wage and salaried work wage are deflated using the monthly, state-level price index for agricultural laborers from the Indian Labour Bureau. Exposure is an indicator for whether district has at least one contiguous early-phase neighbor. Individual-level controls (age, age squared, gender, indicators for literacy, land possessed, and caste group), rural fraction of district population, year-quarter and district fixed effects are included in all regressions. All estimates are computed using sampling weights. Standard errors are clustered at district-level to control for intra-district correlation. ***Significant at 1% level, **Significant at 5% level, *Significant at 1% level.

Table 9: Exposure induced Spillovers in pre-study period (Placebo Test)

	Log deflated casual wage	Log deflated salaried wage	Unemployment Days	NLFP Days	Work Days
Post-1994*Exposure	-0.045	-0.023	0.075*	0.013	0.183
	(0.137)	(0.064)	(0.043)	(0.147)	(0.221)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	12,936	8,423	165,267	165,267	165,267

Note: Each column represents results from a separate regression. The sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 1993 to June 1994 and from July 1999 – March 2000. Unemployment days, non-labor force participation days and work days are calculated using time allocation responses for the week preceding interview. Log casual wage and log salaried wage is the log of earnings per day worked for people who report working in casual labor and salaried work respectively. Daily casual wage and salaried work wage are deflated using the monthly, state-level price index for agricultural laborers from the Indian Labour Bureau. Exposure is an indicator for whether district has at least one contiguous early-phase neighbor. Individual-level controls (age, age squared, gender, indicators for literacy, land possessed, and caste group), rural fraction of district population, year-quarter and district fixed effects are included in all regressions. All estimates are computed using sampling weights. Standard errors are clustered at district-level to control for intra-district correlation. ***Significant at 1% level, **Significant at 5% level, *Significant at 1% level.

Table 10: Exposure Intensity induced Spillovers in pre-study period (Placebo Test)

	Log deflated casual wage	Log deflated salaried wage	Unemployment Days	NLFP Days	Work Days
Post-1994*Exposure	0.044	0.104	0.021	0.222	-0.132
	(0.132)	(0.152)	(0.077)	(0.190)	(0.286)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	12,936	8,423	165,267	165,267	165,267

Note: Each column represents results from a separate regression. The sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 1993 to June 1994 and from July 1999 – March 2000. Unemployment days, non-labor force participation days and work days are calculated using time allocation responses for the week preceding interview. Log casual wage and log salaried wage is the log of earnings per day worked for people who report working in casual labor and salaried work respectively. Daily casual wage and salaried work wage are deflated using the monthly, state-level price index for agricultural laborers from the Indian Labour Bureau. Exposure is an indicator for whether district has at least one contiguous early-phase neighbor. Individual-level controls (age, age squared, gender, indicators for literacy, land possessed, and caste group), rural fraction of district population, year-quarter and district fixed effects are included in all regressions. All estimates are computed using sampling weights. Standard errors are clustered at district-level to control for intra-district correlation. ***Significant at 1% level, **Significant at 5% level, *Significant at 1% level.

Table 11: Exposure induced Spillovers to Late Districts in Low Coverage States

	Log deflated casual wage	Log deflated salaried wage	Unemployment Days	NLFP Days	Work Days
Post*Exposure	0.135***	-0.001	0.015	0.066	-0.081
	(0.042)	(0.080)	(0.051)	(0.114)	(0.124)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	8,029	8,013	69,352	69,352	69,352

Note: Each column represents results from a separate regression. The sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 2004 to June 2005 and from July 2007 – March 2008. Unemployment days, non-labor force participation days and work days are calculated using time allocation responses for the week preceding interview. Log casual wage and log salaried wage is the log of earnings per day worked for people who report working in casual labor and salaried work respectively. Daily casual wage and salaried work wage are deflated using the monthly, state-level price index for agricultural laborers from the Indian Labour Bureau. Exposure is an indicator for whether district has at least one contiguous early-phase neighbor. Individual-level controls (age, age squared, gender, indicators for literacy, land possessed, and caste group), rural fraction of district population, year-quarter and district fixed effects are included in all regressions. All estimates are computed using sampling weights. Standard errors are clustered at district-level to control for intra-district correlation. ***Significant at 1% level, **Significant at 5% level, *Significant at 1% level.

Table 12: Exposure Intensity induced Spillovers to Late Districts in Low Coverage States

	Log deflated casual wage	Log deflated salaried wage	Unemployment Days	NLFP Days	Work Days
Post*Exposure	0.183*	-0.020	0.031	-0.079	0.048
	(0.104)	(0.091)	(0.083)	(0.218)	(0.238)
Controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	8,029	8,013	69,352	69,352	69,352

Note: Each column represents results from a separate regression. The sample consists of late phase individuals between the ages of 15 and 59, interviewed from July 2004 to June 2005 and from July 2007 – March 2008. Unemployment days, non-labor force participation days and work days are calculated using time allocation responses for the week preceding interview. Log casual wage and log salaried wage is the log of earnings per day worked for people who report working in casual labor and salaried work respectively. Daily casual wage and salaried work wage are deflated using the monthly, state-level price index for agricultural laborers from the Indian Labour Bureau. Exposure is the ratio of the number of contiguous early-phase neighbor to number of contiguous neighbors. Individual-level controls (age, age squared, gender, indicators for literacy, land possessed, and caste group), rural fraction of district population, year-quarter and district fixed effects are included in all regressions. All estimates are computed using sampling weights. Standard errors are clustered at district-level to control for intra-district correlation. ***Significant at 1% level, **Significant at 5% level, *Significant at 1% level.