Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee *

Clément Imbert † and John Papp ‡

November 7, 2014

This version: September 2014. First version: October 2011.
Accepted for publication in the *American Economic Journal: Applied Economics*

Abstract

We estimate the effect of a large rural workfare program in India on private employment and wages by comparing trends in districts that received the program earlier relative to those that received it later. Our results suggest that public sector hiring crowds out private sector work and increases private sector wages. We compute the implied welfare gains of the program by consumption quintile. Our calculations show that the welfare gains to the poor from the equilibrium increase in private sector wages are large in absolute terms and large relative to the gains received solely by program participants.

*Thanks to Abhijit Banerjee, Robin Burgess, Anne Case, Denis Cogneau, Angus Deaton, Dave Donaldson, Esther Duflo, Erica Field, Maitreesh Ghatak, Marc Gurgand, Reetika Khera, Rohini Pande, Martin Ravallion, Sandip Sukhtankar and Dominique Van de Walle, as well as numerous seminar and conference participants for very helpful comments. Clément Imbert acknowledges financial support from CEPREMAP and European Commission (7th Framework Program). John Papp gratefully acknowledges financial support from the Fellowship of Woodrow Wilson Scholars at Princeton.

†Oxford University, Department of Economics, Manor Road, Oxford, UK, clement.imbert@economics.ox.ac.uk.

‡johnhpapp@gmail.com

1 Introduction

Recent studies have shown that policy interventions in developing countries have important effects on non-participants. Food distribution policies affect consumer prices (Jayachandran et al., 2010), and direct cash transfers can increase the consumption of non-beneficiaries through risk-sharing networks (Angelucci and Giorgi, 2009). The impact of policy interventions on labor market equilibrium, however, has received little attention. This is despite the fact that short-term manual labor ("casual labor") is an important income source for the poor (Banerjee and Duflo, 2007) and that even non-labor market interventions such as cash transfers and infrastructure creation have been shown to have important impacts on labor supply (Ardington et al., 2009; Dinkelman, 2011).

The first objective of this paper is to document the differential trends in wages and employment within districts where the flagship Indian anti-poverty program, the National Rural Employment Guarantee Act (NREGA), was rolled out first relative to those districts where the program was introduced later. We focus on how these differential trends may be used to estimate the causal impact of the program. The second objective is to use the resulting estimates along with a model of rural labor markets to calculate how the welfare gains from the program are distributed across the population. We compare the gains due to the estimated equilibrium rise in wages to the gains due solely to participation in the program for poor and rich households.

The NREGA provides a particularly good opportunity to study the labor market impacts of a large workfare program. Started in 2006, the NREGA provides short-term manual work at a wage comparable to or higher than the market rate. According to government administrative data, in 2010-11 the NREGA provided 2.3 billion person-days of employment to 53 million households making it the largest workfare program in operation today.[1] Further, the program was introduced at the district-level, an administrative unit large enough to reasonably approximate a distinct labor market (Rosenzweig, 1978; Topalova, 2010).

Assessing the labor market impact of large-scale policy interventions is complicated by the fact that a plausible counterfactual for areas affected by the interventions rarely exists and by the fact that even large-scale programs are often introduced within an area too small to be considered a distinct labor market.[2] We exploit the fact that the program was introduced gradually throughout India starting with the poorest districts in early 2006 and

[1] Figures are from the official NREGA website `nrega.nic.in`.

[2] The well-studied Mexican Progresa program for example was rolled out at the village level (Angelucci and Giorgi, 2009).

extending to the entire country by mid-2008. We document changes in employment and wages in districts that received the program between April 2006 and April 2007 relative to those that received it after April 2008.

We show that the introduction of the workfare program is correlated with a substantial increase in low-wage, low-skilled public employment and a roughly equivalent fall in private sector work (waged, self-employed or domestic work). We find daily wages of casual laborers increase in early districts relative to late districts. The estimated 4.7% increase in wages and 1.5% decrease in private employment imply a 0.31 elasticity of labor demand, which is consistent with existing evidence (Binswanger and Evenson, 1980).

Given that poorer districts were more likely to be selected as early phase districts, the differential changes in employment and wages that we document may in part reflect differential trends in early and late phase districts unrelated to the program. A number of facts suggest the employment and wage results are indeed due to the program. First, these results are concentrated during the agricultural off-season when most NREGA employment is provided. Second, the results are concentrated in the seven states that account for the majority of employment generated by the program (we call them "star" states in the analysis). Third, the results are robust to controlling for district characteristics (including early phase selection criteria) which could predict changes in labor market outcomes. Though we find evidence of a positive trend in casual wages in early relative to late districts before the program was implemented, this trend disappears once we control for district characteristics, which suggests that our specification effectively deals with selection. The results are also robust to controlling for pre-program changes in outcomes for each district and for state-level time effects.

Our second objective is to use the wage and employment estimates combined with household-level data on consumption, casual labor supplied, and casual labor hired to calculate how the welfare gains from the increase in wages are distributed across rural households. We show the rise in wages redistributes income from richer households (net buyers of labor) to poorer households (net suppliers of labor). We then use individual-level data on program wages and participation to estimate the magnitude of the direct gains from participation in the program. Our estimates show that for households in the bottom three consumption quintiles, the estimated welfare gain due to the wage change represents 31% of the total welfare gain from the program. Further we find that households in the richest quintile are actually made worse off by the program as a result of the increase in wages. Our calculations also suggest that the NREGA cost of 241 Rs per rural household is larger than the welfare gains even to households in the poorest quintile (112 Rs). A complete cost-benefit analysis, however, is beyond the scope of this paper, as it would include the productivity benefits of NREGA infrastructures.

The results contribute to the literature in three ways. First, we document the wage and employment trends surrounding the implementation of a particular, widely adopted anti-poverty policy. Government hiring by public works programs may crowd out private sector work and therefore lead to a rise in equilibrium private sector wages (Ravallion, 1987; Basu et al., 2009). The empirical evidence, however, on the equilibrium impacts of workfare programs is limited. The few existing studies include two concurrent studies, which present evidence consistent with the NREGA raising unskilled wages (Azam, 2012; Berg et al., 2013). Zimmermann (2013) finds no significant impact on wages, but confidence intervals are sufficiently large to include our own estimate.

Second, we modify the theoretical framework presented in Deaton (1989) and Porto (2006) in order to quantify the extent to which the equilibrium impacts on the labor market both benefit and hurt different segments of the population. This framework allows us to estimate the welfare impact of a policy using empirical estimates of its aggregate effect on wages. A similar methodology could be used to assess the equilibrium impacts of other policy interventions in developing countries which affect labor supply (Ardington et al., 2009; Dinkelman, 2011).

Finally, the results contribute to the literature on the structure and functioning of labor markets in developing countries (Rosenzweig, 1980; Behrman, 1999) as well as the broader literature that uses the impact of policy interventions to infer how markets operate (Card and Krueger, 1992). Specifically, the rise in casual wages following the implementation of the program is hard to reconcile with a naïve model of "surplus labor" in which self-employed members of poor households could be hired with no effect on private sector wages (Sen, 1966). Our findings are in line with Rosenzweig and Foster (2010)'s recent argument that surplus labor in rural India is not due to a lack of employment opportunities.

The following section describes the workfare program in more detail. Section 3 proposes a model of rural labor markets which provides a framework for estimating the distributional effects of the program. Section 4 presents our data and empirical strategy. Section 5 presents the main empirical results. Section 6 uses these results to estimate the welfare gains due to the program and Section 7 concludes.

2 The Workfare Program

The National Rural Employment Guarantee Act (NREGA), passed in September 2005, entitles every household in rural India to 100 days of work per year at a state-level minimum wage. In 2010-11 the NREGA provided 2.3 billion person-days of employment to 53 million households.[3] The India-wide budget was Rs. 345 billion (7.64 billion USD) representing 0.6% of GDP. The act was gradually introduced throughout India starting with 200 of the poorest districts in February 2006, extending to 130 districts in April 2007, and to the rest of rural India in April 2008. The NREGA sets out guidelines detailing how the program is to be implemented in practice. Whether and how these guidelines are actually followed varies widely from state to state and even from district to district (Sharma, 2009; Dreze and Khera, 2009; Institute of Applied Manpower Research, 2009; The World Bank, 2011). Drawing from existing field studies, we provide an overview of how the act operates in practice.

2.1 Poverty Reduction through Employment Generation

One of the chief motivations for the act is poverty reduction through employment generation. In this respect, the NREGA follows a long history of workfare programs in India (see Appendix Section A.1). Although a nominal goal of the act is to generate productive infrastructure, The World Bank (2011) writes, "the objective of asset creation runs a very distant second to the primary objective of employment generation...Field reports of poor asset quality indicate that [the benefit from assets created] is unlikely to have made itself felt just yet." Indeed, the act explicitly bans machines from worksites and limits material, capital and skilled wage expenditure to 40% of total expenditure. Wages paid for unskilled work are borne entirely by the central government while states can pay at most 25% of expenditure on materials, capital and skilled wages. Together, these restrictions create a strong incentive to select projects that require mainly unskilled labor.

Well-targeted, well-implemented NREGA projects may have had a positive impact on agricultural productivity in the medium run. Deininger and Liu (2013) find that in Andhra Pradesh low caste households' income increased three years after irrigation works were made on their land. The evidence, however, reviewed by The World Bank (2011) suggest that NREGA infrastructures did not substantially improve productivity in the initial years of program implementation. Our study focuses on these initial years.

[3] Figures are from the official NREGA website `nrega.nic.in`.

2.2 Short-term, Unskilled Jobs

The work generated by the program is short-term, unskilled, manual work such as digging and transporting dirt. Households with at least one member employed under the act in agricultural year 2009-10 report a mean of only 38 days of work and a median of 30 days for *all* members of the household during that year. The jobs provided by the program are similar to private sector casual labor jobs. In fact, India's National Sample Survey Office (NSSO), which collects the main source of data used in this paper, categorizes employment under the NREGA as a specific type of casual labor. Out of those who report working in public works in the past week, 45% report they usually or sometimes engage in casual labor, while only 0.2% report that they usually or sometimes work in a salaried job.[4] The similarity of these public sector jobs and casual labor jobs motivates our focus on casual wages in the empirical analysis.

2.3 Wages and Payment

Wage rates are set at the state level, and NREGA workers are either paid a piece-rate or a fixed daily wage. Under the piece-rate system, which is more common, workers receive payment based on the amount of work completed (e.g. volume of dirt shoveled). The resulting daily earnings are almost always below the state-set wage levels. Theft by officials reduces the actual payment received.[5] Despite the fact that actual daily earnings often fall short of stipulated wage rates, NREGA work appears to be more attractive than similar private sector work available to low-skill workers. Based on a nationally representative India-wide survey during the agricultural year 2007-08, both male and female workers report earning an average of Rs. 79 per day for work under the act.[6] Reported earnings are 12% higher than the average daily earnings for casual workers (National Sample Survey Office, 2010). These figures may actually understate the attractiveness of NREGA work for the typical rural worker if search costs or other frictions drive the private sector wage rate above the marginal value of time (Walker and Ryan, 1990).

[4] Authors' calculations based on NSS Round 66 Employment and Unemployment Survey. The Employment surveys are described in detail in Section 4.1.

[5] Based on a survey in the state of Orissa of 1499 individuals who show up as working in the government administrative data, only 821 both exist and report having worked (Niehaus and Sukhtankar, 2013a). Of these 821, most received less than the stipulated minimum wage.

[6] Authors' calculations based on NSS Employment and Unemployment Survey Round 64.

2.4 Employment, Rationing and Awareness

Perhaps a more direct way to assess whether NREGA work is more attractive than available work is to ask people. The studies that ask find high levels of unmet demand (Dreze and Khera, 2009; Imbert and Papp, 2014). Although the act stipulates a minimum employment guarantee of 100 days of work per household per year, actual employment falls well short of the 100 day guarantee, even for households that report wanting to work the full 100 days.

One may naturally wonder, if the act guarantees 100 days and households want 100 days, why workers do not simply demand 100 days of work. However, as The World Bank (2011) summarizes: "In practice, very few job card holders formally apply for work while the majority tend to wait passively for work to be provided." Even those who demand work are not guaranteed work. During agricultural year 2009-10, an estimated 19% of households reported attempting to get work under the act without success.[7] The rationing of demand for NREGA work is one reason that across Indian states, the number of NREGA days provided is only weakly correlated with poverty (Dutta et al., 2012).

2.5 Seasonality and Cross-State Variation in Implementation

The above generalizations mask considerable state and even district variation in the implementation of the program (Dreze and Khera, 2009; Dreze and Oldiges, 2009). Figure 2 shows cross-state differences in public employment provision, as measured by the fraction of days spent on public works by rural adults in 2007-08 according to National Sample Survey. Consistent with anecdotal evidence and administrative data, seven states are top performers: rural adults in Andhra Pradesh, Chattisgarh, Himachal Pradesh, Madhya Pradesh, Rajasthan, Uttarkhand and Tamil Nadu spend more than 1% of the year on public works. Throughout the paper, we will refer to these seven states as “star states”. Differences in NREGA implementation are explained by some combination of political will, existing administrative capacity, and previous experience in providing public works (see Appendix A.2).

Public employment provision is also highly seasonal. Local governments start and stop works throughout the year, with most works concentrated during the first two quarters of the year prior to the monsoon. The monsoon rains make construction projects difficult to undertake, which is likely part of the justification. Field reports, however, document government attempts to stop works during the rainy season so they do not compete with the labor needs of farmers (Association for Indian Development, 2009). Figure 3 illustrates the

[7] Authors’ calculations using NSS Employment and Unemployment Survey Round 66.

variation in public employment provision between the dry and the rainy season. It shows the fraction of days spent on public works by rural adults in each quarter of 2007-08, according to National Sample Survey. Rural adults spend on average 1.5% of their time on public works during the first six months of the year, and less than 0.5% during the last six months, when the monsoon rains have come.

3 Model

We present a model to clarify how an increase in public sector hiring will impact aggregate employment and wages. We use the framework to trace out the equilibrium distributional impact of a workfare program across households. The model draws heavily from Deaton (1989) and Porto (2006) except that we focus on the labor market rather than the market for consumption goods.

In rural labor markets in developing countries, where a large part of the labor force is engaged in self-employment or domestic work, the opportunity cost of time may be lower than the market wage (Datt and Ravallion, 1994). The framework we use for calibration allows each household's opportunity cost of time to be less than the market wage.

3.1 Households

Consider an economy consisting of a continuum of households indexed by i. Household i operates a production function $F_i(D_i)$ where D_i is labor used (demanded) by the household. We assume households differ in their production function by a productivity factor A_i, so that for each household $F_i(D_i) = A_i G(D_i)$, with $G'(\cdot) > 0$ and $G''(\cdot) < 0$. $A_i \in [\underline{A}, \overline{A}]$ reflects differences in productive assets owned by the households (e.g. land), which we consider as exogenous.

Households have utility function $u(c_i, l_i)$ over household consumption c_i and leisure l_i. We assume the function is increasing and concave in both arguments. Let y_i denote non-labor income and π_i profits from home production. Let $\widetilde{W_i}$ be the shadow wage, i.e. the price of labor for household i, which could be lower than the market wage W. Let L_i^s denote household total labor supply and D_i denote household total labor demand. Households choose L_i^s, D_i and c_i to solve:

$$\max_{c_i, L_i^s, D_i} u(c_i, T - L_i^s)$$
$$\text{s. t. } c_i = y_i + \widetilde{W}_i L_i^s$$
$$y_i = \pi_i = A_i G(D_i) - \widetilde{W}_i D_i$$

The solution to this problem depends on the productivity factor A_i and the shadow price of labor $\widetilde{W}_i$. Let us first consider the case where labor markets are perfect and the market wage is the relevant price of labor for all households ($\widetilde{W}_i = W \ \forall i$). Household production and labor supply decisions are separable, and households equalize the marginal productivity of labour to the market wage: $A_i G'(D_i^*) = W$. The most productive households (e.g. large landholders), with high A_i, are net buyers of labor ($D_i^* > L_i^{*s}$) and the least productive ones, with low A_i, are net sellers of labor ($D_i^* < L_i^{*s}$).

Now suppose due to labor market frictions (e.g. job search costs) a wedge $p \in [0, 1]$ exists between the returns to one unit of wage labor for workers (pW) and its costs for employers (W). In this case, high productivity households are net labor buyers and set $A_i G'(D_i^*) = W$ while low productivity households are net labor sellers and set $A_i G'(D_i^*) = pW$. Households with intermediate productivity levels do not participate in the market and set $A_i G'(D_i^*) \in [pW, W]$. This model makes clear that the opportunity cost of time may be lower than the market wage for poorer households as in Benjamin (1992).[8]

3.2 Public works

Now suppose the government hires workers for public works projects. Motivated by the evidence on rationing of public works employment discussed in Section 2.4, we assume the government provides public works employment at wage $W_g > W$. The government must therefore determine the amount of employment to provide to each household, denoted by L_i^g. Total public employment provided is $L^g = \int_i L_i^g di$.

Throughout, we will assume that households use the shadow wage as the relevant opportunity cost of time, rather than the government wage. This will be the case as long as households that work in public works spend some time working on their own farm or on others' farms. Given that periods of public works employment for the typical worker are

[8] For simplicity, we abstract from differences in family size across households by assuming that total time is the same for all households. However, if separability does not hold then family size will affect the amount of labor used on the farm (see for example Benjamin (1992)).

quite short (often under thirty days per year), this assumption seems reasonable.

The household's maximization problem remains the same except for the additional source of income from public employment:

$$\begin{aligned}
\max_{c_i, L_i^s, D_i} & \; u(c_i, T - L_i^s) \\
\text{s. t. } c_i &= y_i + \widetilde{W}_i L_i^s \\
y_i &= \pi_i + (W_g - \widetilde{W}_i) L_i^g \\
\pi_i &= A_i G(D_i) - \widetilde{W}_i D_i
\end{aligned}$$

Because of the assumption that public employment is rationed and that the shadow wage is the relevant opportunity cost of time, the government wage from public sector work W_g only enters through its impact on non-labor income.

Public hiring (a change in L_g) reduces labor supply to the private sector, and therefore wages must rise to equate supply and demand. This argument is straightforward if labor markets are perfect ($p = 1$), and also applies to the case with search frictions ($p < 1$). In both cases, private employment (the sum of wage employment and self-employment) falls. Appendix A.3.4 presents the formal derivation.

This result contrasts with some versions of "surplus labor" models in which low productivity households are isolated from the market, and hence the government can hire workers from these households without any effect on the market wage (Sen, 1966; Rosenzweig, 1988). An increase in private sector wages caused by the employment guarantee would be inconsistent with the predictions of these models.

Models of imperfect competition also yield different conclusions. If employers have market power (Binswanger and Rosenzweig, 1984) then government hiring may actually increase private sector wages *and* employment (Basu et al., 2009). Our empirical analysis of the effect of the employment guarantee on employment and wages will allow us to directly test this prediction.

In the special case where $p = 1$, we can compute the elasticity of labor demand as the ratio of the percentage change in the wage divided by the percentage change in employment. See Appendix A.3.4 for details.

3.3 Impact on Household Welfare

Let the expenditure function corresponding to the dual of the utility maximization problem above be given by $e(\widetilde{W}_i, u_i)$. The expenditure function gives the total income required to achieve utility level u_i given a shadow wage $\widetilde{W}_i \in [pW, W]$. Since this is a one-period model, expenditure equals income, so we can write:

$$e(\widetilde{W}_i, u_i) \quad = \quad \pi_i(\widetilde{W}_i) + \widetilde{W}_i T + (W_g - \widetilde{W}_i)L_i^g + z_i \tag{1}$$

where z_i is exogenous income, $e(\widetilde{W}_i, u_i)$ is the expenditure or total income required to achieve utility level u_i and $\pi_i(\widetilde{W}_i) + \widetilde{W}_i T + (W_g - \widetilde{W}_i)L_i^g + z_i$ is total income.

A change in L_g may have two effects for household i. First, depending on the allocation rule, it may increase L_i^g, the time spent on public works by members of the household. Second, as we discussed in the previous section an increase in government hiring may increase the market wage W and hence the shadow wage $\widetilde{W}_i$.

For fixed z_i, when L_g changes, Equation 1 will no longer hold because the expenditure required to achieve the same utility will change and because the household's available income will change. Appendix A.3.3 derives the change in z_i required to maintain equality in Equation 1 and therefore maintain the same utility level, following a small change in L_g:

$$\begin{aligned} -dz_i \quad &= \quad (L_i^s - L_i^g - D_i)\widetilde{W}_i \frac{dW/W}{dL_g} + (W_g - \widetilde{W}_i)dL_i^g \\ &= \quad \textit{Net Casual Labor Earnings} \times \frac{dW/W}{dL_g} + (W_g - \widetilde{W}_i)\frac{dL_i^g}{dL_g} \end{aligned} \tag{2}$$

We interpret $-dz_i$ as the amount of money that a social planner would have to take from household i in order for the household to have the same level of utility before and after the implementation of the program. In this sense, it is a measure of the welfare effect of the program and is often referred to as the compensating variation (Porto, 2006).[9]

[9] The impact on welfare is not the same as the impact on consumption. In Appendix A.3.5, we derive the impact of the program on consumption of household i. The key difference compared with Equation 2 is that the impact on consumption includes the change in consumption due to the change in labor supply from the change in income.

4 Data and Empirical Strategy

With the theoretical framework above in mind, we next describe how we estimate the impact of NREGA on employment and wages and discuss the possible threats to our empirical strategy.

4.1 Data

Our primary source of data is the nationally representative Employment and Unemployment survey carried out by the National Sample Survey Organization ("NSS Employment Survey"). We use village-level data from the 2001 census aggregated to the district-level, as well as data on agricultural yield, rainfall, political cycles and roads built under a national rural roads construction program (PMGSY) to construct district-level controls, which are described in detail in Appendix A.4. For the calibration in Section 6, we use the ARIS-REDS data set, described in Appendix A.4.3.

Our identification strategy relies on changes at the district-level. Districts are administrative units within states. Because the workfare program is applicable only to persons living in rural areas, we drop districts that are completely urban and only use data for persons located in rural areas. Our sample includes districts within the twenty largest states of India, excluding Jammu and Kashmir. We exclude Jammu and Kashmir since survey data is missing for some quarters due to conflicts in the area. The remaining 497 districts represent 97.6% of the rural population of India. Appendix A.4 details how we adjust the data to account for district splits and merges. The median district in our sample had a rural population of 1.37 million in 2008 and an area of 1600 square miles.

We use four rounds of the NSS Employment Survey, which is stratified by urban and rural areas of each district. Surveying is further divided into four sub-rounds each lasting three months. Although the sample is not technically stratified by sub-round, the NSSO states that it attempts to distribute the number of households surveyed evenly within each district sub-round. We discuss in detail in the next section the extent to which this goal is accomplished in practice. The NSSO over-samples some types of households and therefore provides sampling weights (see National Sample Survey Organisation (2008) for more details). All statistics and estimates computed using the NSS data are adjusted using these sampling weights.

The NSS Employment Survey is conducted on an irregular basis, sometimes with a small and sometimes with a large sample. To enhance precision and ensure that our sample is representative, we only use years with a large sample ("thick rounds"). We use data spanning

July 2004 to June 2005 to form the pre-program period. For the post-program period, we use data spanning July 2007 to June 2008. In the placebo analysis, we use data from July 1999 to June 2000 to test for differential trends in outcomes before the program was implemented. Finally we use data from July 2009 to June 2010 to document trends in outcomes after the program had been introduced to all districts.

4.2 Construction of Outcomes

Our main outcomes are individual measures of employment and wages. We construct the employment measures as follows. The NSS Employment Survey includes detailed questions about the daily activities for all persons over the age of four in surveyed households for the most recent seven days. We restrict the sample to persons aged 18 to 60. We then compute for each person the percentage of days in the past seven days spent in each of four mutually exclusive activities: private sector work, public works, not in the labor force, and unemployed. Private sector work includes waged work, self-employment and domestic work. Domestic work could arguably be categorized as not in the labor force. However, given that most households engage in small-scale agriculture, many activities could equally well be categorized as domestic work or self-employment. In the context of the model presented in Section 3, we believe both domestic work and self-employment naturally fall under the definition of private sector work.

Our wage measure is computed as follows. The NSSO makes the distinction between two types of waged work depending on the duration and formality of the relationship with the employer: salaried work is long-term and often involves a formal contract while casual work is temporary and informal. The NSSO asks individuals who worked in casual labor over the past seven days their total earnings from casual labor. For each individual we compute the average earnings per day worked in casual labor (the "casual wage").

Although the NSSO makes an effort to survey villages within each district throughout the year, in a few occasions no households were surveyed in some district-quarters. Even if households were surveyed, if none of the surveyed adults worked in casual labor, we do not have a measure of wages for that district-quarter. Table A.1 presents the number of non-missing observations for each district-quarter for the employment and wage outcomes. The fraction of districts with no measure of employment for a given quarter is relatively low: it is always below 1% in the survey years 2004-05 and 2007-8. For wage outcomes, this fraction is higher, but still low (2-4%) in the survey years 2004-05 and 2007-8. Missing observations are more of a concern for the survey year 1999-2000, which we only use for the placebo analysis. Appendix A.4 provides further discussion.

4.3 Empirical Strategy

Our empirical strategy compares changes in districts that received the program earlier to changes in districts that received the program later. The program was first introduced in 200 districts in February 2006, extended to 130 districts in April 2007, and finally to the rest of rural India in April 2008.[10] From our sample of 497 districts, our analysis compares the 288 districts selected to be part of the first two phases ("early" districts) to the 209 districts which received the program in 2008 ("late" districts). We use for our pre-period July 2004 to June 2005, and for our post-period July 2007 to June 2008. Both periods contain one full year.[11]

Early phase districts were selected to have lower agricultural wages, a larger proportion of "backward" castes and lower agricultural output per worker. These targets were balanced with the goal of spreading early phase districts across states. As a result, some early phase districts in richer states rank significantly better based on the three indicators than late phase districts in poorer states. Further, political considerations seem to have played some role in the selection of early districts (Gupta, 2006). Figure 1 shows the geographic distribution of early and late districts across India. Early districts are relatively well spread out, though there is a concentration of early districts in northern and eastern India, where rural poverty is higher. Because early districts were purposefully selected based on variables that are correlated with labor market outcomes, a simple comparison of early and late districts is unlikely to be informative of the program impact. For this reason, we compare changes over time in early districts relative to late districts.

These difference-in-differences estimates will be biased if outcomes in early districts are trending differentially from outcomes in late districts. We are able to partly address this concern by including controls meant to capture differential changes across districts. We control for pre-program measures of caste composition, agricultural wages and agricultural output

[10] Prior to the official start date in February 2006, the government launched a pilot program known as the Food for Work Program in November 2004 in 150 of the initial 200 districts. Confirming existing field observations (Dreze, 2005), we find little evidence of an increase in public works during this pilot period.

[11] Late districts technically received the program in April 2008. We use the entire survey round July 2007 to June 2008 both to increase sample size and so that we can observe effects throughout the whole agricultural year. Even in the second quarter, we find a significant differential rise in public works in early relative to late districts, likely due to the fact that public works employment did not start immediately in late districts in April 2008.

per worker, which were the three criteria used for the selection of early phase districts.[12] The rest of the district-level controls are shown in Table 1 and include pre-program measures of poverty, literacy, population density, labor force participation, workforce composition and land irrigation. We interact these time-invariant controls with a dummy for post-program status to pick up trends correlated with the controls. We also include time-varying controls: annual percentage deviation from average rainfall, its square, and a dummy variable for the one year preceding a state or local election. Since outcomes may respond differently to these variables in early phase districts our specification allows the effect of time-varying controls to differ in early and late phase districts.[13] We control for the number of kilometers of road completed in the district over the last year under a national rural road construction program started in 2001, the Pradhan Mantri Gram Sadak Jozna (PMGSY).

Migration between early and late phase districts is unlikely to be a major concern for our analysis. Rural to rural inter-district migration for employment is limited. Out of all adults (18 to 60) living in rural areas, only 0.4% percent report having migrated to a different rural district for employment within the past year.[14] Low levels of migration are similarly documented in Munshi and Rosenzweig (2009) and Topalova (2010). A higher number, 1.9% of rural adults, report having migrated from rural to urban areas. Imbert and Papp (2014) present evidence that the NREGA reduces short-term migration from rural to urban areas in a group of villages in northwest India. Since urban areas are excluded from our analysis, a drop in rural to urban migration induced by the NREGA would not bias our estimates of the impact on wages. It would, however, imply the estimated effect of the NREGA on rural labor markets may have been mitigated by migration flows and that urban labor markets too may have been affected by the NREGA.

4.4 Regression Framework

We estimate variations of

$$Y_{idt} = \beta T_{dt} + \delta Z_d \times \mathbf{1}_{\{t>2006\}} + \gamma X_{dt} + \lambda X_{dt} \times \mathbf{1}_{\{Early\}} + \alpha H_i + \eta_t + \mu_d + \varepsilon_{idt}$$

where Y_{idt} is the outcome (e.g. earnings per day worked) for individual i surveyed in district d in quarter t, T_{dt} is a dummy equal to one for early phase districts in the post

[12] These controls are not redundant with the program dummy because the selection of early districts was not based entirely on these criteria and because the Planning Commission used measures of agricultural wages and output from the 1990s, more than a decade older than our controls (Planning Commission, 2003).

[13] In particular, wages may be more pro-cyclical in early phase districts, which are poorer (Jayachandran, 2006).

[14] Authors' calculations using NSS Employment and Unemployment Survey Round 64.

period (July 2007 to June 2008), Z_d are time-invariant district controls, $\mathbf{1}_{\{t>2006\}}$ is a dummy variable equal to one after 2006, X_{dt} are time-varying district controls, $\mathbf{1}_{\{Early\}}$ is a dummy variable equal to one for observations within early districts, H_i are individual controls, η_t are year-quarter fixed effects, and μ_d are district fixed effects. All estimates are adjusted for correlation of ε_{idt} over time within districts by clustering at the district-level. For most specifications, we include interactions of T_{dt} with other variables such as season dummies or dummies for whether the district is in a star state, in order to exploit the variation in public employment provision across seasons and states.

Because we are interested in the impact of the program on the labor market equilibrium, the relevant level of analysis is not the individual but the district. We re-weight individual observations so that the sum of all weights within a district-quarter is constant over time for each district and proportional to the rural population (see Appendix A.4.4 for details). Individual controls are used only to ease concerns that our estimates of the program impact on wages are driven by worker selection. For example the average private sector wage might increase mechanically if the program is more likely to hire low-wage workers. The individual-level controls include dummy variables for age group, education level, gender, caste, religion and marital status (see Section A.4 in Appendix for more details).

5 Results

We next present descriptive statistics for early and late districts. We then turn to our empirical estimates of the effect of the program on public employment, private sector work and wages.

5.1 Summary Statistics

Table 1 presents the pre-period means for the controls used for early and late districts as well as districts in star states and non-star states. As expected given the criteria used to choose early districts, early districts are poorer based on every measure (literacy, poverty, share of low caste population, fraction of the labor force in agriculture). Star states, on the other hand, seem to be slightly richer and employ a greater fraction of agricultural workers than other states, with a larger share of tribal (ST) population. Recall from Section 2.5 that star states are states which implemented the program better than other states.

Table 2 presents the pre-period means for the outcomes used in the paper for early and late districts as well as districts in star states and non-star states. The allocation of days between private sector work, public sector work, unemployment and not in the labor force is similar in early and late districts. As expected given the stated selection criteria used by the government, casual labor earnings per day are 13-24% lower in early phase districts prior to the introduction of the program. The main difference between star states and the rest of India is a smaller fraction of time spent on domestic work.

Our empirical strategy compares trends in outcomes in early and late districts before and after the implementation of NREGA. Figures 4, 5 and 6 present the trends of our main outcomes, public works, private sector work and wages for the dry season in early and late phase districts. Figure 4 shows that before NREGA, early and late phase districts had similarly low levels of public employment. In 2007-08, time spent on public works increased sharply in early districts. Interestingly, public employment provision in late districts does not catch up completely with early districts after the program is extended to all of India by 2009-10.[15] Figure 5 shows that time spent doing private sector work fell in 2007-08 in early relative to late districts and in 2009-10 in late relative to early districts. The magnitude of the drop is similar to the observed increase in public employment. Finally, Figure 6 shows that wages in early phase districts were lower but increasing relative to late phase districts even before NREGA was implemented. Relative wage growth appears to have accelerated in 2007-08, when NREGA was rolled out in early phase districts, and decelerated in 2009-10, when the program was extended to the rest of India.

5.2 Change in Time Allocation

We divide daily activities into four mutually exclusive categories: public works, private sector work (including casual labor, salaried work, domestic work and self employment), unemployment and not in the labor force. The results for our main specification using these outcomes are presented in columns one to four of Table 3.[16] We find a strong and significant impact of the program on the fraction of total time spent working in casual public employment. Public works employment increased by 1.17 percentage points during the dry season and 0.46 percentage points during the rainy season. Hence the results confirm that the rise in public works is concentrated during the dry season. The one percentage point increase in public employment during the dry season masks considerable heterogeneity among states. We explore this heterogeneity further in Section 5.4. During the dry season, the rise in public

[15] The lack of catch-up by late districts is why we chose not to make use of the potential second difference-in-differences estimate comparing late districts and early districts from 2007-08 to 2009-10 in our main specification. We discuss the results of this second difference-in-differences in section 5.5 below.

[16] Table A.2 in the appendix presents the results without controls.

employment is offset by a fall in private sector work rather than time spent outside the labor force or unemployment. We cannot reject that private employment falls one-for-one with public employment generation.

Perhaps surprisingly, unemployment does not appear to fall in early districts relative to late districts. A possible explanation is that after the introduction of the program, unemployed workers are more likely to report that they are available or looking for work while working at home or on the farm. As a result the fall in private sector work may in part represent a fall in disguised unemployment or private sector work with close to zero productivity (see Section 3.2). Another explanation is more driven workers take up private and public sector work, while the others remain unemployed. This does not seem to be the case: the proportion of casual workers who report that they have worked on public works during the last seven days is the same for those who report an unemployment spell and those who do not (2.3% in NSS Employment Survey Round 66).

5.3 Change in Private Sector Wages

The model presented in Section 3.2 predicts that the fall in private sector work during the dry season be matched with a rise in wages. Column five of Table 3 presents the results for our main specification using deflated log casual earnings per day as the dependent variable. The estimates for the dry season show that deflated daily earnings rise by 4.7 log points more in early relative to late districts. The specification includes district-level controls to control for potential differential trends in wages and worker-level controls to account for possible change in worker composition.[17] This result suggests that NREGA increased wages for unskilled labor. This finding is inconsistent with "surplus labor" models in which low productivity households are isolated from the market, and hence the government can hire workers from these households without any effect on the market wage (see Sen (1966) and Rosenzweig (1988) for a discussion).

One may wonder whether the magnitude of the wage increase is reasonable given the fall in private sector work. To explore this question, we assume labor markets are competitive so that changes in the wage are due to shifts along a labor demand curve. We can then use the estimate of the increase in the wage of 4.73% and the fall in private sector work to compute a labor demand elasticity. Rural prime-age adults in early phase districts spend on average 89.3% of their time in private sector work (Table 2). Therefore our estimate from

[17] Table A.2 presents the results with district controls but without worker controls. The estimated impact on wages is slightly lower, which suggests that worker selection biases our estimate downwards, not upwards.

the previous section implies that private sector work declined by $1.31/.893 = 1.46\%$. As discussed in Section 3.2, if labor markets are competitive, the elasticity of labor demand is given by the ratio of the change in private employment to the change in the wage. Hence our estimate of the elasticity of labor demand is $\hat{\epsilon_d} = \frac{1.46}{4.73} \approx 0.31$, which lies within the 0.25 to 0.40 range estimated by Binswanger and Evenson (1980) for farm employment in India.

5.4 Star States

We next document the changes in labor market outcomes for early districts in the few states that provided most of the NREGA employment (see Section 2.5). Star states are by definition selected based on their implementation of the program. As a result, it is possible that even conditional on controls, labor market outcomes in these states would have changed differentially absent the program. This important caveat notwithstanding, we believe documenting the trends in employment and wages of early districts in star states as compared to late districts is important. If the employment and wage changes were concentrated in states where the NREGA was not well implemented, it would cast doubt on the validity of our empirical strategy.

Table 4 presents our main specification with the program dummy interacted with whether the district is in one of the star states as well as a dummy for the rainy or dry season. The results in column one confirm that the field studies are correct in labeling these states as star states. While time spent on public employment in early districts of star states rises by 3.1 percentage points in the dry season, there is no increase in public employment in early districts of non-star states.

Columns two through four show that the fall in private sector work documented for all of India is concentrated within the early districts of star states during the dry season. The estimates are consistent with a one-for-one crowding out of private employment by public sector work. Neither unemployment nor not in the labor force seem to be affected by the program. Column five further shows that in star states, daily casual earnings increase by a strongly significant 8.98% in the dry season. Consistent with the all India estimate, the implied labor demand elasticity is equal to $\frac{3.07}{0.893*8.98} = 0.38$. During the rainy season in star states, wages increase by an insignificant 4.58%. The coefficients for other states are on the order of one to two percent.

5.5 Robustness Checks

A primary concern is that the differential change in employment and daily earnings documented above for early relative to late districts may represent changes unrelated to the program. That the effects are concentrated during the dry season and in states where the program is best implemented suggests the results are due to the program. As a further check, Table 5 presents a similar specification to the one in Table 4 except that the sample is composed of years 1999-00 and 2004-05 and the program dummy is set to one for early districts in 2004-05. In other words, we estimate the differential changes across early and late districts prior to the program.

As expected, we do not find any differential increase in public employment in early relative to late phase districts prior to the implementation of the program. As we saw from Figure 6 daily casual earnings did increase in early relative to late phase districts between 1999-00 and 2004-05. However, once we control for district characteristics using our preferred specification, the point estimates are small and insignificant. Finally, even with controls, we find a significant decrease in time spent in private sector work and an insignificant increase in unemployment and in time spent outside of the labor force in early as compared to late districts between 1999-00 and 2004-05.

The analysis of pre-existing trends suggests that labor markets in early and late phase districts either were on different long-term paths or experienced different seasonal shocks between 1999-00 and 2004-05. The inclusion of controls in our main specification seems to be effective in dealing with most of the effect of selection into early phases of the program. As a further test we include changes in the outcomes between 1999-00 and 2004-05 as controls in our main specification; our results are not affected (see Appendix Table A.4).

Economic shocks or policy changes concurrent to NREGA roll out represent another important threat to our identification strategy. Since the states are the relevant level for many policy decisions (e.g. industrial policy, infrastructure programs) and are integrated economically, it seems natural to test whether our results hold when we only compare early and late districts within each state. We do this by including in our main specification a dummy for each state interacted with a dummy for 2007-08. The results presented in Table A.6 are close to our main estimates: employment on public works rises by 1.1 percentage points, private sector work falls by 1.8 percentage points and wages for casual labor increase by 4.2%.

As a final test of the parallel trend assumption, we use the second difference-in-differences based on changes in outcomes in late and early phase districts between 2007-08 and 2009-10 during which the program was extended to late phase districts. The results are presented in Table A.4. We find a small and insignificant increase in public employment in late as compared to early districts; this is due to the fact that public employment provision continued to expand in early districts. We also find that private employment decreases and unemployment increases in late relative to early districts as the NREGA is extended to all of India. This may be due to the fact that workers in late districts are more likely to declare being unemployed once the employment guarantee is implemented. Finally, we find a small and insignificant increase in casual wages in late relative to early districts, which suggests our main results are not driven by a long-term rise in wages in early relative to late districts.

5.6 Alternative specifications

Three concurrent studies by Azam (2012), Berg et al. (2013) and Zimmermann (2013) estimate the impact of the NREGA on labor market outcomes. In order to better understand how our results relate to their findings, we estimate specifications that are similar to theirs using our data. We provide an overview of the results and leave many of the details to Section A.5.

First, we follow Azam (2012) and estimate the program effect separately for men and women (results are presented in Appendix Table A.7). As Azam (2012) we find a stronger effect for women and an insignificant effect for men. However, when we include district controls, the effect for men increases and becomes significant and the effect for women drops and becomes insignificant. This suggests that part of the difference in wage trends between male and female workers in early districts relative to late districts may be shocks or trends unrelated to the program.

Second, we follow Berg et al. (2013) and redefine the treatment variable as the number of months since the program was launched in each district. The estimates suggest the program had a positive effect on time spent in public works, a negative effect on time spent in private sector employment, and a positive effect on casual wages. Adding district-specific trends changes the magnitude of the coefficients but not their sign or their statistical significance.

Berg et al. (2013)'s approach yields similar conclusions to ours: public works increase, private employment falls, and casual wages rise. In the first half of 2008, the NREGA had been in place for two years in first phase districts and for one year in second phase districts. Based on the monthly estimates, wages increased by 6.48% and 3.24% in first and second phase districts respectively. Using population weights (60% and 40%), the estimated average impact is 5.1% in early districts, which is close to our own 4.7% estimate.

Third, we follow Zimmermann (2013) and use a regression discontinuity design to identify the effect of the program on employment and wages. The selection of early districts was based on a backwardness ranking detailed by Planning Commission (2003). Within each state, and taking the number of early districts as given, one can use each district's backwardness rank to predict its assignment to early or late phases. One can then estimate the effect of the program by comparing 2007-08 outcomes in early and late phase districts close to the cut-off, controlling for the backwardness rank.

Allowing for different slopes on each side of the cut-off, we find a positive but insignificant effect of the program on time spent in public works (0.51 and 0.35 percentage points for the linear and quadratic specification respectively), a negative but insignificant effect on time spent in private sector work (-0.8 and -1.5 percentage points), and positive effects on private sector wages (6% and 11%). These estimates are reasonably close to those of our preferred specification and never significantly different from them. The standard errors of the estimated coefficients are large.

6 Estimating the Distributional Impact

The previous analysis suggests the workfare program increased government work and led to an increase in wages for private sector casual laborers. Recall from Section 3 that the compensating differential for household i given by Equation 2 is

$$-dz_i = \textit{Net Casual Labor Earnings}_i \times \frac{dW/W}{dL_g} + (W_g - \widetilde{W_i})dL_i^g \tag{3}$$

We use the estimates from the previous section combined with pre-program household-level data. We focus on the first half of the year, the off-season of agriculture, when most of the employment is generated by the program.

6.1 Gains and Losses from Wage Change

The first term of Equation 3 ($\textit{Net Casual Labor Earnings}_i \times \frac{dW/W}{dL_g}$) is the change in welfare due to the equilibrium change in the wage. To estimate this term, we use 4.7% for the wage change ($\frac{dW/W}{dL_g}$) based on the estimates in Table 3.

Net casual labor earnings is more difficult to estimate because in the NSS Employment Survey we only observe casual labor earnings, not payments. We use the 1999-00 ARIS/REDS data set, which is a nationally representative survey of households in rural India. The ARIS/REDS survey includes questions on total casual earnings as well as total payments to hired casual laborers (see Appendix A.4.3 for more detail). We first compute the share of labor costs paid by each consumption quintile by dividing the sum of labor payments made in each quintile by total labor earnings. These shares do not sum to one because casual labor earnings reported by rural households may come from urban employers. We then multiply each share by total casual labor earnings reported in the NSS Employment Survey to obtain the estimates of casual labor payments by consumption quintile given in row seven of Table 6.

We observe casual labor earnings directly in the NSS Employment Survey, and these earnings are reported in the third row of Table 6. Net casual earnings (row eight) are given by total casual earnings (row three) less total casual payments (row seven). As expected, net casual earnings decrease as we move from the bottom to top quintiles. The resulting net gain from the wage change is 4.7% multiplied by net labor earnings for each quintile (row ten).

6.2 Direct Gains from Participation

We next quantify the second term in Equation 3. The term $(W_g - \widetilde{W_i})dL_i^g$ is the direct gain for program participants from working for the program. The welfare gain due to program participation is $(W_g - \widetilde{W_i})\Delta L_g$. Ideally, we would estimate ΔL_g using a direct measure of how many days households in each consumption quintile worked for the program. However, since we measure employment in all types of public works projects and not only employment provided by the program, we instead estimate the change in public works by quintile using our main specification with the program dummy interacted with a dummy for each consumption quintile. That is, we estimate:

$$Y_{idt} = \sum_q \beta_q T_{dt} \times D_{idt}^q + \gamma X_{dt} + \lambda X_{dt} \times \mathbf{1}_{\{Early\}} + \delta Z_d \times \mathbf{1}_{\{t>2006\}} + \alpha H_i + \eta_t^q + \mu_d^q + \varepsilon_{idt}$$

where Y_{idt} is the fraction of time spent on public works by individual i at date t in district d. D_{idt}^q is a dummy variable equal to one if individual i belongs to consumption quintile q. Quintiles are defined separately for each year of data. T_{dt} is a dummy for program districts in the post period (July 2007 to June 2008), X_{dt} are time-varying district controls, Z_d are time-invariant district controls, H_i are individual controls, η_t^q are year-quarter-quintile fixed effects, and μ_d^q are district-quintile fixed effects.[18]

[18] We also estimate the effect of the program on employment and wage outcomes for different consumption quintiles by regressing these outcomes on an interaction of the program dummy with a dummy for each

The estimates of β_q for each quintile provide an estimate of the increase in public works (ΔL_g) for each quintile. These estimates are presented in row 11 of Table 6. As compared to our main specification, this method of estimating the increase in public works employment relies on the additional assumption that the composition of each consumption quintile did not change differentially in early and late phase districts and was not affected by the program. Given the short time lag between the pre and post-program periods, and given the relatively small size of the income transfer due to the program, we believe that large changes in the distribution of consumption are unlikely.

We estimate W_g using daily earnings for program participants. Based on the NSS 2007-08 Employment Survey, average daily earnings for program participants were 15% higher than average casual daily earnings in early districts. This figure likely understates the initial public-private wage gap, since private wages have moved closer to the government wage as a result of the program. The estimated wage increase following program implementation between 2004-05 and 2007-08 is 4.7%. Hence, for the calibration we set the government wage to be $20\% \approx 15\% + 4.7\%$ higher than the mean casual wage in 2004-05.

As discussed in Section 3, participants' outside option $\widetilde{W}_i$ may be lower than the market wage. Datt and Ravallion (1994) use a survey of participants in a similar Indian workfare program in the state of Maharashtra and conclude that forgone income represents 20-30% of the earnings from the workfare program. We adopt their estimate for the purpose of our calibration and assume the shadow wage $\widetilde{W}_i$ is on average 30% of the market wage (equivalently 25% of the public sector wage).[19] The implied direct transfer $(W_g - \widetilde{W}_i)\Delta L_g$ under this assumption is presented in Row 14 of Table 6.

6.3 Comparing Equilibrium and Direct Gains

Figure 7 presents the estimated gain due to the change in wages, the gain due to participation in the program assuming an outside option equal to 30% of the market wage, and the sum of the two for each quintile. For the three poorest quintiles, the equilibrium wage effect is of comparable magnitude to the gains from participation; approximately a third of the total gain is due to the increase in wages. For the richest quintile, the increase in labor costs more

quintile and by including district-quintile and time-quintile fixed effects. Regression results are shown in Table **??**.

[19]Since non-farm opportunities were more scarce in rural areas in the 1990s, 30% is likely a lower bound of the opportunity cost of time. Using counterfactual questions asked to NREGA workers in Bihar, Murgai et al. (2013) estimate that foregone earnings are 34% of public works wage.

than offsets the gains from participation resulting in a welfare loss for these households.[20]

The numerical estimates plotted in Figure 7 are presented in Table 6. Row 15 presents the total estimated gain for each consumption quintile. Row 16 further shows that the fraction of the total gain due to the equilibrium change in wages is between 22% and 42% of the total gain from the program for the three poorest quintiles. Finally, row 17 expresses the total gain from the program as a fraction of total expenditure: although richer households lose from the program, the impact is less than one percent of total expenditures.

6.4 Discussion

Our calibration results depend on the validity of the theoretical framework outlined in section 3. We discuss here some of the assumptions of the framework presented above, and how the results might change if those assumptions do not hold.

Imperfect Competition: We assume labor markets are competitive. If employers have market power then Equation 2 would capture the welfare impact of the program for labor suppliers but not for labor buyers (see Appendix A.3.6 for more details). However, the results presented in Section 5 show an increase in wages and a simultaneous *fall* in private employment. This contradicts predictions of imperfect competition models and lends support to our assumption that labor markets are competitive.

Changes in Worker Productivity: To the extent the program increases wages by changing worker productivity, Equation 2 will not capture the full welfare impacts of the program. Though there is limited existing evidence, the discussion in Section 2.1 suggests it may be reasonable to assume that infrastructure created by the program did not have large effects on worker productivity within the first two years of implementation.[21] To the extent that these effects exist, our framework will underestimate the welfare gains for households that hire labor.

Fiscal cost of the program: Our model implicitly assumes the NREGA is funded from the outside (e.g. by taxes levied on urban taxpayers). In practice, the central government bears 90% of the cost of NREGA. It derives 56% of its tax revenues from corporate and income tax ("direct taxes"), and 44% of its revenues from customs, excise, and value added

[20] When computed over the whole year, the monthly welfare gains from the program are 30% lower than the gains for the dry season only, but the relative magnitudes of the direct and indirect effects remain the same.

[21] Worker productivity may have also increased through other channels. For example, the increased income due to the program may allow workers to make investments in their health leading to higher productivity (Strauss, 1986), or higher wages may have led employers to substitute away from labor towards labor-saving capital.

taxes ("indirect taxes").[22] Since most of the rural economy is informal, and agriculture is exempt from corporate and income taxes, few rural households pay direct taxes. Indirect taxes, however, may affect both rural and urban households (Jha and Srinivasan, 1989). If we assume these taxes are progressive, because rich households consume more goods from the formal (taxable) sector, our estimate may understate the distributive impact of the program.

6.5 Cost-benefit analysis

An important question is how the welfare gains of the program compare with the costs. A complete cost-benefit analysis of the NREGA would require estimating many factors beyond the scope of this paper, such as the productivity of the infrastructure generated. We may gain some insight, however, by comparing the welfare gains as estimated in the previous section with the government expenditure. From January to June 2008, total monthly NREGA expenditure in early districts was 241 Rs. per rural household (167 Rs. on unskilled labor alone).[23] Hence, the cost of the NREGA per rural household is much higher than the estimated welfare gains even to the poorest consumption quintile (112 Rs.).

One reason for the large difference between the fiscal cost of the program and the benefits accrued to rural households is foregone income. Murgai et al. (2013) compare the poverty impacts of the NREGA with a hypothetical cash-transfer scheme given to all rural households without any work requirement and conclude that the cash-transfer is more cost-effective. Another reason for the gap between costs and benefits of the program is widespread corruption. Imbert and Papp (2011) compare administrative data on person-days provided under NREGA with estimates using the NSS survey data on days spent on public works in 2007-08. We find that only 42 to 56% of NREGA days are independently confirmed by survey data. This finding suggests that leakages severely limit the extent of redistribution achieved by NREGA (Niehaus and Sukhtankar, 2013b).

[22] These figures are for the 2011-12 financial year (`http://dor.gov.in/revenue_ctc`).

[23] In order to compute monthly NREGA expenditures per rural household we used monthly progress reports available on NREGA official website (nrega.nic.in) and census 2001 data, adjusted for annual population growth in each state between 2001 and 2011 (censusindia.gov.in).

7 Conclusion

We provide some of the first evidence on the equilibrium impacts of workfare programs in a developing country context. Like many social programs in developing countries, workfare programs involve a transfer to the rural poor funded by (mostly urban) taxpayers. We show that through their effect on labor markets, workfare programs trigger a redistributive effect within rural areas, from households which are net labor buyers to households which are net labor sellers. Further, we show that these redistributive effects are quantitatively significant. Under reasonable assumptions, the increase in equilibrium wages represents a third of the total welfare gain for the poor.

Our analysis also suggests that the fiscal cost of the program is much higher than the estimated welfare gains to rural households. This result must be interpreted with caution, given that we do not account for the productivity benefits of the generated infrastructure. As discussed in Section 2.1, field reports indicate that the quality of NREGA infrastructures is poor overall (The World Bank, 2011), so that zero productivity seems a natural benchmark. Once rigorous empirical evidence on the effect of NREGA infrastructure is available, we hope our calibration will provide a useful framework which can be extended to include these effects in the welfare analysis.

References

Afridi, F. (2008). Can Community Monitoring Improve the Accountability of Public Officials. *Economic and Political Weekly*, 35–40.

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

Ardington, C., A. Case, and V. Hosegood (2009). Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa. *American Economic Journal: Applied Economics 1*(1), 22–48.

Association for Indian Development (2009). Key Observations on NREGA work in Andhra Pradesh. Available at http://aidindia.org/main/content/view/y88/279/.

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

Banerjee, A. and E. Duflo (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives 21*(1), 141–167.

Basu, A. K., N. H. Chau, and R. Kanbur (2009). A Theory of Employment Guarantees: Contestability Credibility and Distributional Concerns. *Journal of Public economics 93*(3-4), 482–497.

Behrman, J. R. (1999). *Handbook of Labor Economics*, Volume 3, Chapter 43, pp. 2859–2939. Elsevier Science.

Benjamin, D. (1992). Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models. *Econometrica 60*(2), 287–322.

Berg, E., S. Bhattacharyya, R. Durgam, and M. Ramachandra (2013, October). Can Rural Public Employment Schemes Increase Equilibrium Wages? Evidence from a Natural Experiment inIndia.

Binswanger, H. P. and R. E. Evenson (1980). Estimating labor demand functions for Indian agriculture. Center Discussion Paper, Economic Growth Center, Yale University No 356.

Binswanger, H. P. and M. R. Rosenzweig (Eds.) (1984). *Contractural Arrangements, Employment, and Wages in Rural Labor Markets in Asia.* Yale University Press.

Card, D. and A. Krueger (1992). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review 84*(4), 772–793.

Datt, G. and M. Ravallion (1994). Transfer Benefits from Public-Works Employment: Evidence for Rural India. *The Economic Journal 104*(427), 1346–1369.

Deaton, A. (1989). Rice Prices and Income Distribution in Thailand: A Non-Parametric Analysis. *The Economic Journal 99*(395), 1–37.

Deininger, K. and Y. Liu (2013). Welfare and poverty impacts of india's national rural employment guarantee scheme. Technical Report 6534, World Bank Policy Research Working Paper.

Dinkelman, T. (2011, December). The effects of rural electrification on employment: New evidence from south africa. *American Economic Review 101*(7), 3078–3108.

Dreze, J. (2005). Loot For Work Programme. *The Times of India*.

Dreze, J. and R. Khera (2009). The Battle for Employment Guarantee. *Frontline 26*(1).

Dreze, J. and C. Oldiges (2009). Work in Progress. *Frontline 26*(4).

Dutta, P., R. Murgai, M. Ravallion, and D. van de Walle (2012). Does India's employment guarantee scheme guarantee employment ? World Bank Policy Research working paper No 6003.

Fetzer, T. (2013, October). Can workfare programs moderate violence? evidence from india. Manuscript.

Gupta, S. (2006). Were District Choices for NFFWP Appropriate? *Journal of Indian School of Political Economy 18*(4), 641–648.

Imbert, C. and J. Papp (2011). Estimating Leakages in India's Employment Guarantee. In R. Khera (Ed.), *Battle for Employment Guarantee*. Oxford University Press.

Imbert, C. and J. Papp (2014). Short-term Migration and India's Employment Guarantee. Manuscript.

Institute of Applied Manpower Research (2009). All-India Report on Evaluation of NREGA: A Survey of Twenty Districts.

Jayachandran, S. (2006). Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy 114*(3), 538–575.

Jayachandran, S., J. Cunha, and G. D. Giorgi (2010). The Price Effect of Cash Versus In-Kind Transfers. CEPR Development Economics Workshop Working Paper.

Jha, S. and P. V. Srinivasan (1989). Indirect taxes in india: An incidence analysis. *Economic and Political Weekly 24*(15), pp. 811–830.

Khera, R. (Ed.) (2011). *The Battle for Employment Guarantee*. Oxford University Press.

Munshi, K. and M. Rosenzweig (2009). Why is Mobility in India so Low? Social Insurance, Inequality and Growth. NBER Working Paper 14850.

Murgai, R., M. Ravallion, and D. Van de Walle (2013, October). Is workfare cost-effective against poverty in a poor labor-surplus economy? Policy Research Working Paper 6673, World Bank.

National Sample Survey Office (2010). Employment and Unemployment Sitation in India 2007-08. NSS 64th Round, July 2007 to June 2008, Appendix B.

National Sample Survey Organisation (2008). Employment and Unemployment Situation in India 2005-06 NSS 62nd Round July 2005 - June 2006. Technical report, Ministry of Statistics and Programme Implementation Government of India.

Niehaus, P. and S. Sukhtankar (2013a, November). Corruption Dynamics: The Golden Goose Effect. *American Economic Journal: Economic Policy 5*(4), 230–69.

Niehaus, P. and S. Sukhtankar (2013b). The marginal rate of corruption in public programs: Evidence from india. *Journal of Public Economics 104*(C), 52–64.

Planning Commission (2003). Identification of Districts for Wage and Self Employment Programmes. Report of the Task Force.

Porto, G. G. (2006). Using Survey Data to Assess the Distributional Effects of Trade Policy. *Journal of International Economics 70*(1), 140–160.

Ravallion, M. (1987). Market Responses to Anti-Hunger Policies: Effects on Wages, Prices and Employment. WIDER Working Paper 28.

Rosenzweig, M. (1980). Neoclassical Theory and the Optimizing Peasant: An Econometric Analysis of Market Family Labor Supply in a Developing Country. *Quarterly Journal of Economics 94*(1), 31–55.

Rosenzweig, M. (1988). Labor Markets in Low-Income Countries. *Handbook of Development Economics 2*.

Rosenzweig, M. R. (1978). Rural Wages, Labor Supply, and Land Reform: A Theoretical and Empirical Analysis. *American Economic Review 68*(5), 847–861.

Rosenzweig, M. R. and A. D. Foster (2010, October). Is there surplus labor in rural india? Working Papers 991, Economic Growth Center, Yale University.

Sen, A. K. (1966). Peasants and dualism with or without surplus labor. *Journal of Political Economy 74*(5), pp. 425–450.

Sharma, A. (2009). *NCAER-PIF Study on Evaluating Performance of National Rural Employment Guarantee Act*. National Council of Applied Economic Research.

Solon, G., S. J. Haider, and J. Wooldridge (2013, February). What are we weighting for? NBER Working Papers 18859, National Bureau of Economic Research, Inc.

Strauss, J. (1986). Does Better Nutrition Raise Farm Productivity. *Journal of Political Economy 94*(2), 297–320.

The World Bank (2011). Social Protection for a Changing India. Available at http://www.wds.wordbank.org/curated/en/2011/01/14087371/social-protection-changing-india-vol-1-2-executive-summary.

Topalova, P. (2010). Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty and Inequality from India. *American Economic Journal: Applied Economics 2*(4), 1–41.

Walker, T. S. and J. G. Ryan (1990). *Village and Household Economies in India's Semi-Arid Tropics*. Johns Hopkins University Press.

Zimmermann, L. (2013, October). Why guarantee employment? evidence from a large indian public-works program.

Figure 1: Map of Early and Late Districts

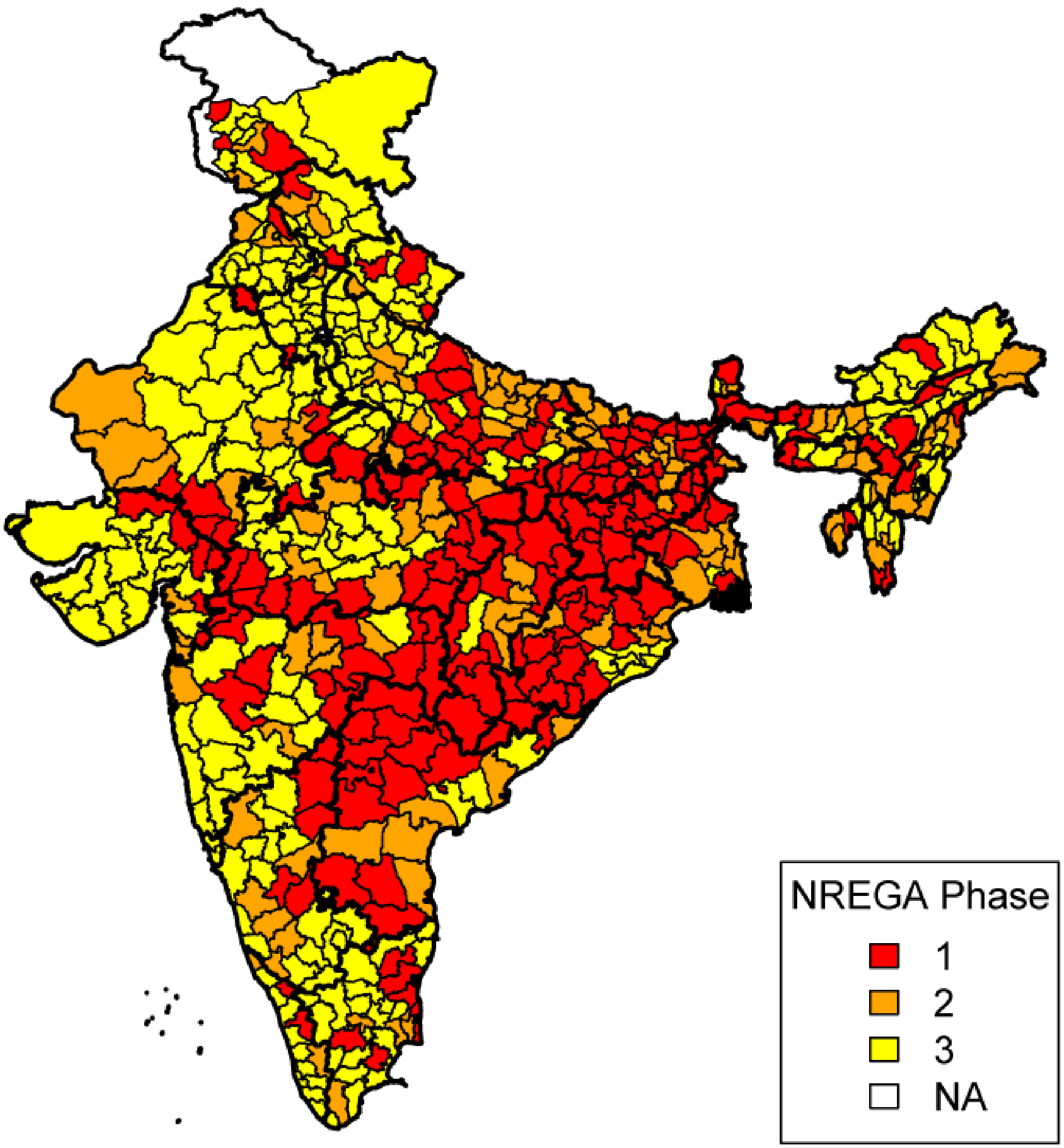
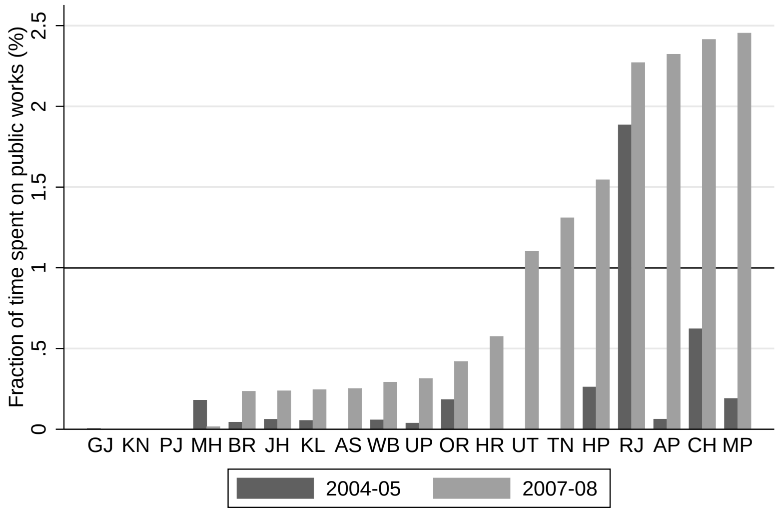
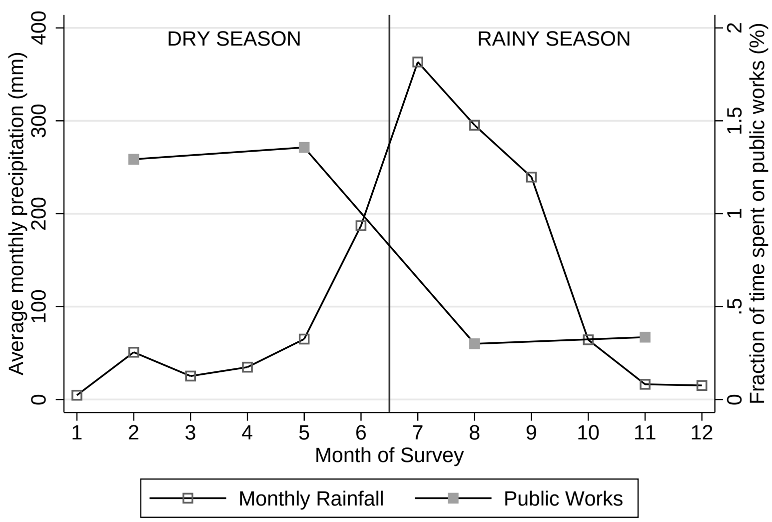


Figure 2: Heterogeneity in employment provision across States



Source: National Sample Survey (NSS) 2004-05 and 2007-2008. Rural prime-age adults in early phase districts only. All states above 1% are 'Star states' in the analysis.

Figure 3: Seasonality in employment provision



Early phase districts only. Source: National Sample Survey (NSS) 2007-08 and Tropical Rainfall Measuring Mission (NASA - JAXA) for the year 2007-08

Figure 4: Difference in Time spent on Public Works during the Dry Season between Early and Late Phase Districts.

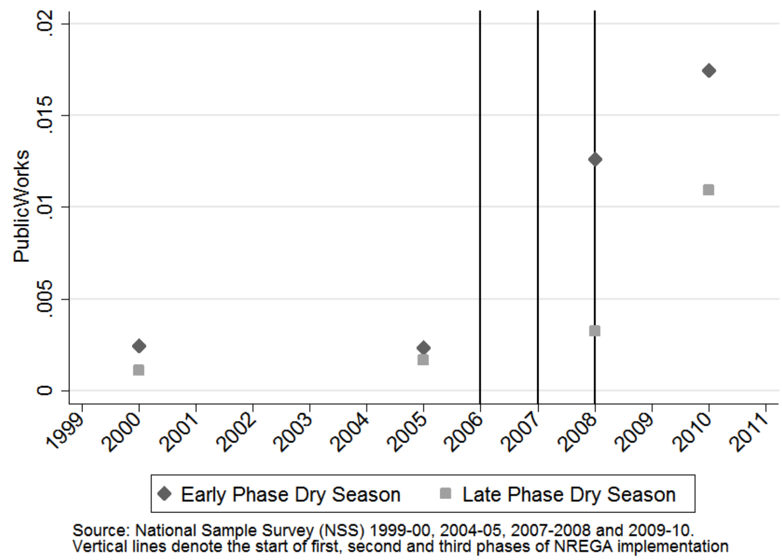


Figure 5: Difference in Time spent on Private Sector Work during the Dry Season between Early and Late Phase Districts.

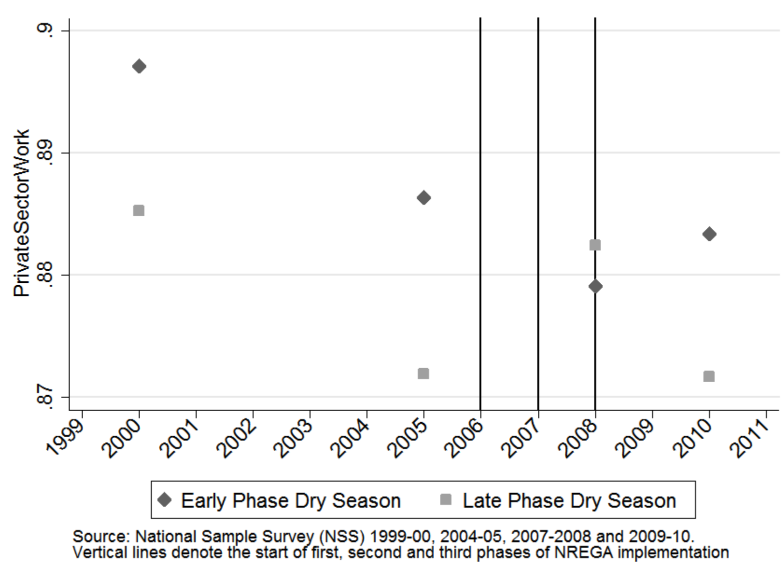


Figure 6: Difference in log Deflated Casual Daily Earnings during the Dry Season between Early and Late Phase Districts.

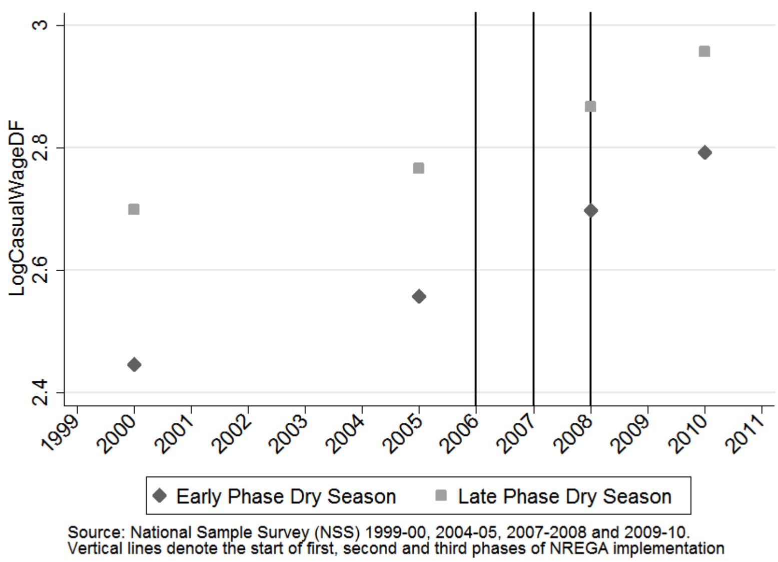
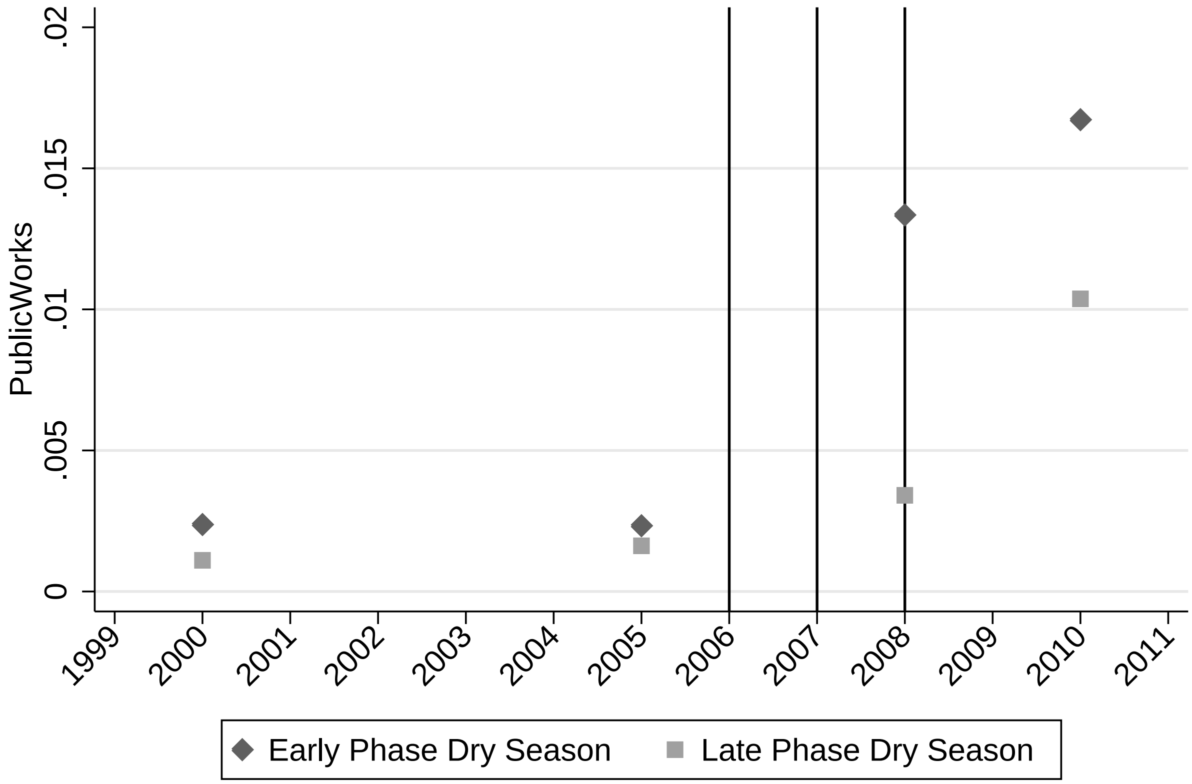


Figure 7: Welfare Gains by Expenditures Quintiles



Source: National Sample Survey (NSS) 1999-00, 2004-05, 2007-2008 and 2009-10.
Vertical lines denote the start of first, second and third phases of NREGA implementation

Table 1: District Controls Summary Statistics

	Early (1)	Late (2)	p-value (3)	Star States (4)	Other States (5)	p-value (6)	Source (7)	Time-varying? (8)
Fraction Scheduled Castes (SC)	0.187	0.173	0.07	0.184	0.181	0.65	2001 Census	No
Fraction Scheduled Tribes (ST)	0.134	0.048	0.00	0.144	0.083	0.00	2001 Census	No
Agricultural Productivity Per Worker (normalized)	-0.107	0.164	0.00	-0.064	0.023	0.08	Min of Ag	No
Log Daily Wage for Agricultural Casual Labor	3.519	3.781	0.00	3.578	3.638	0.29	NSS 2004	No
Poverty Rate	0.321	0.209	0.00	0.231	0.299	0.00	NSS 2004	No
Population Density (per sq. km)	481	406	0.01	241	544	0.00	2001 Census	No
Literacy Rate	0.555	0.650	0.00	0.591	0.592	0.96	2001 Census	No
Female Labor Force Participation Ratio	0.380	0.368	0.39	0.499	0.322	0.00	2001 Census	No
Male Labor Force Participation Ratio	0.634	0.629	0.22	0.659	0.621	0.00	2001 Census	No
Fraction Ag Casual Laborers	0.195	0.162	0.00	0.226	0.164	0.00	NSS 2004	No
Fraction Non-Ag Casual Labor	0.046	0.062	0.00	0.060	0.049	0.01	NSS 2004	No
Fraction Cultivators	0.284	0.259	0.03	0.323	0.254	0.00	NSS 2004	No
Fraction Non-Ag Business	0.093	0.095	0.59	0.093	0.094	0.79	NSS 2004	No
Fraction Salaried Work	0.042	0.070	0.00	0.058	0.051	0.05	NSS 2004	No
Fraction Labor Force in Agriculture	0.758	0.665	0.00	0.771	0.701	0.00	2001 Census	No
Irrigated Cultivable Land per Capita (ha)	0.082	0.117	0.00	0.113	0.088	0.00	2001 Census	No
Non irrigated Cultivable Land per Capita (ha)	0.176	0.171	0.76	0.229	0.151	0.00	2001 Census	No
Cumulative Rainfall (normalized)	0.315	0.234	0.24	0.179	0.329	0.03	IMD	Yes
Election Year	0.368	0.281	0.04	0.340	0.333	0.88	Gov Website	Yes
PMGSY Annual Road Construction (km)	15.8	12.6	0.03	21.4	11.6	0.00	Gov Website	Yes
Number of Districts	288	209		169	328			
Number of Individual Observations	219012	137624		112388	244248			

This table presents means of the controls used in the paper for different samples. Column (1) is restricted to districts that received the workfare program prior to April 2008. Column (2) includes only districts that received the program after April 2008. Column (3) presents the p-values of the Student's t-test of equality of means in Column (1) and (2). Column (4) restricts the sample to star states. Star states include Andhra Pradesh, Chhatisgarh, Himachal Pradesh, Madhya Pradesh, Tamil Nadu, Rajasthan, and Uttarkhand. Column (5) includes districts in non-star states. With the exception of the poverty rate, NSS controls are constructed using NSS Employment survey data from July 2004 to June 2005. Column (6) presents the p-values of the Student's t-test of equality of means in Column (4) and (5). The poverty rate is constructed using Round 61 of the NSS Consumer Expenditure survey. Employment variables from the NSS are computed using the reported usual activity during the past year for adults 18 to 60 only. Literacy and labor force participation are restricted to persons over the age of six. Cumulative rainfall is expressed as the percentage deviation from the cumulative rainfall since the beginning of the monsoon for each district-month from 1975 to 2010. Election year is a dummy variable indicating that state or local (panchayat) elections are to be held in the following year. PMGSY Annual Road Construction is the number of km of roads completed under the PMGSY scheme in the past year. For the Student's t-test in column (3) and (6) standard errors are computed assuming correlation of individual observations over time within each district.

Table 2: Summary Statistics of Outcomes in 2004-05 for Early and Late Districts

	Early (1)	Late (2)	p-value (3)	Star States (4)	Other States (5)	p-value (6)
Time Allocation:						
Public Work (Casual)	0.1%	0.1%	0.90	0.3%	0.1%	0.55
Private Work	89.3%	87.7%	0.56	87.2%	89.3%	0.43
Cultivator	27.2%	26.8%	0.91	30.6%	25.6%	0.22
Non-Ag Self-employed	9.0%	9.1%	0.99	9.0%	9.1%	0.96
Casual Labor	16.8%	15.9%	0.76	18.8%	15.5%	0.31
Salaried Work	4.3%	7.0%	0.20	5.8%	5.1%	0.75
Domestic Work	29.9%	27.4%	0.52	21.3%	32.2%	0.01
Unemployed	5.0%	5.9%	0.60	6.7%	4.8%	0.31
Not in Labor Force	5.6%	6.2%	0.75	5.9%	5.8%	1.00
Log Daily Casual Earnings	3.70	3.90	0.00	3.73	3.80	0.17
Number of Observations	108,201	73,276		58,742	122,735	

This table presents means of the main outcomes used in the paper for different samples. All samples are restricted to persons aged 18 to 60. Column (1) is restricted to districts that received the workfare program prior to April 2008. Column (2) includes only districts that received the program after April 2008. Column (3) presents the p-values of the Student's t-test of equality of means in Column (1) and (2). Column (4) restricts the sample to star states. Star states include Andhra Pradesh, Chhatisgarh, Himachal Pradesh, Madhya Pradesh, Tamil Nadu, Rajasthan, and Uttarkhand. Column (5) includes districts in non-star states. Column (6) presents the p-values of the Student's t-test of equality of means in Column (4) and (5). For the Student's t-test in column (3) and (6) standard errors are computed assuming correlation of individual observations over time within each district.

Table 3: Effect of NREGA on Labor Market Outcomes

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)
Program X Dry	1.174***	-1.306*	0.281	-0.149	0.0473**
	(0.298)	(0.759)	(0.544)	(0.467)	(0.0213)
Program X Rainy	0.460**	0.673	-0.652	-0.481	0.0287
	(0.179)	(0.790)	(0.597)	(0.546)	(0.0240)
Observations	356,636	356,636	356,636	356,636	64,167
District Controls	Yes	Yes	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes

Each column presents results from a separate regression. All regressions include district and year-quarter fixed effects. The sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. Private, unemployed, and not in the labor force are estimates of the percentage of total time spent working in private sector work (including domestic work), unemployed or not in the labor force. Log daily casual earnings is the log of earnings per day worked for people who report working in casual labor. Deflated earnings are deflated using the monthly, state-level price index for agricultural labourers from the Indian Labour Bureau. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. All estimates are computed using weights proportional to district population. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). Worker controls include dummy variables for gender, age group, education levels, caste, religion and marital status. All estimates are computed using sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table 4: Effect of NREGA in States which implement

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Casual Daily Earnings
	(1)	(2)	(3)	(4)	(5)
Program X Dry X Star States	3.132***	-3.071***	0.108	-0.169	0.0898***
	(0.682)	(1.103)	(0.725)	(0.589)	(0.0258)
Program X Rainy X Star States	0.368	0.782	-0.846	-0.305	0.0458
	(0.228)	(0.985)	(0.765)	(0.662)	(0.0285)
Program X Dry X Other States	0.00622	-0.260	0.424	-0.171	0.0181
	(0.168)	(0.809)	(0.608)	(0.509)	(0.0246)
Program X Rainy X Other States	0.0804	1.000	-0.496	-0.584	0.00869
	(0.158)	(0.851)	(0.649)	(0.563)	(0.0267)
Observations	356,636	356,636	356,636	356,636	64,167
District Controls	Yes	Yes	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes

Each column presents the results of a separate regression. All regressions include district and year-quarter fixed effects. The sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. The unit of observation is a person. The outcomes are defined as in Table 3. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). Worker controls are listed in the notes of Table 3. Star states is a dummy variable equal to one for districts within star states. Other states is a dummy variable equal to one for districts that are not in star states. See Table 2 for a description of star states. All estimates are computed using sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and

Table 5: Placebo Treatment

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)
Program X Dry	-0.00116	-1.397**	0.702	0.697	0.00591
	(0.100)	(0.699)	(0.520)	(0.466)	(0.0217)
Program X Rainy	-0.103	-0.851	0.249	0.705*	0.0134
	(0.0806)	(0.625)	(0.492)	(0.425)	(0.0220)
Observations	383,881	383,881	383,881	383,881	67,676
District Controls	Yes	Yes	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes

Each column presents the results of a separate regression. All regressions include district and year-quarter fixed effects. The sample is composed of all adults aged 18 to 60 interviewed from July 1999 to June 2000 and from July 2004 to June 2005. The unit of observation is a person. The outcomes are defined as in Table 3. Program is a dummy variable equal to one for early districts during July 2004 to June 2005. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2005 (post-program). Worker controls are listed in the notes of Table 3. All estimates are computed using sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table 6: Welfare Gains by Expenditure Quintile

		Expenditure Quintile					
		Poorest	Second	Third	Fourth	Richest	Full Sample	Construction
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Household Expenditures and Income								
(1)	Monthly Consumption Per Capita	293.7	400.7	496.9	625.0	1081	540.3	NSS 2004-5
(2)	Total Monthly Consumption	1915	2416	2700	3207	4638	2840	NSS 2004-5
(3)	Total Earnings per Month for Adults doing Casual Labor	831	780	629	469	308	631	NSS 2004-5
(4)	Casual Earnings as Fraction of Household Consumption	0.43	0.32	0.23	0.15	0.07	0.22	NSS 2004-5
(5)	Average Earnings per Day Worked by Adults	39.9	44.1	47.4	51.2	57.3	44.9	NSS 2004-5
Gain from wage change								
(6)	Fraction of Casual Labor Costs Paid by Quintile	4.0%	6.2%	10.3%	14.2%	30.1%	12.8%	NCAER 1999
(7)	Estimated Monthly Labor Cost per Household	125	194	325	447	951	404	(6) x Full (3) x 5
(8)	Net Labor Earnings per Month	707	585	304	22	-643	227	(3) - (7)
(9)	Wage change	4.7%	4.7%	4.7%	4.7%	4.7%	4.7%	Estimated
(10)	Net Income Gain from Wage Change	33.2	27.5	14.3	1.0	-30.2	10.7	(8) x (9)
Gain from Government employment								
(11)	Increase in Days in Public Employment per HH per Month	1.93	0.96	1.23	0.82	0.57	1.10	Estimated
(12)	Average Private Sector Wage	44.9	44.9	44.9	44.9	44.9	44.9	NSS 2004-5
(13)	Average Government Wage	53.8	53.8	53.8	53.8	53.8	53.8	(12) + 20%
(14)	Direct Gain with Outside Option at 30% of Market Wage	77.9	38.8	49.8	33.0	23.1	44.4	(11) x (13)
Total Gain								
(15)	Total Gain with Outside Option at 30% of Market Wage	111.1	66.3	64.1	34.0	-7.1	55.1	(10) + (14)
Gain from Wage Change as Fraction of Total Gain								
(16)	Assuming Outside Option is 30% of Market Wage	29.9%	41.5%	22.3%	3.0%	**	19.4%	(10)/(15)
Total Gain as Fraction of Total Expenditures								
(17)	Assuming Outside Option is 30% of Market Wage	5.8%	2.7%	2.4%	1.1%	-0.2%	1.9%	(15)/(2)

Columns (1) to (5) correspond to different quintiles based on household per capita expenditure. Column (6) is all households. The last column indicates how each figure is obtained. Rows (1) to (5) use data from the NSS 2004-05 Employment Survey to compute averages for each quintile using survey sample weights. The fraction of casual labor costs paid by quintile (sixth row) is computed using data from the 1999-00 ARIS-REDS survey as follows. First we use monthly per capita expenditure to define quintiles. Second, by quintile, we aggregate all wages paid by the household to adult laborers. Third, we aggregate all income from casual labor supplied outside the household by all adults aged 18 to 60. The means in row (6) are obtained for each quintile by dividing total wages paid by total wage income received across all households. The wage change in row (9) is equal to the estimate of the program impact during the dry season from the specificiation in Table 5 with workers controls. The increase in days in public employment per household per month reported in row (11) is obtained from the regressions reported in Table A3. Row (16) , nothing is reported for the fifth quintile because the "gain from wage change" is a loss for this quintile.

FOR ONLINE PUBLICATION ONLY

A Appendix

A.1 History of Public Works Programs in India

India has a long history of providing public works dating back to British rule. Three large-scale public works programs deserve specific mention. First is the Maharashtra Employment Guarantee Scheme passed in 1976 and still in force today. The NREGA is in part based on the design of the Maharashtra EGS.

Second, the Sampoorn Grameen Rozgar Yojana (SGRY) started in 2001 with the purpose of generating employment across India and was still active until 2008. The total allocation to the SGRY was 35 billion Rupees per year from 2004-2008 (Afridi, 2008).

Finally, the National Food for Work Program was introduced as a pilot for the NREGA in 150 of the phase one districts, with an allocation of 60 billion Rupees in fiscal year 2005-06 (Afridi, 2008). As a comparison, during fiscal years 2006-07 and 2007-08, the allocation for the NREGA was 116 billion Rupees. Confirming existing field observations that the National Food for Work Program was poorly implemented and plagued with massive leakages Dreze (2005), we find little evidence of an increase in public works during this pilot period.

A.2 Determinants of Government Employment Provision

The central government funds most of the expenditure for the NREGA (all of labor and 75% of material expenditures). However, the responsibility of implementing the scheme is left to the states and the lower administration levels (districts and village councils). In principle, local officials are meant to respond to worker demand for work, but the process required to provide work requires considerable administrative capacity: selecting public works projects, applying for funds, opening the works, sanctioning expenditures, making payments to workers and suppliers of materials etc. When the scheme started in each district, awareness campaigns also had to be implemented by the administration, sometimes with the help of civil society organizations. Depending on the administrative capacity of each state, NREGA implementation was initially more or less successful.

During the period we study, which is immediately after the launch of the scheme, the states of Andhra Pradesh, Chattisgarh, Himachal Pradesh, Madhya Pradesh, Rajasthan, Tamil Nadu and Uttarakhand, which we call "star states" in the analysis provided significantly more employment than other states (see Figure 2. This was partially due to demand for work in these states. However, very poor states such as Bihar, Jharkhand, Orissa, and Uttar Pradesh where demand should be high saw little employment generation. In this second group of states, lack of administrative capacity and rampant corruption hampered public employment delivery, despite large potential demand (Khera, 2011; Dutta et al., 2012). In the 2009-10 NSS employment survey, workers were asked whether they had, and whether they desired NREGA employment. Using answers to these questions, Dutta et al. (2012) confirm that three years after the scheme started, demand for work is still more rationed in the poorest states of India.

In order to investigate the sources of observed disparities in NREGA implementation across states, we use NSS data to regress time spent on public works by rural adults in 2007-08 on the set of district controls presented in Table 1, and plot state-level averages of the residuals in Figure A.1. The seven states we defined as star states all have higher public employment provision than predicted by the model. This finding is consistent with the view that differences in public employment provision across states are due to supply factors (e.g. political will or administrative capacity) rather than demand factors (e.g. poverty or labor market conditions). The state which has lowest public employment provision compared to the predicted value in Figure A.1 is Maharashtra, which had its own employment guarantee since the 1970s and whose government was reluctant to implement NREGA.

A.3 Theoretical Appendix

A.3.1 Utility maximization

Each household has a utility function $u(c_i, l_i)$ over household consumption c_i and leisure l_i. We assume the function is increasing and concave in both arguments. Let L_i^s denote household total labor supply and D_i denote household total labor demand. Household labor supply L_i^s has two components: family labor used for household production L_i^f and wage work supplied by household members to the market L_i^o. Household labor demand D_i also has two components: family labor L_i^f and hired by the household L_i^h. Households choose L_i^f, L_i^o ,L_i^h and c_i to solve the following maximization problem:

$$\max_{c_i, L_i^f, L_i^o, L_i^h} u(c_i, T - L_i^f - L_i^o)$$
$$\text{s. t. } c_i = pWL_i^o + A_i G(L_i^f + L_i^h) - WL_i^h$$

We further impose that the optimal labor quantities L_i^f, L_i^o ,L_i^h cannot be negative, and both consumption and leisure must be positive ($c_i > 0$ and $T > L_i^f - L_i^o$). We write the Lagrangian:

$$\mathcal{L} = u(c_i, T - L_i^f - L_i^o) + \lambda(pWL_i^o + A_iG(L_i^f + L_i^h) - WL_i^h - c_i)$$

The Kuhn Tucker conditions write

$$\begin{aligned} u'_c - \lambda \leq 0 \quad & and \quad c(u'_c - \lambda) = 0 \\ -u'_l + \lambda pw \leq 0 \quad & and \quad L_i^o(u'_l - \lambda pw) = 0 \\ -u'_l + \lambda A_iG' \leq 0 \quad & and \quad L_i^f(u'_l - \lambda A_iG') = 0 \\ \lambda(A_iG' - W) \leq 0 \quad & and \quad L_i^h(W - A_iG') = 0 \end{aligned}$$

However, we assume that $c_i > 0$ hence the first condition simply yields: $u'_c = \lambda > 0$. We can rewrite the three other conditions using this equality:

$$\begin{aligned} pw \leq \frac{u'_l}{u'_c} \quad & and \quad L_i^o(u'_l - \lambda pw) = 0 \\ A_iG' \leq \frac{u'_l}{u'_c} \quad & and \quad L_i^f(u'_l - \lambda A_iG') = 0 \\ A_iG' \leq W \quad & and \quad L_i^h(W - A_iG') = 0 \end{aligned}$$

There are seven cases to consider depending on whether the optimal L_i^f, L_i^o ,L_i^h are null.

Cases 1 Let us assume that $L_i^o > 0$, $L_i^h > 0$ and $L_i^f > 0$. Then we must have $pw = \frac{u'_l}{u'_c}$ and $W = A_iG'$. However, we also need to have $A_iG' = \frac{u'_l}{u'_c}$. Hence this case is only possible if $p = 1$, i.e. households can be suppliers and buyers of labor at the same time if and only if the labor market is without friction. In the general case with friction, households cannot be on both sides of the market.

Case 2 we assume that $L_i^o > 0$, $L_i^f = 0$ and $L_i^h = 0$. Then we must have that $pw = \frac{u'_l}{u'_c}$, $A_iG' \leq \frac{u'_l}{u'_c}$ and $A_iG' \leq W$. This case is unlikely. Households cannot not choose to supply labor to the market without producing anything on their farm, because for any W one can find a L_i^f small enough so that the marginal productivity of labor will be higher than pW. This is because we assumed that all households are able to produce ($A_i > 0$).

Case 3 we assume that $L_i^0 = 0$, $L_i^f = 0$ and $L_i^h > 0$. Then we must have that $pw \leq \frac{u'_l}{u'_c}$, $A_iG' \leq \frac{u'_l}{u'_c}$ and $A_iG' = W$. This case is also unlikely. Households will not optimally choose to hire workers without supplying any family labor (i.e. reduce their consumption and devote all their time to leisure), because for any W one could find a L_i^f small enough so that the marginal rate of substitution of consumption to leisure will be higher than W.

Case 4 where $L_i^0 = L_i^f = L_i^h = 0$ is not optimal if $A_i > 0$.

Case 5 the household is net supplier of labor ($L_i^0 > 0$, $L_i^f > 0$ and $L_i^h = 0$). Then the marginal productivity on the farm is equal to wage labor earnings, which is less than the market wage (i.e. $\frac{u'_l}{u'_c} = A_i G' = pW \leq W$) .

Case 6 the household is net buyer of labor ($L_i^0 = 0$, $L_i^f > 0$ and $L_i^h > 0$). Then the marginal productivity on the farm is equal to the market wage (i.e. $\frac{u'_l}{u'_c} = A_i G' = W \geq pW$).

Case 7: the household does not participate to the labor market ($L_i^0 = 0$, $L_i^f > 0$ and $L_i^h = 0$). Then the marginal productivity on the farm is equal to the marginal rate of substitution between consumption and leisure. It is lower than the market wage and higher than labor market earnings (i.e. $\frac{u'_l}{u'_c} = A_i G' \in [pW, W]$).

If $p < 1$ only cases 5, 6 and 7 are possible; households are either labor suppliers, labor buyers or they do not participate to the market. If $p = 1$, cases 1, 5 and 6 are possible and case 7 contracts to a single point: households may be labor sellers, labor buyers, or both.

A.3.2 Productivity thresholds

For each value of the wage W, let us consider the value of the productivity factor A_i such that labor supply and labor demand from household i are equal:

$$L_i^s(W, A_i G'(D(W, A_i))) = D_i(W, A_i)$$

Let us denote this value $\phi(W)$. Since $L_Y^s \leq 0$ and $D_A(W, A_i) \geq 0$, $\phi(W)$ exists and is unique. Since $L_W^s > 0$ and $D_W(W, A_i) < 0$, the function $\phi(W)$ is strictly increasing in W .

Proposition 1: A household i is net labor buyer if and only if $A_i > \phi(W)$

Proof: A household with $A_i = \phi(W)$ therefore supplies and demands $D(W, \phi(W))$ labor. Since the marginal cost of hiring labor is W while the marginal value of working in the labor market is $p_i W < W$, the household will always supply labor to its own production function at least up to $D(W, \phi(W))$. Therefore, households with $A_i = \phi(W)$ are neither net labor supplying nor net labor buying households. For $A_i > \phi(W)$, we will have $D(W, A_i) > L^s(W, A_i G'(D(W, A_i)))$, so that the household will be a net labor buyer as long as it can hire labor at W and as long as the marginal value of time is given by W as well.

Since net labor buyers supply labor only to their own farm, this will be the case. Net labor buyers will always face an effective marginal wage of W. Therefore, if $A_i < \phi(W)$, then $D(W, A_i) < L^s(W, A_i G'(D(W, A_i)))$, so that households will not be net buyers of labor.

Proposition 2: A household$i$ is net labor supplier if and only if $A_i < \phi(pW) < \phi(W)$
Proof: A household with $A_i = D(pW, \phi(pW))$ will supply and demand D_w units of labor but because $pW < W$ we have $D(pW, \phi(pW)) < D(W, \phi(W))$ and $\phi(pW) < \phi(W)$. For a household with $A_i < \phi(pW)$, we will have $D(pW, A_i) < L^s(pW, A_i G'(D(pW, A_i)))$, so that the household will be a net labor supplier. Net labor suppliers will always face an effective marginal wage of $p_i W$. For a household with $A_i > \phi(pW)$, we will have $D(p_i W, A_i) > L^s(pW, A_i G'(D(p_i W, A_i)))$, so that the household will not be a net labor supplier.

Proposition 3: For $A_i \in [\phi(pW), \phi(W)]$, household i is neither net supplier or buyer of labor.
Proof: This follows directly from the first two propositions. For $A_i \in [\phi(pW), \phi(W)]$, labor supply and demand D will solve $D = L^s(A_i G'(D), A_i G(D))$. Note that for $A_i \in [\phi(pW), \phi(W)]$, the labor supply and demand will satisfy $A_i G'(D) \in [p_i W, W]$.

Hence the three possible solutions to the utility maximization problem correspond to different values for the productivity factor A_i. The most productive households (e.g. those with most land) are net labor buyers and the marginal productivity on their farm is the market wage. The least productive households (e.g. those with little land) are net labor sellers and the marginal productivity on their farm is equal to wage labor earnings pW. Households with intermediary levels of productivity will not participate to the market (this last case contracts to a single productivity level if $p = 1$.)

A.3.3 Compensating Variation Derivation

Let us first consider households with low productivity levels $A_i < \phi(pW)$. The equation equating expenditure to income writes

$$e(pW, u_i) = \pi_i(pW) + pWT + (W_g - pW)L_i^g + z_i$$

We derive the change in z_i required to maintain the equality, and therefore maintain the same utility level, following a change in L_g. We do this by differentiating Equation A.3.3 with respect to L_g:

$$\frac{de(pW,u_i)}{dL^g} = p\pi_i'(pW)\frac{dW}{dL_g} + pT\frac{dW}{dL_g} + (W_g - pW)\frac{dL_i^g}{dL_g} - pL_i^g\frac{dW}{dL_g} + dz_i$$

By the envelope theorem $\frac{de(pW,u_i)}{dW} = p(T - L_i^s)$ and $\pi_i'(pW) = -D_i$. Using these results and re-arranging yields:

$$\begin{aligned} -dz_i &= (L_i^s - L_i^g - D_i)pW\frac{dW/W}{dL_g} + (W_g - pW)dL_i^g \\ &= \textit{Net Casual Labor Earnings} \times \frac{dW/W}{dL_g} + (W_g - pW)dL_i^g \end{aligned}$$

For households with high productivity levels $A_i > \phi(W)$ the equation equating expenditures to income writes:

$$e(W,u_i) = \pi_i(W) + WT + (W_g - W)L_i^g + z_i$$

Using the same demonstration as before, but replacing p with 1, we find that:

$$\begin{aligned} -dz_i &= (L_i^s - L_i^g - D_i)W\frac{dW/W}{dL_g} + (W_g - W)dL_i^g \\ &= \textit{Net Casual Labor Earnings} \times \frac{dW/W}{dL_g} + (W_g - W)dL_i^g \end{aligned}$$

Finally, for households with intermediary productivity levels $A_i \in [\phi(pW), \phi(W)]$, the equation equating expenditures with revenues writes:

$$e(\widetilde{W}_i, u_i) = \pi_i(\widetilde{W}_i) + \widetilde{W}_i T + (W_g - \widetilde{W}_i)L_i^g + z_i$$

where $\widetilde{W}_i$ is the shadow wage which does not depend on W. The program only affects households welfare through direct participation, and the compensating variation has the simple form:

$$-dz_i = (W_g - \widetilde{W}_i)dL_i^g$$

However, since these households do not buy or sell labor on the market, their net casual labor earnings are zero, and we can also write:

$$-dz_i = Net\ Casual\ Labor\ Earnings \times \frac{dW/W}{dL_g} + (W_g - \widetilde{W_i})dL_i^g$$

Which completes our demonstration.

A.3.4 Impact of Government Hiring on the labor market equilibrium

The market clearing condition imposes that labor supply of households with low productivity and labor demand of households with high productivity are equal. It writes:

$$p\int_{\underline{A}}^{\phi(pW)} [L_i^s(pW) - D_i(pW) - L_i^g]dA_i = \int_{\phi(W)}^{\overline{A}} [D_i(W) - L_i^s(W) + L_i^g]dA_i \tag{4}$$

To determine the impact on wages of public sector hiring we need to differentiate the market clearing condition with respect to L^g. We use Leibnitz integral rule which yields for the left-hand side of equation 4:

$$\begin{aligned}\frac{dp\int_{\underline{A}}^{\phi(pW)}[L_i^s(pW) - D_i(pW) - L_i^g]dA_i}{dL^g} &= [L_i^s(pW,\phi(pW)) - D_i(pW,\phi(pW)) - L_i^g]\phi'\frac{dW}{dL^g} \\ &+p\int_{\underline{A}}^{\phi(pW)}\frac{d[L_i^s(pW) - D_i(pW) - L_i^g]}{dL^g}dA_i\end{aligned}$$

By definition, net labor demand of households with productivity levels $\phi(pW)$ is zero, so that $[L_i^s(pW,\phi(pW)) - D_i(pW,\phi(pW)) - L_i^g] = 0$. Hence the first term is null.

A similar simplification can be made for $\phi(W)$, while differentiating the right-hand side of equation 4. Hence the derivative of 4 with respect to L^g writes:

$$p\int_{\underline{A}}^{\phi(pW)}[\frac{dL_i^s(pW)}{dL^g} - \frac{dD_i(pW)}{dL^g} - \frac{dL_i^g}{dL^g}]dA_i = \int_{\phi(W)}^{\overline{A}}[\frac{dD_i(W)}{dL^g} - \frac{dL_i^g}{dL^g} - \frac{dL_i^s(W)}{dL^g}]dA_i \tag{5}$$

Let us first consider households which are net labor suppliers ($A_i < \phi(pW)$). Public hiring affects labor supply through its effect on the equilibrium wage and through its effect on non-labor income. We decompose the derivative of L_i^s with respect to L^g in two components:

$$\frac{dL_i^s(pW,y_i)}{dL^g} = \frac{dL_i^s(pW,y_i)}{dW}|_{y_i}\frac{dW}{dL^g} + \frac{dL_i^s(pW,y_i)}{dy_i}\frac{dy_i}{dL^g}$$

where $\frac{dL_i^s}{dW}|_{y_i}$ is the derivative of household i's labor supply with respect to the wage holding non-labor income fixed. The slutsky decomposition yields:

$$\frac{dL_i^s(pW,y_i)}{dW}|_{y_i} \quad = p \quad \frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}pL_i^s$$

where $\frac{dL_i^s}{dW}|_u$ is the substitution effect, i.e. the partial derivative of labor supply with respect to the wage holding utility constant. We have that:

$$\begin{aligned}\frac{dy_i^s}{dL_g} &= p\pi_i'(pW)\frac{dW}{dL_g} + (W_g - pW)\frac{dL_i^g}{dL_g} - pL_i^g\frac{dW}{dL_g} \\ &= -pD_i\frac{dW}{dL_g} + (W_g - pW)\frac{dL_i^g}{dL_g} - p\frac{dW}{dL_g}L_i^g\end{aligned}$$

where the second equality follows from the envelope theorem for the profit function $\pi_i'(W) = -D_i$.

Hence, for households with $A_i < \phi(pW)$, we can rewrite the derivative of the labor supply with respect to public hiring as:

$$\frac{dL_i^s(W,y_i)}{dL^g} = p[\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]\frac{dW}{dL_g} + \frac{dL_i^s}{dy_i}(W_g - pW)\frac{dL_i^g}{dL_g}$$

Public hiring affects labor demand only through its effect on the equilibrium wage. Hence the derivative of the labor demand with respect to public hiring writes: $\frac{dD_i(pW)}{dL^g} = pD_i'(pW)\frac{dW}{dL_g}$

Hence, the impact of public sector hiring on the net labor supply of households with

$A_i < \phi(pW)$ is given by the following expression:

$$\begin{aligned}
p \int_{\underline{A}}^{\phi(pW)} [\frac{dL_i^s(pW)}{dL^g} - \frac{dD_i(pW)}{dL^g} - \frac{dL_i^g}{dL^g}]dA_i &= p^2 \int_{\underline{A}}^{\phi(pW)} [\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]\frac{dW}{dL_g}dA_i \\
&+ p \int_{\underline{A}}^{\phi(pW)} \frac{dL_i^s}{dy_i}(W_g - pW)\frac{dL_i^g}{dL_g}dA_i \\
&- p^2 \int_{\underline{A}}^{\phi(pW)} D_i'(pW)\frac{dW}{dL_g}dA_i - p \int_{\underline{A}}^{\phi(pW)} \frac{dL_i^g}{dL_g}dA_i \quad (6)
\end{aligned}$$

Using similar arguments, we can write the impact of public sector hiring on the net labor demand of households with $A_i > \phi(W)$ as:

$$\begin{aligned}
\int_{\phi(W)}^{\overline{A}} [\frac{dD_i(W)}{dL^g} + \frac{dL_i^g}{dL^g} - \frac{dL_i^s(W)}{dL^g}]dA_i &= - \int_{\phi(W)}^{\overline{A}} [\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]\frac{dW}{dL_g}dA_i \\
&- \int_{\phi(W)}^{\overline{A}} \frac{dL_i^s}{dy_i}(W_g - W)\frac{dL_i^g}{dL_g}dA_i \\
&+ \int_{\phi(W)}^{\overline{A}} D_i'(W)\frac{dW}{dL_g}dA_i + \int_{\phi(W)}^{\overline{A}} \frac{dL_i^g}{dL_g}dA_i \quad (7)
\end{aligned}$$

Plugging equations 6 and 7 into 5 and re-arranging yields:

$$\frac{dW}{dL^g} = \frac{E_1 - E_2}{-E_3 + E_4} \quad (8)$$

Where:

$$E_1 = p \int_{\underline{A}}^{\phi(pW)} \frac{dL_i^g}{dL_g}dA_i + \int_{\phi(W)}^{\overline{A}} \frac{dL_i^g}{dL_g}dA_i$$

is the direct crowding out effect of public employment on wage labor (for the poorest households) and self-employment (for the richest households), $E_1 > 0$

$$E_2 = p \int_{\underline{A}}^{\phi(pW)} [\frac{dL_i^s}{dy_i}(W_g - pW)]\frac{dL_i^g}{dL_g}dA_i + \int_{\phi(W)}^{\overline{A}} [\frac{dL_i^s}{dy_i}(W_g - W)]\frac{dL_i^g}{dL_g}dA_i$$

is the effect on aggregate labor supply through non-labor income $E_2 < 0$. Hence $E_1 - E_2$ is positive as long as the income effect is not positive and large.

$$E_3 = p^2 \int_{\underline{A}}^{\phi(pW)} D'(pW)dA_i + \int_{\phi(W)}^{\overline{A}} D'(W)dA_i$$

is the effect on aggregate labor demand through a change in the wage, $E_3 < 0$.

$$E_4 = p^2 \int_{\underline{A}}^{\phi(pW)} [\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]dA_i + \int_{\phi(W)}^{\overline{A}} [\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]dA_i$$

is the effect on aggregate labor supply through a change in the wage. If leisure is not a luxury good, an increase in the wage should increase labor supply, so that $E_4 > 0$. Hence government hiring increases the equilibrium wage because $E_1 - E_2 > 0$, $-E_3 > 0$ and $E_4 > 0$. The effect is stronger when demand is less elastic (small $-E_3$), when labor supply is less elastic to the wage (small E_4).

Assuming that $p = 1$ we obtain the following

$$\frac{dW}{dL^g} = \frac{\int_{\underline{A}}^{\overline{A}} \frac{dL_i^g}{dL_g}dA_i - \int_{\underline{A}}^{\overline{A}}[\frac{dL_i^s}{dy_i}(W_g - W)]\frac{dL_i^g}{dL_g}dA_i}{-\int_{\underline{A}}^{\overline{A}} D'(W)dA_i + \int_{\underline{A}}^{\overline{A}}[\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - D_i - L_i^g)]dA_i}$$

From this equation, we see that an increase in government hiring will raise wages as long as the income effect is not positive and larger than one ($\int_{\underline{A}}^{\overline{A}} L_{y_i}^s(W_g - W)dA_i < 1$). The increase will be larger if demand is less elastic (small $-D'(W)$) or if labor supply is less elastic (small $\int_{\underline{A}}^{\overline{A}} \big(\frac{dL_i^s}{dW}|_u + \frac{dL_i^s}{dy_i}(L_i^s - L_i^g - D_i)\big)dA_i$).

In the special case where $p = 1$, the model indicates how empirical estimates of the impact of government hiring on employment and wages can be used to compute the labor demand elasticity. In a frictionless labor market, the change in aggregate private sector employment can be written as: $\frac{dD}{dL^g} = D'(W)\frac{dW}{dL^g}$, where $D'(W) = \int_{\underline{A}}^{\overline{A}} D_i'(W)dA_i$. Hence, in this framework, we can compute the elasticity of labor demand as the ratio of the percentage change in the wage divided by the percentage change in employment.

A.3.5 Impact on Household Consumption

In this section, we derive the impact of a workfare program on household consumption. The impact on consumption is different from the impact on welfare because it also includes labor supply effects. Household consumption is given by:

$$c_i = \pi_i(\widetilde{W}_i) + \widetilde{W}_i L_i^s(\widetilde{W}_i, y_i) + (W_g - \widetilde{W}_i) L_i^g \tag{9}$$

Assuming a small change in L^g ($\{L_i^g\}$), we totally differentiate 9 to obtain:

$$\begin{aligned}
\frac{dc_i}{dL^g} &= (W_g - \widetilde{W}_i)\frac{dL_i^g}{dL^g} \\
&+ \widetilde{W}_i L_{yi}^s (W_g - \widetilde{W}_i)\frac{dL_i^g}{dL^g} \\
&+ (L_i^s - D_i - L_i^g)\frac{dW}{dL^g} \\
&+ \widetilde{W}_i \Big[\frac{dL_i^s}{dW}|_u + L_{y_i}^s (L_i^s + T - L_i^g - D_i)\Big]\frac{dW}{dL^g}
\end{aligned}$$

The first term is the income gain due to participation in public works. The impact of this increase in income on labor supply is captured by the second term. It is negative if leisure is a normal good. Together, these first two terms yield the increase in consumption that would be observed by matching participants and non-participants in program areas.

The two last terms express the "indirect benefit", i.e. income gains accruing to households through equilibrium effects. The third term is the change in income due to the equilibrium change in the wage (holding labor supply constant). The last term captures the labor supply response due to the change in income from the equilibrium change in the wage. It is composed of a positive substitution effect and an income effect, which could be negative for households that are net buyers of labor.

A.3.6 Imperfect Competition

We assume that the marginal productivity of labor is equal to the wage rate. Some have noted the presence of market power on the part of employers Binswanger and Rosenzweig (1984). If employers have market power then government hiring may actually increase private sector wages *and* employment. We refer the interested reader to Basu et al. (2009), who provide a full analysis. Here, we sketch the main intuition and discuss the implications for the interpretation of the empirical results. A monopsonistic employer with production function $F(L)$ facing an inverse labor supply curve $W(L)$ sets the wage and employment such that:

$$F'(L^*) = W(L^*) + W'(L^*)L^* \tag{10}$$

This is the well-known result that the marginal productivity of labor will be above the wage rate if employers exercise their market power. The extent of the distortion depends on the slope of the labor supply curve ($W'(L)$). If the selection rule used by the government to hire workers under the workfare program shifts $W'(\cdot)$ down (makes labor supply more elastic), then all things equal, L^* must increase to maintain the equality in Equation 10. Since the workfare program also reduces the available workforce, the net effect on private sector work is ambiguous.

For the present analysis, the important issue is whether, given the rise in wages due to the program, Equation 2 still captures the welfare impact of the program under imperfect competition. For labor suppliers, the welfare impact is the same. For labor buyers, however, Equation 2 no longer correctly captures the welfare impact of the program since the welfare impact now depends on how the inverse labor supply function changes, which in turn will be a function of the particular rationing rule used by the government.

A.4 Data Appendix

A.4.1 National Sample Survey Organisation: Employment Surveys

Sample: The main data source used in this paper is the National Sample Survey rounds 55, 61, 64 and 66. These surveys are conducted on an irregular basis roughly every two years. They are "thick" rounds, with a sample size of roughly 70 thousand rural households.[24] The surveys are stratified by urban and rural areas of each district. The survey is conducted from July to June, and in each district, surveying is divided into four sub-rounds each lasting three months. Although the sample is not technically stratified by sub-round, the NSSO states that it attempts to distribute the number of households surveyed evenly within each district sub-round.

[24] Two additional Rounds 60 and 62, have been conducted which we do not use in the analysis, because they are "thin" rounds, with roughly 35 thousand rural households.

Table A.1 presents evidence on how the sample is distributed throughout the years in practice. For employment outcomes, a district is missing in a given quarter if no household was interviewed. From Table A.1 we see that in 1999-00, between 16 and 20 districts were missing per quarter, which is due to two separate issues. First, five districts were not at all surveyed, second, despite NSSO effort to distribute interviews in a given district during the whole year a few districts have been surveyed in some quarters only. This methodological issue was fixed in the later rounds, as can be seen in Table A.1. In 2004-05 and 2007-08, which are the years we use in our main specification, we have observations for almost all 497 district-quarters but one or two. In 2009-10, four districts are missing because they could not be matched unambiguously with 2007-08.

For casual wages, a district is missing in a given quarter if no household was surveyed or if no prime-age adult reported doing casual work in the past week. As a result the proportion of missing observations is larger for wages than for the employment variables. The fraction of missing observations is as high as seven percent for the first quarters of the survey year 1999-00, but not more than four percent for the years 2004-05 and 2007-08. One might worry that by reducing private employment the program may increase the probability that a district is missing in a given quarter. However, this does not seem to be a major concern given that the fraction of early districts among non-missing observations is constant across quarters.

Outcomes: Our main outcomes are individual measures of employment and wages, which are constructed as follows. The NSS Employment Survey includes detailed questions about the daily activities for all persons over the age of four in surveyed households for the most recent seven days. We compute for each person the fraction of days in the past seven days spent in each of four mutually exclusive activities: non-government work, public works, not in the labor force, and unemployed. Individuals who worked in casual labor over the past seven days are asked their total earnings from casual labor. For each individual we compute average earnings per day worked in casual labor. We perform a similar computation using days spent doing salaried labor to construct our measure of daily salaried earnings.

Individual controls For the purpose of our analysis of the impact of NREGA on casual labor earnings, we include workers characteristics as controls in our main specification. Individual controls include dummy variables for age groups (31-40, 41-50, 51-60), education levels (below primary, primary, middle, secondary or higher), caste (ST, SC, OBC), religion (Muslim, Other), gender (Female) and marital status (Married). The omitted category is single, illiterate, hindu males, aged 18-30 and belonging to general caste.

A.4.2 District Controls

Table 1 provides a list of district controls and their sources. Here, we describe how the district controls are constructed.

Census A number of the district controls are computed from the primary census abstract of 2001. In all cases, we use information for rural areas only, which we then aggregate to the district level. We compute "fraction of scheduled tribes" and "fraction of scheduled castes" by dividing by total population. "Population density" is obtained by dividing total population by total area. "Literacy rate," "male labor force participation ratio" and "female labor force participation ratio" are respectively computed by dividing the number of literate persons, of male workers and of female workers respectively by total population aged six and over. "Fraction of labor force in agriculture" is obtained by dividing the number of rural individuals who report working as cultivators or agricultural laborers as their main or secondary occupation by the total number of workers. Finally, we use information from the census village directory to compute "irrigated cultivable land per capita" and "non irrigated cultivable land per capita."

Agricultural Productivity: We compute agricultural productivity per worker for each agricultural year in each district using two sources of data. First, the Ministry of Agriculture publishes yearly data on output and harvest prices of 36 grain and cash crops in every district [25]. This allows us to compute the value of agricultural production for every district-year. Second, we use National Sample Survey data to estimate the number of (self employed and wage) workers active in agriculture for every district-year. NSS survey years match exactly the Ministry of Agriculture definition of agricultural years (July-June). Hence, dividing output value by the number of agricultural works yields agricultural productivity per worker for each NSS survey year.

Rainfall To control for monthly rainfall at the district level over the period 1999-2010, we use data from the Tropical Rainfall Measuring Mission (TRMM), which is a joint mission between NASA and the Japan Aerospace Exploration Agency (JAXA). The TRMM Muli-Satellite Precipitation Analysis provides rainfall data for every three hours at a resolution

[25] Data is available at `http://eands.dacnet.nic.in/`.

of 0.25 by 0.25 degree grid-cell size. Rainfall measurement are made by satellite and calibrated using monthly rain gauge analysis data from the Global Precipitation Climatology Project (GPCP).[26] The data is then scaled up to obtain mean monthly rainfall for every cell. The match between TRMM data and Indian districts was made by Thiemo Fetzer (Fetzer, 2013). On average there are 6 grid-cells per district. We compute cumulative rainfall in each district-month as the sum of rainfall since July 1st, and express it as percentage deviation from the 1998-2011 mean for this district-month.

Other district controls "Pre-election year" is a dummy for whether state assembly or Panchayati Raj (local) elections are to be held in the following year. To construct this control, we used online reports from the Electoral Commission of India[27] and from the State Election Commissions of each states. “PMGSY Road Construction” is an estimate of the number of km of road built under the national rural roads construction program Pradhan Mantri Gram Sadak Yozna. We use online reports on each road built under the scheme to compute for each district quarter the average number of km completed per quarter over the last five quarters.[28]

A.4.3 ARIS-REDS Household Hired Labor

For our calibration exercise in Section 6, we require estimates of labor hired by households, information which is not available in the NSS Employment Surveys. For this reason, we use the ARIS-REDS survey data, collected by the National Council of Applied Economic Research (Delhi) in 1999-00.[29] The ARIS-REDS survey covers a nationally representative rural sample of Indian households, with detailed information on household expenditures, on household members’ employment income and on operating costs of households’ farm and non-farm businesses.

For each household, we sum all income earned by prime-age household members from casual labor and total labor costs for farm and non-farm businesses. For each consumption quintile, we then compute the total casual payments as a fraction of total casual earnings for all households across all quintiles. Let e_t^q and p_t^q denote casual earnings and casual payments, respectively, for households in consumption quintile q at date t. We compute for each quintile $f_{2000}^q = \frac{p_{2000}^q}{\sum_q e_{2000}^q}$. The resulting fractions are reported in the sixth row of Table 6. As expected

[26]Data is available at `http://trmm.gsfc.nasa.gov/`
[27]`http://www.eci.nic.in/ecimain1/index.aspx`
[28]`http://pmgsy.nic.in/`
[29]`http://adfdell.pstc.brown.edu/arisredsdata/readme.txt`

the fraction of total casual earnings paid by households in the lower quintiles is much lower than the fraction paid by households in the upper quintiles. These fractions sum to less than one across consumption quintiles because some casual labor earnings come from urban employers.

In order to estimate casual labor payments by households of each consumption quintile in 2004-05, we make the assumption that casual labor payments made by each consumption quintile as a fraction of total earnings is constant over time, i.e. $f^q_{2005} = f^q_{2000}$. We then multiply total casual labor earnings from the NSS Employment Survey by the fractions in row six for each consumption quintile to obtain our estimate of casual labor payments by quintile: $\widehat{p^q_{2005}} = f^q_{2005} * \sum_q e^q_{2005}$. Our estimates are shown in row seven of Table 6

A.4.4 Weighting

The NSSO provides sample weights which ensure that the weighted mean of each outcome is an unbiased estimate of the average of the outcome for the population National Sample Survey Office (2010). For the purpose of our analysis, we re-weight observations so that the sum of all weights within each district is constant over time and proportional to the rural population of the district as estimated from the NSS Employment Surveys. When we use NSSO survey weights without reweighting, the results are almost identical to our main results (results not shown). As compared to using ordinary least squares without any weighting, our approach allows us to make sure that our results are not driven by smaller districts with few observations for casual wages. More concretely, let w_i be the weight for person i, and let Ω_{dt} be the set of all persons surveyed in district d at time t. Then the new weight for person i is $w_i \times \frac{\omega_d}{\sum_{i\in\Omega_{dt}} w_i}$ where ω_d is the population weight for district d.

We also present estimates of our main specification without using any sample weight (see A.3). The estimated wage effects increase, which suggests that smaller districts experience larger changes in wages. Perhaps surprisingly, standard errors decrease slightly as compared to the estimation with sample weights. Whether the use of weights enhances precision or not depends on the variance structure of the error term (Solon et al., 2013). On the one hand, smaller districts have fewer observations per quarter, hence taking into account differences in sample size across districts may increase precision. On the other hand, since labor market outcomes are highly correlated within districts, the district error may represent a large share of the variance of the error term, and the use of weights may harm precision. Following Solon et al. (2013) suggestion, we implemented a Breush Pagan test by regressing the squared error term on the inverse of the district sample size. The test confirms both the presence of heteroskedasticity, and the importance of the district error in the variance of the error term, with the latter effect dominating the former.

A.4.5 Construction of District Panel

During the period covered by the analysis, some districts split while other districts merged together. Constructing the district panel requires matching districts both over time as well as across data sets. Fortunately, the NSS district definitions for surveying stayed constant from 2004 to 2008, despite splits and merges. We therefore use the NSS district definitions from this period and match other data sets to these. We first match the NSS 1999-2000 to 2004-05 and 2007-08 data. All districts could be matched between the two surveys but for five districts missing in 1999-00. However about fifty of them had split between 1999-00 and 2005-05. We adopt the following procedure If a given district has split in x districts (x is most of the time equal to two, sometimes three), we duplicate observations from that district x times so that one set of observation can be matched with one of the newly created district. In order to keep the total weight of that district constant, we divide each weight in the 1999-00 data-set by x. We next match the NSS 2009-10 data: all districts but four could be matched unambiguously with districts in NSS 2004-05 and 2007-08 data. In two occurrences, two districts were split to create a third one, making it impossible to match observations from the new districts to a specific district. We remove these districts from 2009-10 data. We further match NSS data with Census 2001 survey, NREGA phases 2005, ARIS-REDS 1999-00 survey, PMGSY road construction data from 2001 to 2010

A.5 Alternative specifications

Berg et al. (2013) and Zimmermann (2013) estimate the labor market impact of the program using empirical strategies different from ours. In this section, we describe how we apply their strategy to our data and compare the resulting estimates with their findings.

A.5.1 Berg et al. (2013)

Berg et al. (2013), use monthly wage time series over the whole 2000-2010 period to estimate the effect of the program using two alternative specifications. The first is a difference-in-differences strategy similar to ours. The second is a trend break model. In order to compare our results with Berg et al. (2013), we use all four survey years (from 1999-00 to 2009-10) from the NSS data and estimate the effect of the program using two different specifications.

First, We estimate our main specification without controls and using the four rounds:

$$Y_{idt} = \beta T_{dt} + \eta_t + \mu_d + \varepsilon_{idt}$$

This specification estimates the program impact using two difference-in-differences. The impact of the program is identified based on the difference between changes in outcomes in early districts and in late districts between 2004-05 and 2007-08 and the difference between changes in outcomes in late districts and in early districts between 2007-08 and 2009-10. The results are presented in the first column of table A.8. We find a significant increase in time spent on public works, a significant decrease in private sector work, and a significant 3.4% increase in casual wages. Findings from this specification are consistent with our main results and our estimates for the 2007-08 to 2009-10 period presented in Table A.5.

Second, we redefine the treatment variable T_{dt} as the number of months since the program was launched in district d. We also include a district specific time trend δ_d and estimate the following equation:

$$Y_{idt} = \beta T_{dt} + \delta_d t + \eta_t + \mu_d + \varepsilon_{idt}$$

This specification identifies the program effect as a break in trends when the program was launched. As Table A.8 shows, the estimates provide strong evidence that the program had a positive effect on time spent in public works, a negative effect on time spent in private sector employment, and a positive effect on casual wages, with an estimated effect of 0.27% per month. Adding district-specific trends changes the magnitude of the coefficients but not their sign or their statistical significance. These results are close to Berg et al. (2013).

A.5.2 Zimmermann (2013)

Zimmermann (2013) uses a regression discontinuity design to identify the effect of the program on employment and wages. The selection of early districts was based on a backwardness ranking made by the Planning Commission for an earlier program (Planning Commission, 2003). Hence within each state, and taking the number of early districts as given, one can use each district's backwardness rank to predict its assignment to early or late phases. One can then estimate the effect of the program by comparing 2007-08 outcomes between early and late phase districts close to the cut-off, controlling for the backwardness rank.

The identifying assumption of this regression discontinuity framework is that absent the program, districts to the left and the right of the cut-off would have had the same labor market outcomes. An important threat to this strategy is manipulation of the assignment of districts to early and late phases. In this context, it seems unlikely that the backwardness index itself was manipulated, since it was defined years before NREGA was invented. However, the number of early districts in each state (and hence the state level cut-off) was the result of an intense political bargain, and is unlikely to be exogenous (Gupta, 2006). We hence have some concerns regarding the validity of the regression discontinuity approach.

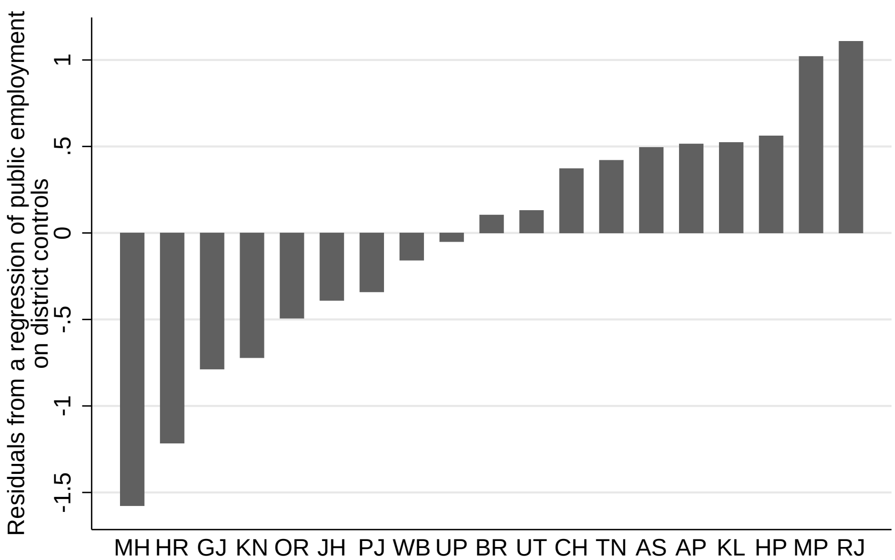
We first assess whether the algorithm accurately predicts whether a district is in early or late phases. Since the ranking is only available for 17 states, Himachal Pradesh and Uttarkhand are excluded from the sample. The prediction is accurate for 95% phase 1 districts, 81% of phase 2 districts and 84% of phase 3 districts. This suggests that the Planning Commission ranking was not perfectly followed for the assignment of districts into implementation phases. Political considerations likely explain why there was imperfect compliance, and why the regression discontinuity design is "fuzzy" (Gupta, 2006).

We follow Zimmermann (2013) and control for the outcome level at baseline (in 2004-05) Y_{ds}^{05} and state fixed effects μ_s in the specification. We also restrict the sample to phase two and three districts (phase 1 districts are far from the cut-off) and estimate different polynomials of the district rank R_{ds} to the left and to the right of the state specific cutoff κ_s. If Y_{ids}^{08} denotes the outcome for individual i in district d and state s in year 2007-08, the estimating equation is:

$$\begin{aligned} Y_{ids}^{08} &= \beta T_{ds} + \delta_0 Y_{ds}^{05} + \delta_1 R_{ds} * (R_{ds} > \kappa_s) + \delta_2 R_{ds} * (R_{ds} < \kappa_s) \\ &+ \delta_3 (R_{ds})^2 * (R_{ds} > \kappa_s) + \delta_4 (R_{ds})^2 * (R_{ds} < \kappa_s) + \mu_s + \varepsilon_{id} \end{aligned}$$

Table *A.9* presents the estimated program impact using this approach. We focus here on the flexible specification, which allows for a different slope to the right and to the left of the cutoff, which is Zimmermann (2013)'s preferred specification. We find a positive but insignificant effect of the program on time spent on public works (0.51 and 0.35 percentage points for the linear and quadratic specification respectively), a negative but insignificant effect on time spent doing private sector work (-0.8 and -1.5 percentage points), and positive effects on private sector wages (6 and 11%). These estimates are reasonably close to those of our preferred specification, and never significantly different from them. The estimation is however very noisy, and none of these estimates is significant.

Figure A.1: Unexplained heterogeneity in employment provision across States



Source: National Sample Survey (NSS) 2007-2008.
The sample is composed of all rural adults in early phase districts.
We compute state-level averages of residuals from a regression of time spent on public works on district controls presented in Table 1. All estimates are computed using sample weights.

Table A.1: Balance of District Panel

	Q3 Jul-Sep (1)	Q4 Oct-Dec (2)	Q1 Jan-Mar (3)	Q2 Apr-Jun (4)
Employment Variables				
1999-00	478	478	483	482
2004-05	497	496	494	495
2007-08	496	497	495	497
2009-10	495	495	494	495
Casual Wages				
1999-00	462	465	474	473
2004-05	478	480	479	481
2007-08	480	483	487	483
2009-10	475	475	476	479

Each cell shows the number of districts with non-missing observations per district-quarter. There are 497 districts in the panel. The NSS attempts to survey an equal number of villages in each districts during each quarter. During thick rounds (1999-2000, 2004-05, 2007-08, 2009-10), this is generally possible. Casual wages are only available for district-quarters during which at least one respondent reports working in casual labor. Five districts were not surveyed in 1999-2000.

Table A.2: Main Specification estimated without controls

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)	(6)
Program X Dry	0.964***	-1.947***	1.623***	-0.640*	0.0353*	0.0460**
	(0.246)	(0.640)	(0.452)	(0.369)	(0.0197)	(0.0232)
Program X Rainy	0.206***	-0.00801	0.653	-0.851**	0.00496	0.0215
	(0.0768)	(0.586)	(0.455)	(0.391)	(0.0198)	(0.0260)
Observations	356,636	356,636	356,636	356,636	64,167	64,167
District Controls	No	No	No	No	No	Yes
Worker Controls	No	No	No	No	No	No

Each column presents the results of a separate regression. All regressions include district and year-quarter fixed effects. In columns 1 through 5, the sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. The unit of observation is a person. The outcomes are defined as in Table 3. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). Worker controls are listed in the notes of Table 3. All estimates are computed using sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table A.3: Main Specification estimated without Sample Weights

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)
Program X Dry	1.110***	-1.402**	0.449	-0.157	0.0679***
	(0.254)	(0.639)	(0.470)	(0.376)	(0.0207)
Program X Rainy	0.466***	0.265	-0.206	-0.526	0.0563**
	(0.159)	(0.624)	(0.496)	(0.371)	(0.0219)
Observations	356,636	356,636	356,636	356,636	64,167
District Controls	Yes	Yes	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes

Each column presents the results of a separate regression. All regressions include district and year-quarter fixed effects. In columns 1 through 5, the sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. The unit of observation is a person. The outcomes are defined as in Table 3. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). Worker controls are listed in the notes of Table 3. All estimates are computed without sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table A.4: Main Specification controlling for changes in outcomes between 1999-00 and 2004-05

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)
Program X Dry	1.163***	-1.666**	0.329	0.235	0.0543**
	(0.297)	(0.731)	(0.505)	(0.443)	(0.0210)
Program X Rainy	0.457***	0.274	-0.630	-0.0305	0.0386
	(0.174)	(0.740)	(0.548)	(0.536)	(0.0235)
Observations	356,636	356,636	356,636	356,636	64,167
District Controls	Yes	Yes	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes

Each column presents the results of a separate regression. All regressions include district and year-quarter fixed effects. Each regression also includes the change in average outcome in the district between 1999-00 and 2004-05. The sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. The outcomes are defined as in Table 3. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). Worker controls are listed in the notes of Table 3. All estimates are computed using sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table A.5: Changes in outcomes in late relative to early phase districts between 2007-08 and 2009-10, when the program is extended to late phase districts

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)
Program	0.00569***	-0.0106**	0.00838**	-0.00297	0.0340**
	(0.00139)	(0.00488)	(0.00362)	(0.00299)	(0.0171)
Program X Dry	0.00948***	-0.0203***	0.0132***	-0.00191	0.0489**
	(0.00247)	(0.00612)	(0.00412)	(0.00344)	(0.0191)
Program X Rainy	0.00190**	-0.000908	0.00353	-0.00403	0.0191
	(0.000741)	(0.00547)	(0.00409)	(0.00396)	(0.0190)
Observations	356,636	356,636	356,636	356,636	64,167
District Controls	No	No	No	No	No
Worker Controls	No	No	No	No	No

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)
Program	0.00804***	-0.00710	-0.00101	0.000695	0.0496**
	(0.00215)	(0.00662)	(0.00482)	(0.00435)	(0.0202)
Program X Dry	0.0114***	-0.0163**	0.00349	0.00201	0.0568***
	(0.00291)	(0.00734)	(0.00504)	(0.00448)	(0.0207)
Program X Rainy	0.00425**	0.00336	-0.00612	-0.000810	0.0412*
	(0.00172)	(0.00747)	(0.00551)	(0.00541)	(0.0232)
Observations	356,636	356,636	356,636	356,636	64,167
District Controls	Yes	Yes	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes

Each column presents the results of a separate regression. All regressions include district and year-quarter fixed effects. Each regression includes 1999-00 to 2004-05 changes in outcome in each district. The sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. The outcomes are defined as in Table 4 and 5. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). Worker controls are listed in the notes of Table 5. Star states is a dummy variable equal to one for districts within star states. Other states is a dummy variable equal to one for districts that are not in star states. See Table 2 for a description of star states. All estimates are computed without sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table A.6: Main Specification controlling for state specific time effects

	Public	Private	Unemployed	Not in Labor Force	Log Deflated Daily Casual Earnings
	(1)	(2)	(3)	(4)	(5)
Program X Dry	1.083***	-1.864**	0.833	-0.0520	0.0424**
	(0.285)	(0.743)	(0.544)	(0.465)	(0.0198)
Program X Rainy	0.333	0.170	-0.127	-0.376	0.0101
	(0.208)	(0.785)	(0.581)	(0.531)	(0.0224)
Observations	356,636	356,636	356,636	356,636	64,167
District Controls	Yes	Yes	Yes	Yes	Yes
Worker Controls	No	No	No	No	Yes

Each column presents the results of a separate regression. All regressions include district and year-quarter fixed effects, as well as a dummy for each state interacted with a dummy for 2007-08.B41 The sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. The outcomes are defined as in Table 3. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). Worker controls are listed in the notes of Table 3. All estimates are computed using sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table A.7: Program effect on wages for men and women

	Log Deflated Casual Wages		
	Whole sample	**Male Workers**	**Female Workers**
	(1)	(2)	(3)
Program	0.0202	0.0177	0.0481***
	(0.0127)	(0.0184)	(0.0152)
Program X Dry	0.0353**	0.0321**	0.0630***
	(0.0151)	(0.0151)	(0.0235)
Program X Rainy	0.00496	0.00285	0.0347
	(0.0152)	(0.0153)	(0.0225)
Observations	44,278	19,889	64,167
District Controls	No	No	No
Workers Controls	No	No	No
	Whole sample	**Male Workers**	**Female Workers**
	(4)	(5)	(6)
Program	0.0403***	0.0421**	0.0292
	(0.0166)	(0.0221)	(0.0201)
Program X Dry	0.0488***	0.0516***	0.0381
	(0.0162)	(0.0177)	(0.0250)
Program X Rainy	0.0304	0.0313	0.0186
	(0.0187)	(0.0199)	(0.0277)
Observations	44,278	19,889	19,889
District Controls	Yes	Yes	Yes
Workers Controls	Yes	Yes	Yes

Each column presents results from a separate regression. All regressions include district and year-quarter fixed effects. The sample is composed of all adults aged 18 to 60 interviewed from July 2004 to June 2005 and from July 2007 to June 2008. Program is a dummy variable equal to one for early districts during July 2007 to June 2008. Dry is a dummy variable equal to one for the first two quarters of the year. Rainy is a dummy variable equal to one for the second two quarters of the year. District controls are listed in Table 1. District controls that do not vary over time are interacted with a dummy for 2007-08 (post-program). All estimates are computed using sampling weights. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.

Table A.8: Trend break specification

	PUBLIC WORKS		
	(1)	(2)	(3)
Program	0.00425***		
	(0.00133)		
Months in the Program		0.000126***	6.75e-05***
		(2.39e-05)	(2.21e-05)
Observations	692,651	692,651	692,651
District Trends	No	No	Yes
	Private Sector Work		
	(1)	(2)	(3)
Program	-0.00973**		
	(0.00421)		
Months in the Program		-0.000168***	-0.000148***
		(4.81e-05)	(4.34e-05)
Observations	692,651	692,651	692,651
District Trends	No	No	Yes
	Log Deflated Casual Wages		
	(1)	(2)	(3)
Program	0.0363**		
	(0.0141)		
Months in the Program		0.00352***	0.00268***
		-0.000136	-0.000115
Observations	125,339	125,339	125,339
District Trends	No	No	Yes

Each column presents results from a separate regression. All regressions include district and year-quarter fixed effects. The sample is composed of all adults aged 18 to 60 interviewed from July 1999 to June 2000, from July 2004 to June 2005, from July 2007 to June 2008 and from July 2009 to June 2010. Program is a dummy variable equal to one for early districts during July 2007 to June 2010 and for late districts during July 2009 to June 2010. "Months in the Program" is equal to the number of months since NREGA was launched, i.e. February 2006, April 2007 and April 2008 for districts in first, second and third phase respectively. The specification is described in Section A.5 in Appendix. No control is included. All estimates are computed using weights proportional to district population. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and *

Table A.9: Regression discontinuity

Panel A	Dependent Variable: Public Sector Work				
	(1)	(2)	(3)	(4)	(5)
Predicted NREGA	0.00367	0.00309	0.00507	0.00238	0.00346
	(0.00250)	(0.00315)	(0.00357)	(0.00334)	(0.00524)
Observations	1,063	1,063	1,063	1,063	1,063
Panel B	Dependent Variable: Private Sector Work				
	(1)	(2)	(3)	(4)	(5)
Predicted NREGA	-0.00611	-0.00416	-0.00849	-0.00483	-0.0151
	(0.00642)	(0.00881)	(0.0112)	(0.00939)	(0.0183)
Observations	1,063	1,063	1,063	1,063	1,063
Panel C	Dependent Variable: Log Deflated Daily Casual Earnings				
	(1)	(2)	(3)	(4)	(5)
Predicted NREGA	-0.126***	-0.00295	0.0630	-0.00517	0.113
	(0.0261)	(0.0366)	(0.0474)	(0.0410)	(0.0885)
Observations	872	872	872	872	872
Baseline Control	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Linear Slope	No	Yes	Yes	Yes	Yes
Quadratic Slope	No	No	No	Yes	Yes
Flextible Specification	No	No	Yes	No	Yes

Each column presents results from a separate regression. A unit of observation is a district-quarter. The sample is composed of all adults aged 18 to 60 interviewed from July 2007 to June 2008 living in second and third NREGA phase districts. Predicted NREGA is a dummy variable equal to one if the district rank according to the Planning Commission Backwardness Index is lower than the state specific cut-off for early phases. The specification is described in Section A.5 in Appendix. Flexible Specification allows for different slopes to the right and to the left of the cutoff. All estimates are computed using weights proportional to district population. Standard errors in parentheses are adjusted for correlation of the errors at the district level. ***, **, and * indicate significance at the 1, 5, and 10% levels.