Why Guarantee Employment? Three Essays on the World's Largest Public-Works Program

by

Laura V. Zimmermann

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2014

Doctoral Committee:

Professor Jeffrey Andrew Smith, Co-Chair
Associate Professor Dean C. Yang, Co-Chair
Assistant Professor Raj Arunachalam
Professor Brian P. McCall

ACKNOWLEDGEMENTS

I thank my dissertation committee members Jeffrey Smith, Dean Yang, Brian McCall and Raj Arunachalam for detailed feedback on my entire dissertation.

The individual chapters of my dissertation benefitted substantially from valuable comments and suggestions by Achyuta Adhvaryu, Manuela Angelucci, Arnab Basu, Prashant Bharadwaj, John Bound, Charles Brown, Adi Dasgupta, Taryn Dinkelman, James Fenske, Kishore Gawande, Devesh Kapur, Gaurav Khanna, Julien Labonne, David Lam, Abhiroop Mukhopadhyay, Susan Parker, Jacob Shapiro, Mel Stephens, Oliver Vanden Eynde, and participants of the SOLE Conference in Boston, the PAA Annual Meeting in New Orleans, Pacific Conference for Development Economics 2013, the Centre for Studies of African Economies Conference 2013, the Workshop on India's Maoist Insurgency at Princeton University, the University of Michigan Economic Development Seminar, the University of Michigan Informal Development Seminar, and the University of Michigan Labor Lunch. Melissa Trebil provided excellent research assistantship for Chapter 2.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Why Guarantee Employment? Three Essays on the World's Largest Public-Works Program

by

Laura V. Zimmermann

Co-Chairs: Jeffrey Andrew Smith, Dean C. Yang

The Indian government started implementing the world's largest public-works program, the National Rural Employment Guarantee Scheme (NREGS), in 2006. With a legal guarantee of 100 days of public employment per year for all rural households, NREGS is the flagship example of the resurgence of interest in public-works programs in developing countries in the last couple of years. This may be surprising given the available evidence of the limited success of such schemes in developed countries and of severe implementation problems in developing countries. My dissertation therefore seeks to understand what benefits NREGS created that may have made its introduction optimal from a government's viewpoint despite the known drawbacks. Chapter 1 reconstructs the government algorithm that the Indian government used to assign districts to implementation phases, which can then be exploited in a regression-discontinuity design to analyze the program impacts empirically. The chapter focuses on the labor-market impacts of NREGS and sets up a household time-allocation model that allows households to take up the scheme both as an alternative form of employment and as a safety net after bad economic shocks. The empirical results suggest that the program provides a safety net but has no large other labor-market impacts. Chapter 2 uses the same empirical strategy to study the impact of the program on the Maoist conflict, which the Indian prime minister referred to as India's biggest security challenge. The results show that violence increases in the short run, which is driven by a rise in police-initiated attacks. The paper discusses and tests the implications of a number of potential theories, but finds that the empirical patterns are most consistent with the introduction of NREGS making civilians more willing to share information on insurgents with the police, which in turn improves the police's effectiveness in tracking down rebels. Chapter 3 analyzes the impacts of NREGS during the next general elections. My empirical results using a regression-discontinuity design suggest that both the government parties and incumbents of any party benefit from the scheme, but that these effects are concentrated in areas with high implementation quality and longer exposure to the scheme.

CHAPTER I

Why Guarantee Employment? Evidence from a Large Indian Public-Works Program

1.1 Introduction

Public-works programs are popular policy tools to help households cope with economic shocks in countries around the world. But while the interest in such schemes has declined heavily in many developed countries, recent years have seen a resurgence of such initiatives in developing countries[1]: the World Bank, for example, funded public-works programs in 24 countries between 2007 and 2009. In contrast to earlier schemes, many of the recent programs emphasize more long-term anti-poverty and safety net goals rather than viewing jobs in government projects simply as a way of reducing temporary unemployment. Public-works programs now often work as a predictable safety net that households know they have access to when they experi-

[1]See e.g. Lal et al. 2010. For an overview of public-works programs in developing countries see e.g. Zimmermann (forthcoming). Subbarao et al. (2013) provide an extensive account of public-works programs and recent developments in middle- and low-income countries.

ence a negative economic shock, but can also provide an additional source of income for underemployed workers unable to find a job. Both of these functions have the potential to reduce poverty by ensuring a larger and less variable stream of income for the poor.

How well public-works programs can fulfill such anti-poverty goals in practice is still a debated question, however, and as of now there is little empirical evidence on the causal labor-market impacts of public-works programs in developing countries.[2] Estimating the impacts of public-works programs on labor-market outcomes is often challenging because many programs are rolled out non-randomly and without a comparable control group, making a causal analysis difficult. The experience from developed countries suggests that such government initiatives often prove unable to raise workers' human capital and are in danger of crowding out private-sector jobs.[3] Additionally, concerns about the implementation quality of government initiatives due to problems with corruption and rationing potentially limit the economic benefits for workers in developing countries.

This paper analyzes the labor-market impacts of the largest public-works program in the world, one of the few instances where the program effects can be rigorously evaluated. India's National Rural Employment Guarantee Scheme (NREGS) provides a legal guarantee of up to 100 days of public-sector employment annually for each rural household (about 70 percent of the Indian population) that can be taken up at any point during the year. This feature makes NREGS one of the most ambitious anti-poverty schemes in developing countries, which is also reflected in high annual expenditures on the scheme of typically around 1 percent of Indian GDP. The program

[2]See e.g. Basu (forthcoming), Besley and Coate (1992), and Datt and Ravallion (1994) for some examples of theoretical and empirical analyses. Most of the existing empirical literature on the topic lacks a credible causal identification strategy, however.

[3]For an overview of public-works programs in developed countries see e.g. Kluve (2010). In developing countries, public-works program could have lower private-sector impacts if demand for temporary public employment tends to be high at times when there are few other jobs available (Subbarao 1997).

aims to work as an additional source of income for underemployed workers in rural labor markets and as a safety net for the rural poor after bad economic shocks.

NREGS was phased in across India in three steps between 2006 and 2008 in a highly non-random manner that prioritized economically underdeveloped districts. This feature makes the use of empirical strategies like a difference-in-difference approach unattractive since the parallel trend assumption can be shown to be violated.[4] Instead, I rely on a regression-discontinuity design to estimate the causal impact of the employment guarantee scheme on labor-market outcomes. To carry out this research design, I uncover the algorithm the government used to assign districts to treatment phases and reconstruct the algorithm values that can then be used as a running variable.

To provide some intuition for the expected empirical impacts of NREGS, I set up a household time-allocation model: households choose to allocate their time between a private-sector job and self-employment, where the latter is assumed to be the generally preferred but riskier occupation. Once NREGS is introduced, it is allowed to function both as a third form of employment and as a safety net for self-employed households after a bad economic shock. The model implies, among a number of other testable predictions, that the safety-net function of the employment guarantee scheme affects a household's optimal time allocation even when no negative shock occurs. The availability of a safety net makes self-employment a less risky occupation than before, which indirectly subsidizes such activities and reduces the need for households to work in the safer casual private sector.

The empirical results support the model's predictions about NREGS functioning as a safety net, whereas there is little evidence for the program providing an alternative form of employment. NREGS has very limited overall labor market impacts for

[4]See appendix Figure C.0.7 and the discussion in appendix D for more details. Figure C.0.7 plots the trends in private employment for the baseline data, for example, and shows that the parallel trend assumption does not hold.

both men and women: the program does not lead to a substantial increase in public-sector or total employment and does not raise private-sector casual wages. Consistent with the safety-net function, there is evidence that take-up of the program increases after bad rainfall shocks, and that men leave the private sector even in the absence of a negative shock. Despite this evidence that NREGS provides some insurance to rural households, I do not find any large positive impacts on other outcomes of interest such as household expenditures.

Overall, my paper suggests that we need a more comprehensive understanding of household options and optimal behavior to correctly evaluate the labor-market impacts of public-works programs than is implicit in the existing literature and policy debate. One of the main components that makes recent public-works programs in developing countries different from earlier initiatives is that they are often conceptualized as long-standing schemes rather than as short-term interventions. This means that households know that they will have access to public employment after bad economic shocks even before the shock actually occurs, and can therefore re-optimize their time allocation to reflect this reduction in risk. These indirect effects of the program may have substantial welfare implications that are typically not captured in the debate on the net benefits of public-works programs, and may mean that programs are not ineffective in altering the living situation of the poor despite low actual take-up. In the case of India, the employment guarantee scheme potentially indirectly subsidizes self-employment, which may have large long-term impacts for rural households.

My paper therefore extends the general literature on public-works programs in developing countries. The empirical results are consistent with some existing research that finds that the insurance function of such government schemes often seems to dominate the direct income benefits, although this literature in general does not consider the time-allocation impacts that should arise even in the absence of a shock (see e.g. Dev 1995, Subbarao 1997, Subbarao et al. 2013). Some other papers document changes in time allocation very similar to the results I find in this paper, but do not link these patterns to a broader conceptual framework or to the safety-net function of public-works programs (see e.g. Berhane et al. 2011, Gilligan et al. 2008). In general, the existing literature is dominated by propensity-score matching and difference-in-difference estimators and has traditionally focused very heavily on targeting and take-up of public-works programs rather than on their broader labor-market impacts.

My paper also contributes to the active literature on the Indian employment guarantee scheme in two respects: the existing recent papers on the labor-market impacts of the program use difference-in-difference approaches and concentrate on showing that there are important heterogeneous treatment effects with respect to variables such as seasonality and implementation quality (Azam 2012, Berg et al. 2012, and Imbert and Papp 2013). Given the non-random rollout of the scheme, the often substantial effects reported in these papers could be due to the violation of the parallel trend assumption. My regression-discontinuity analysis, which does not require the parallel trend assumption, clarifies that the overall direct impacts of NREGS seem to be small, although it confirms the conclusion of the other papers that the overall impact masks important heterogeneous treatment effects. Additionally, I show both theoretically and empirically that it is important to analyze substitution effects between different forms of non-public employment to fully capture the labor-market impacts of the employment guarantee scheme, which is not done in the existing literature. This also complements existing research on how rural labor markets in India are affected by unanticipated productivity shocks as in Jayachandran (2006).

The rest of this paper is structured as follows: section 1.2 provides some background information on the characteristics of NREGS. Section 1.3 sets up a theoretical model of a household's time optimization problem that generates a number of predictions about the labor market impacts of NREGS. Section 1.4 describes the rollout of the program and how it can be used in a regression discontinuity framework, while section 1.5 discusses the data and the empirical specifications. Section 1.6 presents the main results and some extensions. Section 1.7 concludes.

1.2 Background

1.2.1 Program Characteristics

The National Rural Employment Guarantee Scheme (NREGS) is one of the most ambitious government development programs in the world.[5] It is based on the National Rural Employment Guarantee Act (NREGA) that legally guarantees each rural household up to 100 days of public-sector work a year at the minimum wage. There are no formal eligibility rules other than that the household lives in a rural area and that their members are prepared to do manual work at the minimum wage. Households can apply for work at any time of the year, and men and women are paid equally. At least one third of the NREGS workforce in a village is required to be female.

NREGS projects are supposed to advance local development primarily through drought-proofing, flood prevention and irrigation measures, and need to be carried out without the help of contractors or machines. Paid wages are the state minimum wage for agricultural laborers, although NREGA specifies a floor minimum wage.[6] At the introduction of the scheme, this floor wage was 60 rupees per day, but it has been raised over time. In most states wages are paid on a piece-rate basis where the

[5] The program was renamed the Mahatma Gandhi National Rural Employment Guarantee Scheme in 2009. The original name is still widely used especially in the academic literature on the program, however. For more details on the scheme see e.g. Dey et al. (2006), Government of India (2009), and Ministry of Rural Development (2010).

[6] In practice, most states have minimum wages that are higher than the national floor wage, so that the NREGS wage is state-specific.

rates are supposed to be adjusted such that a typical worker working for 8 hours will earn the minimum wage. Wages must be paid within 15 days of the day the work was performed, and are supposed to be given out on a weekly basis.

1.2.2 Implementation and Effectiveness of the Program

How well the ambitious features of NREGS work in reality has been of great interest to researchers, NGOs and the press right from the beginning of the scheme. Qualitative and descriptive research suggests that NREGS is implemented well enough to generate substantial benefits for the poor, for example during the agricultural off-season and after idiosyncratic shocks, and has improved women's access to jobs with reasonable wages and working conditions. At the same time, however, such studies also stress widespread practical limitations and violations of the provisions in the National Rural Employment Guarantee Act: muster rolls are often faulty and include ghost workers, for example, and wages are often paid with long delays and may not conform to the state minimum wage. Additionally, many local governments seem to lack the technical expertise to propose useful local projects. Big landowners have also repeatedly complained about labor shortages and demanded NREGS work be banned during the peak agricultural season (Centre for Science and Environment 2008, Institute of Applied Manpower 2007, Khera 2009, Khera and Nayak 2009, NCAER-PIF 2009, Samarthan Centre for Development Support 2007).

Varying levels of NREGS implementation quality are also documented in a number of economics papers that typically focus on individual Indian states: Johnson (2009a) looks at the impact of rainfall shocks on the take-up of NREGS in the Indian state Andhra Pradesh, and finds that participation in public-works projects increases when rainfall is lower than expected, so that NREGS seems to provide a safety net for rural households. Deininger and Liu (2013) find that NREGS increases nutritional intake and household assets in the same state, whereas the analysis in Johnson (2009b) shows that the working of NREGS in Andhra Pradesh does not seem to be strongly affected by the parties in power at the local level.

But while these papers suggest that NREGS works well in Andhra Pradesh, other research documents that this is not the case in all parts of India: Niehaus and Sukhtankar (2013a, 2013b) analyze the existence and characteristics of corruption in the implementation of NREGS in the Indian state Orissa, and find that an increase in the minimum wage was not passed through to workers. Dutta et al. (2012) use nationally representative data from 2009/10 to study the effectiveness of reaching the target population. They find that demand for NREGS often far outstrips supply and that the rationing of projects is especially common in poorer states. Despite these shortcomings, Klonner and Oldiges (2012) find some evidence of substantial reductions in poverty using a difference-in-difference approach. NREGS has also been credited with reducing rural-urban migration and improving children's education outcomes (Afridi et al. 2012, Ravi et al. 2012).

Some existing economics papers also analyze the impact of the program on rural labor markets. Imbert and Papp (2013) use a difference-in-difference approach to look at the program's impact on wages and employment, comparing early-NREGS districts to the districts that had not yet received the program in 2007/08 and therefore function as control districts. They find that NREGS increases employment by 0.3 days per prime-aged adult and private-sector wages by 4.5 percent, with the impacts concentrated during the agricultural off-season. Azam (2012) also uses a difference-in-difference approach, and finds that public-sector employment increases by 2.5 percent while wages for males and females increase by 1 and 8 percent, respectively. In a variation of the difference-in-difference design, Berg et al. (2012) analyze the impact of NREGS on agricultural wages by using monthly information on agricultural wages from 2000 to 2011. The results in the paper suggest that agricultural wages have increased by about 5 percent in districts with a high implementation quality, but that it takes between 6 and 11 months after program roll-out for these wage effects to be realized.

The difference-in-difference strategy requires these papers to make the parallel-trend assumption that labor market outcome trends would have been similar in early and late NREGS districts in the absence of the program. Given the non-random rollout of the program according to poverty criteria this is a strong assumption, however, which could substantially affect their results.[7] Appendix figure C.0.7 shows, for example, that the parallel trend assumption is violated for private casual employment in the baseline data, and the assumption also fails for other labor-market outcomes. The regression discontinuity approach used in this paper, on the other hand, does not require such an assumption and therefore provides a cleaner empirical identification of the impacts of NREGS.

Additionally, these papers do not consider potential substitution effects between various categories of non-public employment, which could arise if the introduction of NREGS induces households to re-optimize their time-allocation decisions. To have a clearer understanding of the overall expected empirical impacts of the scheme, it is useful to set up a simple theoretical model of a household's optimization problem.

[7] Imbert and Papp (2013) discuss, for example, that wages in treatment and control districts were on different trends prior to the introduction of NREGS. They attempt to address potential concerns about the internal validity of their difference-in-difference estimates by including extensive district-level controls.

1.3 A Model of the Household Optimization Problem

1.3.1 The Baseline Model without NREGS

The model describes a household's optimal time allocation in a one-period setting.[8] Before NREGS is introduced, a household can first choose to allocate the total time of their household members, T, between working for a big landowner as an agricultural laborer in the private casual sector, l, and working on the family farm, f.[9] After this decision has been made, a weather shock is realized which determines the payoff from farm work.[10] The period ends, and the household earns the fixed wage w in the private sector, and income y for the time spent in farming.[11] The household derives utility both from the time spent working in self-employment on the family farm, and from the total income earned in both activities during the period.[12] The

[8] A more detailed discussion of the model as well as proofs for the model predictions are provided in Appendix B.

[9] Implicit in this setup is the assumption that a household has perfect control over l or, put differently, that the household can always get a job in the private sector at wage w for the desired duration. One period in this framework is thought of as an agricultural year, which includes peak times like planting and harvesting. While views about the structure of Indian rural labor markets differ substantially (see e.g. Kaur 2011 and Basu 2002), theoretical papers like Basu (2002) assume that landlords hire agricultural laborers competitively during the harvesting season.

[10] To fix ideas, the shock in this model is referred to as a weather shock. The model can accommodate all types of shocks that make self-employment more uncertain than private-sector employment, including health or other idiosyncratic shocks. If anything, the model's simplifying assumption that wages are fixed is more likely to hold in such cases. In the NSS data used for my empirical analysis, most households own some land. 53 percent of men self-identify as engaging in family employment as the main occupation, and about 40 percent of men live in households that are self-employed in agriculture.

[11] The fixed-wage assumption is consistent with the cross-sectional relationship between private wages and rainfall for rainfall shocks up to 5 standard deviations at baseline in the data. The analysis controls for the mean and standard deviation of rainfall in a district. For rainfall shocks that are larger than 5 standard deviations, the wage is increasing in the rainfall shock. Assuming that the private-sector wage is constant is a simplifying assumption. All that is needed for the model predictions to go through is that private-sector employment is less risky relative to self-employment.

[12] The assumption that households derive utility from working in self-employment ensures that the optimization problem has an interior solution. The qualitative predictions of the model are not affected by relaxing this assumption, however. Bandiera et al. (2013) show that less poor workers are more likely to be self-employed than the poorest, and that an intervention that relaxes credit constraints and improves skills for the very poor leads to substantial increases in the self-employment rate. Banerjee et al. (2011) report similar results. These findings are consistent with a

utility function is additively separable in these components, with weight α given to the utility from self-employment. The probability density function of y is *g(y)*.

At the beginning of the period, a household's optimization problem is

$$\max_{l} \alpha v(T-l) + (1-\alpha)E[u((T-l)y + lw)]$$

with $u' > 0, u'' < 0$, $v' > 0, v'' < 0$. This leads to the first-order condition

$$\alpha v'(T-l) = (1-\alpha)\int u'((T-l)y + lw)(w-y)g(y)dy \tag{1.1}$$

(1.1) pins down the optimal proportion of time l spent working in the private sector implicitly. Intuitively, the expected marginal utility from being self-employed needs to equal the expected marginal utility from working in the private sector.

An interesting extension is how the optimal proportion of time spent in private employment changes with the variability of the weather-shock distribution. As I show in Appendix B, assuming that the distribution of y in district B is a mean-preserving spread of the distribution in district A, it can be shown that households spend more in time in casual private employment in the riskier district B than in less risky district A, and therefore substitute away from the risky activity farming.

1.3.2 The Model with NREGS

After NREGS is introduced, the program can be used both as an alternative source of employment regardless of the weather shock, and as an insurance tool after bad weather shocks. This alters the baseline model in two ways. The household now first makes a time-allocation decision among three alternatives: working in the private casual sector (l), working on the family farm (f_1), and taking up a NREGS job (n_1).

general preference for self-employment, which is also in line with anecdotal evidence from developing countries.

After this decision has been made, as before a weather shock is realized that affects the payoff from farm work. The time originally allocated to farm work,f_1, can then be split between actually working on the farm (f_2) and taking up public employment in a NREGS project (n_2) instead. After this decision, the period ends and the payoffs are realized. As before, the payoff from farm employment is y and the private-sector wage is w. The NREGS program wage is $\overline{w}$. The household again derives utility from the time spent in self-employment and from the total income earned.

The new household optimization problem at the beginning of the period is now given by

$$\max_{l,n_1} E[\alpha v(T-l-n_1-n_2^*)+(1-\alpha)u((T-l-n_1-n_2^*)y+n_2^*\overline{w}+lw+n_1\overline{w})]$$

where n_2^* is the best-response function of n_2 given y since the household can optimize the time spent working for NREGS and actually working on the family farm after the weather shock has occurred and y has been realized. Once a household chooses the fraction of time to spend on NREGS employment after the weather shock has occurred, l, n_1, and y are fixed. The household therefore chooses n_2 to maximize

$$\max_{n_2} \alpha v(T-l-n_1-n_2)+(1-\alpha)u((T-l-n_1-n_2)y+n_2\overline{w}+lw+n_1\overline{w})$$

Leading to the first-order condition

$$\alpha v'(T-l-n_1-n_2)=(1-\alpha)u'((T-l-n_1-n_2)y+n_2\overline{w}+lw+n_1\overline{w})(\overline{w}-y) \quad (1.2)$$

Define the shock y_0 as the shock at which the first-order condition implies n_2^*=0. Then the first-order condition traces out the best-response function n_2^* for all weather shocks that imply a farming income of y_0 or less. For all larger values of y, the optimal n_2 is zero.

Knowing n_2^* and the distribution of y, at the beginning of the period the household needs to decide how much time to spend in the private sector, in NREGS employment, and in anticipated farming. A household will never work in both private-sector work l and in ex-ante NREGS employment n_1, but will work in the job that pays more. This is because l and n_1 are perfect substitutes for a household in terms of their contribution to household utility. Both are safe sources of employment that need to be committed to before the weather shock is realized. A household therefore maximizes utility by choosing the alternative that pays a higher wage. Define j as the amount of time spent working in the activity that pays the higher wage, such that

$$j = \begin{cases} n_1 & w \leq \overline{w}; \\ l & w > \overline{w}. \end{cases}$$

And define $\widetilde{w}$ analogously as the corresponding wage.

Working in the fact that the optimal n_2 is zero at large positive shocks, the first order condition of the household maximization problem is

$$\frac{\alpha}{1-\alpha}\left[\int\limits_{y\leq y_0} v'(T-j-n_2^*)(1+\frac{\partial n_2^*}{\partial j})g(y)dy + v'(T-j)\right]$$

$$-\int_{y>y_0} u'((T-j)y + j\widetilde{w})(\widetilde{w}-y)g(y)dy$$
$$= \int_{y\leq y_0} u'((T-j-n_2^*)y + n_2^*\overline{w} + j\widetilde{w})(\widetilde{w}-y+(\overline{w}-y)\tfrac{\partial n_2^*}{\partial j})g(y)dy \quad (1.3)$$

It can be shown that a sufficient condition for the existence of a solution is that the Arrow-Pratt measure of absolute risk aversion is high 'enough'.[13]

[13]See appendix for the proof. This condition does not depend on the sign of $\frac{\partial n_2^*}{\partial j}$, which is ambiguous. Intuitively, how the time allocated to the ex-post NREGS employment responds to an increase in the time allocated to precautionary activity j depends on the attractiveness of the wage for j relative to the NREGS wage $\overline{w}$ and y. In other words, j only functions well as a precautionary savings tool if the paid wage in that activity is not too low relative to the payoffs that can be achieved through NREGS employment and farming after the weather shock is realized. A sufficient condition for j and n_2^* being substitutes for shocks $y \leq y_0$ is $\widetilde{w} \geq \overline{w}$.

A couple of predictions about the impact of NREGS follow from the model setup under reasonable assumptions and are derived in Appendix B. The appendix also discusses the impact of a couple of extensions on the model, including the NREGS cap of 100 days, implementation problems, and private-sector wage variability.

1. If NREGS is predominantly used as a new form of employment regardless of the shock, NREGS employment rises, private-sector employment falls and the impact on farm employment is ambiguous.

2. If NREGS is predominantly used as a safety net after negative shocks, then the program has two effects

   - Ex post effect: NREGS employment is higher after a negative shock.
   - Ex ante effect: if no negative shock occurs, NREGS employment is low overall. Private employment decreases and farm employment increases.

Intuitively, the first set of predictions arises because NREGS introduces a more attractive form of safe employment that can be used as a risk-mitigation measure, which directly crowds out private-sector work. Since private-sector and NREGS employment are perfect substitutes in the model, this effect requires that the NREGS wage is higher than the private wage. The impact on farm employment is theoretically ambiguous: a higher wage in the safe form of employment makes working there more attractive, but households can now accumulate the same amount of money from safe employment as before in less time. The new optimal time-allocation pattern therefore depends on the magnitude of $\overline{w}$ relative to y and on the household's degree of risk aversion. The larger the implementation problems of the program are, for example in terms of rationing or underpayment of wages, the less likely can NREGS function as a new general form of employment.

The second set of predictions follows from the fact that NREGS as a safety net tool makes self-employment less risky than before since it can be taken up after bad shocks. This reduces the need for a household to insure against adverse shocks by working in the private sector. Households therefore spend less time in private employment and more time being self-employed. NREGS take-up will be low unless a bad shock is actually realized.

Both of these channels can potentially interact as well: in riskier districts, the probability of a large negative shock occuring is higher than in less risky districts. Knowing this, households in risky districts will decrease their time spent in safe employment less than households in less risky districts if the NREGS cap of 100 days of employment and any implementation problems mean that insurance is incomplete after large negative shocks. At the same time, however, households in riskier districts may also have a higher demand for NREGS as a safe form of employment to mitigate the higher risk they face from self-employment in general than they would in a less risky district.

1.3.3 Wage Impacts of NREGS

The model assumes that the private-sector wage is fixed and does not change in response to workers spending less time in the private sector to work on their own farms. This is clearly a simplifying assumption. How the private-sector wage changes after the introduction of NREGS depends on the industry structure of local labor markets and on the composition of the workforce, but there is little consensus in the existing literature about the best way of modelling the Indian casual private sector.[14] In a standard perfectly competitive setup where employers pay workers their marginal product and the marginal product is decreasing in the number of workers employed, for example, a decrease in the supply of labor because of NREGS will lead to a

[14]The models in Basu (2002) and Basu (forthcoming), for example, are built on the existence of two types of workers: those with long-run contracts, and those with short-run contracts. While the papers cite some evidence of the existence of such long-run contracts in some parts of India, other papers like Kaur (2012) argue that daily labor contracts are the norm in Indian rural labor markets. Imbert and Papp (2013) focus heavily on small farmers with simultaneous labor supply and demand decisions.

higher marginal product of labor for the remaining workers and therefore to higher wages, which in turn attenuates the negative impact NREGS has on private-sector employment. Wages should also rise if the public-works program practically enforces the existing minimum wage laws.

Wages could also fall under certain conditions, although such a scenario in general requires much more detailed assumptions about local structures and the shape of the production function. Suppose, for example, that each worker gets paid their marginal product, but that the marginal product is independent of the number of workers employed. There is heterogeneity in terms of a worker's productivity, with higher-productivity workers deriving more utility from self-employment (a higher α in terms of the model). NREGS will then make farming more attractive for high-productivity workers than for lower-productivity workers, which changes the composition of the workforce to consist of a higher percentage of low-productivity workers than before. Since a worker's marginal product is independent of the number of workers employed, wages for a worker of a given productivity will remain unchanged. Due to the change in the composition of the workforce, the average wage paid in the private sector will fall, however.

The impact NREGS has on private-sector wages is therefore an empirical question.

1.4 Program Rollout and Regression Discontinuity Design

1.4.1 Program Timeline and Details of the Rollout

The National Rural Employment Guarantee Act (NREGA) was passed in the Indian Parliament in August 2005. NREGS came into force in February 2006 in the first 200 districts. The scheme was then extended to the rest of the country in two steps: an additional 130 districts received the program in April 2007, and all remaining rural districts started NREGS in April 2008 (Ministry of Rural Development 2010). I will refer to the district phases as Phase 1, Phase 2, and Phase 3, respectively.

This phasing in of the employment guarantee scheme allows the empirical analysis of the program's labor market impacts by using a regression discontinuity (RD) design since the government assigned districts to implementation phases based on an algorithm. Unfortunately, the criteria used in the algorithm are not explicitly explained in the official documents on the program and the algorithm values are not directly publicly available. To be able to construct the running variable required for the regression-discontinuity design I therefore uncover and reconstruct the government algorithm. I do this by combining information from a number of government documents on NREGS, earlier development programs and other government reports. The algorithm had been used to determine the treatment status of earlier programs, and its use in the case of NREGS is confirmed by a former member of the Indian Planning Commission.[15]

Treatment assignment for each implementation phase was made according to a two-step algorithm: In the first step, the number of eligible districts was allocated to states according to the proportion of India's poor living in a given state. In the second step, districts within states were then supposed to be chosen based on an existing development ranking of districts, with poor districts receiving the program first.

While the algorithm values themselves are not available directly, knowledge of the procedure allows their reconstruction. The development index values used in the second step of the algorithm are publicly available from a Planning Commission report (Planning Commission 2003). The exact headcount poverty ratio values used in the first step are not known with complete certainty since three new states were

[15]More detailed information on the algorithm can be found in Appendix A.

created after the data used to calculate the values was collected. The poverty measure therefore needed to be adjusted. The Planning Commission published such revised values in 2009 (Planning Commission 2009), so those values are likely to be very close to the values used in the NREGS treatment assignment decisions, and are therefore used in this paper.

Table 1.1 provides an overview of how well the algorithm predicts NREGS receipt in the first and second phase for 17 major Indian states for all districts with non-missing development rank information.[16] The first column provides the number of non-missing rank districts per state, whereas columns 2 and 3 show the actual number of NREGS treatment districts for each state in Phase 1 and Phase 2, respectively. Columns 4 and 5 give the success rate of the proposed algorithm in predicting the treatment status of districts in Phases 1 and 2. The prediction success rate is calculated as the percent of treated districts of a given phase where actual and predicted treatment status are the same.

Table 1.1 shows that the overall prediction success rate of the proposed algorithm is about 84 percent in Phase 1 and about 82 percent in Phase 2, so there is some slippage in treatment assignment in both phases.[17] The prediction success rates are considerably higher than the ones that would be expected from a random assignment of districts, which are 40.27 percent for Phase 1 and 37.45 percent for Phase 2, respectively. The table also reveals that there is considerable heterogeneity in the performance of the algorithm across states, but that the algorithm performs well in

[16]Rank data is available for 447 of 618 districts in India. Data for the index creation was unavailable in some states, in most cases because of internal stability and security issues during the early 1990s when most of the data was collected. A former member of the Planning Commission says that in these states state governments may have had considerable say in district allocation, so in the absence of a general rule treatment status in these states is likely to be endogenous. I therefore exclude these states from my analysis. Rank data in the 17 major Indian states is complete for all districts classified as rural by the Planning Commission in their report, so there is no endogeneity in the availability of data in these states. Urban districts in the Planning Commission report are districts that either include the state capital or that have an urban agglomeration of more than one million people.

[17]Prediction success rates for Phase 2 are calculated after dropping Phase 1 districts from the analysis.

almost all of the 17 states.[18] Potential threats to internal validity are discussed in the next section.

1.4.2 Regression Discontinuity Design

Given the treatment algorithm's two-step procedure, the generated cutoffs that can be used for a regression discontinuity (RD) analysis are state-specific. Two cutoffs can be empirically identified: the cutoff between Phase 1 and Phase 2, corresponding to Phase 1 treatment assignment, and the cutoff between Phase 2 and Phase 3, which is equivalent to the Phase 2 rollout of the program. Since the dataset that I will be using in my empirical analysis was collected at a time when NREGS had been rolled out to Phase 1 and Phase 2 districts, but not yet to Phase 3 districts, only the cutoff between Phase 2 and Phase 3 can be used to analyze the impact of the government program. I therefore focus on this cutoff in the remainder of this paper.

Treatment cutoffs differ by state, so for the empirical analysis ranks are made state-specific and are re-centered so that a district with a normalized state-specific rank of zero is the last district in a state to be eligible for receiving the program in Phase 2. The data are then pooled and the treatment effects are estimated at the common discontinuity. Negative numbers are assigned to districts with lower ranks than the cutoff rank, whereas positive numbers are assigned to the districts that are too developed to be eligible and that will function as control districts.

Figure 1.1 shows the number of observations at each state-specific rank for Phase 2 district assignment. It reveals that all 17 states used in the analysis have at least one district receiving NREGS in Phase 2, but that only a few states have districts further away from the 0 cutoff.[19] Figure 1.1 reports observations based on the predicted

[18] As for the general sample, at the state level the relationship between predicted and actual treatment is usually much tighter than the one that would be predicted by random assignment of districts. The main exception to this are the Phase 2 assignments for the states Bihar, Jharkhand and West Bengal, since all remaining districts in those states are treated in Phase 2. In this case, random and algorithm-based assignment therefore yield the same results.

[19] While this pattern mostly reflects that there are only few states with a large number of districts,

NREGS receipt of the algorithm. As Table 1.1 shows, however, actual program receipt does not completely follow this assignment. Therefore, the empirical identification strategy is a fuzzy RD design. The fundamental assumption of the RD design is that districts that were just poor enough to receive the program, and districts that were just too rich to be included are similar to each other in terms of unobserved characteristics, so that outcome differences are solely attributable to the introduction of the employment guarantee scheme.

In order for the RD design to be valid, districts must have imperfect control over their treatment status in a given phase (Lee 2008). This implies that states and districts should not have been able to manipulate either the index variable used to rank districts, or the quotas allocated to states.[20] Otherwise, districts close to the cutoff on either side are not plausibly similar to each other in terms of unobserved characteristics, but differ on characteristics such as perceived benefit from the program or political influence.

That states or districts were able to manipulate the poverty index seems unlikely. First, the index was constructed based on somewhat dated available information: the Planning Commission used data from the early to mid-1990s for the ranking of districts, rather than collecting current information from districts. This limits the possibility to strategically misreport information. Second, the ranking had originally been used to target earlier development programs for especially poor districts, although with lower cutoffs of 100 and 150 districts, which implied lower state-specific cutoffs as well. So if districts were able to act strategically, the incentive would have

a number of states are also fully treated after Phase 2 assignment so that they have no Phase 3 districts and therefore no positive-rank districts in Figure 1.1. The results in this paper are robust to dropping fully-treated states from the analysis.

[20]The all-India number of treatment districts in each phase, 200 and 130, do not seem to have been chosen to accommodate state or district demands for a certain number of treatment districts. 200 was the number of districts the Planning Commission suggested for an earlier development program which never really took off. The number 130, on the other hand, seems to have been adopted because a number of states that had received many NREGS districts in the first phase had only few untreated districts left that could be treated in Phase 2.

been to be among the 150 poorest districts, but not among the 200 poorest districts used for NREGS in the first phase, and certainly not among the 330 poorest districts that received NREGS in either Phase 1 or Phase 2. Third, the creation of the index from the raw data by the Planning Commission is done in a transparent way. The Planning Commission report outlines the exact procedure with which the index was created, and also lists the raw data for all districts, so that the composite index as well as the district ranking can be perfectly replicated.

Figures 1.2 and 1.3 look more closely at the distribution of index values over ranks. Figure 1.2 shows the relationship between the poverty index value and the assigned rank by the Planning Commission for all 447 districts for which data is available. Across India, the distribution of poverty index values is smooth and continuous across ranks. As the chosen cutoffs are state-specific, Figure 1.3 plots the relationship between the Planning Commission's index and the normalized state-specific ranks for the Phase 2 cutoff. For most states, the poverty index values seem pretty smooth at the cutoff of 0. Overall, these patterns suggest that manipulation of the underlying poverty index variable is not a serious concern.

Manipulation of the criterion used for the allocation of treatment districts across states also seems unlikely: The state headcount ratios are calculated from mid-1990s information that had long been available at the time of NREGS district assignment. Additionally, I use 2001 Census information on the states' rural population to calculate the poverty prevalence measures, which also was widespread publicly-available information at the time. Again, it was therefore probably impossible for Indian states or districts to exert political influence on the treatment status of individual districts by manipulating the data.[21]

[21]This does not mean, however, that actual treatment assignment was not subject to political pressures, since Table 1.1 reveals that compliance with the proposed algorithm is substantially lower than 100 percent. It can be shown that deviations from the algorithm are correlated with the party affiliation of members of parliament from the same district. This finding is in line with research like Gupta (2006) who analyzes the correlation of political party affiliation and treatment status in an earlier district development program. This program most likely also used the two-step

Given that I do not have access to the actual poverty-prevalence measure used in the algorithm, my reconstructed values introduce measurement error into predicted treatment status if the Indian government used different values to make state allocations. While this potentially makes the regression discontinuity design fuzzier than it really is, it should not introduce systematic bias into the calculations since I am using the best available source of estimates.

Another way of analyzing whether manipulation is likely to be a problem is to test whether there are discontinuities at the cutoff in the baseline data: if the RD specification is valid, we would expect baseline outcomes to be smooth at the cutoff if treatment and control districts near the cutoff are indeed similar on observables and unobservables in the absence of treatment. Table 1.2 reports the results of such tests for all of the labor-market outcome variables used in this paper as well as for three other outcomes for which data is available at baseline (years of education, area of land owned and log per capita expenditure) for all parametric specifications of the RD estimator used in this paper. The estimates show that the vast majority of the 64 coefficents are not statistically significant. The only variable for which a discontinuity at the cutoff is quite consistently found is the years of education variable, although the magnitude of the effect for both men and women is small. For women, two coefficients are significant among the outcome variables, but this pattern is not consistent across empirical specifications. Again, widespread manipulation of treatment assignment seems unlikely based on these results. To control for any baseline differences in outcomes as well as to soak up residual variance, the main results in this paper control for the baseline outcome variable, however, even though the estimated coefficients are not substantially affected by the exclusion of the baseline controls.

With the fuzzy RD design used in this paper, we need to verify that there is indeed

algorithm proposed in this paper, however, which is not taken into account in Gupta's paper and could potentially affect the results in substantial ways.

a discontinuity in the probability of receiving NREGS at the cutoff values for Phase 2 NREGS districts. Figure 1.4 shows this graphically for the normalized state-specific cutoff for Phase 2. It plots the probability of receiving NREGS in the given phase for each bin, as well as fitted quadratic regression curves and corresponding 95 percent confidence intervals on either side of the cutoff. The graph shows that the average probability of receiving NREGS jumps down at the cutoff. This suggests that there is indeed a discontinuity in the probability of being treated. Figure 1.4 also shows that compliance with the algorithm is relatively low directly at the normalized cutoff of zero, which could for example be a function of measurement error in the first step of the algorithm. In a robustness check of my main results, I therefore drop observations right around the cutoff in an application of the donut hole RD approach.

1.5 Data and Empirical Specification

1.5.1 Data and Variable Creation

The data used in this paper comes from household surveys collected by the National Sample Survey (NSS) Organisation. These surveys are representative of the Indian population, and drawn from the population in a two-stage stratified sample design. In the first stage, villages are selected, and individual households within these villages are sampled in the second stage. The dataset that can be used to analyze the impact of NREGS on wages and employment is the 64th round of NSS data, which was collected from July 2007 to June 2008. It has a sample size of about 120,000 households and interviews were carried out over the course of a year in four sub-rounds, each spanning three months. By this time, NREGS had just been rolled out to Phase 2 districts in April 2007. Phase 3 districts received the program in April 2008, although general delays in implementation suggest that Phase 3 districts can be treated as control districts even for the last three months of the survey.[22] To analyze the labor market impacts of NREGS by using an RD design, I therefore focus on the state-specific cutoffs between Phase 2 and Phase 3 and drop Phase 1 districts.

The dataset collects wage and employment information as well as a number of socio-demographic characteristics. Additionally, a sample of households are interviewed in a given district in every sub-round, if possible. While the household data is strictly cross-sectional, this means that at the district level it is possible to generate a sub-round panel with up to four observations per round. I will exploit this feature of the data empirically by aggregating individual-level information up to the district level for each sub-round separately.

Consistent with other NREGS papers, I restrict my sample to individuals of prime age (18-60 years) who are living in rural areas and have at most secondary education. The NSS employment module asks detailed questions about an individual's work status in the last 7 days. I use these questions to create various employment and wage variables, focusing on casual jobs. Employment measures are dummy variables equal to 1 if an individual worked at all in a public-sector job, a private-sector job or in family employment in the past 7 days, respectively, and 0 otherwise. I add up the value of wages received in cash and kind for private-sector casual jobs and divide it by the amount of time spent in that type of work to create a daily private wage for workers. I then aggregate the labor market measures up to the district-sub-round level using sampling weights. Data from the 61st round (July 2004-June 2005) is used as baseline information.

1.5.2 Empirical Specification

The preferred way of estimating the treatment effect at the cutoff in an RD design is to restrict the sample to observations close to the cutoff and to then run separate

[22]See e.g. Imbert and Papp (2013). The results reported in this paper are qualitatively the same when these potentially contaminated control group observations are excluded.

local linear regressions on both sides (Lee and Lemieux 2010). The difference of the regression lines at the cutoff then provides the estimate of the treatment effect. In choosing which observations are 'close' to the cutoff, researchers need to trade off concerns about precision and bias: The larger the window of observations used in the regressions, the more precise the estimates are likely to be since the number of observations is higher. At the same time, however, this implies that observations further away from the cutoff are used, which may bias the estimate of the treatment effect at the cutoff.

This trade-off is of particular relevance in the case of NREGS where the number of districts is limited so that there are few districts 'close' to the cutoff. To get an idea of how bad the bias introduced by using observations further away from the cutoff is, appendix figures C.0.1 to C.0.6 non-parametrically plot the relationship between the running variable and three outcomes of interest for men and women separately. The graphs show the averaged outcomes of all district observations with a given state-specific rank and also include the estimated regression function for a quadratic function on both sides of the cutoff. The graphs show that a quadratic function fits the data quite well in all specifications, and that the estimated regression lines for public employment are even well approximated by a linear function. These patterns suggest that using the whole sample of Phase 2 and Phase 3 districts and estimating the treatment effect at the cutoff using linear and quadratic functions of the running variable is not a bad approximation of the underlying data. That a larger bandwidth may be plausible is also supported by Figure 1.3, which showed that the underlying poverty index is smooth at the cutoff. F-tests also reject the null hypothesis that higher-order polynomials add important flexibility to the model. More flexible models also tend to be unstable, although the estimated coefficients are often qualitatively similar to the quadratic results.

My overall preferred empirical specification therefore uses quadratic regression curves estimated on either side of the cutoff (referred to as 'quadratic flexible slope' in the result tables). As a robustness check, all my results also report the estimates using a quadratic function constrained to have the same slope on either side of the cutoff, and corresponding flexible and constrained linear regression lines. Additionally, I also report the estimates of the main results when using a linear flexible regression function, but restricting the sample to observations closer to the cutoff, in Table 1.7.

The equation below shows the regression equation for the most flexible specification:

$$y_{ijk} = \beta_0 + \beta_1 rank_{ij} + \beta_2 rank_{ij}^2 + \beta_3 nregs_{ij} + \beta_4 nregs * rank_{ij} + \beta_5 nregs * rank_{ij}^2 + \beta_6 baseline\ y_{ij} + \eta_j + \epsilon_{ijk}$$

where the subscripts refer to individual i in district j in season k, y is an outcome variable of interest, *rank* is a district's rank based on the state-specific normalized rank, and η are state fixed effects.

The main results report the intent-to-treat effect of NREGS, so *nregs* is an indicator variable equal to 1 if a district is predicted to have received NREGS in Phase 2 according to the state-specific algorithm, and zero otherwise. A corresponding appendix table reports the treatment-on-the-treated estimates where actual NREGS receipt is instrumented with predicted NREGS receipt. The coefficient of interest is β_3. In all empirical specifications, standard errors are clustered at the district level.[23] Results are reported for men and women separately.

The above specification uses the commonly employed technique of re-centering the treatment cutoffs and pooling the data to estimate the treatment effect at a single cutoff. An alternative approach is a meta-analysis as used, for example, in Black et al. (2007): the treatment effect is estimated for each cutoff separately, and the

[23]The results from reweighting observations by their 2001 Census population size are qualitatively very similar to these results and therefore not presented here. This extension takes into account that district-averages will be more precisely estimated in large districts than in small ones since the individual-level data is representative of the Indian population. At the same time, however, such a specification assumes that there is no district heterogeneity in treatment effects.

estimates are then combined to a single estimate afterwards by using appropriate weights. I report the results for such an analysis in Table 1.8 for a simple average and a population-weighted average of the state treatment effects. These estimates also take into account that the covariance between the state-specific estimates may not be zero.

Lastly, there is the choice of the running variable. While the specification above uses the state-specific rank as the running variable, an alternative would be to use the poverty index instead. Treatment assignment was made according to the state-specific rank, however: the first step of the government algorithm determines the size of the treatment group in a given state, which is then filled with the poorest districts according to their rank. The relevant distance of a district from the cutoff is therefore its rank and not its index value, since in many cases a district could have a very different poverty index value without altering its rank or distance from the cutoff. Additionally, the plotted conditional mean function using the rank variable is flatter than the one using the index values, suggesting that a larger bandwidth is less problematic when using the rank variable. I report the estimates of the main results using the state-specific index variable as a robustness check in appendix table C.0.4.

1.5.3 Summary Statistics

Table 1.3 presents baseline wage and employment summary statistics for districts separately by phase for men and women respectively. As the table shows, early NREGS districts have lower baseline wages for men than later districts, consistent with the idea that NREGS was rolled out to poorer districts first. The daily wage of a typical male casual worker of prime age with at most secondary education in an average Phase 2 district is about 53 rupees, whereas the corresponding wage is about 66 rupees in Phase 3 districts. Private-sector daily wages are very similar to overall casual daily wages, and there is no substantial difference between public-sector and private-sector wages.

In general, however, it is very uncommon to work in the public sector in all districts: 0.4 percent of workers work in the public sector in a typical Phase 2 district in the week prior to the survey, and the corresponding number for Phase 3 districts is 0.2 percent. In contrast, in all districts about 30 percent of males work in private casual jobs, and about 58 percent work in a family business or on the family farm. The remainder are males who are unemployed or out of the labor force in a given week. Table 1.3 also shows that the situation for Indian women is qualitatively similar to that of men, but that women are about half as likely to work in casual jobs of any kind or in family employment as men.

1.6 Results

1.6.1 Main Results

Figures 1.5 to 1.10 and Tables 1.4 and 1.5 present the main results of the impact of NREGS for men and women separately. The figures show the RD design graphically for the probability of being employed in a public works program or in casual private-sector work in the past 7 days and for the log private-sector daily casual wage earned in the past week. One scatter point represents the average residual outcome value in a given season. The residuals come from a regression of the outcome variable on the baseline outcome variable and state fixed effects to make the graphs comparable to the results in the tables. The regression lines are quadratic in the running variable, with the slope allowed to differ between the two sides of the cutoff. The graphs also plot the 95 percent confidence intervals. With the exception of Figure 1.6, none of the figures show a statistically significant discontinuity at the cutoff value zero, suggesting that the impacts of NREGS on the Indian rural labor market are limited for districts near the state-specific cutoff. Figure 1.6, on the other hand, reveals that private employment for men in NREGS districts is statistically significantly lower than that in control districts.

Tables 1.4 and 1.5 focus on different empirical specifications of the RD design in more detail. Whereas these tables present the intent-to-treat effects, appendix table C.0.1 shows the corresponding IV results for the same specifications. In all tables, one observation is a district in a specific season. Table 1.4 shows the main results for men and women in Panel A and Panel B, respectively. Each row presents the impact of NREGS on the outcome variables of interest for a different parametric functional form of the running variable. Panel A looks at the estimates for men and column 1 demonstrates again that NREGS does not have a large impact on public-sector casual employment: the typical estimate is positive but small in magnitude and statistically insignificant. The coefficient in the first row of column 1, for example, suggests that being in a NREGS district increases the average rural prime-aged man's probability of having had a public-works job in the last 7 days by 0.12 percentage points. This translates into an increase of 17.4 percent since mean public employment is only 0.69 percent, but the effect is statistically insignificant.

Column 2 of Panel A reveals that the NREGS impact on private casual employment is negative and statistically significant at the 10 percent level. The estimated coefficients suggest that NREGS lowers private-sector casual employment for men by about 3-5 percentage points across specifications, which translates into a percentage change of 11-16 percent. The impact of NREGS on the probability of being in family employment in column 3 is positive and of about the same absolute magnitude as the estimates in column 2, although imprecisely estimated. The overall impact of NREGS on total employment is small, negative, and statistically insignificant.

Panel A also shows the results for the log daily private casual wage. The outcome variable in column 5 is the average district log wage earned, conditional on having earned a positive daily wage. Since column 2 provides some evidence of private employment changes, any wage impacts of NREGS in the conditional log wage should be seen as a potential combination of changes in the selection of workers into private employment and of wage changes of workers conditional on workforce composition. According to column 5, the impact of NREGS on private wages is small and statistically insignificantly different from zero. If anything, the results point to a decrease in the private-sector wage. The estimated coefficient in the first row of column 5, for example, suggests that the average log private wage for men employed in casual private-sector work decreased by 0.4 percent in treatment districts relative to control districts at the cutoff.

Panel B shows the corresponding results for women. As column 1 demonstrates, the impact of NREGS on the probability of being employed in a public-works project for women is typically positive and of a similar magnitude as the one for men, although the estimates are again small in magnitude and statistically insignificantly different from zero. Column 2 shows that the impact of NREGS on casual private-sector employment for women is negative and typically smaller than for men, although the confidence intervals are wide. In contrast to men, however, the total employment coefficients are positive, although they are again very imprecisely estimated. Additionally, Panel B suggests that NREGS has no large-scale effects on private-sector wages for casual work for women and a positive, but statistically insignificant impact on family employment.

These results show that the general impacts of NREGS on labor-market outcomes seem to be limited, although the coefficients are often imprecisely estimated. There is no statistically significant increase in public employment and the empirical analysis can rule out public-employment increases larger than one percentage point. The employment guarantee has also not led to upward pressure on the private-sector wage. If anything, private-sector wages fall, which rules out that NREGS enforces existing minimum wage laws or increases competition in local labor markets that forces employers to substantially raise the private-sector wage. In contrast, there is some evidence of male workers leaving the private casual sector. Taking the imprecise estimates on public, family, and total employment at face value, the analysis suggests that most men switch from private to family employment, which would be consistent with the ex ante effect of NREGS functioning as a safety net: The availability of NREGS as a safety net after bad economic shocks would then lower the relative riskiness of family employment and therefore lead men to leave the private casual sector even when no shock occurs. On the other hand, the estimates provide no support for the idea that NREGS is predominantly taken up as a new alternative form of employment. Overall, Table 1.4 therefore implies that NREGS has not had large labor-market impacts. At best there is some role for NREGS as a safety net.

Whether the employment guarantee scheme can function as a safety net after bad economic shocks can also be tested more directly. If this is true, we should see an increase in the take-up of public employment after a negative shock, which was the ex post effect in the model. Table 1.5 reports the results of such an analysis: the specification focuses on districts during the agricultural off-season (January to June), but considers rainfall shocks that occurred at the beginning of the previous agricultural season in the months July to September, which roughly corresponds to the monsoon season. This gives the rainfall shock some time to feed through to household incomes. The main treatment variable is interacted with an indicator variable equal to one if a district experienced a negative rainfall shock (so lower rainfall than expected based on average rainfall in the district) in the agricultural main season.

As column 1 of Panel A shows, NREGS take-up for men is indeed statistically significantly higher after such an adverse shock, with interaction effects of around 3 percentage points. The sum of the main effect on NREGS and the interaction effect with the negative shock is also always statistically significantly different from zero at conventional levels, implying that that the NREGS impact in bad rainfall shock areas is also statistically significantly different from zero. This higher take-up of the employment guarantee after bad rainfall shocks confirms the take-up effects found in Johnson (2009a) for Andhra Pradesh. The magnitude of the effect is similar for women, as reported in Panel B, but imprecisely estimated. Again, there is little evidence of large employment or wage impacts, however: the employment guarantee scheme does not lead to a net increase in employment in NREGS districts even after a bad rainfall shock. Taken at face value, the statistically significant increase in public employment after a bad rainfall shock for men in treatment districts comes at the cost of private-sector employment rather than providing work to unemployed workers, although the coefficients on private employment are estimated imprecisely.

Taken together, Tables 1.4 and 1.5 therefore support the idea that there are no large benefits for workers from the introduction of NREGS, but that the safety net feature of the program plays some role. As appendix table C.0.2 shows for the male sample, there is also no evidence that the employment guarantee scheme has had a large impact on per-capita expenditures, the total wage or remittances received in the past year. In results not reported here, I also find no effect of the program on the variance of household expenditures.

The theoretical model also implied that we should expect heterogeneous treatment effects depending on the riskiness of a district's farming income distribution, and that safety net and employment functions of NREGS may interact in those districts. If NREGS does indeed mostly work through providing a safety net after bad economic shocks, we should see that these impacts are especially pronounced in higher-risk districts. According to the model, public employment should increase more in high-risk districts because they are more likely to have experienced a bad economic shock, and private-sector employment may or may not decrease more strongly in high-risk districts depending on the magnitude of negative income shocks relative to the maximum amount of income that can be earned from NREGS, with family employment mirroring the private-sector employment effects. Additionally, public employment in higher-risk districts may also increase more because of a higher demand for safe employment. Since households know that they live in a more risky area, the demand for a buffer stock may be higher, especially if NREGS cannot provide full insurance after bad economic shocks and if there is rationing in NREGS or private-sector employment so that the buffer-stock demand for work cannot be met by just working in the higher-paying alternative as in the model. So households may demand NREGS employment even in the absence of a negative shock in addition to private-sector work.

Table 1.6 reports the empirical results that analyze the importance of risk heterogeneity. The table shows the estimates for men where NREGS treatment is interacted with an indicator variable equal to 1 if a district has a higher than median variance in rainfall at baseline, which proxies for income volatility. The results for women are qualitatively similar to those for men, but the coefficients are typically smaller and not statistically significantly different from zero. To make the estimates comparable to the ones in Table 1.5, the sample is again restricted to the agricultural off-season observations, although the results are very similar when using the full sample. The regressions also control for expected rainfall. Column 1 of Panel A shows that public employment is statistically significantly higher in high-variance districts. The interaction effect for private employment is typically positive, although not statistically significant, and the family employment results mirror these effects. Additionally, the private-sector wage in column 5 is substantially higher in risky districts, although only statistically significant at the 5 percent level in the complete sample and not in the off-season sample reported in Table 1.6. Total employment is statistically significantly higher in riskier districts, on the other hand, which is consistent with the idea that there is a higher demand for employment in these districts. The sum of the main effect and the interaction effect for the public and total employment outcomes is not statistically significant at conventional levels. This means that while the impact of the employment guarantee is larger for these outcome variables in riskier districts than in less-risky districts, the impact of the program in risky areas is not statistically significantly different from zero.

To analyze whether the increase in public employment is due to a higher demand for NREGS as a safety net or as an alternative form of employment in riskier districts, in a specification not reported in this paper I re-run the public-employment regressions from Table 1.6 for men, but control for the last monsoon rainfall. This specification at least to some extent controls for the demand for insurance after adverse shocks. The impact on public employment is still statistically significantly higher in riskier NREGS districts than in low-risk treatment districts when controlling for monsoon rainfall, and the magnitudes are very similar to the estimated coefficients when last-season rainfall is excluded. This suggests that the public-employment effects in Table 1.6 to a large degree reflect risk-mitigation motives. Overall, the magnitude of public-employment changes is still small relative to the private- and family-employment coefficents, however, potentially implying that the insurance mechanism is more important.

Tables 1.4 to 1.6 therefore suggest that the overall impacts of NREGS are small, but that there are instances where the program works as a safety net and, in risky districts, as an alternative form of employment.

1.6.2 Robustness Checks

A couple of alternative specifications can be used to test the robustness of the main results. One check is to change the sample restrictions: the main results keep all Phase 2 and Phase 3 districts in the analysis, which potentially biases the estimates since observations far away from the treatment cutoff can influence the estimate of the treatment effect at the cutoff. Table 1.7 therefore reports the analogous results to Table 1.4 for a linear flexible specification and three more restrictive definitions of the sample to observations with state-specific ranks between -5 and 5, -4 and 4, and -3 and 3, respectively. As the results show, the qualitative pattern of the results from Table 1.4 persists, although, consistent with the tradeoff between precision and bias, the coefficients tend to be more imprecisely estimated than before.

A second potential concern about the reported estimates is that they may be heavily affected by measurement error: since the exact numbers used to determine the number of treatment districts assigned to states are not known, my choice of the most plausible values introduces measurement error right around the state-specific cutoff values. To test how sensitive the estimates are to this, I re-estimate Table 1.4 without the districts right around the cutoff by excluding districts with a normalized rank between -1 and 1. This approach is typically referred to as the donut-hole approach.[24] The results of the donut hole approach are reported in appendix Table C.0.3 and are very similar to those in Table 1.4.

A different way of checking the robustness of the main results is to use a meta-analysis approach to estimate the treatment effect at the cutoff. So far, the impact of NREGS was estimated by re-centering the state-specific cutoffs and then pooling the data to estimate the overall treatment effect. An alternative method is to estimate the treatment effect separately for each state, and to then combine those estimates in a meta-analysis. The results of this analysis are reported in Table 1.8 for men. The results for women tend to be qualitatively similar, but, as in the other tables, smaller and less precisely estimated. Since the number of observations for an individual state is often small and more flexible specifications are often highly collinear with the treatment variable, Table 1.8 only reports the results for two of the four specifications

[24]See e.g. Almond and Doyle (2011) for a similar application. Applying this approach has its disadvantages as well: First, the regression discontinuity design relies on estimating the treatment effect in the neighborhood of the cutoff, so dropping the observations closest to the cutoff weakens the fundamental assumption that districts close to the cutoff on either side are similar to each other in terms of all characteristics except the treatment status of NREGS. Second, dropping observations reduces the sample size.

from Table 1.4. The first two rows report the results based on a simple average of the state-specific treatment effects, whereas the last two rows weight the state-specific treatment effects by the state population. The results in Table 1.8 show that the empirical patterns are robust to this alternative estimation technique.

Two additional robustness checks are provided in the appendix: table C.0.4 re-estimates Table 1.4 using the state-specific index value instead of the rank as the running variable, whereas Table C.0.5 estimates the impacts of NREGS at the individual rather than the district level.[25] Overall, results are again qualitatively similar to those reported in Table 1.4 and again suggest that the labor market impacts of NREGS are limited.

Lastly, the main results are also robust to a number of other specifications not reported here, like the exclusion of the baseline outcome variables, the inclusion of additional control variables and the exclusion of potentially contaminated Phase 3 districts due to the timing of data collection.[26]

1.6.3 Discussion

Overall, the results suggest that NREGS only has a very limited direct influence on the Indian rural labor market, although in a number of empirical specifications the effects are not precisely enough estimated to rule out more substantial effects. Instead, NREGS seems to work as an insurance tool that reduces the riskiness of family employment relative to private-sector work, even though the risk heterogeneity results suggest that buffer stock considerations are not completely absent, either.

The safety net does not seem to generate substantial welfare benefits in the form of

[25] The individual observations are weighted using sampling weights. Since the data at the individual level are cross-sectional, we cannot control for the baseline outcome variable in the same way as before. The regressions reported in Table C.0.5 do not control for any baseline outcomes, but the results are robust to controlling for the baseline district average in the outcome.

[26] Phase 3 districts received NREGS in April 2008, whereas the data was collected between July 2007 and June 2008 and Phase 3 districts are treated as controls throughout in the main specifications.

higher per-capita expenditures, however. One potential explanation for this finding is that such effects may take longer to be realized. The analysis in this paper is limited to the first year of NREGS implementation because of data limitations and since the program is rolled out to control districts afterwards. Medium- to long-term benefits of NREGS can therefore not be captured. Even if there are no household expenditure impacts, however, the program may have substantial welfare implications through the occupational changes and may therefore alter the unobserved utility households derive from employment.

Maybe most surprising is the fact that a large-scale public-works program like NREGS does not seem to significantly increase the working-age population's probability of having held a public-works job in the past 7 days. Mean public employment is only 0.69 percent for men and 0.53 percent for women in Phase 2 and Phase 3 districts. So while some of the estimated coefficients are equivalent to large increases in public employment in percentage terms, statistical power is not big enough to precisely estimate such small effects. The estimates in Table 1.4 imply that the empirical analysis can rule out increases in public employment above 1 percentage point.

While the theoretical model suggests that we should not expect large increases in public employment if NREGS is mainly used as a safety net rather than as an additional form of employment in a typical year, one potential alternative explanation for these small effects is the time frame of the household survey. Since employment information is based on a 7-day recall window, it is by design much noisier than employment histories over a longer time horizon, although there should be no issues with recall error. It is therefore useful to compare the prevalence of NREGS employment in the household survey data to the employment numbers based on administrative data. While some papers have documented that administrative records are exaggerating the effectiveness of NREGS due to corruption issues at least in some Indian states (see e.g. Niehaus and Sukhtankar 2013a, 2013b), the administrative records should provide an upper bound on NREGS impacts.

According to administrative records, the employment guarantee scheme provided 1.4 billion person-days of employment in 1.78 million projects in the 330 Phase 1 and Phase 2 NREGS districts in 2007-2008.[27] 61.15 percent of this employment was given to women. The average daily wage paid was 75 rupees (about $1.8). This means that in a typical week, the scheme generated 83677 workdays of employment in 104 projects in the average district. With an average prime-aged district population of 1.10 million people, this translates into 0.0764 NREGS workdays per week per person. In the NSS data, the number of public-works workdays in Phase 1 and Phase 2 districts are 0.0789 for prime-aged adults, or about 4 days of public employment per person per year. This means that the NREGS employment generated for the chosen sample of prime-aged adults in this paper is in the same ballpark as that suggested by administrative sources, and is low at the local level: the implied weekly number of NREGS workdays per prime-aged adult in the average district would be 0.9615, for example, if we assume that 50 percent of workers have a NREGS job for 100 days per year. These back-of-the-envelope calculations therefore support the public employment results in this paper in that generated employment opportunities seem to be relatively modest at the local level.[28]

This conclusion runs counter to the results obtained in most of the difference-in-difference papers that analyze the impact of NREGS on wages and employment and typically find substantial wage effects. I discuss this issue in more detail in Appendix D and find that the overall results of my paper do not directly contradict the DID results of other papers, but mostly reflect different choices about sample composition

[27]The NREGS year starts on April 1, whereas the NSS household survey data starts in July, so the overlap of both data sources is not perfect.

[28]Another way of scaling the public-employment impacts is to calculate the annual increase in NREGS employment implied by the regression results. Taking the RD estimates for public employment from the Phase 2 vs Phase 3 regressions of Table 4 literally, they imply a 6 percentage point increase in public employment per year. According to administrative data, the average person worked 42 days in that year. This implies that about 1 percent of a district's population had a NREGS job at some point over the course of the year.

and the main empirical specification: Analyzed for a general sample of the working-age population the overall labor-market impacts of NREGS are relatively modest. This is also confirmed by appendix table C.0.7, which provides the estimates of a DID approach for the sample used to generate the RD results and finds no statistically significant wage effects. As appendix figure C.0.7 shows, however, the parallel trend assumption underlying the DID approach is violated for private employment at baseline, and this is also true for a number of other labor-market outcomes not reported here. This implies that the regression-discontinuity estimates, which do not require this assumption, provide the more believable program effects.

1.7 Conclusion

Using a regression discontinuity design, this paper has analyzed the impacts of the Indian National Rural Employment Guarantee Scheme (NREGS) on the rural labor market. The results suggest that the overall direct effects on the labor market are small, although many of the coefficients are so imprecisely estimated that larger effects cannot always be ruled out. The general qualitative pattern is robust across a range of different empirical specifications, however: the introduction of the public-works scheme at best only leads to small increases is public employment and, if it affects it at all, lowers the private-sector wage. There is some evidence that workers drop out of the private sector and move into family employment. The NREGS employment impacts are also statistically significantly higher in high-risk districts than in low-risk districts and after a negative rainfall shock.

Overall, these results suggest that NREGS is ineffective at raising private-sector casual wages through increased competition in rural labor markets or a better enforcement of minimum wage laws. The program seems to work better at providing a safety net for rural populations, although this does not translate into substantial improvements in other variables like per-capita expenditures, at least in the short run. The results, while imprecise, are also consistent with NREGS indirectly subsidizing self-employment activities by making them less risky. NREGS here mainly functions as an insurance tool after bad economic shocks rather than as a way to accumulate precautionary savings.

Given the large size of a program like NREGS with expenditures of about 1 percent of Indian GDP, the results raise the question whether the provided welfare benefits are large enough to warrant the existence of such an ambitious scheme, at least in its current form, or whether the money would be more effectively spent on other anti-poverty measures. In the presence of widely documented implementation problems like rationing of NREGS jobs, the program may disproportionately benefit the poor who have the option of becoming self-employed rather than the most economically vulnerable households with few employment alternatives. Broader welfare benefits will therefore depend heavily on improving implementation quality, although some other research on NREGS also suggests that wage impacts may take more time to materialize than could be analyzed in this paper.

Table 1.1: Predictive Success of Algorithm for Major Indian States

.

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

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

Table 1.2: Baseline Tests

.

Specification	employment public	employment private	employment family	employment total	log private wage	education	land	log per capita expenditure
Panel A: men								
Linear	-0.0006	-0.0188	0.0077	-0.0111	0.0596	-0.16*	83.97	-0.0015
	(0.0024)	(0.0187)	(0.0212)	(0.0201)	(0.0398)	(0.09)	(123.03)	(0.0314)
Linear Flexible Slope	-0.0007	-0.0187	0.0077	-0.0109	0.0596	-0.16*	80.19	-0.0019
	(0.0024)	(0.0187)	(0.0212)	(0.0199)	(0.0397)	(0.09)	(118.21)	(0.0314)
Quadratic	-0.0009	-0.0155	0.0088	-0.0069	0.0527	-0.17*	31.01	-0.0116
	(0.0023)	(0.0187)	(0.0210)	(0.0194)	(0.0396)	(0.09)	(118.39)	(0.0315)
Quadratic Flexible Slope	-0.0013	-0.0365	0.0297	-0.0070	0.0805	-0.04	51.60	-0.0248
	(0.0040)	(0.0265)	(0.0278)	(0.0277)	(0.0542)	(0.11)	(147.20)	(0.0403)
N	1063	1063	1063	1063	1007	1063	1063	1063
outcome mean	0.0025	0.3109	0.5529	0.8663	4.0352	3.32	1099.63	6.34
Panel B: women								
Linear	0.0018	0.0005	0.0459	0.0503	0.0608	-0.17*	53.70	-0.0037
	(0.0012)	(0.0132)	(0.0303)	(0.0336)	(0.0494)	(0.09)	(130.69)	(0.0317)
Linear Flexible Slope	0.0018	0.0003	0.0457	0.0500	0.0609	-0.17*	49.72	-0.0041
	(0.0012)	(0.0130)	(0.0302)	(0.0333)	(0.0495)	(0.09)	(126.00)	(0.0317)
Quadratic	0.0018	-0.0011	0.0420	0.0450	0.0615	-0.18**	-3.91	-0.0133
	(0.0012)	(0.0129)	(0.0298)	(0.0330)	(0.0494)	(0.09)	(123.27)	(0.0319)
Quadratic Flexible Slope	0.0047**	-0.0170	0.0278	0.0183	0.1324**	-0.11	-3.70	-0.0265
	(0.0020)	(0.0162)	(0.0394)	(0.0440)	(0.0645)	(0.11)	(155.16)	(0.0400)
N	1063	1063	1063	1063	656	1063	1063	1063
outcome mean	0.0018	0.1400	0.3059	0.4480	3.6807	2.34	1134.90	6.35

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is a district in a given season in the baseline data (July 2004-June 2005). An employment outcome is the proportion of working-age adults (18-60 years) with at most secondary education in rural areas working in a given type of employment in the last 7 days. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment.

Table 1.3: Summary Statistics for Districts at Baseline by Phase (Men and Women)
.

	Men				Women			
	phase 2		phase 3		phase 2		phase 3	
		N		N		N		N
private employment	0.2975	396	0.2938	668	0.1397	396	0.1332	668
family employment	0.5810	396	0.5271	668	0.2559	396	0.3281	668
public employment	0.0038	396	0.0015	668	0.0028	396	0.0013	668
daily wage (total)	52.75	387	65.71	645	38.19	306	45.93	504
daily wage (private)	52.77	386	65.78	645	37.69	303	45.76	497
daily wage (public)	53.44	18	63.54	22	53.42	12	52.32	17

Note: An observation is a district with non-missing Planning Commission index value in a given season in the baseline data (July 2004-June 2005). Summary statistics are calculated from aggregated and weighted individual NSS data.

Table 1.4: NREGS Impact: Wages and Employment (Men and Women)

.

Specification	employment public	private	family	total	log private wage
Panel A: men					
Linear	0.0012	-0.0351*	0.0253	-0.0069	-0.0041
	(0.0038)	(0.0208)	(0.0247)	(0.0185)	(0.0377)
Linear Flexible Slope	0.0011	-0.0351*	0.0256	-0.0068	-0.0041
	(0.0038)	(0.0208)	(0.0244)	(0.0185)	(0.0377)
Quadratic	0.0007	-0.0369*	0.0292	-0.0055	-0.0070
	(0.0038)	(0.0204)	(0.0243)	(0.0187)	(0.0375)
Quadratic Flexible Slope	0.0018	-0.0522*	0.0302	-0.0165	-0.0196
	(0.0045)	(0.0273)	(0.0331)	(0.0231)	(0.0500)
N	1063	1063	1063	1063	1007
outcome mean	0.0069	0.3279	0.4846	0.8195	4.1212
Panel B: women					
Linear	0.0013	-0.0035	0.0166	0.0140	0.0041
	(0.0044)	(0.0166)	(0.0259)	(0.0301)	(0.0660)
Linear Flexible Slope	0.0013	-0.0034	0.0161	0.0137	0.0038
	(0.0044)	(0.0166)	(0.0256)	(0.0298)	(0.0663)
Quadratic	0.0015	-0.0020	0.0108	0.0101	0.0050
	(0.0045)	(0.0165)	(0.0255)	(0.0296)	(0.0660)
Quadratic Flexible Slope	-0.0026	-0.0073	0.0340	0.0263	-0.0706
	(0.0043)	(0.0210)	(0.0334)	(0.0385)	(0.0925)
N	1063	1063	1063	1063	656
outcome mean	0.0053	0.1309	0.2285	0.3647	3.6488

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is a district in a given season. An employment outcome is the proportion of working-age adults (18-60 years) with at most secondary education in rural areas working in a given type of employment in the last 7 days. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment.

Table 1.5: NREGS Impacts and Safety Net (Men and Women)

.

	employment				log private
Specification	public	private	family	total	wage
Panel A: men					
Linear	-0.0047	-0.0291	0.0400	0.0057	0.0365
	(0.0090)	(0.0302)	(0.0330)	(0.0262)	(0.0512)
NREGS*negative shock	0.0285*	-0.0212	-0.0129	-0.0026	-0.0605
	(0.0148)	(0.0336)	(0.0411)	(0.0319)	(0.0702)
Linear Flexible Slope	-0.0050	-0.0288	0.0397	0.0101	0.0365
	(0.0090)	(0.0303)	(0.0331)	(0.0260)	(0.0513)
NREGS*negative shock	0.0288*	-0.0219	-0.0124	-0.0019	-0.0610
	(0.0147)	(0.0337)	(0.0414)	(0.0317)	(0.0712)
Quadratic	-0.0058	-0.0282	0.0397	0.0057	0.0326
	(0.0090)	(0.0302)	(0.0330)	(0.0262)	(0.0511)
NREGS*negative shock	0.0286*	-0.0214	-0.0128	-0.0026	-0.0583
	(0.0147)	(0.0337)	(0.0412)	(0.0319)	(0.0708)
Quadratic Flexible Slope	-0.0057	-0.0381	0.0389	-0.0051	-0.0056
	(0.0107)	(0.0404)	(0.0458)	(0.0326)	(0.0677)
NREGS*negative shock	0.0299**	-0.0223	-0.0085	0.0021	-0.0595
	(0.0152)	(0.0337)	(0.0414)	(0.0316)	(0.0717)
N	532	532	532	532	504
outcome mean	0.0115	0.3380	0.4681	0.8176	4.1786
Panel B: women					
Linear	-0.0053	0.0150	0.0011	0.0052	-0.0148
	(0.0081)	(0.0232)	(0.0285)	(0.0342)	(0.0801)
NREGS*negative shock	0.0240	-0.0271	0.0104	0.0158	-0.0071
	(0.0166)	(0.0287)	(0.0401)	(0.0458)	(0.1191)
Linear Flexible Slope	-0.0020	0.0101	0.0173	0.0193	0.0175
	(0.0088)	(0.0249)	(0.0306)	(0.0361)	(0.0873)
NREGS*negative shock	0.0245	-0.0277	0.0127	0.0178	-0.0128
	(0.0166)	(0.0287)	(0.0397)	(0.0454)	(0.1188)
Quadratic	-0.0053	0.0150	0.0011	0.0052	-0.0148
	(0.0081)	(0.0232)	(0.0285)	(0.0342)	(0.0801)
NREGS*negative shock	0.0240	-0.0271	0.0104	0.0158	-0.0071
	(0.0166)	(0.0287)	(0.0401)	(0.0458)	(0.1191)
Quadratic Flexible Slope	-0.0163*	0.0100	0.0404	0.0304	-0.0215
	(0.0094)	(0.0284)	(0.0385)	(0.0433)	(0.1032)
NREGS*negative shock	0.0280	-0.0277	0.0067	0.0149	-0.0049
	(0.0172)	(0.0278)	(0.0396)	(0.0456)	(0.1201)
N	532	532	532	532	321
outcome mean	0.0093	0.1282	0.2114	0.3489	3.7233

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the district level in parentheses. *negative shock* is a dummy variable equal to 1 if there was a negative deviation of rainfall from expected rainfall during the last monsoon season. Sample is restricted to agricultural off-season.

Table 1.6: NREGS Impacts and Risk: Wages and Employment (Men)

.

Specification	public	employment private	family	total	log private wage
Linear	-0.0056	-0.0447*	0.0335	-0.0241	-0.0153
	(0.0082)	(0.0243)	(0.0274)	(0.0313)	(0.0552)
NREGS*risky	0.0169**	0.0160	-0.0151	0.0527**	0.0765
	(0.0072)	(0.0200)	(0.0223)	(0.0258)	(0.0550)
Linear Flexible Slope	-0.0055	-0.0461	0.0282	-0.0240	-0.0505
	(0.0082)	(0.0350)	(0.0389)	(0.0306)	(0.0593)
NREGS*risky	0.0167**	0.0026	0.0341	0.0525**	0.0780
	(0.0073)	(0.0309)	(0.0319)	(0.0256)	(0.0553)
Quadratic	-0.0064	-0.0351	0.0158	-0.0244	-0.0180
	(0.0081)	(0.0325)	(0.0353)	(0.0313)	(0.0548)
NREGS*risky	0.0168**	0.0021	0.0347	0.0527**	0.0756
	(0.0073)	(0.0308)	(0.0319)	(0.0257)	(0.0552)
Quadratic Flexible Slope	-0.0033	-0.0483	0.0169	-0.0333	-0.0633
	(0.0092)	(0.0418)	(0.0479)	(0.0372)	(0.0729)
NREGS*risky	0.0166**	0.0025	0.0339	0.0521**	0.0773
	(0.0072)	(0.0309)	(0.0319)	(0.0254)	(0.0551)
N	532	532	532	532	504
outcome mean	0.0115	0.3380	0.4681	0.8176	4.1786

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the district level in parentheses. Log private wage in column 5 is conditional on private employment. Risky districts are those with an above-median variance of rainfall. Regressions control for expected rainfall.

Table 1.7: NREGS Impacts and Different Sample Restrictions: Wages and Employment (Men and Women)

.

Specification	public	employment private	family	total	log private wage
Panel A: men					
Linear flexible	-0.0014	-0.0431	0.0058	-0.0342	-0.0747
(-5$\leq$ rank $\leq$ 5)	(0.0052)	(0.0350)	(0.0393)	(0.0272)	(0.0589)
N	543	543	543	543	522
Linear flexible	0.0014	-0.0551	0.0215	-0.0238	-0.0568
(-4$\leq$ rank $\leq$ 4)	(0.0084)	(0.0348)	(0.0436)	(0.0315)	(0.0615)
N	463	463	463	463	445
Linear flexible	-0.0034	-0.0690*	0.0319	-0.0298	-0.1120
(-3$\leq$ rank $\leq$ 3)	(0.0055)	(0.0401)	(0.0531)	(0.0368)	(0.0805)
N	375	375	375	375	358
Panel B: women					
Linear flexible	-0.0001	0.0105	0.0200	0.0304	0.0105
(-5$\leq$ rank $\leq$ 5)	(0.0045)	(0.0278)	(0.0368)	(0.0448)	(0.1010)
N	543	543	543	543	363
Linear flexible	0.0033	0.0107	0.0613	0.0776	0.0697
(-4$\leq$ rank $\leq$ 4)	(0.0077)	(0.0305)	(0.0434)	(0.0518)	(0.0989)
N	463	463	463	463	311
Linear flexible	-0.0035	0.0109	0.1045**	0.1111*	-0.0421
(-3$\leq$ rank $\leq$ 3)	(0.0049)	(0.0352)	(0.0454)	(0.0563)	(0.1124)
N	375	375	375	375	251

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is a district in a given season. An employment outcome is the proportion of working-age adults (18-60) with at most secondary education in rural areas working in a given type of employment in the last 7 days. The log private wage in column 5 is conditional on private employment. Sample restrictions apply to the re-centered state-specific rank variable.

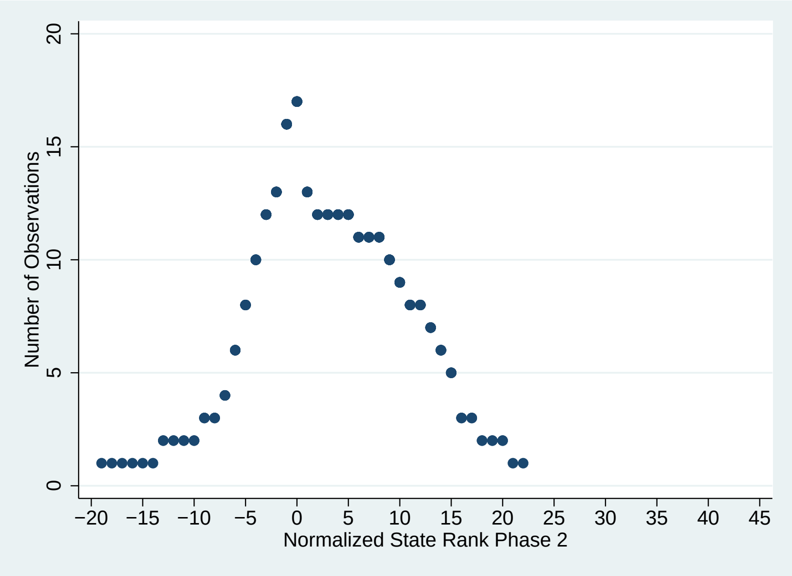
Table 1.8: NREGS Impact (Meta-Analysis): Wages and Employment (Men)

.

Specification	employment public	 private	 family	 total	log priv. wage
Linear (simple av.)	-0.0021	-0.0348**	0.0302	-0.0067	0.0153
	(0.4926)	(0.0283)	(0.1113)	(0.6368)	(0.6214)
Quadratic (simple av.)	0.0029	-0.0738***	0.0693***	-0.0017	-0.0156
	(0.3468)	(0.0001)	(0.0037)	(0.9108)	(0.6865)
Linear (pop. weighted)	-0.0016	-0.0299*	0.0374*	0.0059	0.0111
	(0.5661)	(0.0632)	(0.0606)	(0.6802)	(0.7301)
Quadratic (pop. weighted)	-0.0003	-0.0501***	0.0616***	0.0113	-0.0059
	(0.9297)	(0.0051)	(0.0067)	(0.4549)	(0.8729)
N	951	951	951	951	927

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is a district in a given season. An employment outcome is the proportion of working-age adults (18-60) with at most secondary education in rural areas working in a given type of employment in the last 7 days. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment. NREGS is the predicted treatment status. Treatment effects at the cutoff are estimated separately by state and then combined through a simple average in the *simple av.* specifications, whereas the state-specific estimates are weighted by state popualtion in the *pop. weighted* specifications.

Figure 1.1: Number of observations per state rank for Phase 2



Note: Figure 1.1 excludes Phase 1 districts. Planning Commission ranks are made state-specific and re-centered such that the last district eligible for receiving NREGS in Phase 2 according to the proposed algorithm has a rank of 0. Districts with positive ranks should be ineligible for the program.

Figure 1.2: General Distribution of Index over Ranks

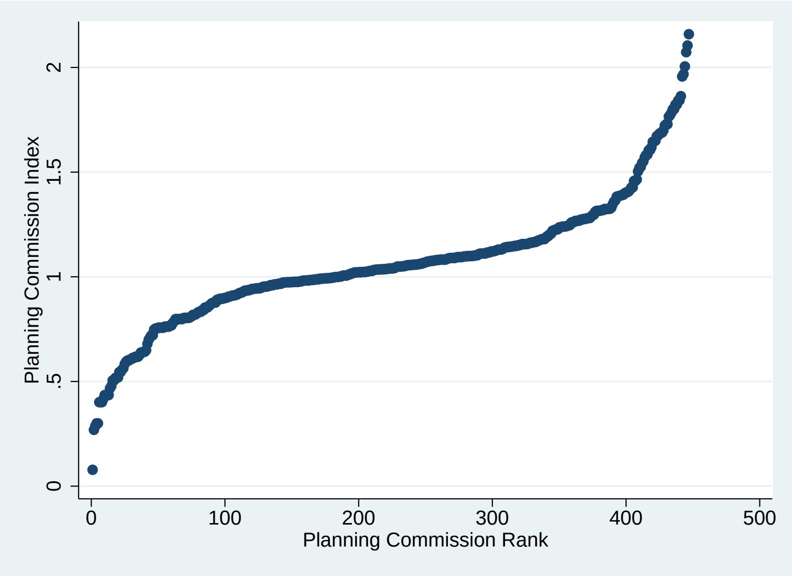


Figure 1.3: Distribution of Index over State-Specific Ranks (Phase 2 vs Phase 3)

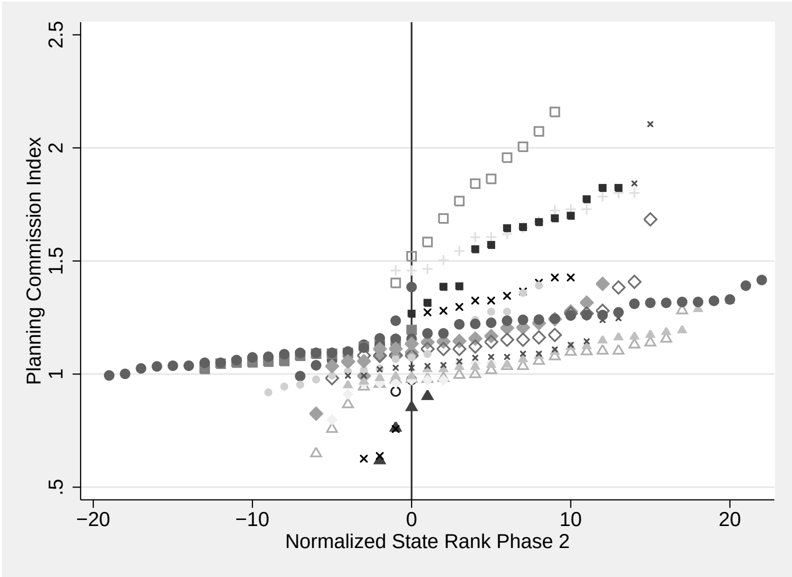
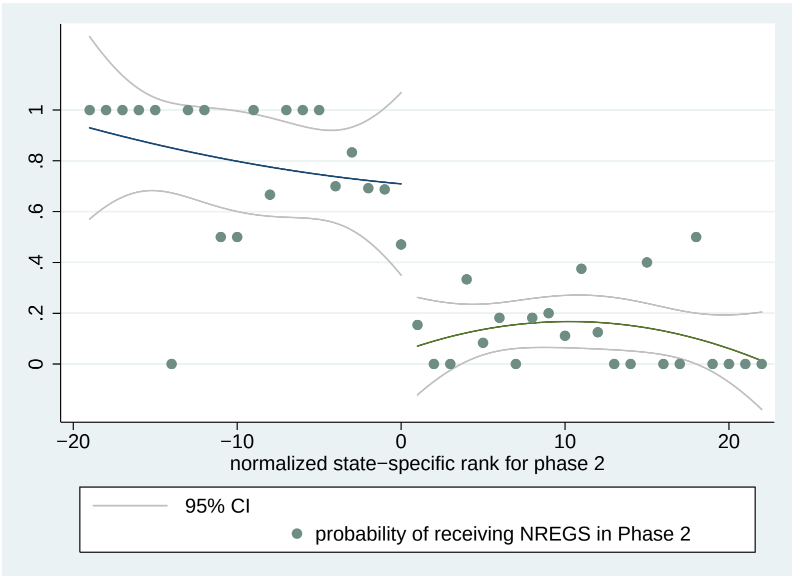


Figure 1.4: Discontinuity of treatment status for Phase 2



Note: Figure 1.4 excludes Phase 1 districts. The used bin size is one (each individual rank).

Figure 1.5: NREGS impact on male public employment

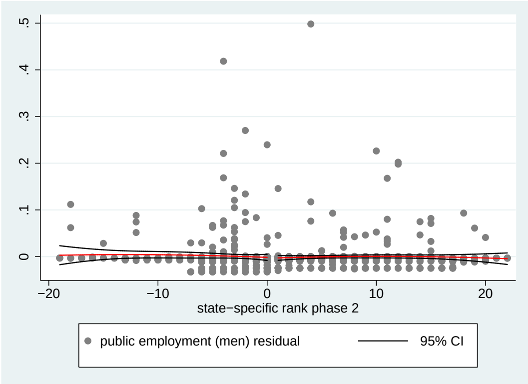


Figure 1.6: NREGS impact on male private employment

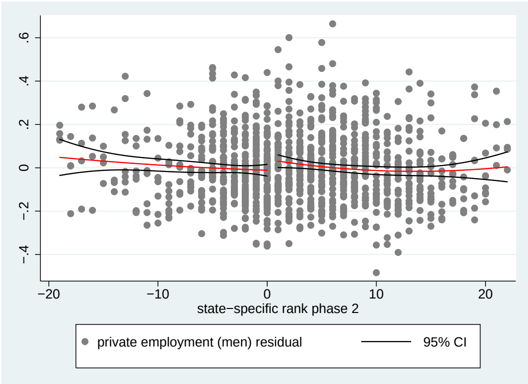
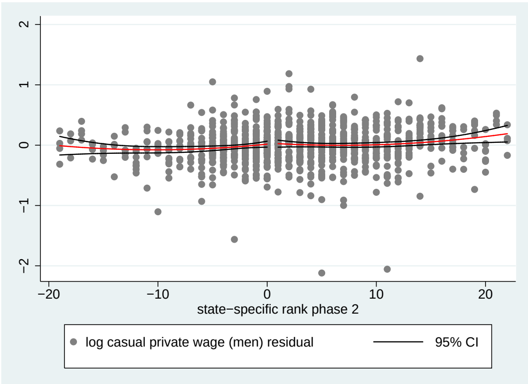


Figure 1.7: NREGS impact on male log private wage



Note: An observation is the residual average district-season-level outcome at a given rank, where the impact of the baseline outcome variable and state fixed effects has been taken out. Fitted curves are quadratic polynomials.

Figure 1.8: NREGS impact on female public employment

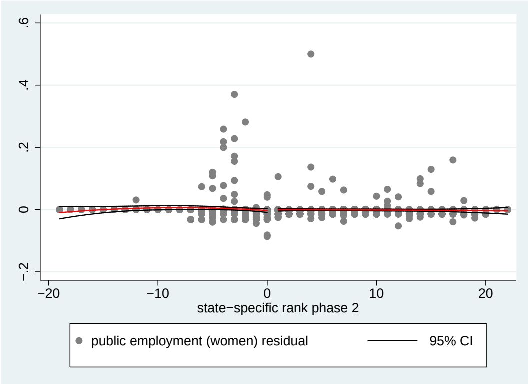


Figure 1.9: NREGS impact on female private employment

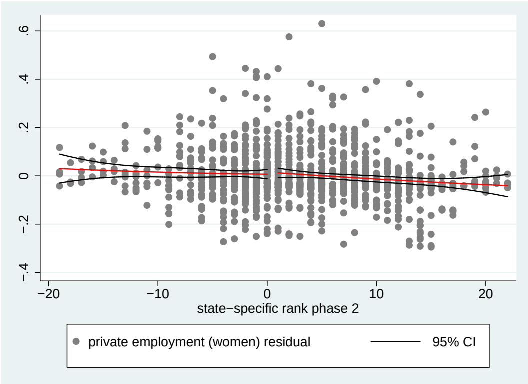
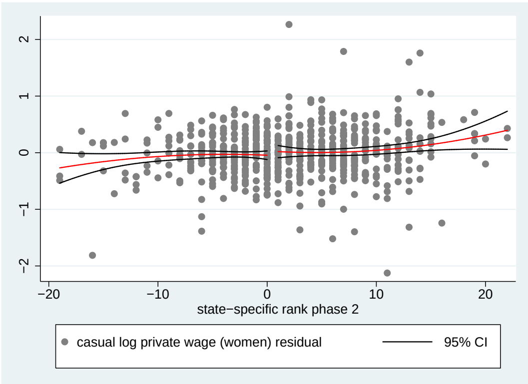


Figure 1.10: NREGS impact on female log private wage



Note: An observation is the residual average district-season-level outcome at a given rank, where the impact of the baseline outcome variable and state fixed effects has been taken out. Fitted curves are quadratic polynomials.

CHAPTER II

Guns *and* Butter? Fighting Violence with the Promise of Development

2.1 Introduction

Internal military conflicts between government troops and insurgents are common in many developing countries.[1][2] Governments have traditionally relied very heavily on military force, but there is a growing awareness that this alone may not be enough to end violence since insurgents often rely on the loyalty of the local population in their guerrilla tactics and recruit members from economically marginalized groups. In such situations, government anti-poverty programs that target conflict areas are increasingly seen as a potential tool for helping reduce conflict intensity by raising

[1]This chapter is joint work with Gaurav Khanna.

[2]Over 20% of countries experienced internal violence over the course of the 1990s (Blattman and Miguel 2010). Cross-country data also shows a high correlation between poverty and conflict (see e.g. Collier and Hoeffler 2007). Miguel et al. (2004) and Miguel and Satyanath (2011) find that economic growth, instrumented by rainfall shocks, has a negative impact on conflicts. For recent microeconomic studies on the relationship between development and conflict, see e.g. Do and Iyer (2007), Murshed and Gates (2005), and Humphreys and Weinstein (2008).

the opportunity cost of being an insurgent and improving the willingness of civilians to support the government.[3] At the same time, however, such programs may increase violence, for example if the resources flowing into conflict areas make territorial control of these locations more attractive for insurgents.[4]

What effect government programs have on internal conflict intensity is therefore an empirical question. In this paper, we analyze the impact of the world's largest public-works program, the Indian National Rural Employment Guarantee Scheme (NREGS), on the incidence of insurgency-related violence in the country. NREGS is based on a legal guarantee of 100 days of public-sector employment to all rural households (about 70 percent of the population) willing to work at the minimum wage, and annual expenditures on the scheme amount to around one percent of Indian GDP. While the program's main goals are to generate labor market opportunities and to improve local infrastructure, one of the expectations of the government was to reduce incidents of violence by Naxalites, a group of Maoist insurgents that ia a major internal security threat in India.

A large program like NREGS could in principle affect insurgency-related violence through a number of channels. One of the main potential mechanisms is the citizen-support channel, where the introduction of a large anti-poverty program makes civilians more likely to assist the police in their fight against insurgents. To clarify the expected empirical impacts from this channel, we set up a theoretical model in which government troops and insurgents fight over territory. NREGS improves the relationship between the government and the people and makes civilians more willing to provide information and other kinds of assistance to the police, which in turn increases the police's efficiency in tracking down rebels. In contrast to the existing literature, the model predicts that the program should lead to an increase in police

[3] See e.g. Grossman (1991) for an opportunity cost model and Akerlof and Yellen (1994) for a model of civilian support in the context of street-gangs.

[4] See e.g. Hirshleifer (1989), Grossman (1991), Skaperdas (1992).

initiated attacks against Maoists, and retaliatory attacks by the Maoists on civilians informants. It also predicts that violence should increase in the short run, but decline only in the longer run when the presence of insurgents has become weaker. In addition to the theoretical predictions implied by the citizen-support channel, we also discuss the existing qualitative evidence on the nature of the Maoist conflict and the implied empirical patterns of a number of alternative theories.

We then test these predictions empirically. NREGS was rolled out non-randomly in three implementation phases, with poor districts being treated earlier. The government used an algorithm to assign districts to phases which generates state-specific treatment discontinuities and allows the use of a regression discontinuity design to analyze the empirical impact of the program. The results show that for districts that received the program, it leads to about 914 more fatalities in about 368 more incidents over the following year. Consistent with the hypothesis that the police are cracking down on the Maoists, we find that most attacks are initiated by the police, and that the insurgents are the most affected group, whereas there is little impact on police casualties. There is also some evidence of an increase in the number of attacks by insurgents on civilians. Most of these impacts are concentrated in the short run, and the results are robust across a number of different specifications.

Taken together, these empirical patterns are consistent with the predictions of our theoretical model and the existing qualitative and quantitative evidence on the Maoist conflict in the literature, but are difficult to explain with a number of alternative theories. Overall, the empirical results therefore suggest that anti-poverty programs may improve the relationship between the government and the local population, which in turn may help reduce the intensity of internal conflict in the longer run. This provides a non-negligible additional benefit of government programs which, at least initially, may occur even for programs that are prone to implementation problems.

Our paper contributes to a growing literature on the impact of government programs on internal violence in developing countries. While the impact of development programs on conflicts has been studied in a few different contexts, the empirical patterns and explanations differ substantially. Berman, Shapiro and Felter (2011), for example, look at the impact of a development program in Iraq. They find that violence decreases after the program is introduced, and argue that this is due to increased information-sharing by civilians. In the Philippines, Crost, Felter and Johnston (forthcoming) show that insurgents sabotage government programs because they fear that the success of such programs would encourage civilians to switch support away from the insurgents. In other contexts, recent papers find both positive and negative impacts of government programs on internal conflicts that are typically consistent with more than one explanation (see e.g. Nunn and Qian 2012, Crost, Felter and Johnston 2012, Dube and Vargas 2012).

Given this heterogeneity across countries, a deeper understanding of how exactly government programs affect internal violence is of high policy relevance. The Indian case allows a cleaner analysis of the underlying mechanism than is possible in many other contexts since a number of potential channels are not relevant here. There is no evidence that insurgents sabotage NREGS projects directly, for example, probably because the projects are difficult to sabotage since the program hardly creates any assets or destroyable infrastructure. This also means that NREGS does not provide appropriable assets that would be attractive to gain control over.

The program may increase the opportunity costs of supporting the Maoists by providing employment opportunities and a safety net, however. This would ideally lead to a fall in violence on the introduction of NREGS. Fetzer (2013) finds that NREGS attenuates the relationship between rainfall shocks and Maoist violence, which is also consistent with this idea. But the opportunity-cost channel cannot explain the overall effect of the employment guarantee scheme that we find in this paper, and therefore cannot be the dominant channel. Additionally, Vanden Eynde (2011) argues that the link between violence and rainfall-shocks arises out of the citizen-support channel since he proposes a model where Maoists try to prevent the local population from becoming police-informants in bad economic times, by increasing their attacks on civilians suspected to work for the government. Fetzer's (2013) results could therefore also be explained by the idea that NREGS increases the willingness of civilians to help the government at all times, and not just during bad economic times, breaking the link between rainfall fluctuations and violence. In any case, the empirical patterns in our paper as well as qualitative evidence on the Maoist conflict point to the citizen-support channel as the most plausible mechanism to explain the impact of NREGS on internal violence.

Dasgupta, Gawande and Kapur (2014) use a difference-in-difference approach and find that NREGS increases Maoist violence in the year of implementation, but then leads to lower violence in the year after. This long-run reduction in violence is concentrated in Andhra Pradesh, a state with high implementation quality relative to other areas. The authors attribute this effect to rising opportunity costs. Again, the empirical patterns are also consistent with the citizen-support channel explanation, however, which predicts a short-run increase and a long-run decrease in violence, whereas the higher conflict intensity in the short run is difficult to explain with an opportunity-cost story.

Overall, the results in this paper therefore contribute to our understanding of the mechanisms underlying the impact of government programs on internal violence, and help reconcile some of the seemingly contradictory findings in the recent literature. The results in Crost, Felter and Johnston (forthcoming), for example, are consistent with the insurgents seeking to prevent the citizen-support channel from working, by sabotaging the projects under the program. While our paper suggests that the citizen-support channel may well lead to an initial increase in violence, it predicts that violence would eventually fall as in Berman, Shapiro, and Felter (2011). This also suggests that dynamic patterns are important for understanding the impact of government programs on violence intensity, which are typically neglected in the existing literature.

The remainder of this paper is structured as follows: Section 2.2 discusses potential hypotheses regarding the impact of NREGS on violence and sets up a theoretical model of the citizen-support channel. Section 2.3 provides some background on the Maoist movement and NREGS, whereas section 2.4 describes the empirical strategy and the data. Section 2.5 presents the main results as well as some extensions and robustness checks, and section 2.6 concludes.

2.2 Theoretical Model and Alternative Theories

To clarify the expected empirical effects of NREGS on Maoist-related violence according to different theories, we first set up a theoretical model that incorporates the importance of citizens assisting the government by sharing information, although in practice this can also include other forms of assistance. We then discuss alternative explanations and their empirical predictions. While there are some models that stress the importance of citizen support, such as the models in Berman, Shapiro and Felter (2011) on counterinsurgency in Iraq and the Akerlof and Yellen (1994) study on street-gangs, our model differs from those in a number of respects that fit our context better. First, we allow insurgents to fight for territory, whereas the rebels' goal in the Berman, Shapiro and Felter model is only to impose costs on the government. Second, in our model civilians make their information-sharing decisions before rather than after the government and the insurgents move. Lastly, we consider how aggregate violence patterns may change dynamically.

2.2.1 A Citizen-Support Model

The model describes the optimal strategies in the conflict by three players, the government, the Maoists, and the civilians. In the Indian context, the employment guarantee scheme was implemented across the country and prioritized poor districts regardless of their internal security condition in the assignment algorithm. Therefore, the decision about whether, and if so, how much, to invest in anti-poverty programs like an employment guarantee scheme is taken to be exogenous.[5] There are L identical locations in the country where the government fights for territorial control with the insurgents. In each location, the probability that the government gains control of the territory $p(m, v, i)$, depends positively on the amount of governmental military action m, negatively on the amount of Maoist violence inflicted upon the police v, and positively on the amount of information that the police has i.

Civilians first choose[6] how much information i to share with the government to maximize their expected utility

$$EU_C = b(i) - c_C(r(i)) + p(m, v, i)u(y_G + g) + (1 - p(m, v, i))u(y_N) \tag{2.1}$$

where $b(i)$ is the utility derived from the benefits of sharing information with the government, which may include both monetary and non-monetary components. $c_C(r(i))$ measures the disutility from sharing information because Maoists may retaliate against civilians for sharing information according to the known retaliation function *r(i)*. y_G and y_N are the benefits civilians receive when their location is under government control or Naxalite control at the end of the period, respectively, and *u(.)* is the utility function for these benefits. g is the extra benefit to citizens from

[5] In other contexts, a number of economic and political economy factors will enter the government's objective function in addition to anticipated internal security benefits.

[6] The results are not sensitive to the order of moves as long as the rebels and the police move simultaneously, and the government expenditure is decided on before the civilians move. The order used in this model is related to the context - where civilians first decide on providing tip offs to the government, and the police then act on the information

governmental programs like an employment guarantee scheme.

Overall, civilians therefore take into account both costs and benefits that arise directly from providing assistance to the government and the benefits provided by whoever is in power at the end of the period, which is also influenced by the level of information.

After civilians have made their decision, government troops and the insurgents simultaneously decide on their actions. The police decides how much military action m to take against the Maoists to maximize the expected utility

$$EU_G = p(m, v, i) - c_G(m) \tag{2.2}$$

For simplicity, the government's expected utility from territorial control is assumed to equal the probability $p(.)$ that the government gains control, whereas the disutility from military action is given by $c_G(m)$.

At the same time, the Naxalites determine how many attacks v to plan against the government, maximizing their expected utility

$$EU_N = [1 - p(m, v, i)] - c_N(v) \tag{2.3}$$

where $c_N(v)$ are the costs incurred from violence level v and $1 - p(.)$ is the probability that the Maoists will be in control of the location at the end of the period. Additionally, the Maoists retaliate against civilians for working as police informers, where retaliation $r(i)$ increases with the amount of information and assistance provided to the police.[7]

Together, the decisions by government actors and insurgents determine the level of

[7]The retaliation is used to prevent further sharing of information, which is something that is not captured in this one-period model. It is also possible to model retaliation as a decision taken simultaneously with the civilian's information sharing in order to capture the value of the 'threat' of retaliation.

violence in a given location and the endogenous probability that the government gains territorial control. At the end of the period, a location either becomes controlled by the government or remains contested, and payoffs are realized. In the next period, the process is repeated in all locations that remain contested, whereas there is no further violence in government-controlled areas. Cost functions $c_C(.)$, $c_G(.)$, and $c_N(.)$ are increasing and convex.

The model can be solved by backward induction to find the pure-strategy subgame perfect Nash equilibrium. Once civilians have decided on the amount of information $i*$ to share with the government, the government maximizes its expected utility, taking $i*$ and the violence level v chosen simultaneously by the Maoists as given. The first-order condition of (2) is therefore given by

$$\frac{\partial p(m, v, i^*)}{\partial m} - c_G'(m) \leq 0 \tag{2.4}$$

This equation pins down the best response function of military action $m*$ for every potential violence level v chosen by the insurgents. Since $c_G(.)$ is convex in m whereas *p(.)* is concave in m, a unique maximum exists according to the Intermediate Value Theorem that satisfies the second-order conditions.

Similarly, the first-order condition for the Maoists is given by

$$\frac{-\partial p(m, v, i*)}{\partial v} - c_N'(v) = 0 \tag{2.5}$$

which implicitly traces out the best-response function of $v*$ for every potential government violence level m. Assuming that $p(.)$ is decreasing and convex in v, this once again satisfies the second-order conditions.

In equilibrium, both actors make correct predictions about the level of violence chosen by the other player, leading to a Nash equilibrium in each subgame given the level of i^* where the best-response functions for the two players intersect. Assuming that $p_{mv} = p_{vm} < 0$,[8] it can be shown that $\frac{dm*}{dv} < 0$ and $\frac{dv*}{dm} > 0$ using the Implicit Function Theorem, which guarantees the existence of a unique Nash equilibrium.

We assume that government military action is more effective with access to more information $p_{mi} > 0$, while more information could make Maoist attacks against the police less effective $p_{vi} \leq 0$. This, in turn, implies that according to the Implicit Function Theorem $\frac{dm*}{di} > 0$ and $\frac{dv*}{di} \geq 0$, so more shared information by the civilians leads to higher levels of violence by both conflict parties.[9]

Civilians decide how much information to share with the government at the beginning of the period, knowing the best response and equilibrium violence-level functions, which leads to the first-order condition

$$b'(i) - c_C(r(i)) + \frac{dp}{di}[u(y_G + g) - u(y_N)] \leq 0 \tag{2.6}$$

where $\frac{dp}{di} = \frac{\partial p}{\partial m}\frac{dm}{di} + \frac{\partial p}{\partial v}\frac{dv}{di} + \frac{\partial p}{\partial i}$.[10] By the implicit function theorem, $\frac{di^*}{dg} = \frac{-\frac{dp}{di}u'(y_G+g)}{SOC} > 0$. This implies that civilians will assist the police with more information or assistance when they receive governmental programs like NREGS.[11]

As discussed above, a higher level of shared information increases m^* and v^*. Additionally, Naxalites also retaliate more against civilians since $r(i)$ is increasing in i. This means that overall violence in a given location rises after the introduction of the government program, and the impact will be greater for districts that do a better job of implementing the program.

The equilibrium decisions by civilians, insurgents and the government determine

[8]Intuitively, a given actor's effectiveness of violence on control over a location becomes lower the higher the violence of the other conflict party.

[9]For the police, the military action is complementary to the amount of information, and thus increases with more information. The rebels, however, fight harder to hold on to territory that is slipping away.

[10]Assistance and information increases the probability of police control if $\frac{\partial p}{\partial m}\frac{dm}{di} + \frac{\partial p}{\partial i} > -\frac{\partial p}{\partial v}\frac{dv}{di}$.

[11]While the police force in practice largely consists of local officers whereas NREGS is a national program, implementation quality largely depends on local institutions. Zimmermann (2013b) finds, for example, that both the government parties as well as local incumbents regardless of party affiliation benefit from NREGS in areas where the program is implemented well in the 2009 general elections. This suggests that the people are aware that local institutions and personnel matter.

the probability p^* that the government will gain control in a given location, or district, at the end of the period. Since all locations are identical, in expectation the number of contested territories ℓ_t will decrease over time according to the relationship

$$\ell_t = (1 - p^*)\ell_{t-1} \tag{2.7}$$

After the conflict has lasted τ periods, the number of contested location is therefore: $\ell_\tau = (1 - p^*)^\tau \ell_0$. Given the simplifying assumption in this model that violence in a location stops once the government gains control, the number of territories decreases over time until the war ends in period T when $\ell_T = 0$.

The improved information flow increases the equilibrium probability that the government gains control in a location, which will speed up the end of the conflict. With a higher government success probability more locations will fall under government control in a period than before, leading to the fewer contested territories in the next period. While violence in a given location has gone up, this effect means that the aggregate violence, averaged across locations, will fall over time as the government wins the war more quickly than it otherwise would have had.

Overall, the model therefore generates a number of testable predictions about the impact of a government program like NREGS on the incidence of conflict. First, the introduction of NREGS increases insurgency-related violence in the short-run. In the longer run, violence falls. Second, the program increases the government's effectiveness in tracking down insurgents, so there are more police-initiated attacks. This also implies that insurgents should be more likely to die or to be injured or captured than before. Furthermore, civilians may be more affected by violence than before if the Maoists retaliate against them for sharing information.

2.2.2 Alternative Theories

The citizen-support channel provides only one theory of the impacts the introduction of a large anti-poverty program like NREGS may have on insurgency-related violence. A number of alternative theories from the broader literature on the relationship between development and conflict are relevant in this context and often generate very different empirical predictions.

While our model implies that violence should rise initially, two established theories in the literature predict a fall in the incidence of conflict. The first theory considers the problem of credible commitment: In a situation where the government has been unable to credibly commit to economic development and in which insurgents fight for better economic conditions, a program like NREGS may solve this problem and 'complete the contract' (Powell 2006) between rebels and the government. The costs of dismantling NREGS may be high because of constitutional obligations and the political popularity the scheme may garner, and thereby force the government to keep implementing the program.[12] In this case, the introduction of NREGS gives insurgents fewer reasons to continue their struggle, and violence should fall. This also implies that we should expect to see a lower number of attacks by Maoists on the police, but not necessarily any changes in police action or attacks on civlians, assuming that NREGS is actually effective as an anti-poverty program.[13]

The second theory that predicts a fall in violence after the introduction of NREGS is an opportunity-cost story: If the program provides jobs and other welfare benefits from participation in the scheme, the program will increase the opportunity cost of being a Maoist (Grossman 1991). Naxalite supporters should therefore drop out of the

[12]Chapter 3 of this dissertation shows that the duration of NREGS receipt may have had an important influence on national election outcomes, for example.

[13]Additionally, there may be asymmetric information about the dismantling costs of NREGS, and the rebels may expect the government to renege on the promise (Dal Bo and Powell 2007). The Indian government has a long history of development programs that were often temporary and largely ineffective, which may affect present expectations on the longevity of this scheme and could further weaken any conflict-reducing impacts.

organization to take advantage of these improved economic opportunities, and rebels should find it harder to recruit new soldiers, both of which decrease the strength of the insurgents and their ability to inflict violence (see e.g. Grossman 1991 for such a model). We would then expect to see a fall in the number of insurgents, which may lead to a decrease in the amount of overall violence.[14] Again, for this explanation to apply NREGS needs to actually generate economic benefits.

In contrast to these two theories, one of the most widespread theories in the literature suggests, however, that we should expect violence to increase after the introduction of NREGS. This theory focuses on the idea of competition for resources (see e.g. Hirshleifer 1989, Grossman 1991, Skaperdas 1992): If NREGS increases the wealth of a region, this creates a larger resource pie that is worth fighting over. Contest models that focus on this channel usually predict that when resources rise in a region in equilibrium more effort will be put into fighting rather than production. Again, this presupposes that NREGS generates resources that can be appropriated through violence. We would also expect both rebel attacks against police forces and police-initiated attacks against the insurgents to increase as more assets are created, but there is little reason to expect an increase in violence against civilians.

Another mechanism that would predict an increase in violence is that NREGS may put a spotlight on treatment areas, encouraging the police to increase their efforts of cracking down on crime in these regions because of external pressures for the program to be seen as running smoothly. As NREGS is a big program that has garnered a lot of attention in the media, this could incentivize the state and district leaders to put pressure on the police to work harder than before to ensure a good image of their districts in the press, for example. This increased police effort would imply the same pattern as the citizen-support channel, with an increase in violence and especially of

[14]This idea is also closely related to work on economic inequality and group formation in the conflict literature. Grossman (1999) argues, for example, that incentives such as wages, opportunities to loot and protection from danger are often used to motivate participation. In this view, economic inequality may lead to conflict because there is more to gain from victory (Fearon 2007).

police-initiated attacks. However, the spotlight theory should encourage the police to crack down on other forms of crime and violence as well to make the security situation in their district look good, which is a prediction we will be able to test empirically. Moreover, under the citizen-support channel insurgents have a reason to retaliate against civilians, whereas there is no such motive under the spotlight theory.

While most of these different theories about the impact of government programs on internal violence can be disentangled by focusing on the implied patterns of changes in violence and of the most heavily affected groups, available qualitative and quantitative evidence on the nature of the Maoist conflict in India and the working of the employment guarantee scheme also provides useful information on the attractiveness of the various mechanisms in the Indian context.

2.3 Background

2.3.1 The Naxalite Movement

According to the Government of India, the Naxalite movement is one of India's most severe threats to national security. In 2006, Prime Minister Manmohan Singh famously referred to it as "the single biggest internal security challenge ever faced by our country".[15] Members of the movement are typically called Naxalites or Maoists, although official government documents often refer to affected districts as Left-Wing Extremism (LWE) districts.[16]

Naxalites have been operating since 1967 when landlords attacked a tribal villager in the small village Naxalbari in West Bengal and triggered an uprising. By the early

[15] Hindustan Times, April 13, 2006: Naxalism biggest threat: PM

[16] There is some debate in the literature about the correct way of addressing the insurgents. Mukherji (2012) argues, for example, that the insurgents should be referred to as Maoists rather than Naxalites since the organizations that grew out of the original Naxalite movement of the 1960s mostly reject the actions of the Communist Party of India (Maoist) (CPI(M)) that is largely responsible for the violence in recent years. A number of Naxalite organizations even refer to the CPI(M) as terrorists.

1970s, the movement had spread to Andhra Pradesh, Bihar and Orissa but splintered into more than 40 groups. In 2004, the two biggest previously competing Naxalite groups joined hands to form the Communist Party of India (Maoist). This is believed to have substantially exacerbated India's problem with the Naxalites and to have driven the recent growth in violence in the country (Lalwani 2011). The Indian Home Ministry believed the movement to have around 15000 members in 2006, to control about one fifth of India's forests, and to be active in 160 districts (Ministry of Home Affairs 2006). Figure 2.1 shows all the districts that experienced at least one Maoist incident between January 2005 and March 2008, the period studied in this paper, in black, dark grey and light grey. As can be seen, Naxalite-affected districts are concentrated in the eastern parts of India. These areas are often referred to as the Red Corridor.

The Naxalites' main goal is to overthrow the Indian state and to create a liberated zone in central India, since they believe that the Indian government neglects the lower classes of society and exclusively caters to the elites. Traditionally, the primary targets were landlords and upper caste landowners, but attacks have recently turned into larger and better-planned strikes on government institutions and personnel (Singh undated). One common target of attacks are infrastructure projects such as railways, public buses and telecommunication towers (Ramana 2011). The strongholds of the Naxalites continue to be in tribal areas that are typically resource-rich but chronically underdeveloped. Substantial mining activities that lead to the displacement of the tribal population and the absence of economic development opportunities are believed to explain the continuing popularity of the movement among the local population (see e.g. Borooah 2008). Naxalites finance themselves mainly through levies on industries and forest contractors, both in return for protection as well as in payment for engaging in illegal tree-felling or mining (Sundar 2011).

The Indian government has been fighting the Maoists since the 1960s, but decades of using force have been largely unsuccessful in suppressing the movement. While India officially subscribes to a population-centric approach to counterinsurgency, which relies on a mixture of force and winning the support of the local population by taking care of their grievances, a number of researchers note that India traditionally relies almost exclusively on military strength to fight the Naxalite movement (see e.g. Banerjee and Saha 2010, Lalwani 2011). The main responsibility in this fight rests with the civil and paramilitary forces of the state police in the affected areas, although they are often supported by central paramilitary battalions. While some states have been more successful at suppressing violence than others, the central government generally blames the overall failure to contain the Naxalites on inadequate training and equipment, as well as on poor coordination between police forces of different states. Many observers also refer to the often widespread disregard for local perceptions, however, as well as the sometimes excessively brutal nature of police force behavior that also affects many civilians. These destroy not only the trust of the local population in the Indian state but also the opportunity for police forces to take advantage of information on insurgents and other forms of assistance provided by the people (Bakshi 2009, Lalwani 2011, Sundar 2011).

Both Maoists and security forces believe that civilians have a lot of information on the insurgents, so pressures on the local tribal population (also called adivasis) to pick a side and cooperate with one of the conflict parties is high. The Naxalites' continued survival depends on help from civilians who hide them and provide them with resources and information. Maoist insurgents often warn the local population not to provide shelter or information to police forces, for example, and instead ask them to keep track of government personnel and their actions.[17] The government, on the other hand, often also does not seem to regard civilians as neutral, and some

[17]There is some evidence that insurgents in turn provide civilians with some help, for example in the form of teaching them more effective farming techniques (Mukherji 2012). Naxalites also claim to protect civilians from exploitation by large mining conglomerates (Borooah 2007).

experts claim that an important percentage of incarcerated adivasis are in jail due to false accusations of being Maoist supporters (Mukherji 2012).

In addition to these pressures, adivasis also face economic incentives to join the conflict: Their knowledge of local conditions in the often remote forest areas is very valuable for both insurgents and government troops. In areas of chronic underdevelopment with few employment opportunities, working for one of the conflict parties therefore allows the poor to earn some income (Mukherji 2012).[18]

In consequence, many adivasis are involved in the conflict as tacit supporters, informants and recruited fighters on both sides, and switching sides once conditions change is not uncommon.[19] Economic opportunities or changes in the perception of which side has the upper hand seem to affect behavior: In many states, Maoists who surrender to the police and provide information on their organization receive some land and money to start a new life, for example.[20] Vanden Eynde (2011) also shows that Naxalite violence against civilians increases after negative rainfall shocks, which is consistent with his theoretical model in which Maoists try to prevent the local population from being recruited as government informants during bad economic times. A number of instances where Maoists left leaflets after killing civilians, in which they accused the victims of being police informers, are also in line with the idea that Maoists retaliate against civilians believed to support the government.[21]

In light of this situation, the view that military force alone may not be effective in solving the Naxalite problem in the long run seems to have grown in recent years: In 2007, for example, Prime Minister Manmohan Singh said ‘Development and in-

[18]Many low-rank Maoists directly involved in encounters with security personnel as well as an important portion of the police force consist of young tribals, for example.

[19]See e.g. Mukherji (2012)

[20]See e.g. www.satp.org

[21]See e.g. www.satp.org for the following press releases from 2007: *“Cadres of the CPI-Maoist shot dead a 45-year-old shopkeeper at Sringeri in the Chikmagalur District, suspecting him to be a Police informer...Before fleeing, they left behind pamphlets with a message that read: ‘Let us expose informers and teach them a befitting lesson.’” “Two brothers were killed at Tamba village by the CPI-Maoist cadres on suspicion of being Police informers... More than 20 Police informers have reportedly been killed in the last one year in Jharkhand.”*

ternal security are two sides of the same coin. Each is critically dependent on the other.'[22] He also noted that many Maoist recruits come from economically deprived and marginalized groups of society. The central government has therefore shown a growing interest in increasing economic development in underdeveloped areas of the country through anti-poverty programs, in the hope that an improvement in the local population's situation would lead to a reduction in Naxalite violence (Ramana 2011). The National Rural Employment Guarantee Scheme is by far the most ambitious and largest anti-poverty program introduced by the Indian government, and some case studies suggest that the program may have indeed helped reduce violence in certain areas.[23]

This change in conflict intensity also seems to hold more generally: Maoists have been losing ground in a number of Indian states. They are now mostly non-existent in Andhra Pradesh and have lost influence in Bihar and even their stronghold states Jharkhand and Chhattisgarh. The Maoists seem to be forced to move out of many traditional areas of Maoist control and to retreat into the Dandakaranya forest area where their headquarters are assumed to be (Mukherji 2012).

Improved access to information seems to have played an important role in this development: The Indian Home Secretary Gopal K. Pillai said in 2010, for example, that the intelligence gathering system of the police has improved over the last couple of years, making police forces more successful at catching Maoists.[24] These developments are also recognized by the insurgents, who are accusing the government of turning the local population into police informers and of using surrendered Maoists as sources of information.[25]

---

[22]The Indian Express, December 20, 2007: Divide, uneven growth pose threat to our security: PM

[23]See oneworld.net (2011) for a case study of Balaghat in Madhya Pradesh.

[24]Summary of a lecture given by Gopal K. Pillai on March 10, 2010, which is available at: http://www.idsa.in/event/EPLS/Left-WingExtremisminIndia

[25]According to a press report from 2007, for example: *"The CPI-Maoist reportedly issued a press release at Chintapalli village in the Visakhapatnam District, blaming the Police for turning the Girijans (local tribals) into informers by spending huge amounts of money... (and) that surrendered*

2.3.2 NREGS

The National Rural Employment Guarantee Scheme (NREGS) is often referred to as the largest government anti-poverty program in the world.[26] The scheme provides an employment guarantee of 100 days of manual public-sector work per year at the minimum wage to all rural households. The legal right to this employment is laid down in the National Rural Employment Guarantee Act (NREGA) that was passed in the Indian Parliament in August 2005. Under the scheme, all households can apply for work at any time of the year as long as they live in rural areas and their members are prepared to do manual work at the minimum wage. Wages are to be paid within 15 days after the work was performed, otherwise the worker is eligible for an unemployment allowance.[27] While the minimum wage is state-specific, NREGA specifies a floor minimum wage which was Rs. 60 per day at the introduction of the program. It has been raised over time, and was Rs. 120 per day in 2009.

NREGS was rolled out non-randomly in accordance with a poverty ranking across the country in three phases: 200 districts received the scheme in February 2006 (Phase 1), whereas 130 districts started implementation in April 2007 (Phase 2). Since April 2008, the scheme operates in all rural districts in India (Ministry of Rural Development 2010).[28]

Many of the poorest Indian districts are also those heavily affected by Naxalite violence, as can be seen from Figure 2.1. The figure shows Red Corridor districts predicted to receive NREGS is Phase 1, Phase 2, and Phase 3 in black, dark grey, and light grey, respectively, and reveals that a large proportion of Maoist-affected

---

Maoists are helping the Police, were not leading a normal life and were always with the Police who provided them with all luxuries and used them in combing operations..."

[26] The program was renamed to Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS) in 2009, but the abbreviations NREGS and NREGA (for the Act on which it is based) have stuck in the academic literature on this topic.

[27] For more details on the scheme see e.g. Dey et al. (2006), Government of India (2009), and Ministry of Rural Development (2010).

[28] The scheme only excludes districts with a 100 percent urban population, and is active in 99 percent of Indian districts.

districts are poor enough to be assigned to the first implementation phase. One potential concern with development programs in these areas is that the presence of local governments is relatively weak and Naxalites may hinder or prevent the working of these schemes in their fight against the government. Naxalites are indeed known to have blocked a number of government development programs (Banerjee and Saha 2010). Interference from Naxalites does not seem to be a large-scale problem for the working of NREGS, however: In contrast to a number of other government schemes, chief secretaries from the seven states most heavily affected by Naxalite violence believe NREGS to work relatively well in their districts.[29] Case studies in Jharkhand, Chhattisgarh and Orissa also come to the conclusion that the Naxalites are not blocking NREGS projects, with the exception of road construction projects which Maoists claim are built for military counterinsurgency purposes (Banerjee and Saha 2010). In our manual coding of all incidents of Maoist violence used in this paper there was also no incident that targeted a NREGS worksite, providing further evidence against the hypothesis that Naxalites sabotage the program in important ways.

An emerging literature suggests that implementation issues may substantially limit the effectiveness of the program, with widespread rationing of NREGS employment especially in poorer states and corruption in the form of underpayed wages and ghost workers (Dutta et al. 2012, Niehaus and Sukhtankar 2013b). Drawing on field reports about the working of NREGS on the ground, the five states Andhra Pradesh, Chhattisgarh, Madhya Pradesh, Rajasthan and Tamil Nadu are often referred to as 'star states' because of the higher implementation quality of the program in those areas (Dreze and Khera 2009, Khera 2011).[30]

[29]The Times of India, April 14, 2010: Naxals backing NREGA?

[30]The proportion of households issued job cards (which are free cards required to be able to apply for NREGS employment) in the first year of implementation on average is also substantially higher in star states than in the rest of the country, suggesting a much higher level of awareness of NREGS even if households may not necessarily have received employment or known much about the detailed provisions of the scheme.

A number of papers have also focused on analyzing the impact of the employment guarantee scheme on rural labor markets in India. Using difference-in-difference approaches, empirical analyses often suggest low overall benefits but positive impacts on public employment and private-sector wages in the agricultural off-season, in areas with high implementation quality, and among casual workers (Azam 2012, Berg et al. 2012, Imbert and Papp 2013). Chapter 1 of this dissertation uses a regression-discontinuity framework and finds that NREGS is primarily used as a safety net rather than as an additional form of employment and does not lead to an overall increase in public-sector employment, the casual private-sector wage or household income. Taken together, the empirical literature on NREGS therefore suggests that while there may be important heterogeneous impacts, overall NREGS does not raise the opportunity cost of other occupations in the traditional sense of offering a better paid job, although the program may affect opportunity costs through occupational changes induced by the safety net (Chapter 1).

The available information on the implementation of NREGS also helps rule out that the competition for appropriable resources explains the program effects on violence, since NREGS creates hardly any appropriable assets in practice. A breakdown of project categories reveals that NREGS focuses on drought-proofing measures and does not generate a lot of infrastructure improvement or physical assets.[31]

Since we focus on a similar time interval as the existing literature in our empirical analysis, the impacts of the scheme on violence in this paper are therefore not driven by any substantial household income increases or by a fight for the control of appropriable resources. Given the documented implementation problems, it is also unlikely that a credible commitment to development can explain the program effects. Instead, the effects could be due to changes in the opportunity cost of being a Maoist

[31] According to Ministry of Rural Development (2010), for example, the breakdown of projects for the financial year 2008-09, was as follows: 46% water conservation, 20% provision of irrigation facility to land owned by lower-caste individuals, 18% land development, 15% rural connectivity (roads), 1% any other activity.

supporter or the promise of government benefits since households may expect future benefits from the program or see NREGS as a signal of the government's commitment to the fight against poverty. This last effect could change the people's perception of the government and their expectation of which side will eventually win the conflict. This may be particularly strong if the police become more effective at tracking down insurgents because of improved information and assistance from civilians.

In terms of the potential mechanisms discussed in the theory section, qualitative and quantitative evidence on the working of NREGS therefore points to the opportunity cost, citizen support and spotlight theories as the potential channels. Given the nature of the conflict and the importance attached to information-gathering and civilian assistance, the citizen-support channel may be especially important if NREGS improves the relationship between the government and the civilians. To test these different explanations empirically, we exploit the roll-out of the program in a regression-discontinuity design.

2.4 Identification Strategy, Data and Empirical Specification

2.4.1 NREGS Roll-out and the Assignment Algorithm

The Indian government used an algorithm to determine which districts would start implementing the program in which phase. Chapter 1 of this dissertation reconstructs the algorithm from available information on the NREGS roll-out and institutional knowledge about the implementation of development programs in India. The algorithm has two stages: First, the number of treatment districts that are allocated to a given state in a given phase is determined. It is proportional to the prevalence of poverty across states, which ensures inter-state fairness in program assignment.[32] Second, the specific treatment districts within a state are chosen based on a development ranking, with poor districts being chosen first.

We use this procedure in our empirical analysis. The 'prevalence of poverty' measure used in the first step of the reconstructed algorithm is the state headcount ratio times the rural state population, which provides an estimate of the number of below-the-poverty-line people living in a given state and shows how poverty levels compare across states. In the first step of the algorithm, a state is therefore assigned the percentage of treatment districts that is equal to the percentage of India's poor in that state. For the calculations, we use headcount ratios calculated from 1993-1994 National Sample Survey (NSS) data.[33]

The development index used to rank districts within states comes from a Planning Commission report from 2003 that created an index of 'backwardness', which is a term often used in India to refer to economic underdevelopment. The index was created from three outcomes for the 17 major states for which data was available: agricultural wages, agricultural productivity, and the proportion of low-caste individuals (Scheduled Castes and Scheduled Tribes) living in the district (Planning Commission 2003).[34] Data on these outcomes was unavailable for the remaining Indian states, and it is unclear whether a comparable algorithm using different outcome variables was used for them. We therefore restrict our empirical analysis to these 17 states.

[32]In practice this provision also ensures that all states (union territories are usually excluded from such programs) receive at least one treatment district.

[33]We use the rural state headcount ratios from Planning Commission (2009), since the original headcount ratio calculations do not have estimates for new states that had been created in the meantime. Since these are official Planning Commission estimates, they are likely to be closest to the information the Indian government would have had access to at the time of NREGS implementation. NSS is a nationally representative household survey dataset. The newest available information on headcount ratios at the time would have been the 1999-2000 NSS data, but that dataset was subject to data controversies and therefore not used.

[34]The purpose of the index was to identify especially underdeveloped districts for wage and self-employment programs and, as mentioned above, it was used in pre-NREGS district initiatives, although those programs were much less extensive than NREGS and usually envisioned as temporary programs.

Districts were ranked on their index values. In a deviation from the algorithm, 32 districts that were on a government list of being heavily affected by Maoist violence all received NREGS in Phase 1 regardless of their poverty rank. We drop these districts from the sample for the empirical analysis.

Because of the two-step procedure of the algorithm, the resulting cutoffs for treatment assignment in a given phase are state-specific. Since implementation proceeded in three phases, two cutoffs can be empirically identified: the cutoff between Phase 1 and Phase 2, and the cutoff between Phase 2 and Phase 3. These cutoffs correspond to the Phase 1 and Phase 2 NREGS roll-out, respectively. We exploit both cutoffs in a regression discontinuity framework.

Since treatment cutoffs differ by state, ranks are made phase- and state-specific for the empirical analysis and are normalized so that a district with a normalized state-specific rank of zero is the last district in a state to be eligible for receiving the program in a given phase.[35] This allows the easy pooling of data across states, since the treatment effect can then be measured at a common discontinuity in each phase. Negative numbers are assigned to districts with lower ranks than the cutoff rank, whereas positive numbers are assigned to the districts that are too developed to be eligible according to the district ranking and will function as control districts in our empirical analysis.

The overall prediction success rate of the assignment algorithm is 84 percent in Phase 1 and 82 percent in Phase 2. The prediction success rate is calculated as the percent of districts for which predicted and actual treatment status coincide.[36].

[35]Rank data in the 17 major Indian states is complete for all districts classified as rural by the Planning Commission in their report, so there is no endogeneity in the availability of data in these states. Urban districts in the Planning Commission report are districts that either include the state capital or that have an urban agglomeration of more than one million people. Rank data is available for 447 of 618 districts in India. Data for the index creation was unavailable in some states, in most cases because of internal stability and security issues during the early 1990s when most of the data was collected. We exclude these states from the analysis.

[36]Prediction success rates for Phase 2 are calculated after dropping Phase 1 districts from the analysis.

This means that there is some slippage in treatment assignment in both phases, and there is considerable heterogeneity in the performance of the algorithm across states. Nevertheless, the algorithm performs quite well in almost all states and the prediction success rates are also considerably higher than the ones that would be expected from a random assignment of districts, which are 40.27 percent for Phase 1 and 37.45 percent for Phase 2 at the national level, respectively. Overall, this suggests that the proposed algorithm works well for predicting Phase 1 and Phase 2 district allocations.

To achieve internal validity, the RD framework crucially relies on the idea that beneficiaries were unable to perfectly manipulate their treatment status, so that observations close to the treatment cutoff value are plausibly similar on unobservables and differ only with respect to their treatment status (Lee 2008). In the case of the two-step RD, this means that districts should not have been able to manipulate their predicted status under the algorithm in either step.

This seems plausible: As mentioned above, the headcount poverty ratio used to calculate the number of treatment districts for a state in the first step of the algorithm used data from the mid-1990s, which had long been available by the time the NREGS assignment was made.[37]

Like the information used for the first step, it also seems unlikely that it was possible to tamper with the data used in the second step of the algorithm: The 'backwardness' index was constructed from outcome variables collected in the early to mid-1990s, eliminating the opportunity for districts to strategically misreport information. Additionally, the suggestion of the original Planning Commission report had been to target the 150 least developed districts, but NREGS cutoffs were higher than this even in Phase 1 (200 districts received it in Phase 1). Therefore, districts would

[37]The algorithm also uses state rural population numbers from the 2001 Census to transform headcount ratios into absolute numbers, but those figures were also long publicly available at the time. The RD may be potentially fuzzier than it really is because of some potential for measurement error introduced into the algorithm at this step since the exact numbers the government used in this step are not known, but this should not introduce systematic bias into the empirical analysis.

have had an incentive to be among the 150 poorest districts but not to be among the 200 or 330 least developed districts for Phase 1 and Phase 2, respectively. Lastly, the Planning Commission report lists the raw data as well as the exact method by which the development index was created, again eliminating room for districts to manipulate their rank. Overall, it therefore seems like manipulation of the rank variable is not a major concern.[38]

Figures 3.1 and 3.2 look more closely at the distribution of index values over state-specific ranks. They plot the relationship between the Planning Commission's index and the normalized state-specific ranks for the Phase 1 and Phase 2 cutoffs, respectively. For most states, the poverty index values seem smooth at the cutoff of 0, again suggesting that manipulation of the underlying poverty index variable is not a big concern.

Another way of analyzing whether manipulation is likely to be a problem is to test whether there are any discontinuities at the cutoffs in the baseline data. If districts close to the cutoff are really similar to each other, so that outcome differences are just due to the different treatment status, we should not find impacts in the baseline data. Table 2.2 presents the results of such an analysis for the main outcome variables used in this paper (the number of violence-affected individuals, of fatalities, major incidents and total incidents) for the time period before NREGS was rolled out to any phase.[39] Overall, Table 2.2 suggests again that manipulation is not an important problem.[40]

Finally, we need to verify that there really is a discontinuity in the probability of receiving NREGS at the state-specific cutoff values. Figures 3.3 and 3.4 show

[38]This does not mean that actual treatment assignment was not subject to political pressures since compliance with the algorithm is often below 100 percent. It can be shown that deviations from the algorithm are correlated with party affiliation.

[39]Since Phase 2 received the program later than Phase 1, the pre-treatment phase for Phase 2 in theory lasts an additional year. As we later show explicitly, however, it looks like Phase 2 was affected by spillover effects from Phase 1 districts before it actually received NREGS.

[40]Our main results also include the baseline outcome variable as a regressor, which controls for any baseline differences and should soak up some of the residual variance.

the probability of receiving NREGS in a given phase for each bin, as well as fitted quadratic regression curves and corresponding 95 percent confidence intervals on either side of the cutoff. The graphs demonstrate that the average probability of receiving NREGS jumps down at the discontinuity, although this discontinuity is much stronger in Phase 2 than in Phase 1. This suggests that there is indeed a discontinuity in the probability of being treated with the employment guarantee scheme at the cutoff.

2.4.2 Data and Variable Creation

The primary source of data used in this paper comes from the South Asian Terrorism Portal (SATP). This is a website managed by a registered non-profit, non-governmental organization called the Institute of Conflict Management in New Delhi. The Institute provides consultancy services to governments and does extensive research on insurgency-related activities. The SATP aggregates and summarizes news reports on Naxalite-related incidents, and such summaries usually contain the location of the incident (district), the date of the incident, the number of casualties (Naxalites, civilians, or police), and the number of injuries, abductions or surrenders. The source also codes the incident as 'minor' or 'major.'

In many cases, the party initiating an incident can be identified from the newspaper description, and we manually code up these details for the incidents in our sample. Events are labelled as police-initiated, Maoist initiated against the government or Maoist incidents against civililans.

An example of a police-initiated attack in the newspaper reports is this incident: *"Chhattisgarh, 2006: July 7 Central Reserve Police Force personnel raided a CPI-Maoist hideout under Basaguda police station in the Dantewada district and shot dead seven Maoists."* A typical Maoist attack aimed at government troops reads like this: *"West Bengal, 2006: February 26 Cadres of the CPI-Maoist detonate a* *landmine blowing up a police vehicle that killed four persons, including two security force (SF) personnel, at Naakrachhara in the West Midnapore district."* Incidents like the following one are labelled as a Maoist-initiated attack against civilians: *"Andhra Pradesh, 2007: November 16 Three persons, including two migrant tribals from the Bastar region of Chhattisgarh, were killed by cadres of the CPI-Maoist in Narsingpet village of Chintoor mandal in Khammam district. Before the killing, the Maoists reportedly grilled them in the presence of the villagers by organising a panchayati (village level meeting) and branded them as police informants."*

Using this information we construct violence-intensity variables at the district-month level, with 'no incidents' being coded as 0. If some information is unclear, we verify the information by searching for the source news reports. We use data between January 2005 (the earliest time for which data is available on the website) and March 2008, since the districts in the final phase started receiving NREGS in April 2008. This gives us data before and after implementation of the program, with about two years of post-treatment data for Phase 1 districts and a year's worth of after-NREGS data for Phase 2 districts. This dataset is then merged with information on the poverty rank from the 2003 Planning Commission Report.

As is common with this kind of dataset, there are certain limitations to using it: The number of Naxalites killed or injured is difficult to verify and may be incomplete, and security forces may have an incentive to overstate their accomplishments by inflating the numbers. This concern is mitigated to some degree by the fact that police are required to disclose names and ranks of the Maoists killed to validate their reports and that we make use of incidents reported in the media rather than administrative data.

The potential data quality limitations introduce measurement error into the analysis, but will not systematically bias our regression discontinuity (RD) results unless reporting standards on Naxalite violence are systematically correlated with the predicted treatment status of NREGS according to the government algorithm. Such a correlation may occur if the spotlight theory holds and the police face external pressures to perform better in treatment areas than in control districts. In the empirical results section, we provide evidence against the spotlight theory by analyzing which parties initiated incidents and by looking at other types of violence and crime.

Table 2.1 shows some summary statistics for our primary variables of interest. Our dataset records 1458 incidents, covering a total of 2030 fatalities. 267 of these incidents were coded by the SATP source as 'major'. Furthermore, in this 39-month period, 2545 people were either injured, abducted or surrendered to the police. On average, in any given red-corridor district, there are about 0.44 deaths a month related to Naxalite activities and about 0.32 incidents a month.

We also collect data on the police force from the Indian Ministry of Home Affairs, which contains state-level information on the number of police officers, police posts and stations, as well as some other measures of police strength. District-level data on other types of crimes also come from the Ministry of Home Affairs, and monthly wage data for agricultural laborers comes from the Agricultural Wages in India (AWI) datasets.

2.4.3 Empirical Specification

Since NREGS was rolled out based on an algorithm that assigned state-specific ranks to districts, these ranks can be used as a running variable in an RD framework. In an ideal case, we would restrict the data to observations in the close neighborhood of the cutoff and estimate the treatment effect at the cutoff using local linear regressions. As the number of observations near the cutoff is limited, however, we are also using all available observations that are relevant for a given cutoff. Such an increase of the bandwidth will increase the precision of the estimates because of a larger sample size, but potentially introduces bias since observations far away from the cutoff can influence the treatment effect estimated at the cutoff (Lee and Lemieux 2010).

We address this concern in three ways: First, all result tables show the estimated coefficients for three different parametric specifications (linear, linear with slope of regression line allowed to differ on both sides of the cutoff, quadratic). F-tests reject the null hypothesis that higher-order polynomials add important flexibility to the model, and the quadratic flexible specification is always outperformed statistically by the linear flexible specification.[41] Second, while our results use all districts of the treatment and control phase in a given specification, we test the robustness of our main estimates by restricting the sample to observations closer to the cutoff. Third, Figures 2.6 to 2.9 show the non-parametric relationships between the main outcome variables of interest and also plot quadratic polynomial regression curves. Similar to the summary statistics, they show that insurgency-related violence intensity is low in many districts. We therefore also test the robustness of our results to the use of a Count Data models, like the zero-inflated Poisson model.[42]

Since the algorithm only generates a fuzzy RD, we use a two-stage least squares specification where actual NREGS receipt in a given phase is instrumented with predicted NREGS treatment according to our algorithm, although intent-to-treat effects of the main results are reported in the appendix. To increase the precision of our estimates, we control for the baseline outcome variable.[43] To ensure that the RD results are not affected by observations far away from the cutoff, we run results

[41]More flexible models also tend to be unstable in the second stage of the two-stage least squares estimation procedure, although the estimated coefficients are often qualitatively similar to the quadratic results.

[42]Our results are also robust to the choice of the running variable. Our analysis uses the poverty rank as the running variable, whereas an alternative would be to use the underlying poverty index values. We use the rank since it is the variable that treatment is based on since the first step of the algorithm specifies the size of the treatment group. This implies that what determines a district's distance from the cutoff is its rank rather than its poverty index value, since in many situations a district could have a very different index value without altering its treatment likelihood. Additionally, the conditional mean function of the outcomes of interest is flatter when using the rank rather than the index, which means that a large bandwidth is less problematic when using the rank variable. Our results are robust to using the index value as a running variable instead, however, even though they tend to be less precise.

[43]Controlling for month or year fixed effects leaves the results unaltered.

separately by cutoff, and drop the observations that should not affect the treatment effect at a given cutoff: For Phase 1 district assignment, we only include Phase 1 and Phase 2 districts in the analysis, and drop Phase 3 districts. Similarly, for Phase 2 cutoff regressions, we drop Phase 1 observations and only consider the remaining districts.

The equation below shows the regression equation for one of the specifications we run, which is linear in the running variable but does not constrain the coefficients to be the same on either side of the cutoff:

$$y_{ij} = \beta_0 + \beta_1 rank_i + \beta_2 nregs_i + \beta_3 nregs * rank_i + \beta_4 baseliney_{ij} + \epsilon_{ij}$$

where the subscripts refer to district i in month j since one observation is a district-month. y is an outcome variable of interest, *rank* is a district's rank based on the state-specific normalized index. The coefficient of interest is β_2, and *nregs* is actual NREGS receipt in a given phase, which is instrumented with predicted NREGS receipt. Standard errors are clustered at the district level.

2.5 Results

2.5.1 Main Results

The main results are presented in Tables 2.3 to 2.5. Table 2.3 shows the impact of NREGS on Maoist incidents for the four main outcome variables: individuals affected (deaths/ injuries/ abductions); deaths; major incidents; and total incidents. Panel B normalizes the variables by the 2001 Census population counts, showing the results per-10 million people. Each panel has three different specifications: linear, linear with a flexible slope, and quadratic. As mentioned above, specifications in this table control for (estimated) police force changes.[44]

[44] Not controlling for police force changes does not change the results substantially. These results are presented in Panel B of Table E.0.4.

Panel A of Table 2.3 shows that violence as measured by all outcome variables increases in Phase 1 districts after NREGS is introduced. Depending on the specification, there is a rise of about 0.55 to 0.75 deaths per month in a given district. At a mean of about 0.44 deaths per month in a Red Corridor district, this amounts to about about a 125% increase from the baseline level. Similarly the number of affected persons (killed, injured, abducted) increase by about 0.56 to 0.73 units per district-month. Since the mean for this variable is 0.99 in Red Corridor districts, this amounts to a rise of 56%. The number of total incidents rises by about 0.22 to 0.27 incidents per month, which is about a 70% increase from the baseline mean. These results are robust across the different parametric specifications. A crude calculation suggests that these effects translate into between 785 and 1071 more fatalities in about 314 to 385 more incidents in the year after implementation. The overall increase in violence matches the first prediction of the citizen-support model and would also be consistent with the spotlight theory, but is not in line with the opportunity cost channel being the dominant mechanism since raising the opportunity cost of becoming a rebel should lead to a fall in violence. Panel B reveals that there are similar impacts once we normalize the variables by Census enumerated population counts.

Figures 2.6 to 2.9 use a quadratic specification to plot the primary variables against the rank variable, and show a sharp discontinuity at the cutoff for all the variables of interest. We can also plot the RD coefficients for each month, to see when the violence starts increasing. Figure 2.10 plots the RD coefficients month by month for the number of incidents, where the first vertical line depicts the time when the Act was passed, and the second vertical line is when Phase 1 was implemented. The figure shows that, across the different specifications, the increase in violence is almost immediate when NREGS is introduced. Similarly Figure 2.11, which plots the monthly RD coefficients for the number of persons affected, shows that while there is an immediate increase in violence, there is also a slight dissipation of effects over time.

Since the data allows us to distinguish between civilians, Maoists and the police force, we can study the impact of NREGS on each of these groups in terms of fatalities, injuries and abductions. In many cases, we can also code up which conflict party initiated the attack. All of this information allows us to test predictions of the theoretical model: According to the citizen-support channel, the police should now have better information to catch the Naxalites, and the insurgents, in turn, may want to retaliate against civilians for helping the police. Thus, we should see an increased number of police-initiated attacks against insurgents and more attacks by Maoists on civilians. This also implies that the bulk of the impact should be concentrated on Naxals and civilians. Table 2.4 reports the empirical results of this analysis, focusing again on Phase 1 districts.

Panel A of Table 2.4 depicts the results for who initiates these attacks and shows that an important part of the increase in violence comes from police-initiated attacks on the Maoists. This is consistent across specifications, and shows a sharp increase in police-initiated attacks in regions that received NREGS. The results also show the Maoists retaliating against civilians, but not a very large increase in Maoist-on-police attacks. The retaliation against civilians is consistent with the citizen-support story since the Maoists are punishing civilians for becoming police informers, but there is no reason to expect retaliation against civilians under the spotlight and competition for resources theories. Additionally, the competition for resources story implies a large increase in the number of Maoist attacks on the police, which is not consistent with the empirical patterns we find.

Panel B of Table 2.4 presents the RD results for fatalities classified by each of these groups. Civilian and police casualty estimates are small and imprecisely estimated, whereas Naxal casualties increase by between 0.3 and 0.4 deaths a month after the introduction of the NREGS, an effect that is also statistically significant at the 5%

level. Appendix Table E.0.1 presents the per-capita results, which again show that the Maoist deaths contribute to most of the new casualties. The police force does not experience a statistically significant increase in fatalities, and the magnitudes are also much smaller. Overall, these results are again consistent with the citizen-support predictions.

Since we have monthly data, it is possible to look at dynamic patterns. Focusing on the Phase 1 implementation group, Table 2.5 divides the post-treatment period into the short run (Panel A) and the medium run (Panel B). The short run is defined as the first 7 months after NREGS eligibility, whereas the medium run is months 8 to 14 months after eligibility. The results show that an important part of the impact occurs in the short run. The impact on the number of affected persons is somewhere between 1.6 and 2.2 times higher, and fatalities are 1.4 to 1.6 times higher in the short run than in the medium run. Again, these empirical patterns are consistent with the citizen-support channel and the spotlight theory, but not with the opportunity cost story, although rising opportunity costs over time once NREGS is implemented better could contribute to the downward trend in violence. These dynamics are also inconsistent with the resource-curse story: if resources are introduced, then they would increase with time, and violence should go up rather than fall.

2.5.2 Extensions

Another test of the plausibility of the spotlight theory is presented in Appendix Table E.0.2, which focuses on the impact of NREGS on other types of violence and crime. If police officers feel an increased pressure to perform better in treatment areas because of increased attention paid to NREGS districts, then we may expect that increased police efforts should apply to other forms of violence and crime as well. Appendix Table E.0.2 provides no evidence of NREGS having a statistically significant impact on crime, however, and the magnitudes tend to be small.

If the citizen-support channel is relevant, we should expect it to be especially important in areas where program awareness and implementation quality are higher. Therefore, the number of police-initiated attacks should be higher in these areas than in the rest of the country. One measure of implementation quality often used in the existing NREGS literature are the so-called 'star states' where, based on field reports, awareness of the program and implementation quality tend to be much higher than in the rest of the country (Dreze and Khera 2009, Khera 2011). In Table 2.6, the NREGS treatment variable is interacted with an indicator variable equal to one if a state is a 'star state', and zero otherwise. As the table shows, police-initiated attacks are indeed higher in star states NREGS districts than in other treatment districts.

If what makes civilians willing to assist the police is mainly the promise of development rather than the actual benefits received from the program, then we may find that civilians change their behavior even in still untreated districts. This effect may occur especially in Phase 2 districts once Phase 1 districts have started implementation, since the people in those districts can take Phase 1 implementation as a signal of the government's commitment to following through with the program and may be aware of the fact that their districts will receive the treatment soon. We would then expect to find positive spillover effects of the program onto Phase 2 districts.

Table 2.7 shows the results of this analysis, and confirms that this effect does indeed hold empirically. At the time that Phase 1 districts have access to NREGS (and other phases do not) there is an increase in violence in Phase 2 districts (Panel A). However, this increase dissipates over time, and by the time Phase 2 is in the spotlight, there is no longer any impact (Panel B). The Phase 2 results are therefore difficult to explain with the spotlight theory. If the police work harder in treatment areas due to increased attention on law and order in NREGS areas, there is no reason for the police in still untreated areas to increase their effort levels.

Overall, the empirical patterns presented in the results section suggest that violence goes up when NREGS is introduced, and does so already in the very short run. A large proportion of the increase is due to police-initiated attacks on Maoists. There is also some retaliation by the Maoists on civilians, but most of the increase in fatalities is explained by rebels dying. All of these empirical patterns are consistent with the theoretical predictions of our model in which civilians are willing to share information and other forms of support with the police after NREGS implementation starts, which allows government troops to crack down more efficiently on the Maoists. They are difficult to explain with alternative explanations like the opportunity cost channel or the spotlight theory, however.

2.5.3 Robustness Checks

A number of checks can be performed to test the robustness of the main results. We focus on our main results, which look at the impact on Phase 1 districts in these additional specifications.

One important concern is that there may be measurement error in the rank variable that is used as the running variable, which may lead to districts right at the cutoff being assigned to the wrong side of the cutoff. We provide a robustness check by using a donut-hole approach that drops the districts with state-level ranks lying between -1 and 1 (the cutoff is at a state-specific rank of 0). These results are presented in Panel A of appendix table E.0.3. They are similar, both in magnitude and statistical significance, to our main results, implying that the estimated treatment effects do not seem to be driven by measurement error of the observations close to the cutoff. Panel B of Table E.0.3 presents the main results when varying the bandwidth by restricting the analysis to observations closer to the cutoff, and once again produces similar results.

Our main results are also robust to a number of other specifications presented in the appendix: Panel A of Table E.0.4 estimates the intent-to-treat (ITT) version of the main results, while Panel B of Table E.0.4 reproduces the main results without controlling for the strength of the police force. Both specifications consistently maintain the main results.[45]

Another potential concern with the main specifications is the nature of the data. All outcomes are count-data outcomes, but we estimate the treatment effects within a normal regression framework rather than using count-data models. Panel A of Appendix Table E.0.5 therefore presents the the results from a Zero-Inflated Poisson Count-Data Model. The Poisson model is the most widely used count-data model (Cameron and Trivedi 2013). Since the data has an excess of zero-values (i.e. no casualties in a given district-month), we use the zero-inflated version of this model. The coefficients are interpreted as the change in the log-counts of the dependent variable on introduction of NREGS, and again show the same qualitative patterns as our main results. The results are also robust to using other count data models like the hurdle model using a Logit-Poisson specification.

Panel B of Table E.0.5 presents the results using a difference-in-difference (DID) approach rather than the RD, which is the most common empirical identification strategy used to study the impacts of NREGS in the literature. We conduct two different types of DID exercises: the Intent-to-Treat (ITT) version, where treatment is assigned based on who should have received NREGS according to the algorithm; and the Actual Treatment version where treatment depends on actually receiving NREGS. While the DID approach estimates the overall average treatment effect on the treated and therefore a different parameter than the RD specifications, the results are again qualitatively similar.

[45] One possible simultaneous change with NREGS implementation could be an increase in the size of the police force. Since we do not have data on the actual police force in a district, we estimate it using state-level information, where any change in the police force for a given state is assumed to be attributable to NREGS districts. In reality, these state-level estimates most likely overemphasize the change in the police force and may therefore provide us with conservative estimates of the impact of NREGS. Our main analysis therefore includes police force controls, although, as Panel B of Table E.0.4 shows, the results are very similar without these controls.

Lastly, we conduct other robustness checks not reported here by re-doing our main results after controlling for rainfall shocks in the current month and in the entire preceding year. We also control for average monthly wages, and find that our results are robust to all these specifications. To the extent that rainfall shocks and wages capture what happens to income in these regions after NREGS is introduced, these results indicate that the opportunity-cost channel is not the driving force behind these results. In other specifications, we also control for the timing of the state elections, in some specifications using not just the election month but also up to 5-months leading up to an election, and our results are unaffected by these controls.

2.6 Conclusion

This paper has analyzed the impact of introducing a large public-works program in India, the National Rural Employment Guarantee Scheme (NREGS), on incidents of left-wing violence by the Naxalite movement. We exploit the fact that the program was phased in over time according to an algorithm that prioritized economically underdeveloped districts in a regression discontinuity (RD) design. The results are robust across a number of different specifications and show a substantial rise in violence in the districts that received NREGS, especially in the short run. Insurgents are the primary affected group as police-initiated attacks rise, with little impact on police force casualties. There is also some evidence for an increase in Maoist-initiated attacks against civilians. The impact is largest among districts that received NREGS in the first phase of the roll-out, but there are positive spillovers of violence to the districts that are next in line to receive the program.

These empirical patterns as well as other available qualitative and quantitative evidence on the conflict are consistent with a model in which the government program makes civilians more willing to support the police because it improves the relationship between the government and the people. In contrast, the results are difficult to explain with a number of alternative theories. Securing the assistance of the local population may therefore be an important factor in internal conflicts and may help answer Max Weber's (1919) question on whether the State should use force *or* development to tackle internal conflict. For the Indian government, which has been trying to fight the Naxalites for over 30 years, as well as for other governments in developing countries the best strategy may be to combine both force *and* development.

It is unclear, however, how successful such strategies are likely to be in the longer run. In the Indian context, a growing body of literature questions the effectiveness of NREGS as a tool for *actual* development due to various implementation problems, although the program still seems to provide some benefits through its safety net function. This implies that at least a part of what may win over the local community initially are anticipated welfare benefits and the *promise* of development rather than substantial actual changes. Such a mere promise of development may not be a credible enough tool to ensure the aid received from the civilian population over a longer period of time, however. Once civilians realize that the program is not delivering on its promises, this may not only stop civilian aid in exchange for the benefits of the program, but may lead to distrust in government programs in general. Therefore, an important component to winning the continued support by the people may be to ensure that government anti-poverty programs are implemented effectively and actually fulfill the promise of development.

.

Table 2.1: Summary Statistics

	Mean	**Mean**	**Total**
	Red Corridor Dist.	**All Dist.**	**All Dist.**
Deaths	0.441	0.116	2030
Inj./Abd./Cap.	0.553	0.146	2545
Affected	0.994	0.262	4575
Major Incidents	0.058	0.015	267
Total Incidents	0.317	0.084	1458
Maoists Killed	0.162	0.043	744
Civilians Killed	0.166	0.044	763
Police Killed	0.114	0.030	523

A unit of observation is a district in a given month and year (i.e. district-month-year). There are a total of 39 months from January 2005 till March 2008. "Red Corridor" districts are districts with Maoist-related incidents. "Inj./Abd./Cap." gives the number of injured, abducted or captured people. "Affected Persons" indicates number of persons killed, injured, abducted or captured. "Major Incidents" indicates number of Major Incidents as coded by the SATP website. "Total Incidents" is number of total Maoist-related incidents.

.

Table 2.2: Baseline Pre-Treatment Results

	Panel A: Non-Normalized			
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.517	0.0666	0.0142	-0.0232
	(0.782)	(0.317)	(0.0319)	(0.110)
R-squared	0.005	0.004	0.008	0.003
Linear Flexible Slope	0.0773	-0.147	-0.00850	-0.0899
	(0.534)	(0.277)	(0.0271)	(0.120)
R-squared	0.004	0	0.001	0
Quadratic	0.198	-0.162	-0.0142	-0.107
	(0.758)	(0.349)	(0.0345)	(0.138)
R-squared	0.006	0	0.003	0
Outcome Mean	0.580	0.263	0.035	0.170

	Panel B: Per Capita			
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	10.34	3.409	0.337	0.821
	(10.39)	(3.673)	(0.375)	(0.969)
R-squared	0.002	0.003	0.005	0.008
Linear Flexible Slope	5.603	1.566	0.157	0.320
	(6.097)	(2.210)	(0.226)	(0.644)
R-squared	0.004	0.005	0.005	0.006
Quadratic	8.537	2.304	0.195	0.457
	(9.529)	(3.384)	(0.347)	(0.926)
R-squared	0.003	0.005	0.006	0.009
Outcome Mean	6.577	3.012	0.360	1.748

Regression looks at impacts for the 13 months pre-treatment.
Panel A contains direct impacts (not normalizing for population). Panel B shows the impact per-10 million people (based on population counts from the 2001 Census).
Controls include baseline averages of each dependent variable and police force changes. Unit of observation is district-month. Regressions contain 2964 observations in 228 clusters.
"Affected Persons" indicates number of persons killed, injured, abducted or captured. "Fatalities" indicates total number of deaths. "Major Incidents" indicates number of 'Major Incidents' as coded by the SATP website. "Total Incidents" is number of total Maoist-related incidents.

.

Table 2.3: Main Results

	Panel A: Non-Normalized			
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.636**	0.594**	0.104***	0.272**
	(0.294)	(0.275)	(0.0402)	(0.112)
R-squared	0.487	0.438	0.380	0.457
Linear Flexible Slope	0.561*	0.548*	0.0843**	0.228**
	(0.300)	(0.301)	(0.0376)	(0.105)
R-squared	0.487	0.439	0.386	0.466
Quadratic	0.723**	0.746**	0.124***	0.274**
	(0.351)	(0.352)	(0.0475)	(0.128)
R-squared	0.487	0.434	0.375	0.456
Outcome Mean	0.580	0.263	0.035	0.170

	Panel B: Per Capita			
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	2.393**	1.852*	0.537**	1.183**
	(1.135)	(1.044)	(0.240)	(0.539)
R-squared	0.526	0.522	0.562	0.683
Linear Flexible Slope	1.906*	1.536	0.391**	0.959*
	(1.047)	(1.113)	(0.187)	(0.541)
R-squared	0.526	0.523	0.563	0.685
Quadratic	2.277*	2.387*	0.602**	1.060*
	(1.209)	(1.275)	(0.246)	(0.607)
R-squared	0.526	0.522	0.562	0.684
Outcome Mean	6.577	3.012	0.360	1.748

Panel A contains direct impacts (not normalizing for population). Panel B shows the impact per-10 million people (based on population counts from the 2001 Census).
Controls include baseline averages of each dependent variable and police force changes. Unit of observation is district-month. Regressions contain 3192 observations in 228 clusters.
"Affected Persons" indicates number of persons killed, injured, abducted or captured. "Fatalities" indicates total number of deaths. "Major Incidents" indicates number of 'Major Incidents' as coded by the SATP website. "Total Incidents" is number of total Maoist-related incidents.

.

Table 2.4: Who Initiates the Attacks and Who is Affected

	Panel A: Who Intiates		
Specification	**Police on Maoist**	**Maoist on Police**	**Maoist on Civilians**
Linear	0.110*	0.0250**	0.0945*
	(0.0610)	(0.0121)	(0.0496)
R-squared	0.133	0.180	0.350
Linear Flexible Slope	0.0889*	0.0218**	0.0626
	(0.0470)	(0.0111)	(0.0394)
R-squared	0.140	0.181	0.357
Quadratic	0.100*	0.0252*	0.0898*
	(0.0607)	(0.0135)	(0.0528)
R-squared	0.136	0.180	0.351
Outcome Mean	0.057	0.028	0.071

	Panel B: Who is Killed		
Specification	**Civilians Killed**	**Police Killed**	**Maoists Killed**
Linear	0.146	0.0456	0.356**
	(0.119)	(0.0677)	(0.176)
R-squared	0.337	0.176	0.204
Linear Flexible Slope	0.132	0.0339	0.306**
	(0.120)	(0.0508)	(0.152)
R-squared	0.337	0.176	0.211
Quadratic	0.168	0.0980	0.394**
	(0.140)	(0.0749)	(0.200)
R-squared	0.337	0.175	0.200
Outcome Mean	0.095	0.079	0.089

Panel A has results for "who initiates the attacks and against whom." Panel B has results for "who is killed"
Controls include baseline averages of each dependent variable and police force changes. Unit of observation is district-month. Regressions contain 3192 observations in 228 clusters.

.

Table 2.5: The Short Run and The Medium Run

	Panel A:	**Short**	**Run**	
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.778**	0.641**	0.123**	0.273*
	(0.383)	(0.322)	(0.0552)	(0.142)
R-squared	0.655	0.599	0.510	0.577
Linear Flexible Slope	0.760*	0.671*	0.113*	0.251*
	(0.420)	(0.364)	(0.0585)	(0.139)
R-squared	0.655	0.598	0.513	0.581
Quadratic	0.967**	0.855**	0.162**	0.311*
	(0.488)	(0.425)	(0.0712)	(0.170)
R-squared	0.654	0.595	0.501	0.571
	Panel B:	**Medium**	**Run**	
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.484*	0.458**	0.0855**	0.273***
	(0.273)	(0.222)	(0.0371)	(0.106)
R-squared	0.359	0.360	0.273	0.370
Linear Flexible Slope	0.396*	0.397*	0.0639**	0.224**
	(0.233)	(0.223)	(0.0285)	(0.0955)
R-squared	0.359	0.361	0.279	0.381
Quadratic	0.445	0.520**	0.0862**	0.240**
	(0.281)	(0.259)	(0.0361)	(0.108)
R-squared	0.359	0.358	0.273	0.376
Outcome Mean	0.580	0.263	0.035	0.170

Panel A shows Short Run impacts (months 1 through 7 after implementations). Panel B shows the Medium Run (months 8 through 14 after implementation).
Controls include baseline averages of each dependent variable and police force changes. Unit of observation is district-month. Regressions contain 1596 observations in 228 clusters.
"Affected Persons" indicates number of persons killed, injured, abducted or captured. "Fatalities" indicates total number of deaths. "Major Incidents" indicates number of 'Major Incidents' as coded by the SATP website. "Total Incidents" is number of total Maoist-related incidents.

.
Table 2.6: Who Initiates the Attacks: Star States vs Non-Star States

Specification	**Police on Maoist**	**Maoist on Police**	**Maoist on Civilians**
Linear	0.0609	0.0294**	0.0735*
	(0.0380)	(0.0141)	(0.0390)
NREG*Star	0.116*	-0.0130	0.0593
	(0.0685)	(0.00899)	(0.0560)
Star States	-0.0136	0.00688*	-0.0305
	(0.0170)	(0.00380)	(0.0269)
R-squared	0.136	0.181	0.351
Linear Flexible Slope	0.0257	0.0228*	0.0244
	(0.0217)	(0.0127)	(0.0200)
NREG*Star	0.363**	0.0312	0.190
	(0.174)	(0.0328)	(0.126)
Star States	-0.0222	0.00597	-0.0372
	(0.0193)	(0.00376)	(0.0279)
R-squared	0.153	0.185	0.362
Quadriatic	0.0592	0.0298**	0.0681*
	(0.0398)	(0.0151)	(0.0401)
NREG*Star	0.116*	-0.0130	0.0602
	(0.0675)	(0.00870)	(0.0548)
Star States	-0.0139	0.00695	-0.0314
	(0.0166)	(0.00429)	(0.0260)
R-squared	0.137	0.181	0.352
Outcome Mean	0.057	0.028	0.071

Star States include Andhra Pradesh, Chhattisgarh, Tamil Nadu, Rajasthan and Madhya Pradesh
Controls include baseline averages of each dependent variable and police force changes. Unit of observation is district-month. Regressions contain 3192 observations in 228 clusters.

.

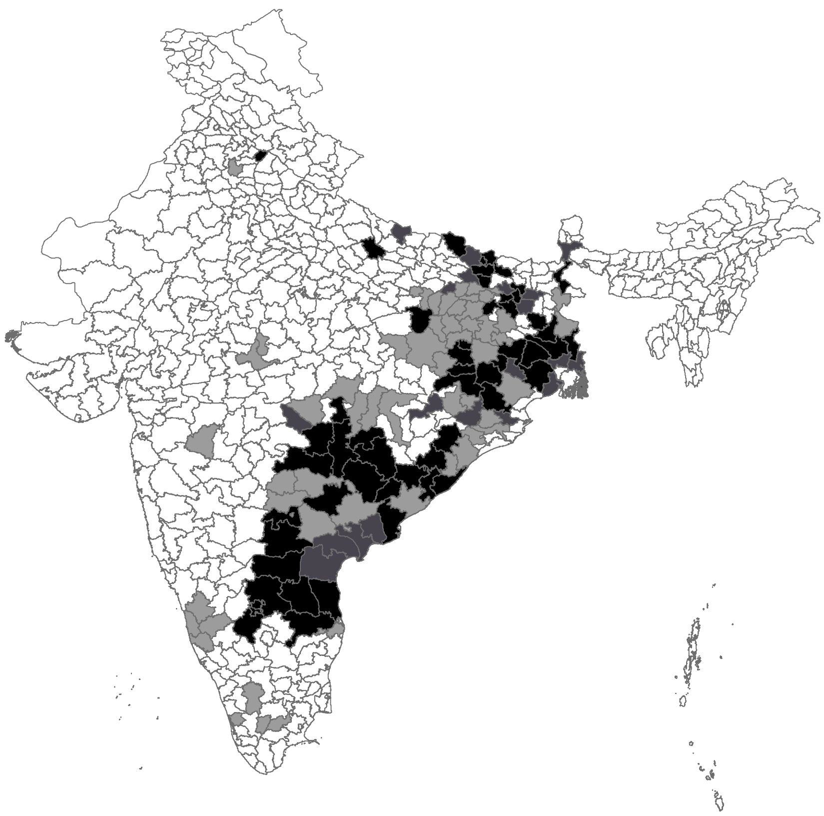
Table 2.7: Phase 2 - While Phase 1 is treated, and then Phase 2 treatment

	Panel A:	**During**	**Phase 1**	**Treatment**
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.131**	0.0825*	0.00345	0.0348***
	(0.0537)	(0.0435)	(0.00434)	(0.0122)
R-squared	0.137	0.141	0.206	0.309
Linear Flexible Slope	0.208*	0.129*	0.0109	0.0434*
	(0.113)	(0.0739)	(0.00910)	(0.0236)
R-squared	0.090	0.103	0.191	0.296
Quadratic	0.183**	0.114*	0.00992	0.0425**
	(0.0925)	(0.0625)	(0.00717)	(0.0199)
R-squared	0.128	0.134	0.203	0.305

	Panel B:	**Phase 2**	**Treatment**	**Period**
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	-0.161	-0.180	-0.00620	0.00399
	(0.194)	(0.168)	(0.00906)	(0.0307)
R-squared	0.055	0.038	0.044	0.184
Linear Flexible Slope	-0.0331	-0.137	-0.00650	0.0327
	(0.232)	(0.173)	(0.00908)	(0.0462)
R-squared	0.055	0.040	0.043	0.178
Quadratic	-0.0702	-0.134	-0.00562	0.0237
	(0.200)	(0.157)	(0.00809)	(0.0377)
R-squared	0.058	0.040	0.044	0.185
Outcome Mean	0.580	0.263	0.035	0.170

Panel A contains impacts on Phase 2 districts during February 2006 and March 2007. During this period, Phase 1 received NREGS, and Phase did not. The regressions contain 3178 observations in 227 clusters.
Panel B shows the impact on Phase 2 districts during April 2007 and March 2008. During this period, Phase 2 also received NREGS. Regressions contain 2497 observations in 227 clusters.
Controls include baseline averages of each dependent variable and police force changes. Unit of observation is district-month.
“Affected Persons” indicates number of persons killed, injured, abducted or captured. “Fatalities” indicates total number of deaths. “Major Incidents” indicates number of ‘Major Incidents’ as coded by the SATP website. “Total Incidents” is number of total Maoist-related incidents.

Figure 2.1: Red Corridor Districts and NREGS Phase



Note: Red corridor districts are all districts that had at least one Naxalite incident in the 39 months of the data used in this paper (January 2005-March 2008). All non-white districts are red corridor districts. Red corridor districts predicted to receive NREGS in the first, second, and third phase based on the algorithm are in black, dark grey, and light grey, respectively.

Distribution of Index and Discontinuities by Phase

Figure 2.2: Distribution of Index - Phase 1

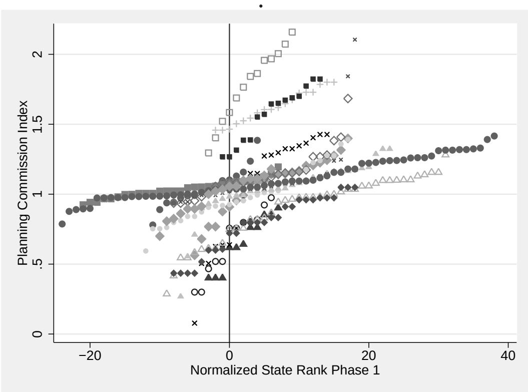


Figure 2.3: Distribution of Index - Phase 2

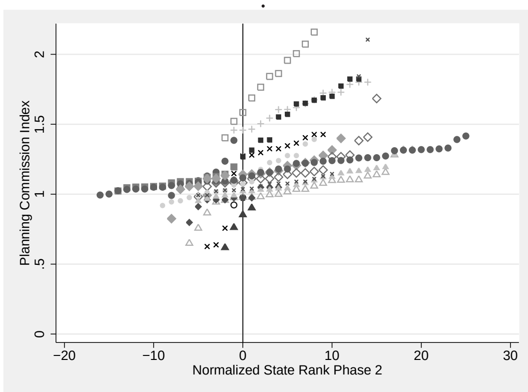


Figure 2.4: Phase 1 Discontinuity

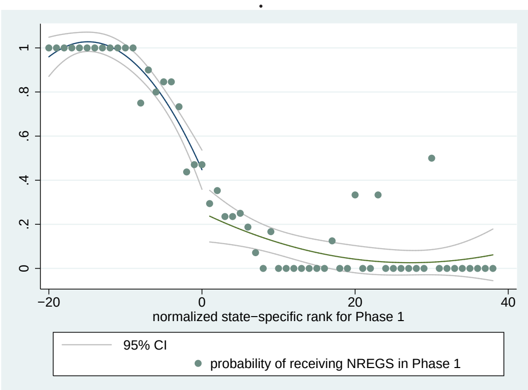
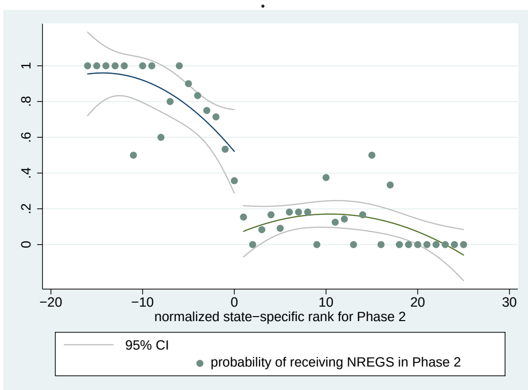


Figure 2.5: Phase 2 Discontinuity



Note: The first row of figures plot the distribution of the index by each state. The second row of figures show the discontinuities in treatment for each phase.

Discontinuities for Main Variables

Figure 2.6: Persons Affected

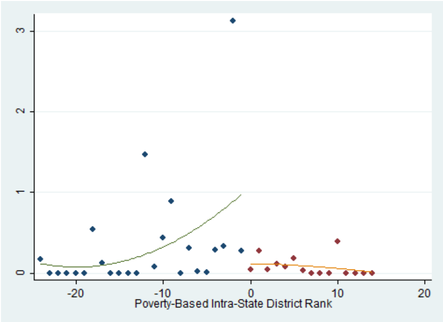
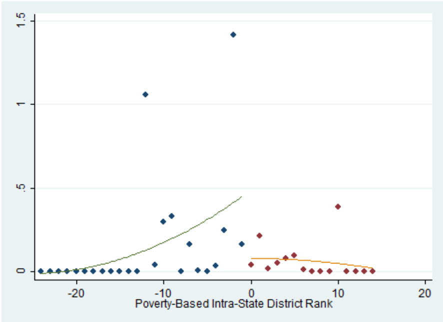
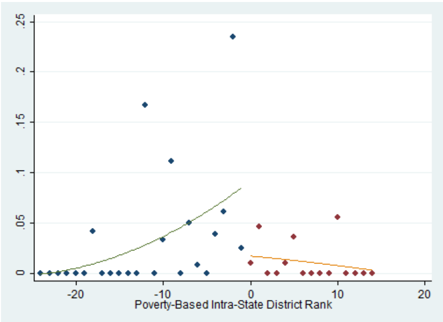


Figure 2.7: Persons Killed



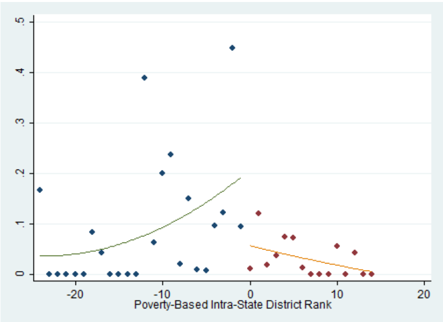
.

Figure 2.8: Major Incidents



.

Figure 2.9: Total Incidents

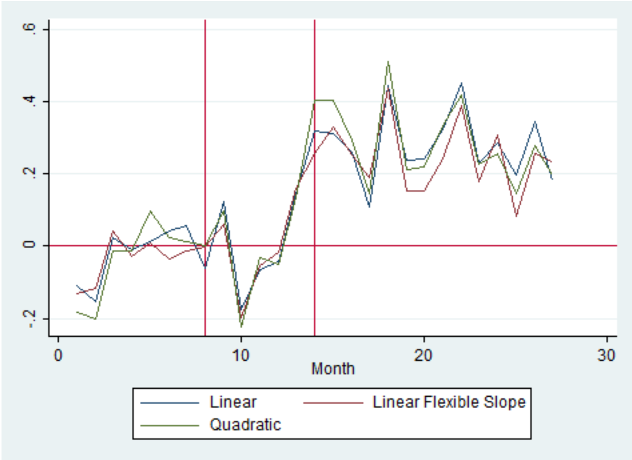


.

.

Note: The bin size for each figure is 1.

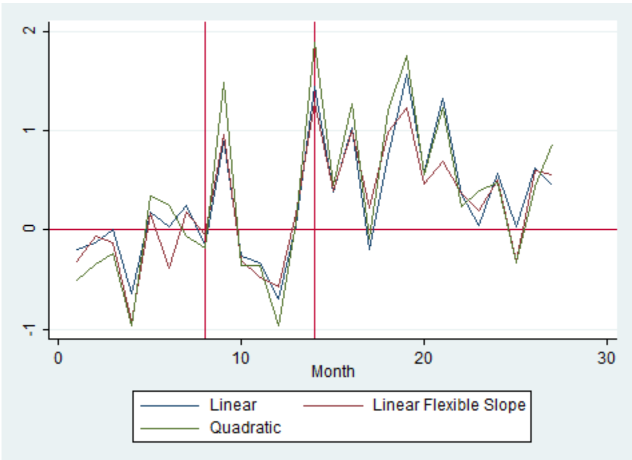
Figure 2.10: Monthly RD Coefficients - Total Number of Incidents



.

Note: Coefficients of month-by-month RD regressions of number of incidents. The first vertical line indicates the passage of the Act in Parliament, and the second vertical line indicates the first month of implementation in Phase 1. Each point on the graph is coefficient for a different regression restricting the sample to the corresponding month.

Figure 2.11: Monthly RD Coefficients - Number of Persons Affected



.

Note: Coefficients of month-by-month RD regressions of number of persons affected. The first vertical line indicates the passage of the Act in Parliament, and the second vertical line indicates the first month of implementation in Phase 1. Each point on the graph is coefficient for a different regression restricting the sample to the corresponding month.

CHAPTER III

Jai Ho? The Impact of a Large Public Works Program on the Government's Election Performance in India

3.1 Introduction

All around the world government parties try to improve their performance in general elections through a number of strategies, including taking credit for good macroeconomic conditions, making promises about future policies, and increasing government spending in areas expected to lead to more votes. Especially the last option allows the ruling parties to exercise substantial control over the number and identity of beneficiaries, and a growing literature studies the success of vote buying, increased public spending in the year of the election, and targeted programs that reach important subgroups of society. In general, these strategies often seem to be much more successful in developing countries than in the developed-country context.[1]

With the increasing importance of large and ambitious anti-poverty programs in a number of developing countries in recent years, two questions of growing interest are whether voters reward government parties for the introduction of such schemes and whether they hold their rulers accountable for the working of the programs.[2] Especially in countries that face large-scale implementation problems with government programs, the overall impact of an anti-poverty scheme on incumbent election outcomes is unclear a priori, since the effect at any given point in time is a combination of at least three effects: the election responses from citizens to awareness of the program, program salience, and their updated expectations of program benefits based on implementation quality. Awareness is likely to be increasing in length of program exposure, salience may be decreasing, and the effect of better knowledge about implementation quality over time could be positive or negative depending on people's initial expectations. The overall electoral impact of government programs is therefore an empirical question and may change depending on the length of the program.

This paper analyzes the impact of the largest public-works program in the world, the Indian National Rural Employment Guarantee Scheme (NREGS), on national election outcomes. NREGS legally guarantees each rural household up to 100 days of manual public-sector work per year at the minimum wage and is supposed to be a completely demand-driven program where households can self-select into employment at any time during the year. A number of field studies and academic papers have highlighted sometimes very severe problems with implementation quality, although

[1]Despite the popularity of some of these strategies among governments around the world and the often held qualitative view that they help boost the incumbents' election performance, the evidence in developed countries is often mixed. In developing countries, on the other hand, most studies point to positive impacts. See e.g. Brender and Drazen (2005, 2008), Drazen and Eslava (2010), Finan and Schechter (2012).

[2]Both conditional cash transfer programs and large and longer-standing public-works programs have been introduced in a number of developing countries around the world in recent years. See e.g. Subbarao et al. (2013) for an overview of public-works programs in developing countries.

some benefits for the rural poor are still realized and there is substantial heterogeneity in the working of the program across India. Despite these shortcomings, the main government party played up the success of the scheme in its general election campaign, and the party itself as well as some experts and members of the popular press believe that NREGS was one of the main reasons for the surprising landslide victory of the government parties in 2009.[3]

The empirical analysis in this paper exploits variation in the timing of the introduction of NREGS across the country and the fact that the Indian government rolled out the program according to an algorithm which allows the use of a regression-discontinuity design. I find that votes for the government parties are lower in districts with a longer exposure to the employment guarantee scheme, with the impact being concentrated in states with low implementation quality. A high implementation quality increases voter turnout and balances this effect out, however, with the additional votes coming at the expense of the main national opposition party. Similar effects hold for constituency-level incumbents of any party affiliation.

The results are consistent with a story in which voters initially anticipate large benefits from the program but realize its practical limitations over time and become less enthusiastic about NREGS. This effect is muted in districts with higher implementation quality where satisfaction with the program is likely to be higher, and implies that voters hold both local-level politicians and national parties accountable for the performance of the employment guarantee. Overall, these results suggest that any electoral benefits of implementing an ambitious anti-poverty program only last if the government resolves practical implementation issues.

The paper makes four main contributions to the existing literature. First, this is the first paper to analyze the importance of implementation quality for the impact of anti-poverty programs on electoral outcomes in developing countries, whereas papers

[3]See e.g. Khera (2010), Ramana (2009).

in the existing literature focus almost exclusively on programs that work very well empirically (see e.g. Baez et al. 2012, De la O 2013, Manacorda et al. 2011). The presence of implementation challenges complicates the expected impacts of the program, however, and is a more realistic scenario for many developing countries. The results in this paper also support existing evidence that better informed voters in developing countries increase the electoral accountability of governments and reduce malpractices (see e.g. Pande 2011). Banerjee et al. (2011) find that voters in the slums of an Indian city who had randomly received information about the performance of the incumbent and two main other candidates had a higher voter turnout and were more likely to vote for more qualified candidates than those in control areas.

Second, the paper provides some insight into the potentially important dynamic effects of voter approval by focusing on comparing election outcomes of areas with a different length of exposure to the program, whereas the large majority of existing studies in developing countries only compares treatment and control groups (see e.g. Manacorda et al. 2011, Baez et al. 2012). All papers find that beneficiaries are more likely to vote for the government parties in the next elections, and the magnitudes are often substantial. Only De la O (2013) compares election effects for areas with longer and shorter exposure to the Mexican conditional cash transfer program Progresa, and finds that electoral support was higher among the early treatment group. In general, the dynamic effects of government programs could be very different, however, especially in the presence of implementation challenges and only slowly growing awareness about many initiatives.

Third, the paper extends our understanding about the external validity of the electoral impact of anti-government programs in developing countries since the existing literature focuses almost exclusively on conditional cash transfer programs in Latin America.[4] Additionally, the nature of the Indian employment guarantee scheme al-

[4]The exiting literature studies conditional cash transfer programs in Brazil (Zucco 2010), Colombia (Baez et al. 2012, and Nupia 2011), Mexico (De la O 2013), the Philippines (Labonne 2013),

lows me to analyze which level in the federal state is held accountable for the program. This is an important consideration in many contexts where the central government pays for the program but where local authorities are responsible for implementing it, and especially where local-level politicians may have a different party affiliation than the government. The results from the Indian context here suggest that a well-implemented program benefits both the government parties and incumbents from any party, which implies that it is not necessarily in an opposition politician's interest to boycott the working of the program. This finding supports existing evidence in the literature that local politicians benefit from centrally funded programs (see e.g. Labonne 2013, Pop-Eleches and Pop-Eleches 2012).

Lastly, my paper extends our understanding of the impacts of the Indian employment guarantee scheme and how it may have impacted the peoples' relationship to the government. While most existing papers on the program focus on the analysis of the economic benefits, the interpretation of the results in this paper is consistent with the explanation in Chapter 2 of this dissertation about the impact of NREGS on insurgency-related violence in India. Focusing on the short-run impacts in the early implementation phases, Chapter 2 finds empirical patterns consistent with NREGS improving the relationship between civilians and government institutions like the police, which makes the police more efficient at tracking down insurgents. Both of these stories stress the anticipation effect of the program in the short run, whereas it may take time for people to experience the program themselves and to find out about the implementation problems.

The rest of this paper is organized as follows: Section 3.2 provides some background information about the working of NREGS and the Indian electoral system. Section 3.3 discusses the rollout of NREGS and the empirical estimation strategy. Section 3.4 presents the data sources and some summary statistics. Section 3.5 dis-

and Uruguay (Manacorda et al. 2011). Pop-Eleches and Pop-Eleches (2012) analyze the impact of a voucher program in Romania.

cusses the results. Section 3.6 concludes.

3.2 Background

3.2.1 NREGS

The National Rural Employment Guarantee Scheme (NREGS)[5] is one of the largest and most ambitious government anti-poverty programs in the world.[6] The scheme is based on the National Rural Employment Guarantee Act (NREGA) which was passed in the Indian parliament in August 2005 and which provides a legal guarantee of up to 100 days of manual public-sector work per year at the minimum wage for each rural household.[7] There are no other eligibility criteria, so households self-select into NREGS work and can apply for work at any time. Men and women are paid equally, and at any given time at least one third of the NREGS workforce is supposed to be female. Wages are the state minimum wage for agricultural laborers, although NREGA specifies a national floor minimum wage.[8] Wages need to be paid within 15 days of the day the work was performed, otherwise the worker is eligible for unemployment allowance. In practice, the focus of NREGS projects is on anti-drought measures and land development.

NREGS was rolled out across India in three phases: 200 districts received the program in February 2006 (Phase 1), 130 additional districts started implementing the scheme in April 2007 (Phase 2), and the remaining districts got NREGS in April 2008 (Phase 3). NREGS is now operating in 99 percent of Indian districts since it excludes

[5]The program was renamed to Mahatma Gandhi National Rural Employment Guarantee Scheme in 2009. Since the abbreviations NREGA and NREGS are more estiablished in the literature, however, I will keep referring to the program as NREGS.

[6]For more details on the scheme see e.g. Dey et al. (2006), Government of India (2009), and Ministry of Rural Development (2010).

[7]The year is the financial year and starts on April 1.

[8]The NREGS minimum wage was originally Rs.60, but has been raised various times since then. It was Rs.120 in 2009.

districts with a 100 percent urban population (Ministry of Rural Development 2010).

NREGS has received widespread attention in the popular press and in academic research. The patterns emerging from a number of papers on the labor-market impacts of the employment guarantee suggest that while there are no large overall benefits of the scheme in terms of increased employment or higher private-sector casual wages, the program provides a safety net and works much better in some states than in others. Imbert and Papp (2013), Azam (2012) and Berg et al. (2012) use a difference-in-difference (DID) approach to look at the program's impact on wages and employment by exploiting the phase-in of the program over time. They find that NREGS seems to have led to higher public employment and increased private-sector wages in the agricultural off-season (Imbert and Papp 2013), in areas with high implementation quality (Berg et al. 2012), and among casual workers (Azam 2012).

All of these papers focus heavily on interesting sources of heterogeneity rather than on the general impacts of the program. In addition, they need to rely on some version of the parallel trend assumption in their analyses, which is likely to be violated in practice because of the non-random roll-out of NREGS. Instead, Chapter 1 of this dissertation reconstructs the government algorithm used during the roll-out of the program and uses a regression-discontinuity framework to look at the overall labor-market impacts of the scheme. The empirical results suggest that NREGS is primarily used as a safety net rather than as an additional form of employment. The introduction of NREGS does not lead to an overall increase in public-sector employment or the casual private-sector wage. The program also does not create any additional jobs in the local economy and has no substantial impact on household expenditures or income. Consistent with the safety net function, however, take-up is higher after bad economic shocks, and workers seem to substitute away from private casual employment and towards other occupations like self-employment, which become relatively less risky with the introduction of a safety net. Johnson (2009a) also finds that NREGS provides a safety net for rural households in the Indian state of Andhra Pradesh since take-up of the program increases after negative rainfall shocks.

Due to the rapid phase-in of the employment guarantee scheme, most of these papers need to focus on the short-run impacts of NREGS, although the analysis in Berg et al. (2012) suggests that it may take some time for outcomes like higher casual private-sector wages to be realized. A growing literature points to large-scale problems with implementation quality, however: Dutta et al. (2012), for example, document widespread rationing of NREGS employment because of excess demand, and show that this is especially common in poorer states. Niehaus and Sukhtankar (2013a and 2013b) analyze the existence and characteristics of corruption in the implementation of NREGS in Orissa, and find that an increase in the minimum wage was not passed through to workers. NREGS seems to work much better in some other states like Andhra Pradesh, where Johnson (2009b) finds that the working of NREGS does not seem to be substantially affected by the specific party in power at the local panchayat level, suggesting that political pressures on NREGS in the state are not pervasive. This state heterogeneity is also routinely found in field reports of the working of the employment guarantee scheme on the ground, where the program seems to work relatively well in the so-called 'star states' (Andhra Pradesh, Chhattisgarh, Madhya Pradesh, Rajasthan, and Tamil Nadu) but faces severe challenges in the rest of the country (see e.g. Khera 2011).

Despite these implementation problems, NREGS has also been found to affect other outcomes in addition to the labor-market impacts. Klonner and Oldiges (2012) find some evidence of substantial reductions in poverty using a difference-in-difference approach, and NREGS has also been credited with improving children's education outcomes (Afridi et al. 2012). Using the regression-discontinuity design, Chapter 2 of this dissertation finds that insurgency-related violence in India increases in the short-run after NREGS is implemented. This pattern seems most consistent with the citizen-support channel, where citizens are willing to share more information about insurgents with the police after the introduction of the employment guarantee scheme because of actual or anticipated benefits of the program, which in turn makes the police more effective at tracking down rebels. This explanation implies that violence intensity should decline in the longer run as the ability of the insurgents to attack declines, which is consistent with results found in Fetzer (2013) and Dasgupta et al. (2014), although both papers attribute their results to a rise in the opportunity cost of being an insurgent after NREGS provides more employment opportunities.

Overall, the existing literature on the impacts of the employment guarantee scheme therefore suggests that the program has generated some economic benefits, and especially in states with a higher implementation quality, but that it may also have improved the relationship between civilians and the government more broadly. If this is true, NREGS may also have influenced the voting behavior of citizens in the next general elections.

3.2.2 India's Political System and the General Elections of 2009

India's electoral system works according to the first-past-the-post system, so the candidate with a plurality of the votes in a given constituency receives the seat in the Lok Sabha, the Indian parliament's lower house. Candidates can therefore be elected with much less than majority support, and the seat allocation can differ substantially from the allocation that would prevail under a proportional representation system. While the first-past-the-post system is often associated with few parliamentary parties, the last two decades have seen the rise of national coalition governments in India: The two big national parties, the Indian National Congress (INC) and the Bharatiya Janata Party (BJP), have created alliances with smaller parties to create working government coalitions called the United Progressive Alliance (UPA) and the National

Democratic Alliance (NDA), respectively (see for example Yadav 1999). Even then governments do not always have a majority of seats in Parliament, however. The UPA government elected in 2004, for example, depended on external support from the Left Front (an alliance of left-wing parties)[9] as well as that of two other parties.[10]

The membership of parties in the UPA and NDA has varied over the years. To the extent that small parties commit to a certain alliance before an election, parties often negotiate seat sharing agreements so that parties do not contest seats in all electoral constituencies and thus minimize vote-splitting between members of the same alliance. Additional negotiations can take place after elections, however, and small parties have also left government coalitions or changed alliances.

The UPA won the general elections in 2004 and followed the previous NDA coalition government. The government coalition included the INC as the main national party as well as 13 smaller parties with mostly regional strongholds.[11] The UPA membership before the 2009 general election differed from this composition after some parties left the coalition and new parties made agreements for the general elections, but since my research question focuses on the electoral benefits of a government program passed in parliament in 2005 I stick to the initial UPA composition for my empirical analysis.

For administrative and security reasons, the general elections of 2009 were held in five phases between April 16 and May 13. The election results were announced on

[9]The Left Front includes the Communist Party of India (Marxist), the Communist Party of India, the Revolutionary Socialist Party, and the All India Forward Bloc.

[10]The Bahujan Samaj Party and the Samajwadi Party.

[11]The small UPA member parties of the 2004 government are: Rashtriya Janata Dal, Dravida Munnetra Kazhagam, Nationalist Congress Party, Pattali Makkal Katchi, Telangana Rashtra Samithi, Jharkhand Mukti Morcha, Marumalarchi Dravida Munnetra Kazhagam, Lok Jan Shakti Party, Indian Union Muslim League, Jammu and Kashmir Peoples Democratic Party, Republican Party of India, All India Majlis-e-Ittehadul Muslimen, Kerala Congress (Times of India, 2006). Before the 2009 general elections, four parties left the government coalition: Telangana Rashtra Samithi, Marumalarchi Dravida Munnetra Kazhagam, Jammu and Kashmir Peoples Democratic Party, and Pattali Makkal Katchi. The empirical results are robust to excluding these parties from the UPA definition. Additional parties joined the UPA for 2009 elections, but I use the 2004 definition for my empirical analysis.

May 16. Election dates are set, and elections monitored, by the autonomous Election Commission of India.[12] Pre-polls had suggested a close race between UPA and NDA with a slight edge for the UPA, so the strong performance of the UPA, and the INC in particular, came as a suprise for most experts (see for example Ramani 2009): The UPA won 262 of the 543 seats (2004: 218)[13], with INC winning 206 seats, an increase of 61 seats relative to the 2004 election results. The NDA, on the other hand, lost support and only won 159 seats (2004: 181 seats).[14] The Left Front also did much worse than predicted and won 79 seats.

The popular press as well as academic experts have advanced a number of hypotheses to explain the unexpectedly strong performance of the UPA, and INC in particular. These include the strong leadership skills of INC leaders Sonia and Rahul Gandhi, the competent and corruption-free image of prime minister Manmohan Singh, as well as intra-party problems in the BJP and regional factors (see for example EPW 2009, Ramani 2009). Many commentators believe, however, that one important factor for the UPA's election success was its focus on welfare policies and other government programs, and specifically NREGS (see e.g. Ramani 2009). INC's manifesto stressed NREGS as one of the main successes of the UPA government, and the party's slogan during the election campaign was *Aam aadmi ke badhte kadam, har kadam par bharat buland* (The common man moves forward, and with his every step India prospers). INC also bought the rights to the title song 'Jai Ho' (May there be victory) of the film Slumdog Millionaire, which tells the story of a boy from the slums who wins the Indian version of the quiz show 'Who Wants to Be a Millionaire?'. This focus on the poor is widely believed to have resonated with the electorate, and INC leaders

[12]The Election Commission had decided in 2006 that NREGS would not be allowed to be extended to more districts after the announcement of elections in any state, and that with very few exceptions employment would need to be provided in ongoing projects during that time. See http://www.righttofoodindia.org/data/ec2006nregacodeofconduct.jpg.

[13]The absolute majority is 272 seats, so the UPA government is still reliant on external support. The UPA received 37.22 percent of the total vote (2004: 35.4 percent).

[14]24.63 percent of the votes (2004: 33.3 percent). NDA's biggest party and INC's main competitor, the BJP, won 116 seats (2004: 138 seats).

have also claimed that the electoral victory was in large part due to NREGS.[15] While such an election campaign strategy had been used repeatedly by INC in the past, experts stress that in contrast to previous campaigns which paid mere lip service to the party's commitment to the situation of the poor, the fact that NREGS was an actual ambitious government program made such claims credible.[16]

To test this hypothesis empirically, I exploit the phase-in of the employment guarantee scheme using a regression-discontinuity design.

3.3 NREGS Rollout and Empirical Strategy

3.3.1 NREGS Rollout

NREGS was rolled out non-randomly in three phases between 2006 and 2008, with poor districts receiving the program earlier than more developed districts. Chapter 1 of this dissertation reconstructs the algorithm that the Indian government used to assign districts to different implementation phases. The algorithm is intended to ensure inter-state and intra-state fairness norms and is therefore a two-step process: In the first step, a quota of treatment districts for each state in a given phase is determined, which is proportional to the prevalence of poverty across states.[17] The 'prevalence of poverty' measure is the state headcount ratio times the rural state population, which provides an estimate of the number of below-the-poverty-line people living in a given state. The headcount ratios are calculated from 1993-1994 National Sample Survey (NSS) data, whereas the rural state population numbers come from the 2001 Indian

[15] See for example Khera (2010).

[16] Indira Gandhi's election campaign slogan for the general elections in 1971 was *Garibi Hatao* (Eradicate poverty), for example. See for example the comments on the election results by political science professors Thachil at casi.ssc.upenn.edu/iit/thachil and Kumbhar at www.mainstreamweekly.net/article1382.html.

[17] In practice this provision also ensures that all states (union territories are usually excluded from such programs) receive at least one treatment district.

census.[18]

In the second step, the state-specific quotas are then filled by choosing a state's poorest districts according to a development ranking. The development index used to rank districts within states comes from a Planning Commission report from 2003 that created an index of 'backwardness', a term often used in India to refer to economic underdevelopment. The index was created from three outcomes for the 17 major states for which data was available: agricultural wages, agricultural productivity, and the proportion of low-caste individuals (Scheduled Castes and Scheduled Tribes) living in the district (Planning Commission 2003). Data on these outcomes was unavailable for the remaining Indian states, and it is unclear whether a comparable algorithm using different outcome variables was used for them. The empirical analysis is therefore restricted to these 17 states. Districts were ranked on their index values. In a deviation from the algorithm, 32 districts that were on a government list of districts heavily affected by Maoist violence all received NREGS in Phase 1 regardless of their poverty rank. I drop these districts from the sample for the empirical analysis.

The two-step algorithm results in state-specific cutoffs for treatment assignment for each implemtation phase. Since there are three implementation phases, two cutoffs can be identified: the cutoff between Phase 1 and Phase 2, and the cutoff between Phase 2 and Phase 3. These cutoffs correspond to the Phase 1 and Phase 2 NREGS roll-out, respectively.

[18]The state headcount ratios used in this paper come from a Planning Commission document from 2009 since the original headcount ratio calculations do not have estimates for new states that had been created in the meantime (Planning Commission 2009). Since these are official Planning Commission estimates, they are likely to be closest to the information the Indian government would have had access to at the time of NREGS implementation. NSS is a nationally representative household survey dataset. The newest available information on headcount ratios at the time would have been the 1999-2000 NSS data, but that dataset was subject to data controversies and therefore not used.

3.3.2 Empirical Strategy

The two cutoffs can be exploited empirically using a regression-discontinuity design. Since the general elections took place in 2009 when all rural districts had access to the program, the phasing in of the program provides variation in the length of time districts had been implementing the scheme. Since treatment cutoffs differ by state, ranks are made phase- and state-specific for the empirical analysis. To be able to pool observations from all states, the state-specific ranks are recentered such that a district with a normalized state-specific rank of zero is the last district in a state to be eligible for receiving the program in a given phase. All eligible districts have negative ranks, whereas the districts that should not receive NREGS in a given phase receive a postive normalized rank value.[19]

The overall prediction success rate of the assignment algorithm is 84 percent in Phase 1 and 82 percent in Phase 2. The prediction success rate is calculated as the percent of districts for which predicted and actual treatment status coincide.[20]. This means that there is some slippage in treatment assignment in both phases, and there is considerable heterogeneity in the performance of the algorithm across states. Nevertheless, the algorithm performs quite well in almost all states and the prediction success rates are also considerably higher than the ones that would be expected from a random assignment of districts, which are 40.27 percent for Phase 1 and 37.45 percent for Phase 2 at the national level, respectively. Overall, this suggests that the proposed algorithm works well for predicting Phase 1 and Phase 2 district allocations.

As there is some slippage in district assignment, the empirical identification strategy is a fuzzy RD design. The fundamental assumption of the empirical strategy is that districts that were just poor enough to receive the program in a given phase

[19] This section follows the information in Chapter 1 of this dissertation closely. See Chapter 1 for more details.

[20] Prediction success rates for Phase 2 are calculated after dropping Phase 1 districts from the analysis.

of the NREGS rollout, and districts that were just too developed to be included in a given phase are similar to each other. Then, discontinuities in the outcome variables for these two types of districts can be solely attributed to the differences in the duration of having had access to NREGS.

One important component of this assumption is that beneficiaries need to have been unable to perfectly manipulate their treatment status (Lee 2008). Otherwise, districts close to the cutoff on either side are not plausibly similar to each other in terms of unobservables, but differ on characteristics such as the perceived benefit from the program. In the case of the two-step RD, this means that districts should not have been able to manipulate their predicted status under the algorithm in either step.

This seems plausible: As mentioned above, the headcount poverty ratio used to calculate the number of treatment districts for a state in the first step of the algorithm used data from the mid-1990s, which had long been available by the time the NREGS assignment was made.[21]

Like the information used for the first step, it also seems unlikely that it was possible to tamper with the data used in the second step of the algorithm: The 'backwardness' index was constructed from outcome variables collected in the early to mid-1990s, eliminating the opportunity for districts to strategically misreport information. Additionally, the suggestion of the original Planning Commission report had been to target the 150 least developed districts, but NREGS cutoffs were higher than this even in Phase 1 (200 districts received it in Phase 1). Therefore, districts would have had an incentive to be among the 150 poorest districts but not to be among the 200 or 330 least developed districts for Phase 1 and Phase 2, respectively. Lastly, the

[21]The algorithm also uses state rural population numbers from the 2001 Census to transform headcount ratios into absolute numbers, but those figures were also long publicly available at the time. The RD may be potentially fuzzier than it really is because of some potential for measurement error introduced into the algorithm at this step since the exact numbers the government used in this step are not known, but this should not introduce systematic bias into the empirical analysis.

Planning Commission report lists the raw data as well as the exact method by which the development index was created, again eliminating room for districts to manipulate their rank. Overall, it therefore seems like manipulation of the rank variable is not a major concern.[22]

Figures 3.1 and 3.2 focus more closely on the distribution of index values over state-specific ranks. Ideally, the assignment variable should be continuous at the cutoff, since discontinuities at the cutoffs are typically taken as signs of potential manipulation. The figures plot the relationship between the Planning Commission's index and the normalized state-specific ranks for the Phase 1 and Phase 2 cutoffs, respectively. For most states, the poverty index values seem smooth at the cutoff of zero, again suggesting that manipulation of the underlying poverty index variable is not a big concern.

Another way of analyzing whether manipulation is likely to be a problem is to test whether there are any discontinuities at the cutoffs in the baseline data, although this is only a relatively crude test in this case since the boundaries of the electoral constituencies were re-drawn before the 2009 general election so that there is no perfect correspondence between the 2009 election outcomes and the baseline election outcomes from 2004. Nevertheless, we should not find impacts in the baseline data if the only factor that is driving any outcome differences in 2009 is the introduction of NREGS. Appendix Tables F.0.1, F.0.2, and F.0.3 present the results of such an analysis for the main outcome variables used in this paper (the vote share and winning constituency variables for INC, UPA, the left front and the BJP, the voter turnout, and the incumbent won and vote share variables) for three different parametric specifications for the 2004 general elections for both cutoffs. They show that only three

[22]This does not mean that actual treatment assignment was not subject to political pressures since compliance with the algorithm is often below 100 percent. It can be shown that deviations from the algorithm are correlated with party affiliation. This is also consistent with Gupta (2006) who analyzes the relationship between rule deviations and party affiliation for an earlier program. That paper ignores the fact that the program most likely also used the two-step algorithm, however, which could substantially affect the results.

of 66 coefficients are statistically significant and that those are not robust across different parametric specifications. Overall, this test suggests again that manipulation is unlikely to be an important problem.

With a fuzzy RD design, we also need to verify that there is indeed a discontinuity in the probability of receiving the program at the state-specific cutoff values for NREGS districts. Figures 3.3 and 3.4 show this graphically for the normalized state-specific cutoffs for Phase 1 and Phase 2, respectively. They plot the probability of receiving NREGS in the given phase for each bin, as well as fitted quadratic regression curves and corresponding 95 percent confidence intervals on either side of the cutoff. The graphs show that the average probability of receiving NREGS jumps down at both discontinuities, although this discontinuity is stronger in Phase 2 than in Phase 1. Both figures suggest, however, that there is indeed a discontinuity in the probability of being treated with the employment guarantee scheme in a given phase at the cutoff.

3.4 Data, Variable Creation and Empirical Specification

3.4.1 Data and Variable Creation

The election outcome data used in this paper are the official general election results of 2009 from the Election Commission of India.[23] For each electoral constituency, the Election Commission lists all candidates, their party affiliation, and the number of votes received per candidate as well as some limited candidate background information like gender, age, and broad caste category. It also gives the number of eligible voters in a given constituency, which allows the calculation of voter turnout.

Election constituencies are created to ensure fair votes-to-seat ratios that are

[23]Data are publicly available at http://eci.nic.in. The Election Commission of India is an autonomous body that schedules and oversees elections.

roughly equal across the country, and therefore do not correspond perfectly to other administrative boundaries in India. Since NREGS was rolled out at the district level, I match each election constituency to the closest appropriate district. The matching is done according to the name of the election constituency, which is usually a major city.[24] To this dataset, I merge information on the poverty index rank from the 2003 Planning Commission Report, district population size from the 2001 Census as well as information on a district's NREGS phase and some NREGS implementation quality information from the official government NREGS website.[25] This NREGS implementation quality information includes the number of individuals and households employed under NREGS in a given district, the number of households reaching the limit of 100 days, and the total person-days generated by NREGS. All these variables are for the financial year 2008-09, which is roughly the year before the general elections take place[26] and the only time span in which districts from all phases had access to NREGS. At the time of the general elections in April 2009, Phase 1 districts had had NREGS for three years, Phase 2 districts for two years, and Phase 3 districts for one year. As an alternative measure of the implementation quality of NREGS I also create an indicator variable equal to 1 if a constituency belongs to what has been called a 'star state', and 0 otherwise. Field reports on the working of NREGS in Khera (2011) identify five states in which NREGS seems to be implemented better than in the rest of the country: Andhra Pradesh, Chhattisgarh, Madhya Pradesh, Rajasthan, and Tamil Nadu.

I create various outcome variables to measure the impact NREGS had on the

[24] All electoral constituencies used in the empirical analysis can be matched non-ambiguously to one district. I use district boundaries from the 2001 Census to make these matches. The procedure of using the constituency name introduces some measurement error since election constituencies can span parts of more than one district, but this is uncommon. Furthermore, assigning such a constituency to the district that the constituency name's town is drawn from minimizes this measurement error since more local election results are not available. In most cases, there is one electoral constituency per district.

[25] The NREGS website is http://nrega.nic.in.

[26] The financial year starts on April 1.

general election results. Since India has a first-past-the-post system, an important outcome is the number of constituencies in which a party received the plurality of the vote since this directly translates to seats won in the Lok Sabha. I therefore create index variables equal to 1 if a given party or alliance won a plurality of votes in a constituency, and 0 otherwise. I also create variables of the received vote share of parties and alliances in a constituency. The empirical analysis focuses on the UPA government coalition and its main party the INC as well as the INC's main national competitor, the BJP, and the Left Front. I also look at voter turnout. In addition to these national level outcome variables, I also look at the election outcomes of incumbents. An incumbent here is an individual who won the 2004 general election in any electoral constituency in India and contested the elections again in 2009.

3.4.2 Empirical Specification

India had 543 elecotral constituencies in the 2009 general elections, but corresponding district Planning Commission rank information needed for the implementation of the regression discontinuity design is only available for 406 of them.[27] As the number of observations near the cutoffs is therefore limited, I use parametric regressions to estimate the impact of NREGS empirically. To test the robustness of the estimates, all main result tables show the estimated coefficients for linear and quadratic regression curves in the running variable with and without constraining the slope of the curves to be the same on both sides of the cutoffs.[28]

I estimate the treatment effect at the discontinuity separately for the two cutoffs. For each cutoff the running variable is the state-specific rank variable for the corresponding implementation phase. I drop the 32 insurgency-affected districts that

[27] Rank information is missing for urban districts as well as for some small states, especially in the North-East.

[28] F-tests reject the null hypothesis that higher-order polynomials add important flexibility to the model. The quadratic flexible specification is always statistically outperformed by the linear flexible specification.

received NREGS in the first phase, since these received the program regardless of their poverty level. The middle category of Phase 2 districts is taken as the reference group, so that the coefficients for Phase 1 and Phase 3 directly provide the estimated treatment effect at the two cutoffs. It is important to note that since the RD design depends on observations being close to the cutoff for identification it is impossible to compare Phase 1 and Phase 3 districts directly, since these will be far apart from each other by design.

The equation below shows the regression equation for one of the specifications we run, which is linear in the running variable but does not constrain the coefficients to be the same on either side of the cutoff:

$$y_{ij} = \beta_0 + \beta_1 rank_i + \beta_2 nregs_i + \beta_3 nregs * rank_i + \epsilon_{ij}$$

where the subscripts refer to constituency i in district j. y is an outcome variable of interest, whereas *nregs* is an indicator variable equal to 1 if a district is predicted to receive NREGS in that phase according to the algorithm, and zero otherwise. *rank* is a district's rank based on the state-specific normalized index. The coefficient of interest is β_2. Since we are dealing with a fuzzy RD rather than a sharp RD, this specification is the intent-to-treat (ITT) specification. Standard errors are clustered at the district level.

The validity of the RD design does not depend on the availability of baseline information, although including such variables as controls can improve the precision of the estimates. Since the electoral constituency boundaries were redrawn between the 2004 and 2009 general elections, and boundaries changed for 499 of the 543 constituencies, it is unclear whether the 2004 outcomes provide good control variables for the 2009 election results. Therefore, my main results do not include control variables.

3.4.3 Summary Statistics

Table 3.1 presents some summary statistics for the sample used for the empirical analysis separately by NREGS phase. As we can see, voter turnout in the 2009 general elections was about 60 percent in Phase 1 and Phase 2 constituencies, although a bit lower in Phase 3 constituencies. The probability of winning a seat is between 33 and 38 percent for INC and weakly increasing in the phase number. This pattern is more pronounced for all the parties belonging to the UPA taken together, where the probability of receiving a plurality of the vote is about 39 percent in Phase 1 constituencies, but about 48 percent in Phase 3 districts. By contrast, the pattern for the INC's main national competitor, the BJP, follows a U-shape, whereas the probability of winning decreases for the Left Front in later NREGS phases. The winning probability of the BJP is around 20 percent, whereas the Left Front's chances are much lower. The corresponding average vote shares for the parties and the UPA reveal that a higher average vote share is often not necessarily associated with a higher probability of winning a seat.

The last two rows provide an overview of the implementation quality of NREGS, giving the number of persons employed under the scheme in the financial year 2008-09 as well as the corresponding number of generated person-days. These data come from administrative sources, since nationally representative information from other sources for the same time span is unavailable. Research by Niehaus and Sukhtankar (2013a and 2013b) has shown that at least in the state of Orissa administrative data vastly overstates actual employment because of corruption, although it less clear how big this problem is nationwide. This means that employment statistics in Table 3.1 should be taken with a grain of salt. They do suggest, however, that employment generation in Phase 3 districts in their first year of implementing the program is lagging significantly behind the areas that have had access to NREGS for a longer period.

3.5 Results

3.5.1 Main Results

Tables 3.2 to 3.7 analyze the impact of NREGS on the 2009 general election outcomes for the national parties as well as the local-level incumbents. Tables 3.2 and 3.3 present the overall impact of the employment guarantee scheme for the UPA coalition government and its main party, the INC, as well as for the two main competitor parties, the BJP and the Left Front. The outcome variables include indicator variables equal to 1 if a given party or coalition won in an electoral consituency and vote share variables that give the achieved share of the vote in a constituency in percent as well as the voter turnout. The reference category in both tables are electoral constituencies in Phase 2 NREGS districts, and the results are reported for three different parametric specifications.

As can be seen from these results, the overall impact of the length of exposure to the employment guarantee scheme is relatively limited, although a number of coefficients are economically significant but imprecisely estimated. In Table 3.2, the estimated coefficients for the INC and UPA outcome variables are typically large and negative, although usually statistically insignificant, implying for example a reduction in the probability that the INC wins in electoral constituencies by about 13 percentage points in the first row for Phase 1 districts relative to Phase 2 districts. The Left Front parties, on the other hand, do consistently better in Phase 1 districts at the cutoff with an about 5 percentage point higher probability of winning a plurality of the votes. A similar picture emerges for the vote share outcome variables, whose coefficients typically have the same sign as their correponding 'won' outcome variables but are imprecisely estimated.

A similar pattern emerges in Table 3.3 for Phase 3 districts relative to Phase 2 districts at the cutoff. Again, districts with a longer exposure to NREGS (2 years versus 1 year) tend to vote less for the government parties, with two out of three coefficients being significant for the government coalition UPA's vote share, although the other coefficients are again imprecisely estimated. There is also some evidence for voter turnout being lower in Phase 3 districts than in Phase 2 districts.

Tables 3.4 and 3.5 reveal that the overall results mask important heterogeneity with respect to implementation quality, however. The main NREGS variable is here interacted with an indicator variable equal to 1 if the district is part of a so-called 'star state' where the employment guarantee scheme is relatively well implemented according to field studies by NGOs and social activists. Table 3.4 shows that there is a large negative impact of NREGS on the winning probability for the government parties in Phase 1 non-star districts that is always statistically significant at at least the 10 percent level. This negative effect of longer exposure to the employment guarantee scheme is reduced and typically even reversed in Phase 1 star states. The reverse pattern is apparent for the main competitor of the government parties, the BJP, where votes in Phase 1 star-state districts are statistically significantly lower than those in non-star Phase 1 districts. Additionally, voter turnout is also statistically significantly higher in star states on the order of 12 percentage points.

These results suggest that citizens in Phase 1 districts, i.e. districts that have had access to NREGS for three years at the time of the elections, are less willing to reward the government parties for the anti-poverty program than citizens in Phase 2 districts unless the program works well. Additional votes for the government in Phase 1 star-state districts cut into the votes for the BJP, but also lead to a higher voter turnout than in non-star states.

The mirror image of this pattern emerges when comparing Phase 3 to Phase 2 districts in Table 3.5, which again implies that voters in districts with longer exposure to NREGS may overall be less willing to vote for the government parties than districts that received NREGS in Phase 3 (although the main effects are here typically imprecisely estimated), but that a high implementation quality in star-states again reverses this effect. As in Table 3.4, star-state districts with longer exposure (Phase 2 districts) see an increase in voter turnout and lower votes for the BJP than non-star Phase 2 districts.

People in districts that had just received the program in the year prior to the elections therefore seem to be the ones most willing to vote for the government, whereas individuals in districts with longer access to the scheme care much more about implementation quality. These patterns are consistent with the idea that people initially may have high expectations of NREGS, leading them to support the government. The more time they have to experience the program themselves or hear about the implementation problems associated with it, however, the less positive they may feel about NREGS and the government.

Tables 3.6 and 3.7 extend the analysis for the national parties to the impacts for incumbents. An incumbent is defined as an individual who won a plurality of votes in an electoral constituency in the 2004 general elections and competed again in the 2009 general election in any electoral constituency. As Table 3.6 shows, incumbents of any party affiliation are statistically significantly more likely to win again in Phase 1 districts relative to Phase 2 districts. While the remaining coefficients are only imprecisely estimated, the sign of the effect goes in the opposite direction for the other cutoff, with Phase 2 district incumbents being less likely to be re-elected.

Table 3.7 reveals that implementation quality is again an important source of heterogeneity, with star-state districts with longer exposure to NREGS being more likely to re-elect incumbents than non-star districts for both cutoffs. This is similar to the corresponding impacts for the national parties in Tables 3.4 and 3.5, and again suggests that voters become more sensitive to implementation quality with longer exposure to the employment guarantee scheme. With the central government paying for most of the costs of NREGS, but with the local level being responsible for implementing the program, it makes sense that we find similar patterns for incumbents and national parties.

Overall, the empirical results suggest that voters hold both the national government parties and the local-level incumbents responsible for the implementation of the employment guarantee scheme. While voters in the last implementation group (Phase 3) seem to be the most positive about NREGS, the people with long exposure to the program, who have presumably had time to experience the program themselves or hear about from others, are less likely to reward the resopnsible parties unless implementation quality is high.

3.5.2 Robustness Checks and Extensions

The robustness of the empirical results is tested in a number of robustness checks reported in the appendix. Tables F.0.4 to F.0.9 report the results of using a donut hole approach. One concern with the regression discontinuity design in this paper is that measurement error in the running variable may lead to the misclassification of some observations close to the cutoff. The donut-hole specifications drop districts with normalized state-specific ranks of -1, 0, and 1 to see whether this issue is a major concern for the estimates, and the tables show that the results are similar to the main results.

The empirical results are also robust to a number of alternative specifications. Tables F.0.10 to F.0.12 present the results using a probit specification instead of a linear probability model for the constituency won outcome variables. Tables F.0.13 to F.0.18 re-estimate the main results using a treated-on-the-treated (TOT) specification where actual receipt of NREGS is instrumented with predicted NREGS receipt according to the algorithm. Lastly, Table F.0.19 shows that the star-state results are also present when using a difference-in-difference approach to estimate the impact of NREGS on election outcomes.

Another potential concern with interpreting the main results is that the star-state heterogeneity may not actually reflect differences in implementation quality, but could be driven by other factors. To test whether there are indeed differences in implementation quality of NREGS between star states and other states Tables F.0.20 to F.0.22 use available information on the allocation of finances and labor-market outcomes at the district level. Table F.0.20 uses administrative information on the allocation of financial resources under NREGS in 2008/2009 on four dimensions: centrally released funds, total available funds, total expenditures, and administrative expenditures. The results show that star-state districts with longer exposure to NREGS do not spend statistically significantly more money than other districts of the same implementation phase and are not allocated more money, although the interaction effects are positive and often economically significant. Star-states have statistically significantly higher administrative costs, which is consistent with the idea that star-state administrations implement the employment guarantee scheme more thoroughly.

Tables F.0.21 and F.0.22 use available data on worker outcomes from administrative data and from National Sample Survey household survey data from 2009/2010. Column 1 in both tables reveals that according to administrative data early star-state districts employ fewer people than other districts from the same implementation phase, which is consistent with a lower number of ghost workers and less corruption in star states that lead to over-reporting of generated employment. The household-survey data in the remaining columns shows that individuals in star-state districts have higher wages and per-capita consumption. The interaction effect for public employment is also positive for longer-exposed star-state districts, although the coefficients are imprecisely estimated. Overall, Tables F.0.20 to F.0.22 are therefore consistent with the idea that star states really have a higher implementation quality of NREGS than districts in other states.

3.6 Conclusion

This paper has analyzed the impact of a major government anti-poverty program in India, the National Rural Employment Guarantee Scheme (NREGS), on the government's performance in the next general elections. Using a regression discontinuity framework, I find that approval for the government parties seems to decline with the length of exposure to the program, with the effect concentrated in districts with low implementation quality. A high implementation quality, on the other hand, balances out this effect and is associated with a higher voter turnout. Votes for the government come at the expense of the main national opposition party, although incumbents of any party affiliation benefit from NREGS in well-implemented areas. The results are consistent with a story in which it takes some time for voters to learn about the implementation quality of NREGS, but where voters realize the practical limitations of the program after having had the scheme for some time and become less enthusiastic about NREGS in areas with substantial implementation challenges.

Overall, the results show that there is an electoral benefit of implementing ambitious anti-poverty programs in India. At the same time, however, the empirical analysis suggests that such electoral payoffs may be relatively short-lived if the government is not really committed to a high-quality implementation of such programs. While election campaigns can serve to inform people about existing government policies and remind them of the benfits of the schemes, such a strategy may not be successful with voters in the longer run. Many voters seem to care about the actual benefits government initiatives like NREGS provide, and not just about a verbal commitment to the fight against poverty. This points to the rise of what political scientists call 'programmatic politics' where patronage networks alone are no longer sufficient to win elections and would therefore imply a deepening of democracy.[29]

A limitation of the paper is that it is impossible to disentangle different theories

[29] See e.g. De la O (2013) for a similar conclusion for the election impacts of a Progresa in Mexico.

of why citizens are more likely to vote for incumbents in areas where the government program works relatively well. The results are consistent with voters viewing the implementation of NREGS as a signal of competence and a commitment to the fight against poverty, which promise more benefits after the election (see e.g. Drazen and Eslava 2010). This seems to be a view implicitly held by a number of commentators on the election success of the Indian government who stressed the fact that NREGS marked the first time that a government had actually implemented a very ambitious anti-poverty program and not mostly relied on lip service of the importance of economic development. The empirical results can also be explained with a model of reciprocity, however, where voters simply reward the government for its past performance (see e.g. Finan and Schechter 2012).

Table 3.1: Summary Statistics

	Phase 1	Phase 2	Phase 3
N	170	106	225
voter turnout	0.6003	0.6087	0.5796
INC win	0.3353	0.3585	0.3778
UPA win	0.3882	0.4151	0.4800
BJP win	0.2176	0.1792	0.2356
Left Front win	0.0294	0.0189	0.0044
INC vote share	26.36	24.83	29.36
UPA vote share	35.95	32.15	37.54
BJP vote share	17.25	17.63	21.96
Left Front share	2.84	1.96	1.23
NREGS employment	114563	120385	58419
NREGS person days	2182372	2224858	1674907

Note: Vote shares given in percent. INC (Indian National Congress), UPA (United Progressive Alliance), BJP (Bharatiya Janata Party) UPA is the name of the government coalition. For the government elected in 2004, the UPA consisted of the following parties: Indian National Congress, Rashtriya Janata Dal, Dravida Munnetra Kazhagam, Nationalist Congress Party, Pattali Makkal Katchi, Telangana Rashtra Samithi, Jharkhand Mukti Morcha, Marumalarchi Dravida Munnetra Kazhagam, Lok Jan Shakti Party, Indian Union Muslim League, Jammu and Kashmir Peoples Democratic Party, Republican Party of India, All India Majlis-e-Ittehadul Muslimen, Kerala Congress Left Front is an alliance of left-wing parties and includes the Communist Party of India (Marxist), the Communist Party of India, the Revolutionary Socialist Party, and the All India Forward Bloc

Table 3.2: Election Results Phase 1 vs Phase 2

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.1272	-.1382*	.0459**	-.0595	-5.45	-6.84*	.35	-2.77	-.0297
(Linear)	(.0818)	(.0836)	(.0228)	(.0630)	(3.48)	(3.51)	(1.47)	(3.06)	(.0288)
NREGS Phase 1	-.0823	-.0942	.0485*	-.0656	-3.63	-5.58	.67	-3.14	-.0015
(Linear flexible)	(.0834)	(.0852)	(.0258)	(.0632)	(3.56)	(3.58)	(1.53)	(3.05)	(.0284)
NREGS Phase 1	-.0460	-.0426	.0529*	-.0658	-1.74	-3.16	.88	-2.94	.0148
(Quadratic)	(.0871)	(.0888)	(.0274)	(.0651)	(3.72)	(3.77)	(1.62)	(3.16)	(.0304)
N	406	406	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table 3.3: Election Results Phase 2 vs Phase 3

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 3	-.0473	-.0150	-.0111	-.0032	-.75	2.04	-1.51	-.02	-.0916***
(Linear)	(.0819)	(.0841)	(.0152)	(.0712)	(3.33)	(3.13)	(1.20)	(3.20)	(.0247)
NREGS Phase 3	.0346	.0751	-.0114	-.0127	2.90	6.24*	-.89	-1.23	-.0405*
(Linear flexible)	(.0901)	(.0934)	(.0153)	(.0809)	(3.59)	(3.23)	(1.25)	(3.56)	(.0244)
NREGS Phase 3	.0619	.1129	.0056	-.0571	2.53	6.22*	-.14	-3.66	-.0166
(Quadratic)	(.0957)	(.0988)	(.0145)	(.0884)	(3.83)	(3.42)	(1.30)	(3.80)	(.0253)
N	406	406	406	406	406	406	406	406	406

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table 3.4: Star State Results Phase 1 vs Phase 2

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.2375***	-.2383***	.0550*	-.0058	-8.34**	-8.72**	-.35	.85	-.0753**
(Linear)	(.0868)	(.0898)	(.0288)	(.0714)	(3.83)	(3.88)	(1.79)	(3.24)	(.0336)
NREGS Phase 1*star	.2913***	.2399**	-.0248	-.1656*	6.78	2.80	2.36	-11.54**	.1324***
	(.1129)	(.1062)	(.0316)	(.0895)	(4.28)	(4.00)	(1.87)	(4.57)	(.0347)
NREGS Phase 1	-.1922**	-.1971**	.0603*	-.0100	-6.40	-7.64*	.08	.56	-.0406
(Linear flexible)	(.0885)	(.0918)	(.0333)	(.0717)	(4.00)	(4.04)	(1.91)	(3.25)	(.0343)
NREGS Phase 1*star	.2559**	.2077*	-.0290	-.1623*	5.26	1.96	2.02	-11.31**	.1052***
	(.1132)	(.1076)	(.0327)	(.0889)	(4.34)	(4.15)	(1.93)	(4.59)	(.0352)
NREGS Phase 1	-.1700*	-.1601*	.0661*	-.0078	-4.99	-5.62	.30	.87	-.0293
(Quadratic)	(.0918)	(.0955)	(.0351)	(.0737)	(4.18)	(4.29)	(2.01)	(3.38)	(.0365)
NREGS Phase 1*star	.2638**	.2081*	-.0293	-.1648*	5.41	1.54	2.09	-11.55**	.1136***
	(.1132)	(.1068)	(.0326)	(.0889)	(4.36)	(4.15)	(1.92)	(4.58)	(.0350)
N	406	406	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise. Star state is an indicator variable equal to 1 if a state was identified as a high implementation quality state, and 0 otherwise. The five star states are Andhra Pradesh, Chhattisgarh, Madhya Pradesh, Rajasthan and Tamil Nadu.

Table 3.5: Star State Results Phase 2 vs Phase 3

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 3	.0717	.0681	-.0151	-.0475	4.56	4.72	-1.42	-4.19	-.0444
(Linear)	(.0926)	(.0957)	(.0191)	(.0798)	(3.76)	(3.62)	(1.54)	(3.54)	(.0296)
NREGS Phase 3*star	-.2576**	-.1104	.0061	.1270	-11.77***	-2.02	-.70	12.69***	-.1281***
	(.1144)	(.1125)	(.0224)	(.0973)	(4.46)	(3.81)	(1.45)	(4.67)	(.0338)
NREGS Phase 3	.1798*	.1795	-.0141	.0262	10.02**	11.02***	-.49	-.70	-.0117
(Linear flexible)	(.1066)	(.1100)	(.0201)	(.0931)	(4.31)	(3.83)	(1.71)	(4.18)	(.0294)
NREGS Phase 3*star	-.3495**	-.1599	.0026	-.0799	-17.68***	-6.87	-1.64	1.91	-.0915**
	(.1711)	(.1718)	(.0238)	(.1470)	(6.62)	(5.05)	(1.80)	(6.90)	(.0458)
NREGS Phase 3	.1481	.1700	.0003	-.0866	6.54	8.11**	-.24	-6.56*	.0132
(Quadratic)	(.1032)	(.1056)	(.0164)	(.0926)	(4.05)	(3.63)	(1.53)	(3.90)	(.0283)
NREGS Phase 3*star	-.2230*	-.0643	.0131	.1093	-10.88**	-.49	-.17	11.61**	-.1020***
	(.1149)	(.1120)	(.0243)	(.0971)	(4.63)	(3.90)	(1.48)	(4.74)	(.0332)
N	406	406	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table 3.6: Incumbent Results

	incumbent won	Phase 1 vs 2 vote share	incumbent won	Phase 2 vs 3 vote share
	(1)	(2)	(3)	(4)
NREGS	.1455**	1.65	.0396	.47
(Linear)	(.0724)	(3.15)	(.0718)	(3.32)
NREGS	.1524**	2.19	.0349	1.35
(Linear flexible)	(.0755)	(3.32)	(.0797)	(3.63)
NREGS	.1719**	2.81	.0579	1.26
(Quadratic)	(.0784)	(3.51)	(.0856)	(3.85)
N	406	406	406	406

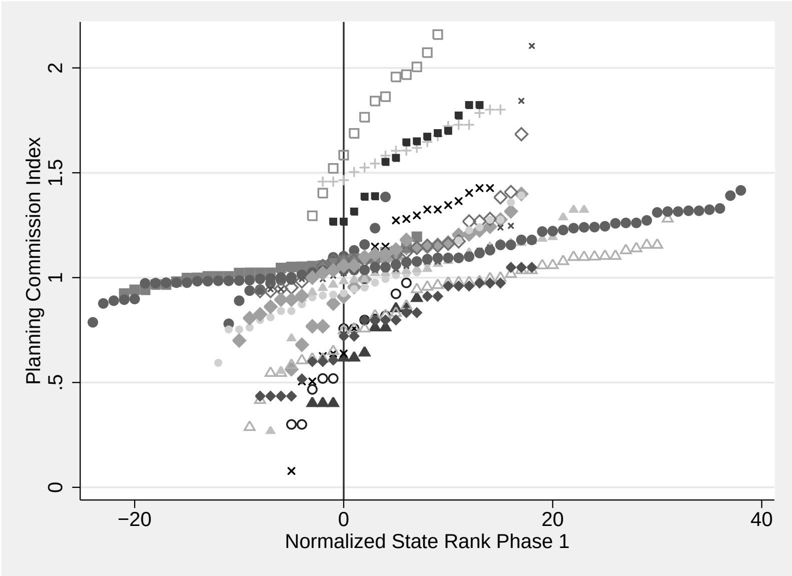
Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise. An incumbent is an individual who won a plurality of votes in his/her constituency in the 2004 general elections and who contested the 2009 elections in any constituency.

Table 3.7: Incumbent Star Results

	incumbent won	Phase 1 vs 2 vote share	incumbent won	Phase 2 vs 3 vote share
	(1)	(2)	(3)	(4)
NREGS	.0767	.43	.1822**	3.45
(Linear)	(.0779)	(3.51)	(.0811)	(3.71)
NREGS*star	.1987*	3.40	-.3884***	-8.05*
	(.1117)	(4.65)	(.1034)	(4.61)
NREGS	.0662	.42	.1433	5.17
(Linear flexible)	(.0828)	(3.82)	(.0961)	(4.21)
NREGS*star	.2901*	7.70	-.3549**	-11.42
	(.1702)	(7.31)	(.1558)	(7.27)
NREGS	.0960	1.57	.1705*	3.57
(Quadratic)	(.0854)	(3.95)	(.0914)	(4.10)
NREGS*star	.1908*	2.94	-.3937***	-8.00*
	(.1132)	(4.70)	(.1051)	(4.69)
N	406	406	406	406

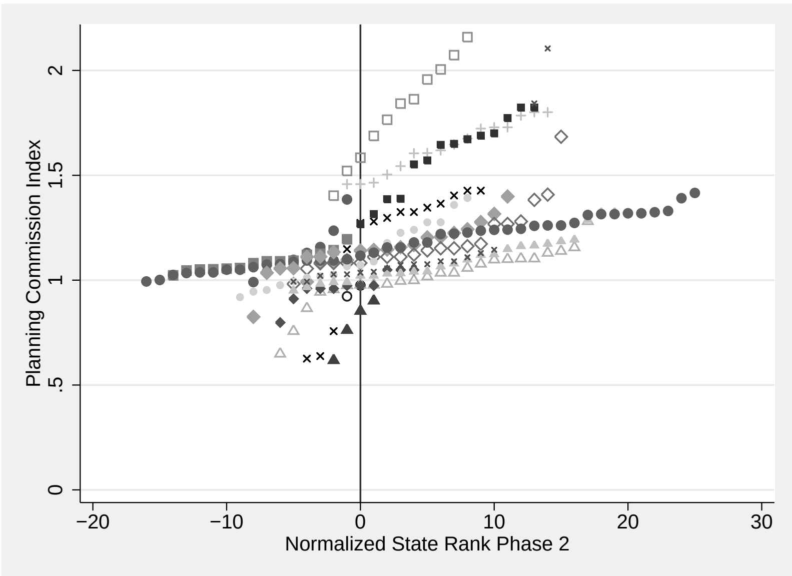
Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Figure 3.1: Distribution of Index over State-Specific Ranks (Phase 1 Treatment Assignment)



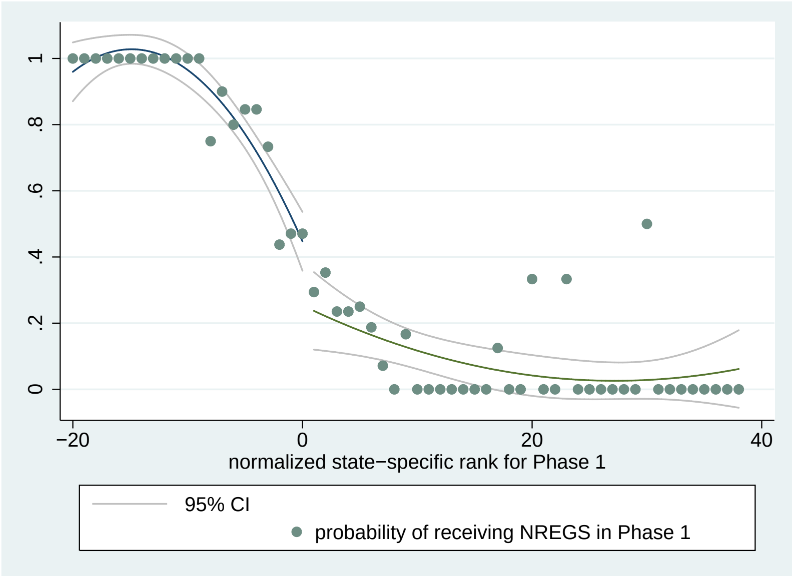
.

Figure 3.2: Distribution of Index over State-Specific Ranks (Phase 2 Treatment Assignment)



.

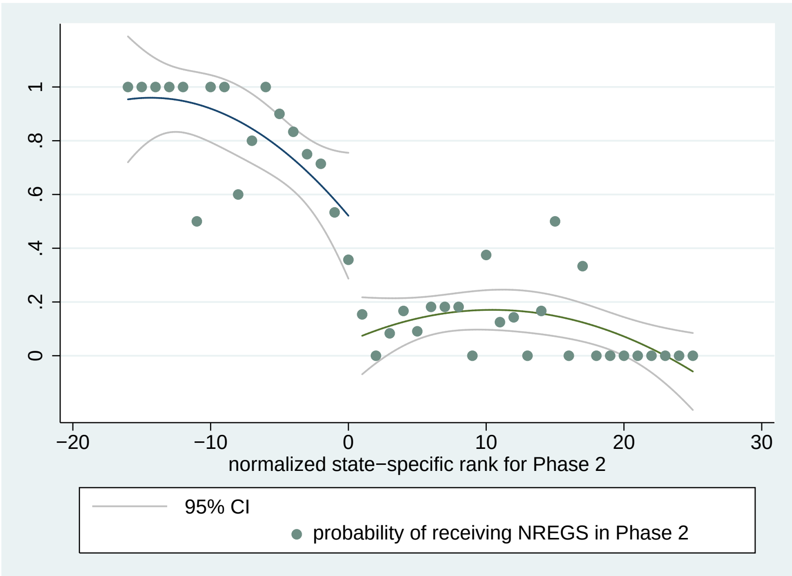
Figure 3.3: Discontinuity of treatment status for Phase 1



.

Note: The used bin size is 1 (each individual rank).

Figure 3.4: Discontinuity of treatment status for Phase 2



.

Note: Figure 3.4 excludes Phase 1 districts. The used bin size is 1 (each individual rank).

APPENDICES

APPENDIX A

Additional Information on the Program Rollout and Used Algorithm (for Chapter 1)

The general government documents state that NREGS was rolled out to the poorest districts first, but do not explicitly define the algorithm the government used to decide which districts would receive the program in which phase. While the actual algorithm is not publicly available, institutional knowledge about existing information at the time and about the workings of earlier development initiatives allows the construction of a plausible algorithm that works in two steps: First, the number of districts allocated to a given state is proportional to the prevalence of poverty across states. This mechanism ensures that the number of districts allocated to a given state is roughly proportional to the percent of India's poor people living in that state.[1] Second, within a state districts are chosen based on a development ranking, so that poor districts are chosen first.

This algorithm is attractive in the Indian context since it takes into account political fairness of resource allocation across states and within states and had been used for earlier development programs: The Indian Planning Commission explicitly states,

[1]In practice this provision also ensures that all states (union territories, which are administrative units directly ruled by the national government, are usually excluded from such programs) receive at least one treatment district.

for example, that this method had been used for treatment assignment of an earlier much smaller and less ambitious temporary government program aimed at less developed districts.[2] A former member of the Planning Commission also confirms that the development ranking was indeed used for NREGS as well and that state allocations were made proportional to the prevalence of poverty across states. Additionally, given the importance of NREGS and the huge political interest and awareness it created among policymakers at all levels as well as NGOs and the press, it seems very likely that the Indian government adhered to these political fairness norms in the allocation of treatment districts for NREGS as well. A number of NGOs and well-known individuals were actively campaigning for the introduction of an employment guarantee scheme like NREGS, and have been closely monitoring the working of the program since its introduction.[3]

The information used to rank districts is well documented and very transparent: In 2003, the Planning Commission published a report that created a 'backwardness index' from data from the early to mid-1990s on three outcomes (agricultural wages, agricultural productivity, and proportion of low-caste individuals living in the district) (Planning Commission 2003). Districts were then ranked based on their index values. The goal of the index at the time was to identify especially underdeveloped districts for wage and self-employment programs and, as mentioned above, it was actually used in pre-NREGS district initiatives, but those programs were much less extensive than NREGS and usually envisioned as temporary programs.

What is less well-documented is the choice of the poverty criterion to determine across-state allocations of treatment districts, since the Planning Commission report that created the development index did not cover this provision of the algorithm,

[2]See e.g. Planning Commission (MLP Division) 2003 for RSVY district assignment.

[3]Jean Dreze and Reetika Khera have been especially involved in NREGS from the beginning. Examples of monitoring include awareness campaigns for workers' rights under NREGS, survey data collection to find out about common challenges and violations of the law, suing governments for NREGA violations, and drawing attention to corruption. See e.g. Samarthan Centre for Development Support 2007.

and other Planning Commission documents that describe the algorithm used for the district assignment of government initiatives just refer to the 'incidence of poverty' as the criterion used, but never explicitly define the term or its operationalization. Given the Indian government's focus on poverty headcount ratios (the percent of people living below the poverty line) in many reports and publications, the best guess of the poverty definition used seems to be the state headcount ratio times the rural state population. This provides an estimate of how many below-the-poverty-line people live in a given state and of how poverty levels compare across states. I therefore use this procedure as the first step of the algorithm, where a state is assigned a percentage of total treatment districts that is equal to the percentage of India's poor living in that state. For the calculations I use the headcount ratios calculated from 1993-1994 NSS data[4], which is nationally representative household survey data that a former member of the Planning Commission says was used to derive state allocations of NREGS districts since the newest available information on poverty at the time from 1999-2000 NSS data was subject to controversies and was therefore not used.

[4]I use the rural state headcount ratios from Planning Commission (2009), since the original headcount ratio calculations do not have estimates for new states that had been created in the meantime. Since these are official Planning Commission estimates, they seem like the best guess of the information the Indian government would have had access to at the time of NREGS implementation.

APPENDIX B

Derivation of Theoretical Results (for Chapter 1)

The Baseline Model without NREGS

The model describes a household's optimal time allocation in a one-period setting. Before NREGS is introduced, a household can first choose to allocate the total time of their household members, T, between working for a big landowner as agricultural laborer in the private casual sector, l, and working on the family farm, f. After this decision has been made, a weather shock is realized that determines the payoff from farm work. The period ends, and the household earns the fixed wage w in the private sector, and income y for the time spent in farming. The household derives utility both from the time spent working in self-employment on the family farm, and from the total income earned in both activities during the period. The utility function is additively separable in these components, with $u' > 0, u'' < 0$, $v' > 0, v'' < 0$, and with weight α given to the utility from self-employment.

At the beginning of the period, a household's optimization problem is

$$\max_{l} \alpha v(T-l) + (1-\alpha)E[u((T-l)y + lw)]$$

Which leads to the first-order condition

$$\alpha v'(T-l) = (1-\alpha)\int u'((T-l)y + lw)(w-y)g(y)dy \tag{B.1}$$

(B.1) pins down the optimal proportion of time l spent working in the private sector implicitly.

Lemma B.1. *There exists a unique optimal private-sector time allocation decision l.*

Proof. The right-hand side of (B.1) is decreasing in l, whereas the left-hand side of (B.1) is increasing in l. By the intermediate value theorem, there must therefore be a unique value of l at which the first-order condition is satisfied. ■

Now consider how the optimal proportion of time spent in private employment changes with the variability of the weather-shock distribution. Suppose that the distribution of y in district B is a mean-preserving increase in spread of the distribution in district A. This means that farming is riskier in district B than in district A.

Proposition B.2. Households spend more in time in casual private employment in riskier districts than in less risky districts.

Proof. As households are risk averse, the expected utility from farming is lower in B, which implies that the right-hand side of (4) is larger in B for given values. The optimal value of l in A is therefore not the optimal time allocation in B. Since the right-hand side of (B.1) is decreasing in l, whereas the left-hand side is increasing in l, the optimal l in district B is larger than in district A. ■

The Model with NREGS

After NREGS is introduced, the program can be used both as an alternative source of employment regardless of the weather shock, and as an insurance tool after bad weather shocks. This alters the baseline model in two ways: The household now first makes a time-allocation decision among three alternatives: working for a big landowner as agricultural laborers in the private casual sector (l), working on the family farm (f_1), and taking up a NREGS job (n_1). After this decision has been made, as before a weather shock is realized that affects the payoff from farm work. The time originally allocated to farm work,f_1, can then be split between actually working on the farm, f_2, and taking up public employment in a NREGS project instead (n_2). After this decision, the period ends and the payoffs are realized. As before, the payoff from farm employment is y and the private-sector wage is w. The NREGS program wage is $\overline{w}$. The household again derives utility from the time spent in self-employment and from the total income earned.

The new household optimization problem at the beginning of the period is now given by

$$\max_{l,n_1} E[\alpha v(T-l-n_1-n_2^*)+(1-\alpha)u((T-l-n_1-n_2^*)y+n_2^*\overline{w}+lw+n_1\overline{w})]$$

where n_2^* is the best-response function of n_2 given y since the household can optimize the time spent working for NREGS and actually working on the family farm after the weather shock has occurred and y has been realized. Once a household chooses the fraction of time to spend on NREGS employment after the weather shock has occurred, l, n_1, and y are fixed. The household therefore chooses n_2 to maximize

$$\max_{n_2} \alpha v(T-l-n_1-n_2)+(1-\alpha)u((T-l-n_1-n_2)y+n_2\overline{w}+lw+n_1\overline{w})$$

leading to the first-order condition

$$\alpha v'(T-l-n_1-n_2)=(1-\alpha)u'((T-l-n_1-n_2)y+n_2\overline{w}+lw+n_1\overline{w})(\overline{w}-y) \quad \text{(B.2)}$$

Lemma B.3. *There exists a unique optimal amount of time spent in n_2 (NREGS employment as ex-post insurance) for a given y.*

Proof. The right-hand side of (B.2) is decreasing in n_2, whereas the left-hand side of (B.2) is increasing in n_2. By the intermediate value theorem, there must therefore be a unique value of n_2 at which the first-order condition is satisfied. ∎

Define the shock y_0 as the shock at which the first-order condition implies n_2=0. Then the first-order condition traces out the best-response function n_2^* for all weather shocks that imply a farming income of y_0 or less. For all larger values of y, the optimal n_2 is zero. Therefore, we have

$$n_2^* = \begin{cases} \text{implied } n_2 \text{ from (5)} & y \leq y_0; \\ 0 & y > y_0. \end{cases}$$

Knowing n_2^* and the distribution of y, at the beginning of the period the household needs to decide how much time to spend in the private sector, in NREGS employment, and in anticipated farming.

Lemma B.4. *A household will work either in private-sector work l or in ex-ante NREGS employment n_1, and will work in the job that pays more.*

Proof. l and n_1 are perfect substitutes for a household in terms of their contribution to household utility. Both are safe sources of employment that need to be committed to before the weather shock is realized. A household therefore maximizes utility by choosing the alternative that pays a higher wage. ∎

Define j as the amount of time spent working in the activity that pays the higher wage, such that

$$j = \begin{cases} n_1 & w \leq \overline{w}; \\ l & w > \overline{w}. \end{cases}$$

and define $\widetilde{w}$ analogously as the corresponding wage.

The household maximization problem can therefore be rewritten as

$$\max_j E[\alpha v((T - j - n_2^*)) + (1 - \alpha)u((T - j - n_2^*)y + n_2^*\overline{w} + j\widetilde{w})]$$

Working in the fact that the optimal n_2 is zero at large shocks, the problem can be rewritten as

$$\max_j \int\limits_{y \leq y_0} [\alpha v(T - j - n_2^*) + (1 - \alpha)u((T - j - n_2^*)y + n_2^*\overline{w} + j\widetilde{w})]g(y)dy$$

$$+ \int_{y > y_0} [\alpha v(T - j) + (1 - \alpha)u((T - j)y + j\widetilde{w})]g(y)dy$$

This leads to the first-order condition

$$\frac{\alpha}{1 - \alpha} \int\limits_{y \leq y_0} v'(T - j - n_2^*)(1 + \frac{\partial n_2^*}{\partial j})g(y)dy + \frac{\alpha}{1 - \alpha} v'(T - j)$$

$$- \int_{y > y_0} u'((T - j)y + j\widetilde{w})(\widetilde{w} - y)g(y)dy$$
$$= \int_{y \leq y_0} u'((T - j - n_2^*)y + n_2^*\overline{w} + j\widetilde{w})(\widetilde{w} - y + (\overline{w} - y)\tfrac{\partial n_2^*}{\partial j})g(y)dy \quad \text{(B.3)}$$

Lemma B.5. *A sufficient condition for the existence of a unique optimal amount of time spent in employment j is that agents are sufficiently risk averse.*

Proof. For an interior solution to be guaranteed, one side of (B.3) should be increasing and the other side decreasing in j. Some algebra shows that signing the partial derivatives on both sides is only possible if the sign of $\frac{\partial^2 n_2^*}{\partial j^2}$ is known. If it is positive, the derivative of left-hand side of (B.3) is positive, whereas each term of the right-hand side derivative is negative as long as

$$-\frac{u''((T-j-n_2^*)y+n_2^*\overline{w}+j\widetilde{w})}{u'((T-j-n_2^*)y+n_2^*\overline{w}+j\widetilde{w})} > \frac{(\overline{w}-y)(-\frac{\partial^2 n_2^*}{\partial j^2})}{(\widetilde{w}-y+(\overline{w}-y)\frac{\partial n_2^*}{\partial j})^2}$$

holds for all possible values of y.

Similarly, if the expression is negative, the derivative of the right-hand side of (B.3) is negative and all terms of the left-hand side derivative are positive as long as

$$-\frac{v''(T-j-n_2^*)}{v'(T-j-n_2^*)} > \frac{-\frac{\partial^2 n_2^*}{\partial j^2}}{(1+\frac{\partial n_2^*}{\partial j})^2}$$

holds for all possible values of y.

Under these conditions, there is a unique interior solution satisfying the first-order condition according to the intermediate value theorem. ■

Since $\frac{-u''(.)}{u'(.)}$ is the Arrow-Pratt measure of absolute risk aversion, these sufficient conditions mean intuitively that an agent needs to be risk averse 'enough'.

Notice how the sufficient conditions for a unique solution do not depend on the sign of $\frac{\partial n_2^*}{\partial j}$, which is ambiguous. Intuitively, how the time allocated to the ex-post NREGS employment responds to an increase in the time allocated to precautionary activity j depends on the attractiveness of the wage for j relative to the NREGS wage $\overline{w}$ and y. In other words, j only functions well as a precautionary savings tool if the paid wage in that activity is not too low relative to the payoffs that can be achieved through NREGS employment and farming after the weather shock is realized. A sufficient condition for j and n_2^* being substitutes for shocks $y \leq y_0$ is $\widetilde{w} \geq \overline{w}$.

A couple of predictions about the impact of NREGS follow from the model setup.

Proposition B.6. If the NREGS wage is high relative to the private-sector wage, the introduction of NREGS completely crowds out private-sector employment.

Proof. This follows directly from Lemma B.3 for $\overline{w} > w$. NREGS as a precautionary savings tool here directly replaces private-sector employment. ■

Proposition B.7. Even if the NREGS wage is low relative to the private-sector wage, the introduction of NREGS reduces the amount of time spent in private-sector employment under reasonable assumptions. Workers spend more time in farm work and, after bad income shocks, in NREGS employment instead.

Proof. This follows from comparing (B.1) and (B.3), where $j = l$ since $\overline{w} < w$. (B.3) can be re-written as

$$\alpha v'(T-l)$$

$$=$$

$$(1\text{-}\alpha)\int_{y\leq y_0} u'((T-l-n_2^*)y + n_2^*\overline{w} + lw)(w - y + (\overline{w}-y)\tfrac{\partial n_2^*}{\partial l})g(y)dy$$
$$+\ (1\text{-}\ \alpha)\int_{y>y_0} u'((T-l)y + lw)(w-y)g(y)dy$$
$$-\ \alpha\int_{y\leq y_0} v'(T-l-n_2^*)(1+\tfrac{\partial n_2^*}{\partial l})g(y)dy \quad \text{(B.4)}$$

The left-hand side of (B.4) is identical to the left-hand side of (B.1), but the first two terms of the right-hand side of (B.4) taken together are lower than the right-hand side of (B.1) since NREGS raises the expected utility at low y outcomes and therefore lowers the expected marginal utility for these shocks.

$\frac{\partial n_2^*}{\partial l}$ is negative since $w \geq \overline{w}$. Assume that n_2^* and l are relatively poor substitutes for each other such that $\frac{\partial n_2^*}{\partial l} > -1$ holds. That the substitutability of the two variables is less than 1 in absolute terms is intuitive since one is a precautionary savings tool whereas the other one functions as ex-post insurance. Then, all three terms of

(B.4) taken together are now smaller than the right-hand side of (B.1). This implies that the old time allocation l is no longer the optimal solution. Since the right-hand side of the equation above is decreasing in l whereas the left-hand side is increasing, this in turn implies that the new optimal l is lower than the old one. ■

Proposition B.8. NREGS take-up is low on average if the program primarily functions as a safety net tool.

Proof. This follows from Propositions B.6 and B.7. If NREGS is primarily used as a precautionary savings measure, NREGS employment crowds out private-sector employment and will be high. If NREGS mainly functions as insurance, then NREGS is only taken up after bad shocks to y, and will therefore be low in the absence of large negative aggregate shocks. ■

Proposition B.9. NREGS employment in riskier districts is higher than in less-risky districts.

Proof. Assume again that district B's distribution of y is a mean-preserving increase in spread of that in district A. This means that the probability of receiving a low draw $y < y_0$ is higher in district B than in district A, which implies that the expected amount of time spent in NREGS employment is higher in B than in A. ■

Proposition B.10. In riskier districts, private-employment decreases more than in less risky districts if NREGS is used as insurance.

Proof. This follows from Proposition B.7, combined with the following observation: The mean-preserving increase in spread of y will reduce the magnitude of the first term of the right-hand side of (B.4), but will increase both the second and the third term in absolute value. At very bad shocks, the absolute value of the third term goes towards infinity as $n_2^* \to 1$, whereas the second term goes towards a fixed value. For these large shocks, the right-hand side of (B.4) therefore decreases more in risky than in lower risk districts, which implies a larger reduction in l in risky than in less risky districts. ■

Extensions: NREGS Cap, Implementation Problems and Private-Sector Wage Variability

So far the model assumes that an agent can perfectly choose the amount of NREGS employment that is optimal for him, be it as a precautionary savings measure n_1 or as a safety net measure n_2. In reality, NREGS employment is officially capped at 100 days per household per year. This makes NREGS less attractive both as a risk-mitigation tool and as an ex post insurance mechanism, and will therefore attenuate the labor market impacts of NREGS predicted by the model. An implication of this feature is also that Proposition 10 may no longer hold: If the restriction on the maximum time spent in NREGS employment means that there is much less insurance after exceptionally bad weather shocks than in the absence of this rule, then households living in risky districts will reduce their time spent in private employment l less than agents in less risky districts.

In addition to the cap on NREGS employment, public-works programs in developing countries are often plagued by implementation problems like rationing of jobs or underpayment of wages due to corruption. This limits the amount of time that can be spent in NREGS employment even further in the case of rationing, and will reduce the actual wage received by program participants in the case of corruption. Both of these changes make NREGS less attractive than in the baseline model and therefore again attenuate the impacts NREGS has on labor-market outcomes.

The model also assumes that the private-sector wage is fixed regardless of the weather shock. If the private-sector wage also depends on the weather, private-sector employment is a less useful tool for risk mitigation than in the model, which increases the negative impacts NREGS has on private-sector employment.

APPENDIX C

Additional Tables and Figures (for Chapter 1)

Table C.0.1: NREGS Impact (TOT): Wages and Employment (Men and Women)

Specification	public	employment private	family	total	log private wage
Panel A: men					
Linear	0.0027	-0.0799	0.0579	-0.0157	-0.0093
	(0.0085)	(0.0508)	(0.0583)	(0.0417)	(0.0847)
Linear Flexible Slope	0.0025	-0.0805	0.0591	-0.0155	-0.0087
	(0.0086)	(0.0507)	(0.0576)	(0.0417)	(0.0853)
Quadratic	0.0017	-0.0875*	0.0696	-0.0130	-0.0165
	(0.0089)	(0.0528)	(0.0608)	(0.0439)	(0.0871)
Quadratic Flexible Slope	0.0082	-0.1056*	0.0603	-0.0328	-0.1203
	(0.0092)	(0.0567)	(0.0631)	(0.0438)	(0.1071)
N	1063	1063	1063	1063	1007
outcome mean	0.0069	0.3279	0.4846	0.8195	4.1208
Panel B: women					
Linear	0.0030	-0.0081	0.0376	0.0318	0.0091
	(0.0100)	(0.0376)	(0.0594)	(0.0683)	(0.1439)
Linear Flexible Slope	0.0030	-0.0078	0.0370	0.0311	0.0063
	(0.0101)	(0.0379)	(0.0596)	(0.0684)	(0.1410)
Quadratic	0.0035	-0.0049	0.0255	0.0239	0.0115
	(0.0106)	(0.0388)	(0.0605)	(0.0698)	(0.1498)
Quadratic Flexible Slope	0.0019	-0.0161	0.0906	0.0792	-0.1220
	(0.0118)	(0.0394)	(0.0687)	(0.0800)	(0.1887)
N	1063	1063	1063	1063	656
outcome mean	0.0053	0.1309	0.2285	0.3647	3.6488

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is a district in a given season. An employment outcome is the proportion of working-age adults (18-60 years) with at most secondary education in rural areas working in a given type of employment in the last 7 days. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment. Actual treatment with NREGS is instrumented with the predicted treatment status. The first-stage F-statistic is 114.80 for the linear, 82.44 for the linear flexible, 100.53 for the quadratic, and 56.82 for the flexible quadratic specifications.

Table C.0.2: NREGS Impacts on Other Outcomes: Expenditures, Total Wage, Remittances (Men)

Specification	log per-capita expenditures	log total wage	log remittances
Panel A: overall sample			
Linear	0.0195	-0.0050	-0.0065
	(0.0346)	(0.0375)	(0.1028)
Linear Flexible Slope	0.0199	-0.0050	-0.0069
	(0.0345)	(0.0375)	(0.1027)
Quadratic	0.0219	-0.0083	-0.0250
	(0.0350)	(0.0372)	(0.1031)
Quadratic Flexible Slope	0.0275	-0.0088	-0.0141
	(0.0488)	(0.0491)	(0.1390)
N	1063	1011	1030
outcome mean	6.4798	4.1267	9.1621
Panel B: rainfall shock			
Linear	0.0115	0.0277	-0.0192
	(0.0391)	(0.0510)	(0.1355)
NREGS*negative shock	-0.0382	-0.0398	0.0181
	(0.0506)	(0.0675)	(0.1673)
Linear Flexible Slope	0.0121	0.0277	-0.0206
	(0.0392)	(0.0510)	(0.1358)
NREGS*negative shock	-0.0394	-0.0394	0.0207
	(0.0501)	(0.0681)	(0.1665)
Quadratic	0.0156	0.0232	-0.0422
	(0.0394)	(0.0504)	(0.1358)
NREGS*negative shock	-0.0389	-0.0378	0.0226
	(0.0502)	(0.0680)	(0.1647)
Quadratic Flexible Slope	0.0435	0.0077	0.0402
	(0.0581)	(0.0678)	(0.1709)
NREGS*negative shock	-0.0476	-0.0419	-0.0267
	(0.0516)	(0.0699)	(0.1658)
N	532	508	514
outcome mean	6.5160	4.1870	9.2775

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is a district in a given season. Parametric regressions with different levels of flexibility are reported. NREGS is the predicted treatment status. The log total wage is conditional on having earned a positive wage.

Table C.0.3: NREGS Impact (Donut Hole Approach): Wages and Employment (Men and Women)

Specification	employment public	employment private	employment family	employment total	log private wage
Panel A: men					
Linear	-0.0001	-0.0408*	0.0511*	0.0132	0.0014
	(0.0040)	(0.0221)	(0.0269)	(0.0182)	(0.0443)
Linear Flexible Slope	-0.0006	-0.0408*	0.0537**	0.0153	0.0036
	(0.0040)	(0.0219)	(0.0265)	(0.0181)	(0.0439)
Quadratic	-0.0008	-0.0431**	0.0578**	0.0167	-0.0007
	(0.0040)	(0.0218)	(0.0269)	(0.0183)	(0.0440)
Quadratic Flexible Slope	0.0010	-0.0462	0.0432	0.0027	-0.0171
	(0.0053)	(0.0283)	(0.0350)	(0.0216)	(0.0554)
N	952	952	952	952	897
outcome mean	0.0062	0.3225	0.4949	0.8236	4.1252
Panel B: women					
Linear	-0.0043	-0.0210	0.0243	-0.0006	-0.0285
	(0.0033)	(0.0168)	(0.0288)	(0.0340)	(0.0678)
Linear Flexible Slope	-0.0046	-0.0206	0.0183	-0.0063	-0.0286
	(0.0032)	(0.0167)	(0.0284)	(0.0335)	(0.0670)
Quadratic	-0.0047	-0.0201	0.0120	-0.0120	-0.0322
	(0.0031)	(0.0167)	(0.0284)	(0.0335)	(0.0680)
Quadratic Flexible Slope	-0.0073*	-0.0300	0.0336	-0.0009	-0.1130
	(0.0041)	(0.0216)	(0.0351)	(0.0398)	(0.0927)
N	952	952	952	952	576
outcome mean	0.0042	0.1275	0.2315	0.3632	3.6489

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is a district in a given season. An employment outcome is the proportion of working-age adults (18-60 years) with at most secondary education in rural areas working in a given type of employment in the last 7 days. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment. Observations with a state-specific rank between -1 and 1 are dropped.

Table C.0.4: NREGS Impact (Index Running Variable): Wages and Employment (Men and Women)

Specification	public	employment private	family	total	log private wage
Panel A: men					
Linear	0.0015	-0.0135	0.0060	-0.0041	-0.0334
	(0.0032)	(0.0174)	(0.0201)	(0.0151)	(0.0324)
Linear Flexible Slope	0.0007	-0.0145	0.0123	0.0005	-0.0311
	(0.0033)	(0.0172)	(0.0197)	(0.0155)	(0.0329)
Quadratic	0.0007	-0.0201	0.0230	0.0055	-0.0324
	(0.0037)	(0.0181)	(0.0213)	(0.0163)	(0.0356)
Quadratic Flexible Slope	0.0000	-0.0353**	0.0338	-0.0183	-0.0044
	(0.0045)	(0.0178)	(0.0250)	(0.0176)	(0.0383)
N	1063	1063	1063	1063	1007
outcome mean	0.0069	0.3279	0.4846	0.8195	4.1212
Panel B: women					
Linear	0.0023	-0.0077	0.0231	0.0164	-0.0379
	(0.0048)	(0.0132)	(0.0221)	(0.0251)	(0.0551)
Linear Flexible Slope	0.0021	-0.0046	0.0284	0.0248	-0.0384
	(0.0050)	(0.0132)	(0.0216)	(0.0243)	(0.0550)
Quadratic	0.0018	0.0012	0.0482**	0.0504*	-0.0412
	(0.0053)	(0.0141)	(0.0222)	(0.0257)	(0.0581)
Quadratic Flexible Slope	0.0006	-0.0191	0.0417*	0.0248	-0.0544
	(0.0057)	(0.0138)	(0.0253)	(0.0294)	(0.0708)
N	1063	1063	1063	1063	656
outcome mean	0.0053	0.1309	0.2285	0.3647	3.6488

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is a district in a given season. An employment outcome is the proportion of working-age adults (18-60 years) with at most secondary education in rural areas working in a given type of employment in the last 7 days. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment. NREGS is the predicted treatment status. The running variable is the normalized state-specific poverty index value of a district rather than its rank.

Table C.0.5: NREGS Impact (Individual Level): Wages and Employment (Men and Women)

Specification	public	employment private	family	total	log private wage
Panel A: men					
Linear	-0.0025	-0.0286	0.0341	0.0031	0.0472
	((0.0039))	(0.0208)	(0.0253)	(0.0185)	(0.0417)
Linear Flexible Slope	-0.0024	-0.0286	0.0339	0.0028	0.0468
	(0.0039)	(0.0209)	(0.0250)	(0.0182)	(0.0416)
Quadratic	-0.0031	-0.0296	0.0391	0.0065	0.0454
	(0.0040)	(0.0205)	(0.0251)	(0.0184)	(0.0418)
Quadratic Flexible Slope	-0.0031	-0.0531**	0.0595*	0.0033	0.0441
	(0.0056)	(0.0252)	(0.0320)	(0.0211)	(0.0480)
N	37224	37224	37224	37224	12062
outcome mean	0.0082	0.3261	0.4756	0.8099	4.0473
Panel B: women					
Linear	0.0009	-0.0025	0.0254	0.0238	-0.0231
	(0.0036)	(0.0171)	(0.0251)	(0.0296)	(0.0528)
Linear Flexible Slope	0.0010	-0.0032	0.0274	0.0252	-0.0220
	(0.0035)	(0.0172)	(0.0250)	(0.0295)	(0.0537)
Quadratic	0.0008	-0.0015	0.0199	0.0192	-0.0257
	(0.0036)	(0.0172)	(0.0248)	(0.0295)	(0.0532)
Quadratic Flexible Slope	-0.0027	-0.0125	0.0409	0.0257	-0.0585
	(0.0041)	(0.0206)	(0.0327)	(0.0381)	(0.0606)
N	41978	41978	41978	41978	5339
outcome mean	0.0046	0.1234	0.2106	0.3385	3.5428

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is a working-age adult (18-60 years) with at most secondary education living in rural areas. The employment outcome variables are indicator variables for having worked in a given employment type in the last 7 days. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment.

Table C.0.6: Seasonality of Wages and Employment (Men and Women)

Specification	public	employment private	family	total	log private wage
Panel A: men					
Linear	-0.0025	-0.0392*	0.0285	-0.0115	-0.0249
	(0.0036)	(0.0229)	(0.0266)	(0.0190)	(0.0416)
NREGS*dry season	0.0073	0.0081	-0.0060	0.0091	0.0404
	(0.0053)	(0.0190)	(0.0208)	(0.0155)	(0.0367)
Linear Flexible Slope	-0.0025	-0.0424*	0.0346	-0.0090	-0.0336
	(0.0034)	(0.0237)	(0.0274)	(0.0194)	(0.0442)
NREGS*dry season	0.0071	0.0144	-0.0176	0.0045	0.0577
	(0.0066)	(0.0235)	(0.0257)	(0.0209)	(0.0463)
Quadratic	-0.0030	-0.0410	0.0324	-0.0100	-0.0283
	(0.0036)	(0.0224)	(0.0262)	(0.0192)	(0.0415)
NREGS*dry season	0.0073	0.0081	-0.0061	0.0090	0.0408
	(0.0053)	(0.0190)	(0.0208)	(0.0155)	(0.0367)
Quadratic Flexible Slope	-0.0004	-0.0654**	0.0470	-0.0148	-0.0378
	(0.0042)	(0.0296)	(0.0357)	(0.0242)	(0.0583)
NREGS*dry season	0.0042	0.0261	-0.0328	-0.0034	0.0335
	(0.0065)	(0.0262)	(0.0297)	(0.0279)	(0.0591)
N	1063	1063	1063	1063	1007
outcome mean	0.0069	0.3279	0.4846	0.8195	4.1212
Panel B: women					
Linear	-0.0017	-0.0121	0.0163	0.0015	0.0923
	(0.0033)	(0.0177)	(0.0283)	(0.0325)	(0.0730)
NREGS*dry season	0.0060	0.0170	0.0012	0.0254	-0.1726***
	(0.0047)	(0.0136)	(0.0186)	(0.0205)	(0.0592)
Linear Flexible Slope	-0.0024	-0.0170	0.0181	-0.0027	0.0845
	(0.0031)	(0.0187)	(0.0294)	(0.0335)	(0.0760)
NREGS*dry season	0.0075	0.0271	-0.0033	0.0331	-0.1569**
	(0.0057)	(0.0170)	(0.0233)	(0.0259)	(0.0724)
Quadratic	-0.0016	-0.0106	0.0105	-0.0024	0.0932
	(0.0034)	(0.0176)	(0.0279)	(0.0321)	(0.0729)
NREGS*dry season	0.0060	0.0170	0.0012	0.0254	-0.1726***
	(0.0047)	(0.0136)	(0.0186)	(0.0205)	(0.0593)
Quadratic Flexible Slope	-0.0025	-0.0253	0.0334	0.0070	0.0159
	(0.0031)	(0.0234)	(0.0378)	(0.0431)	(0.1004)
NREGS*dry season	-0.0001	0.0358	0.0017	0.0387	-0.1668**
	(0.0046)	(0.0195)	(0.0278)	(0.0310)	(0.0838)
N	1063	1063	1063	1063	656
outcome mean	0.0053	0.1309	0.2285	0.3647	3.6488

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. Parametric regressions with different levels of flexibility are reported. The log private wage in column 5 is conditional on private employment. NREGS is the predicted treatment status. July through December roughly correspond to harvesting and planting seasons for most crops in wide parts of the country and are treated as the agricultural main season. January through June are the agricultural off-season or dry season.

Table C.0.7: NREGS Impact: Difference-in-Difference Estimates (Men and Women)

	employment				log private
	public	private	family	total	wage
Panel A: men					
Actual Treatment					
NREGS*post period	0.0083**	0.0060	-0.0344**	-0.0201	0.0100
	(0.0036)	(0.0160)	(0.0173)	(0.0146)	(0.0297)
NREGS	0.0019	-0.0019	0.0319**	0.0317**	-0.0741***
	(0.0018)	(0.0122)	(0.0137)	(0.0135)	(0.0297)
post period	0.0014	0.0147	-0.0555***	-0.0394***	0.0832***
	(0.0009)	(0.0103)	(0.0103)	(0.0101)	(0.0179)
Predicted Treatment					
NREGS*post period	0.0056*	0.0141	-0.0405**	-0.0207	-0.0075
	(0.0031)	(0.0159)	(0.0165)	(0.0144)	(0.0289)
NREGS	-0.0022	-0.0192	0.0404***	0.0190	-0.0664**
	(0.0016)	(0.0121)	(0.0142)	(0.0143)	(0.0283)
post period	0.0022	0.0114	-0.0523***	-0.0387***	0.0900***
	(0.0016)	(0.0104)	(0.0110)	(0.0106)	(0.0188)
N	2126	2126	2126	2126	2014
outcome mean	0.0047	0.3194	0.5188	0.8429	4.08
Panel B: women					
Actual Treatment					
NREGS*post period	0.0075**	0.0035	0.0049	0.0159	-0.0126
	(0.0035)	(0.0109)	(0.0174)	(0.0200)	(0.0461)
NREGS	0.0028	0.0115	-0.0167	-0.0023	-0.0458
	(0.0019)	(0.0102)	(0.0186)	(0.0218)	(0.0369)
post period	0.0007	-0.0104	-0.0793***	-0.0890***	-0.0058
	(0.0005)	(0.0064)	(0.0119)	(0.0131)	(0.0288)
Predicted Treatment					
NREGS*post period	0.0043	0.0073	0.0159	0.0275	-0.0249
	(0.0031)	(0.0104)	(0.0173)	(0.0194)	(0.0451)
NREGS	-0.0001	0.0176*	0.0073	0.0248	-0.1013***
	(0.0014)	(0.0099)	(0.0198)	(0.0224)	(0.0358)
post period	0.0018	-0.0119*	-0.0837***	-0.0939***	0.0004
	(0.0012)	(0.0069)	(0.0122)	(0.0137)	(0.0305)
N	2126	2126	2126	2126	1312
outcome mean	0.0036	0.1354	0.2672	0.4062	3.64

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is a district in a given season. An employment outcome is the proportion of working-age adults (18-60 years) with at most secondary education in rural areas working in a given type of employment in the last 7 days. The log private wage in column 5 is conditional on private employment. NREGS is the actual or the predicted treatment status of a district.

Figure C.0.1: Public employment men

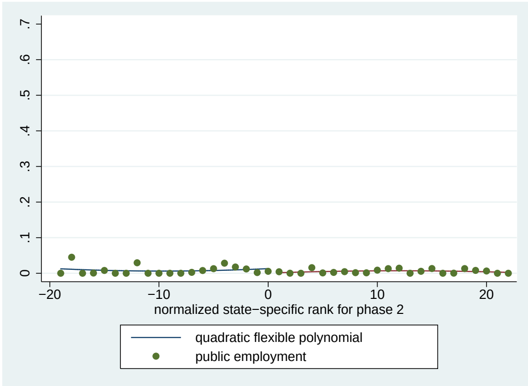


Figure C.0.2: Private employment men

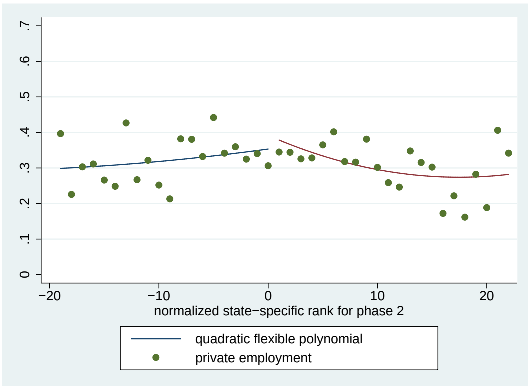
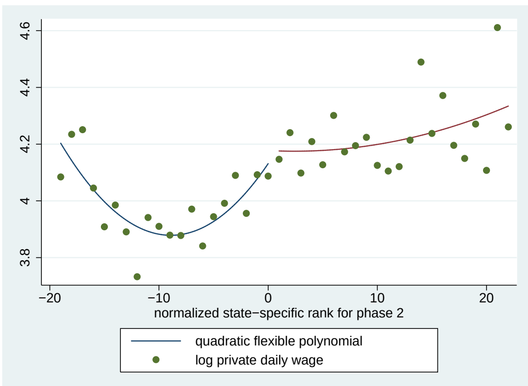


Figure C.0.3: Log daily private wage men



Note: An observation is the average outcome at a given rank, so the bin size is 1. Fitted curves are quadratic polynomials.

Figure C.0.4: Public employment women

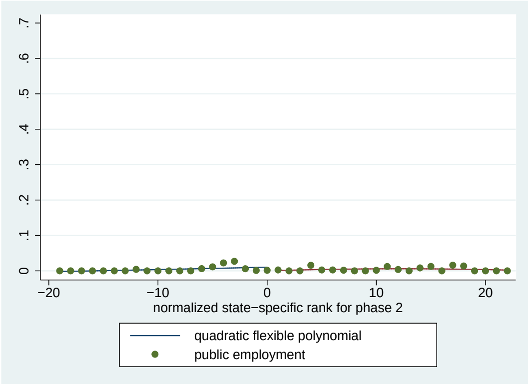


Figure C.0.5: Private employment women

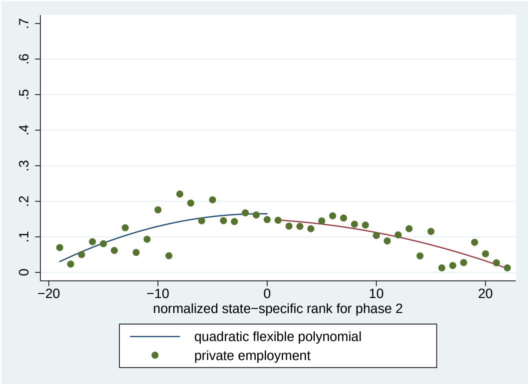
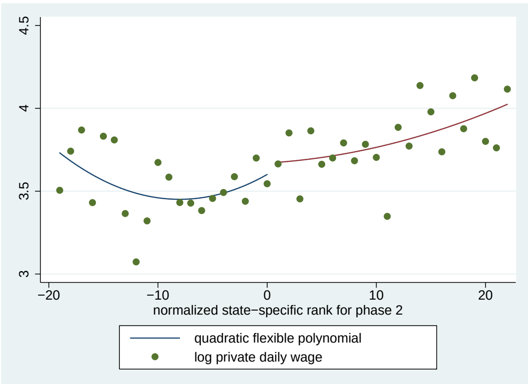
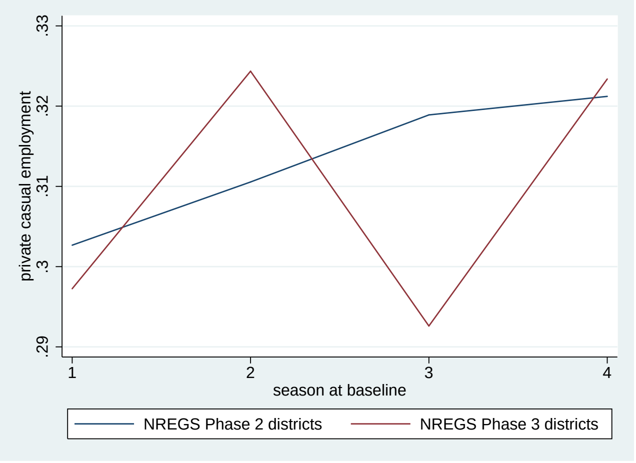


Figure C.0.6: Log daily private wage women



Note: An observation is the average outcome at a given rank, so the bin size is 1. Fitted curves are quadratic polynomials.

Figure C.0.7: Private employment trends for men at baseline



Note: Figure plots average casual private-sector employment in Phase 2 and Phase 3 districts for each sub-round in the baseline data (July 2004-June 2005).

APPENDIX D

RD versus DID estimates (for Chapter 1)

As Figure C.0.7 shows, the parallel trend assumption underlying the difference-in-difference approach is violated since Phase 2 and Phase 3 districts have differential trends in casual private employment at baseline. This may lead to substantial bias in the program impacts. Additionally, difference-in-difference specifications estimate the average treatment effect across all observations rather than the treatment effect for districts close to the treatment cutoff as in the regression discontinuity design. Appendix Table C.0.7 estimates the labor market effects for the sample used for the regression discontinuity analysis using a difference-in-difference approach. Panels A and B report the results using actual treatment as well as predicted treatment status from the algorithm for men and women, respectively. The table shows that in my sample even the difference-in-difference analysis does not generate substantial wage increases. Public-sector employment is statiscally significantly higher in NREGS districts after the program is introduced, although the estimates are again modest relative to the proposed scale of the scheme. Somewhat surprisingly the coefficients for private employment and family work for men are flipped in Table C.0.7 compared to the regression discontinuity results in Table 1.4.

Overall, Table C.0.7 shows that the different results found in this paper are not just due to the choice of the empirical specification. Theoretically, such differences could be driven by differences in the sample composition, the data used, or the chosen specifications. Two impact evaluation papers using DID approaches, Azam (2012) and Imbert and Papp (2013), use almost the same data that is used in this paper, so the data itself is unlikely to be the issue. A replication of the DID results of these two papers (not reported here) allows me to analyze what drives the differences. The results reveal that the choice of the empirical specification seems to be the most important explanation for the differences between Table C.0.7 and the results in Imbert and Papp (2013), although they also use an additional dataset to increase their sample size that improves the precision of the estimates: Their main specification directly estimates the seasonal impacts of NREGS and demonstrates that the impacts are concentrated during the agricultural off-season, whereas my paper focuses on the overall impact of the program.[1] A similar reason explains the wage increases found in the Berg et al. (2012) paper that uses a different dataset on wages than this paper. Their paper, too, does not find important short-run wage increases in the general sample, but documents that wage increases exist in areas with high implementation quality.

The differences between Table C.0.7 and the results in Azam (2012), on the other hand, seem to be mostly due to the choice of the sample: Azam (2012) restricts the sample of casual workers rather than the general working-age population considered in this paper, which explains the higher public employment impacts since NREGS employment should be especially attractive for this group. The wage increases documented for this group of workers is present at the individual level, but disappears once wage impacts are estimated at the district level. Given that who remains a

[1]Estimating the seasonal impact of NREGS with the RD specification for men shows that the results are broadly consistent with their results, although the interaction effects in my specification are statistically insignficant. The RD results still do not show any evidence of wage increases even in the agricultural off-season, however, and the interaction effect for women is even statistically significantly negative. The seasonality results are reported in appendix table C.0.6.

casual worker after the introduction of NREGS may itself be an endogenous decision affected by the program, it is also difficult to generalize these results.

APPENDIX E

Additional Tables and Figures (for Chapter 2)

.

Table E.0.1: Who Initiates the Attacks and Who is Affected (Per Capita)

	Panel A: Who Initiates the Attacks		
Specification	**Police on Maoist**	**Maoist on Police**	**Maoist on Civilians**
Linear	0.975**	0.133	0.491**
	(0.412)	(0.101)	(0.227)
R-squared	0.345	0.194	0.522
Linear Flexible Slope	0.741**	0.153	0.254
	(0.375)	(0.107)	(0.171)
R-squared	0.351	0.194	0.524
Quadratic	0.779*	0.102	0.436**
	(0.424)	(0.0957)	(0.208)
R-squared	0.350	0.194	0.522
Outcome Mean	0.578	0.329	0.753

	Panel B:	**Who is Killed**	
Specification	**Civilians Killed**	**Police Killed**	**Maoists Killed**
Linear	0.531	-0.232	2.051***
	(0.716)	(0.632)	(0.755)
R-squared	0.398	0.206	0.427
Linear Flexible Slope	0.358	-0.148	1.613**
	(0.561)	(0.372)	(0.763)
R-squared	0.398	0.206	0.429
Quadratic	0.595	0.201	2.137**
	(0.706)	(0.637)	(0.897)
R-squared	0.398	0.206	0.426
Outcome Mean	1.099	0.958	0.956

Results are per-10 million people (based on population counts from the 2001 Census).
Panel A has results for ”who initiates the attacks and against whom.” Panel B has results for ”who is killed”
Controls include baseline averages of each dependent variable and police force changes. Unit of observation is district-month. Regressions contain 3192 observations in 228 clusters.

.

Table E.0.2: Other Types of Crime and Violence: Phase 1

	total crimes	murder	kidnapping	theft	burglary	riots
Linear	1,287 (945.6)	13.63 (39.51)	33.93 (30.07)	-30.27 (174.8)	-28.09 (93.75)	-107.2 (115.4)
Linear Flexible	754.8 (924.1)	28.72 (35.76)	28.70 (29.43)	-39.61 (189.1)	-68.79 (93.69)	-86.29 (107.9)
Quadratic	988.9 (1,134)	55.56 (48.31)	29.78 (34.59)	-36.06 (220.3)	-42.57 (115.3)	-122.9 (129.8)
Outcome mean	2768.17	57.23	38.10	325.04	122.18	105.36

Regressions contains 225 observations, where the unit of observation is a district.
Source: Home Ministry of India

.

Table E.0.3: Donut Hole and Varying the Bandwidth

	Panel A:	**Donut**	**Hole**	
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.562*	0.475*	0.0869**	0.287**
	(0.311)	(0.283)	(0.0405)	(0.117)
R-squared	0.494	0.458	0.407	0.484
Linear Flexible Slope	0.429	0.423	0.0598*	0.203**
	(0.280)	(0.311)	(0.0316)	(0.0943)
R-squared	0.495	0.459	0.412	0.498
Quadratic	0.628*	0.615*	0.105**	0.281**
	(0.359)	(0.365)	(0.0453)	(0.127)
R-squared	0.494	0.456	0.404	0.485
	Panel B:	**Varying**	**Bandwidth**	
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
-x < rank ≤ x				
x=10	0.760*	0.853*	0.133**	0.312*
	(0.447)	(0.454)	(0.0614)	(0.161)
R-squared	0.496	0.458	0.422	0.484
x=9	0.851*	0.889*	0.145**	0.316*
	(0.481)	(0.498)	(0.0660)	(0.174)
R-squared	0.499	0.465	0.443	0.487
x=8	0.884*	0.933*	0.152**	0.326*
	(0.523)	(0.539)	(0.0724)	(0.190)
R-squared	0.499	0.468	0.445	0.491
Outcome Mean	0.580	0.263	0.035	0.170

Panel A produces the Donut Hole results - this tackles measurement error by dropping districts closest to the cutoff.
Panel B varies the bandwidth close to the cutoff, where the bandwidth size is "x." The results presented are for the linear specification where the slope is flexible on either side of the cutoff.
Controls include baseline averages of each dependent variable and police force changes.
Unit of observation is district-month.
"Affected Persons" indicates number of persons killed, injured, abducted or captured. "Fatalities" indicates total number of deaths. "Major Incidents" indicates number of 'Major Incidents' as coded by the SATP website. "Total Incidents" is number of total Maoist-related incidents.

.

Table E.0.4: Intent-to-Treat (ITT) and Without Police Controls

	Panel A:	**ITT**		
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.287**	0.270**	0.0470***	0.124**
	(0.133)	(0.123)	(0.0176)	(0.0499)
R-squared	0.489	0.445	0.393	0.480
Linear Flexible Slope	0.317*	0.309*	0.0486**	0.132**
	(0.166)	(0.166)	(0.0211)	(0.0602)
R-squared	0.489	0.445	0.393	0.480
Quadratic	0.335**	0.348**	0.0578***	0.129**
	(0.162)	(0.161)	(0.0215)	(0.0594)
R-squared	0.489	0.445	0.393	0.480
	Panel B:	**Without**	**Police**	
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	0.559*	0.499*	0.0945**	0.240**
	(0.298)	(0.280)	(0.0394)	(0.109)
R-squared	0.486	0.430	0.378	0.446
Linear Flexible Slope	0.541*	0.510	0.0806**	0.209**
	(0.310)	(0.315)	(0.0382)	(0.106)
R-squared	0.486	0.429	0.381	0.452
Quadratic	0.721**	0.738**	0.121**	0.261**
	(0.362)	(0.371)	(0.0486)	(0.133)
R-squared	0.485	0.425	0.371	0.442
Outcome Mean	0.580	0.263	0.035	0.170

Panel A shows the Intent-to-Treat impacts (the reduced form results). Panel B shows the results without controlling for changes to the police force.
Controls include baseline averages of each dependent variable. Unit of observation is district-month. Regressions contain 3192 observations in 228 clusters.
"Affected Persons" indicates number of persons killed, injured, abducted or captured. "Fatalities" indicates total number of deaths. "Major Incidents" indicates number of 'Major Incidents' as coded by the SATP website. "Total Incidents" is number of total Maoist-related incidents.

.

Table E.0.5: Non RD Specifications: Count Data and Difference-in-Differences

	Panel A:	**Count Data**		
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Linear	2.207**	2.443***	2.552***	1.619**
	(1.014)	(0.881)	(0.618)	(0.760)
Linear Flexible Slope	2.125***	2.306***	2.108***	1.105*
	(0.733)	(0.634)	(0.483)	(0.596)
Quadratic	2.503**	2.466***	2.204***	1.368
	(1.066)	(0.723)	(0.551)	(0.931)

	Panel B:	**Difference in**	**Differences**	
Specification	**Affected Persons**	**Fatalities**	**Major Incidents**	**Total Incidents**
Intent-to-Treat	0.223*	0.108	0.0154	0.107**
	(0.132)	(0.0982)	(0.0103)	(0.0512)
R-squared	0.428	0.395	0.359	0.478
Actual Treatment	0.303**	0.0970	0.0227*	0.0556**
	(0.151)	(0.101)	(0.0129)	(0.0260)
R-squared	0.345	0.355	0.319	0.445
Outcome Mean	0.580	0.263	0.035	0.170

Panel A shows the results using a Count Data Model. The model used in this table is the Zero-Inflated Poisson model, where the zeros are predicted using pre-treatment averages of the dependent variable. The results are very similar using Hurdle Models (Logit-Poisson).
Panel B shows the Difference-in-Differences results. The Intent-to-Treat results assigns treatment status to districts who should have received NREGS, whereas the 'Actual Treatment' row assigns treatment status to districts that actually received NREGS.
Controls include baseline averages of each dependent variable. Unit of observation is district-month. Regressions contain 3192 observations in 228 clusters.
"Affected Persons" indicates number of persons killed, injured, abducted or captured. "Fatalities" indicates total number of deaths. "Major Incidents" indicates number of 'Major Incidents' as coded by the SATP website. "Total Incidents" is number of total Maoist-related incidents.

APPENDIX F

Additional Tables and Figures (for Chapter 3)

Table F.0.1: Baseline Pre-Treatment Results (Phase 1 vs Phase 2)

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.0673	-.0711	.0212	-.0390	-7.20*	-7.48**	.38	-.51	-.0327
(Linear)	(.0819)	(.0865)	(.0340)	(.0731)	(3.75)	(3.64)	(1.74)	(3.21)	(.0247)
NREGS Phase 1	-.0372	-.0474	.0275	-.0344	-4.42	-5.01	.76	-.89	-.0151
(Linear flexible)	(.0821)	(.0880)	(.0348)	(.0729)	(3.71)	(3.69)	(1.79)	(3.23)	(.0250)
NREGS Phase 1	-.0039	-.0108	.0315	-.0140	-1.55	-1.74	1.05	-.34	-.0038
(Quadratic)	(.0854)	(.0921)	(.0368)	(.0741)	(3.86)	(3.82)	(1.88)	(3.35)	(.0265)
N	378	378	378	378	378	378	378	378	378

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2004 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.2: Baseline Pre-Treatment Results (Phase 2 vs Phase 3)

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 3	-.0202	-.0016	-.0189	-.0383	-4.38	-3.64	-.40	1.52	-.0685***
(Linear)	(.0813)	(.0895)	(.0266)	(.0766)	(3.37)	(3.16)	(1.44)	(3.47)	(.0225)
NREGS Phase 3	.0692	.0990	-.0076	-.0424	1.35	2.77	.53	.36	-.0312
(Linear flexible)	(.0883)	(.0949)	(.0301)	(.0898)	(3.76)	(3.26)	(1.64)	(3.83)	(.0225)
NREGS Phase 3	.0805	.1279	.0110	-.0658	1.32	4.02	1.43	-1.55	-.0092
(Quadratic)	(.0932)	(.0997)	(.0294)	(.0973)	(3.97)	(3.43)	(1.62)	(4.12)	(.0236)
N	378	378	378	378	378	378	378	378	378

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2004 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise. Star state is an indicator variable equal to 1 if a state was identified as a high implementation quality state, and 0 otherwise. The five star states are Andhra Pradesh, Chhattisgarh, Madhya Pradesh, Rajasthan and Tamil Nadu.

Table F.0.3: Baseline Pre-Treatment Results Incumbent

	incumbent won	Phase 1 vs 2 vote share	incumbent won	Phase 2 vs 3 vote share
	(1)	(2)	(3)	(4)
NREGS	.0791	.96	-.0525	-3.03
(Linear)	(.0817)	(3.35)	(.0833)	(3.45)
NREGS	.0797	1.21	-.0488	-3.08
(Linear flexible)	(.0842)	(3.40)	(.0943)	(3.92)
NREGS	.0966	1.52	-.0671	-3.71
(Quadratic)	(.0879)	(3.53)	(.1003)	(4.19)
N	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2004 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise. An incumbent is an individual who won a plurality of votes in his/her constituency in the 2004 general elections and who contested the 2009 elections in any constituency.

Table F.0.4: Donut Hole Election Results Phase 1 vs Phase 2

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.1830*	-.2057**	.0552*	-.0616	-7.90*	-9.29**	.64	-2.38	-.0343
(Linear)	(.0953)	(.0976)	(.0301)	(.0740)	(4.06)	(4.26)	(1.95)	(3.56)	(.0330)
NREGS Phase 1	-.0915	-.1118	.0657*	-.0741	-3.91	-6.14	1.76	-3.44	.0288
(Linear flexible)	(.1019)	(.1044)	(.0394)	(.0761)	(4.35)	(4.69)	(2.24)	(3.67)	(.0319)
NREGS Phase 1	-.0672	-.0646	.0710*	-.0707	-2.35	-3.45	1.79	-2.71	.0358
(Quadratic)	(.1071)	(.1096)	(.0402)	(.0788)	(4.59)	(4.95)	(2.36)	(3.82)	(.0352)
N	354	354	354	354	354	354	354	354	354

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.5: Donut Hole Election Results Phase 2 vs Phase 3

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 3	-.0893	-.0487	-.0166	.0375	-1.52	4.07	-1.50	.24	-.1106***
(Linear)	(.0903)	(.0934)	(.0189)	(.0761)	(3.73)	(3.61)	(1.38)	(3.51)	(.0282)
NREGS Phase 3	-.0126	.0476	-.0146	.0199	1.87	8.55**	-.64	-1.63	-.0483*
(Linear flexible)	(.0996)	(.1044)	(.0190)	(.0892)	(4.09)	(3.69)	(1.47)	(3.91)	(.0274)
NREGS Phase 3	.0209	.0924	.0034	-.0184	1.80	9.72**	.25	-4.41	-.0227
(Quadratic)	(.1084)	(.1131)	(.0187)	(.0991)	(4.43)	(4.01)	(1.53)	(4.27)	(.0295)
N	375	375	375	375	375	375	375	375	375

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.6: Donut Hole Star State Results Phase 1 vs Phase 2

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.2783***	-.3065***	.0769*	-.0022	-10.09**	-11.71**	.18	1.84	-.0846**
(Linear)	(.1022)	(.1036)	(.0399)	(.0845)	(4.53)	(4.60)	(2.38)	(3.79)	(.0377)
NREGS Phase 1*star	.2451*	.2330**	-.0604**	-.1811*	4.72	4.11	1.75	-13.47***	.1467***
	(.1265)	(.1162)	(.0306)	(.0954)	(4.82)	(4.71)	(2.19)	(4.86)	(.0380)
NREGS Phase 1	-.1941*	-.2299**	.0974*	-.0058	-5.99	-9.20*	1.67	1.17	-.0150
(Linear flexible)	(.1102)	(.1118)	(.0539)	(.0877)	(5.06)	(5.30)	(2.88)	(3.99)	(.0383)
NREGS Phase 1*star	.1981	.1902	-.0719*	-.1791*	2.43	2.70	.91	-13.10***	.1078***
	(.1284)	(.1192)	(.0376)	(.0948)	(4.99)	(5.01)	(2.41)	(4.92)	(.0375)
NREGS Phase 1	-.1844	-.1994*	.1044*	-.0001	-4.95	-7.00	1.67	2.05	-.0144
(Quadratic)	(.1153)	(.1169)	(.0546)	(.0905)	(5.29)	(5.58)	(2.99)	(4.18)	(.0418)
NREGS Phase 1*star	.2108	.1938	-.0705**	-.1818*	2.84	2.39	1.20	-13.55***	.1210***
	(.1281)	(.1180)	(.0353)	(.0949)	(4.99)	(4.96)	(2.35)	(4.91)	(.0378)
N	354	354	354	354	354	354	354	354	354

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise. Star state is an indicator variable equal to 1 if a state was identified as a high implementation quality state, and 0 otherwise. The five star states are Andhra Pradesh, Chhattisgarh, Madhya Pradesh, Rajasthan and Tamil Nadu.

Table F.0.7: Donut Hole Star State Results Phase 2 vs Phase 3

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 3	.0515	.0617	-.0228	-.0102	4.66	7.99*	-1.38	-4.11	-.0582*
(Linear)	(.1022)	(.1061)	(.0237)	(.0849)	(4.19)	(4.15)	(1.76)	(3.86)	(.0346)
NREGS Phase 3*star	-.2804**	-.1237	.0079	.1323	-12.51***	-2.41	-.94	12.89***	-.1357***
	(.1187)	(.1176)	(.0244)	(.1005)	(4.72)	(3.98)	(1.55)	(4.85)	(.0358)
NREGS Phase 3	.1725	.1821	-.0189	.0725	10.62**	14.53***	-.14	-.01	-.0188
(Linear flexible)	(.1174)	(.1217)	(.0248)	(.1016)	(4.84)	(4.28)	(1.99)	(4.60)	(.0334)
NREGS Phase 3*star	-.4431**	-.1831	.0033	-.1300	-21.56***	-7.15	-2.21	-1.66	-.0894*
	(.1881)	(.1950)	(.0270)	(.1612)	(7.73)	(5.63)	(2.04)	(7.50)	(.0515)
NREGS Phase 3	.1276	.1803	-.0041	-.0492	6.60	12.97***	.11	-7.18	.0095
(Quadratic)	(.1165)	(.1199)	(.0207)	(.1027)	(4.66)	(4.18)	(1.77)	(4.38)	(.0337)
NREGS Phase 3*star	-.2483**	-.0738	.0158	.1159	-11.69**	-.31	-.31	11.60**	-.1072***
	(.1194)	(.1172)	(.0265)	(.1000)	(4.94)	(4.05)	(1.58)	(4.94)	(.0350)
N	375	375	375	375	375	375	375	375	375

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.8: Donut Hole Incumbent Results

	incumbent won	Phase 1 vs 2 vote share	incumbent won	Phase 2 vs 3 vote share
	(1)	(2)	(3)	(4)
NREGS	.1184	1.34	.0256	-1.42
(Linear)	(.0833)	(3.49)	(.0821)	(3.71)
NREGS	.1361	2.51	.0504	.18
(Linear flexible)	(.0900)	(3.81)	(.0913)	(4.05)
NREGS	.1588*	3.12	.0594	-.76
(Quadratic)	(.0928)	(4.02)	(.1014)	(4.40)
N	354	354	375	375

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise. An incumbent is an individual who won a plurality of votes in his/her constituency in the 2004 general elections and who contested the 2009 elections in any constituency.

Table F.0.9: Donut Hole Incumbent Star Results

	incumbent won	Phase 1 vs 2 vote share	incumbent won	Phase 2 vs 3 vote share
	(1)	(2)	(3)	(4)
NREGS	.0557	.28	.1761*	2.12
(Linear)	(.0895)	(3.85)	(.0938)	(4.15)
NREGS*star	.1888	3.24	-.4036***	-9.67**
	(.1247)	(4.99)	(.1083)	(4.77)
NREGS	.0533	.80	.1791	5.79
(Linear flexible)	(.0987)	(4.38)	(.1103)	(4.63)
NREGS*star	.3220	11.82	-.4227**	-17.47**
	(.2711)	(9.91)	(.1739)	(8.09)
NREGS	.0898	2.20	.1712	1.86
(Quadratic)	(.1011)	(4.49)	(.1088)	(4.67)
NREGS*star	.1764	2.53	-.4057***	-9.78**
	(.1266)	(5.02)	(.1099)	(4.87)
N	354	354	375	375

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.10: Election Results: Probit specification won constituencies

	Phase 1 vs Phase 2				Phase 2 vs Phase 3			
	INC	UPA	Left	BJP	INC	UPA	Left	BJP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NREGS	-.1271	-.1381*	.0521*	-.0603	-.0506	-.0178	-.0166	-.0042
(Linear)	(.0810)	(.0829)	(.0293)	(.0633)	(.0820)	(.0843)	(.0222)	(.0703)
NREGS	-.0783	-.0916	.0503**	-.0682	.0303	.0710	-.0117	-.0106
(Linear flexible)	(.0819)	(.0841)	(.0254)	(.0653)	(.0884)	(.0925)	(.0213)	(.0762)
NREGS	-.0394	-.0372	.0491*	-.0679	.0609	.1110	.0454*	-.0485
(Quadratic)	(.0869)	(.0888)	(.0255)	(.0675)	(.0940)	(.0975)	(.0264)	(.0811)
N	406	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.11: Star State Results: Probit specification won constituencies

	Phase 1 vs Phase 2				Phase 2 vs Phase 3		
	INC	UPA	Left	BJP	INC	UPA	BJP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NREGS	-.2413***	-.2343***	.0532*	-.0061	.0681	.0649	-.0471
(Linear)	(.0871)	(.0869)	(.0307)	(.0698)	(.0924)	(.0932)	(.0772)
NREGS*star	.2804***	.2379**	.1178**	-.1778*	-.2443**	-.1087	.1224
	(.1056)	(.1059)	(.0487)	(.0981)	(.1050)	(.1104)	(.0924)
NREGS	-.1861**	-.1880**	.0502*	-.0102	.1763*	.1760*	.0260
(Linear flexible)	(.0881)	(.0885)	(.0260)	(.0716)	(.1034)	(.1053)	(.0894)
NREGS*star	.2380**	.2007*	.1117**	-.1743*	-.3365**	-.1597	-.0579
	(.1051)	(.1068)	(.0459)	(.0977)	(.1557)	(.1692)	(.1420)
NREGS	-.1611*	-.1476	.0502*	-.0073	.1415	.1625	-.0777
(Quadratic)	(.0932)	(.0931)	(.0262)	(.0736)	(.0989)	(.0996)	(.0851)
NREGS*star	.2431**	.1964*	.1244**	-.1772*	-.2049*	-.0563	.1055
	(.1053)	(.1058)	(.0489)	(.0976)	(.1059)	(.1116)	(.0933)
N	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.12: Probit Incumbent Main and Star Results won constituencies

	incumbent Phase 1 vs 2		incumbent Phase 2 vs 3	
	main	star	main	star
	(1)	(2)	(3)	(4)
NREGS	.1460**	.0783	.0397	.1809**
(Linear)	(.0716)	(.0794)	(.0719)	(.0809)
NREGS*star		.1794*		-.3767***
		(.1038)		(.0988)
NREGS	.1513**	.0683	.0350	.1452
(Linear flexible)	(.0731)	(.0836)	(.0791)	(.0920)
NREGS*star		.2549*		-.3461**
		(.1501)		(.1585)
NREGS	.1715**	.0983	.0578	.1689*
(Quadratic)	(.0767)	(.0862)	(.0852)	(.0907)
NREGS*star		.1719		-.3823***
		(.1047)		(.1002)
N	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.13: Election Results Phase 1 vs Phase 2 (TOT)

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.3332	-.3621	.1201*	-.1559	-14.28	-17.93*	.92	-7.25	-.0778
(Linear)	(.2239)	(.2283)	(.0626)	(.1717)	(9.83)	(10.12)	(3.82)	(8.29)	(.0768)
NREGS Phase 1	-.2581	-.2933	.1442*	-.1944	-11.34	-17.02	1.91	-9.28	-.0122
(Linear flexible)	(.2543)	(.2596)	(.0797)	(.1989)	(11.13)	(11.58)	(4.48)	(9.53)	(.0854)
NREGS Phase 1	-.1573	-.1457	.1809*	-.2248	-5.94	-10.80	3.01	-10.04	.0508
(Quadratic)	(.3002)	(.3044)	(.1018)	(.2419)	(12.97)	(13.46)	(5.44)	(11.52)	(.1030)
N	406	406	406	406	406	406	406	406	406

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.14: Election Results Phase 2 vs Phase 3 (TOT)

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 3	-.1348	-.0429	-.0315	-.0090	-2.13	5.82	-4.30	-.06	-.2611***
(Linear)	(.2358)	(.2400)	(.0431)	(.2021)	(9.46)	(8.80)	(3.45)	(9.10)	(.0805)
NREGS Phase 3	.1128	.2296	-.0324	-.0380	8.90	18.53*	-2.45	-3.71	-.1068
(Linear flexible)	(.2559)	(.2648)	(.0440)	(.2317)	(10.45)	(9.72)	(3.61)	(10.17)	(.0765)
NREGS Phase 3	.2343	.4272	.0211	-.2161	9.56	23.54	-.54	-13.85	-.0627
(Quadratic)	(.3648)	(.3832)	(.0554)	(.3322)	(15.01)	(14.45)	(4.88)	(14.54)	(.0955)
N	406	406	406	406	406	406	406	406	406

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.15: Star State Results Phase 1 vs Phase 2 (TOT)

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.5399**	-.5550**	.1360*	-.0574	-19.85*	-21.82**	-.31	-.75	-.1611*
(Linear)	(.2488)	(.2501)	(.0697)	(.1791)	(10.68)	(10.68)	(4.32)	(8.24)	(.0909)
NREGS Phase 1*star	.4715**	.3824**	-.0354	-.2875*	10.58	3.64	4.03	-19.86**	.2188***
	(.1943)	(.1794)	(.0572)	(.1695)	(7.51)	(7.63)	(3.11)	(8.86)	(.0617)
NREGS Phase 1	-.5137*	-.5442*	.1725*	-.0863	-18.41	-23.05*	.80	-2.17	-.0993
(Linear flexible)	(.2931)	(.2964)	(.0922)	(.2113)	(12.61)	(12.83)	(5.29)	(9.69)	(.1061)
NREGS Phase 1*star	.4610**	.3781**	-.0499	-.2759	10.01	4.13	3.59	-19.30**	.1942***
	(.2004)	(.1876)	(.0601)	(.1697)	(7.83)	(7.99)	(3.31)	(8.81)	(.0658)
NREGS Phase 1	-.4849	-.4750	.2217*	-.1027	-15.29	-19.28	2.01	-2.15	-.0526
(Quadratic)	(.3507)	(.3504)	(.1195)	(.2581)	(14.98)	(15.40)	(6.47)	(11.89)	(.1254)
NREGS Phase 1*star	.4589**	.3642*	-.0550	-.2771	9.54	3.06	3.50	-19.54**	.1940***
	(.2027)	(.1885)	(.0654)	(.1712)	(7.86)	(8.00)	(3.39)	(8.85)	(.0671)
N	406	406	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.16: Star State Results Phase 2 vs Phase 3 (TOT)

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 3	.1304	.1617	-.0412	-.0987	9.60	12.79	-4.22	-8.28	-.1620**
(Linear)	(.2470)	(.2543)	(.0503)	(.2164)	(10.16)	(9.59)	(4.07)	(9.86)	(.0804)
NREGS Phase 3*star	-.4330**	-.1753	.0065	.2102	-19.44***	-2.26	-1.63	21.15**	-.2373***
	(.1925)	(.1836)	(.0360)	(.1647)	(7.33)	(6.09)	(2.33)	(8.36)	(.0877)
NREGS Phase 3	.3244	.4188	-.0421	-.0998	18.21	25.54**	-2.72	-8.99	-.0435
(Linear flexible)	(.2691)	(.2785)	(.0506)	(.2418)	(11.29)	(10.46)	(4.10)	(10.82)	(.0858)
NREGS Phase 3*star	-.3862**	-.1134	.0063	.2099	-17.36**	.81	-1.27	20.97**	-.2087***
	(.1907)	(.1843)	(.0355)	(.1654)	(7.62)	(6.90)	(2.27)	(8.60)	(.0758)
NREGS Phase 3	.4235	.6052	.0092	-.2609	18.05	30.46**	-1.01	-17.69	-.0133
(Quadratic)	(.3776)	(.4045)	(.0558)	(.3393)	(15.59)	(15.37)	(5.17)	(15.18)	(.1065)
NREGS Phase 3*star	-.3270	-.0149	.0247	.1515	-16.38**	4.13	-.47	17.74**	-.1835**
	(.1994)	(.1962)	(.0414)	(.1695)	(8.22)	(7.75)	(2.20)	(8.58)	(.0727)
N	406	406	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.17: Incumbent Results (TOT)

	incumbent won	Phase 1 vs 2 vote share	incumbent won	Phase 2 vs 3 vote share
	(1)	(2)	(3)	(4)
NREGS	.3812**	4.32	.1130	1.34
(Linear)	(.1939)	(8.22)	(.2076)	(9.48)
NREGS	.4534*	6.40	.0989	4.00
(Linear flexible)	(.2319)	(9.74)	(.2310)	(10.54)
NREGS	.5876**	9.61	.2189	4.76
(Quadratic)	(.2989)	(12.07)	(.3383)	(14.85)
N	406	406	406	406

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise. An incumbent is an individual who won a plurality of votes in his/her constituency in the 2004 general elections and who contested the 2009 elections in any constituency.

Table F.0.18: Incumbent Star Results (TOT)

	incumbent won	Phase 1 vs 2 vote share	incumbent won	Phase 2 vs 3 vote share
	(1)	(2)	(3)	(4)
NREGS	.2490	1.97	.4067*	7.51
(Linear)	(.1875)	(8.54)	(.2411)	(10.25)
NREGS*star	.3545	5.95	-.6325***	-13.20*
	(.2259)	(8.08)	(.1753)	(7.68)
NREGS	.3075	3.98	.3149	8.46
(Linear flexible)	(.2320)	(10.54)	(.2589)	(11.21)
NREGS*star	.3312	5.15	-.6547***	-12.97*
	(.2265)	(8.22)	(.1822)	(7.83)
NREGS	.4281	6.91	.4025	8.60
(Quadratic)	(.2940)	(13.10)	(.3695)	(15.29)
NREGS*star	.3136	4.82	-.6341***	-12.81
	(.2371)	(8.38)	(.1871)	(8.29)
N	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.19: Star State Results DID Specification

	won constituencies				vote share in percent				
	INC	UPA	Left	BJP	INC	UPA	Left	BJP	voter turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NREGS Phase 1	-.0567	-.1398*	.0023	.0860*	-7.90***	-6.56***	1.18	-4.16**	-.0287***
(Linear)	(.0647)	(.0725)	(.0146)	(.0513)	(2.32)	(2.37)	(1.17)	(2.09)	(.0098)
NREGS Phase 1*star	.1635	.1283	.0101	-.1506**	13.55***	14.59***	.29	-4.24	.0772***
	(.1036)	(.1061)	(.0089)	(.0685)	(2.97)	(3.08)	(1.68)	(3.28)	(.0134)
NREGS Phase 3	-.0512	.0373	.0023	.0278	5.37**	1.02	-.98	2.54	.0028
(Linear flexible)	(.0737)	(.0823)	(.0207)	(.0508)	(2.33)	(2.24)	(.96)	(2.28)	(.0085)
NREGS Phase 3*star	.1171	.1768	.0117	-.2648***	1.80	12.03***	-1.42*	9.32**	.0226
	(.1189)	(.1155)	(.0186)	(.0952)	(3.65)	(2.84)	(.72)	(3.71)	(.0153)
N	784	784	784	784	784	784	784	784	784

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported. Vote shares are given in percent. The won variables are indicator variables equal to 1 if a given party received a plurality of the votes in a constituency, and 0 otherwise.

Table F.0.20: Financial Allocation Star State Results (2008/2009)

	Phase 1 vs Phase 2				Phase 2 vs Phase 3			
	centr. rel.	tot. funds	tot. exp.	admin exp.	centr. rel.	tot. funds	tot. exp.	admin exp.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NREGS	2303.70***	2535.78***	2137.52**	22.41	307.07	-184.96	2.87	18.02
(Linear)	(794.32)	(956.14)	(841.97)	(25.83)	(778.10)	(895.54)	(746.14)	(26.05)
NREGS*star	1120.88	1312.40	1167.04	163.83**	-756.08	-795.74	-655.54	-119.57**
	(1857.55)	(1943.82)	(1742.40)	(76.55)	(1869.02)	(1979.81)	(1817.03)	(59.50)
NREGS	1917.42**	2124.01**	1795.82*	13.87	-626.69	-1144.67	-838.37	1.32
(Linear flexible)	(896.50)	(1071.19)	(971.13)	(29.70)	(684.78)	(805.31)	(628.04)	(23.69)
NREGS*star	1423.22	1634.69	1434.48	170.52**	790.59	844.24	310.89	-124.78*
	(1851.20)	(1954.21)	(1760.59)	(76.83)	(2813.59)	(2988.33)	(2709.40)	(71.26)
NREGS	1717.51*	1861.14	1576.65	4.51	93.71	-296.26	-243.04	19.97
(Quadratic)	(974.02)	(1146.35)	(1049.06)	(31.66)	(949.88)	(1064.40)	(888.01)	(27.30)
NREGS*star	1359.73	1587.28	1395.56	171.13**	-852.70	-846.14	-766.90	-118.6815**
	(1860.2)	(1956.87)	(1757.97)	(76.70)	(1911.18)	(2029.08)	(1866.33)	(59.59)
N	406	406	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported.

Table F.0.21: Other Star State Results Phase 1 vs Phase 2

	pers.empl 1	pub empl	priv	fam	log wage	per cap cons
	(1)	(2)	(3)	(4)	(5)	(6)
NREGS Phase 1	72364.88**	.0009	.0093	.0189	-.0921*	-.1654***
(Linear)	(35931.21)	(.0041)	(.0261)	(.0271)	(.0536)	(.0440)
NREGS Phase 1*star	-132107.3***	.0135	-.0245	-.0146	.1829**	.1188**
	(41397.96)	(.0094)	(.0387)	(.0424)	(.0794)	(.0600)
NREGS Phase 1	85938.63**	-.0031	.0268	-.0076	-.0790	-.1377***
(Linear flexible)	(41308.33)	(.0047)	(.0279)	(.0287)	(.0566)	(.0463)
NREGS Phase 1*star	-142731.3***	.0167*	-.0384	.0065	.1722**	.0969
	(44070.06)	(.0096)	(.0387)	(.0422)	(.0808)	(.0615)
NREGS Phase 1	87078.83**	-.0041	.0285	-.0148	-.0737	-.1160**
(Quadratic)	(44201.16)	(.0053)	(.0291)	(.0300)	(.0590)	(.0484)
NREGS Phase 1*star	-138102.6***	.0158*	-.0332	.0007	.1744**	.0964
	(43097.58)	(.0095)	(.0389)	(.0424)	(.0806)	(.0610)
N	406	418	418	418	418	418

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported.

Table F.0.22: Other Star State Results Phase 2 vs Phase 3

	pers.empl	pub empl	priv	fam	log wage	per cap cons
	(1)	(2)	(3)	(4)	(5)	(6)
NREGS Phase 3	-73588.23***	.0029	-.0176	-.0080	.1730***	.1274***
(Linear)	(24819.73)	(.0041)	(.0261)	(.0280)	(.0535)	(.0435)
NREGS Phase 3*star	134425.6***	-.0078	-.0274	.0691*	-.2445***	-.1356**
	(35612.88)	(.0069)	(.0380)	(.0412)	(.0761)	(.0545)
NREGS Phase 3	-77555.1***	-.0010	-.0029	-.0263	.1621***	.1757***
(Linear flexible)	(22451.19)	(.0035)	(.0295)	(.0310)	(.0604)	(.0499)
NREGS Phase 3*star	175921.2***	-.0053	-.0127	.0242	-.1964*	-.1396*
	(48846.17)	(.0084)	(.0572)	(.0635)	(.1154)	(.0808)
NREGS Phase 3	-61627.13***	.0010	.0010	-.0397	.1679***	.1533***
(Quadratic)	(22504.39)	(.0037)	(.0289)	(.0307)	(.0594)	(.0468)
NREGS Phase 3*star	139842.1***	-.0092	-.0139	.0461	-.2480***	-.1168**
	(37054.97)	(.0071)	(.0391)	(.0425)	(.0783)	(.0564)
N	406	406	406	406	406	406

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors clustered at the district level in parentheses. An observation is an election constituency in the 2009 general election. Parametric regressions with different levels of flexibility are reported.

BIBLIOGRAPHY

Afridi, Farzana, Abhiroop Mukhopadhyay, and Soham Sahoo. 2012. 'Female Labour Force Participation and Child Education in India: The Effect of the National Rural Employment Guarantee Scheme' IZA Discussion Paper 6593.

Akerlof, George and Janet Yellen. 1994. 'Gang Behavior, Law Enforcement, and Community Values.' Issue 53 of CIAR Program in Economic Growth and Policy reprint series

Almond, Douglas, and Joseph J. Doyle. 2011. 'After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays.' *American Economic Journal: Economic Policy*, 3(3): 1-34.

Azam, Mehtabul. 2012. 'The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment.' IZA Discussion Paper 6548.

Baez, Javier, Adriana Camacho, Emily Conover, and Romn Zrate. Conditional Cash Transfers, Political Participation, and Voting Behavior. IZA Discussion Paper No. 6870, September 2012.

Bakshi, G D. 2009. 'Left Wing Extremism in India: Context, Implications and Response Options.' Manekshaw Paper No. 9.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2013. 'Can Basic Entrepreneurship Transform the Economic Lives of the Poor?' IZA Discussion Papers 7386.

Banerjee, Abhijit, Selvan Kumar, Rohini Pande, and Felix Su. 2011. 'Do Informed Voters Make Better Choices? Experimental Evidence from Urban India.' Mimeo.

Banerjee, Abhijeet V., Esther Duflo, Raghabendra Chattopadhyay, and Jeremy Shapiro. 2011. 'Targeting the Hardcore Poor: An Impact Assessment.' Mimeo, MIT.

Banerjee, Kaustav and Partha Saha. 2010. 'The NREGA, the Maoists and the Developmental Woes of the Indian State.' *Economic and Political Weekly*, XLV(28): 42-48.

Basu, Arnab. 2002. 'Oligopsonistic Landlords, Segmented Labor Markets, and the Persistence of Tied-Labor Contracts.' *American Journal of Agricultural Economics*, 84(2): 438-453.

Basu, Arnab. Forthcoming. 'Impact of Rural Employment Guarantee Schemes on Seasonal Labor Markets: Optimum Compensation and Workers' Welfare.' *Journal of Economic Inequality*.

Berg, Erlend, Sambit Bhattacharyya, Rajasekhar Durgam, and Manjula Ramachandra. 2012. 'Can Rural Public Works Affect Agricultural Wages? Evidence from India.' CSAE Working Paper WPS/2012-05.

Berhane, Guush, John Hoddinott, Neha Kumar, and Alemayehu Seyoum Taffesse. 2011. 'The Impact of Ethiopia's Productive Safety Nets and Household Asset Building Programme: 2006-2010.' Mimeo, International Food Policy Research Institute.

Berman, Eli, Shapiro, Jacob N., and Joseph H. Felter. 2011. 'Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq.' *Journal of Political Economy*, 119(4): 766-819.

Besley, Timothy, and Stephen Coate. 1992. 'Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs.' *American Economic Review*, 82(1): 249-261.

Besley, Timothy, Rohini Pande, and Vijayendra Rao. 'Just Rewards? Local Politics and Public Resource Allocation in South India.' *World Bank Economic Review*, 26(2): 191-216, 2012.

Black, Dan A., Galdo, Jose, and Jeffrey A. Smith. 2007. 'Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Approach' *American Economic Association Papers and Proceedings*, 97(2): 104-107.

Blattman, Chris and Edward Miguel. 2010. 'Civil War.' *Journal of Economic Literature* 48(1): 3-57.

Borooah, Vani K. 2008. 'Deprivation, Violence and Conflict: An Analysis of 'Naxalite' Activity in the Districts of India. *International Journal of Conflict and Violence*, 2: 317-333.

Brender, A., and A. Drazen. 2005. 'Political Budget Cycles in New versus Established Democracies.' *Journal of Monetary Economics*, 52: 1271-1295.

Brender, A., and A. Drazen. 2008. 'How Do Budget Deficits and Economic Growth Affect Reelection Prospects? Evidence from a Large Panel of Countries' *American Economic Review*, 98: 2203-2220.

Cameron, A. Colin and Pravin K. Trivedi. 2013. 'Regression Analysis of Count Data.' 2nd edition, Econometric Society Monograph No.53, Cambridge University Press.

Collier, Paul and Anke Hoeffler. 2007. 'Civil War', in Handbook of Defense Economics. K. Hartley and T. Sandler eds: Elsevier North Holland.

Crost, Benjamin, Felter, Joseph and Patrick Johnston. 2012. 'Conditional Cash Transfers and Civil Conflict: Experimental Evidence from the Philippines.' Mimeo

Crost, Benjamin, Felter, Joseph and Patrick Johnston. Forthcoming. 'Aid Under Fire: Development Projects and Civil Conflict.' *American Economic Review*

Centre for Science and Environment. 2008. 'An Assessment of the Performance of the National Rural Employment Guarantee Programme in Terms of its Potential for Creation of Natural Wealth in India's Villages.'

Dal Bo, Ernesto and Robert Powell. 2007. 'Conflict and Compromise in Hard and Turbulent Times.' UC Berkeley Department of Political Science Working Paper.

Dasgupta, Adi, Kishore Gawande, and Devesh Kapur. 2014. 'Can Anti-poverty Programs Reduce Conflict? India's Rural Employment Guarantee and Maoist Insurgency,' Mimeo.

Datt, Gaurav, and Martin Ravallion. 1994. 'Transfer Benefits from Public-Works Employment: Evidence from Rural India.' *The Economic Journal*, 104(427): 1346-1369.

De la O, Ana. 2013. 'Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico.' *American Journal of Political Science*, 57(1): 1-14.

Dev, S M. 1995. 'Alleviating Poverty: Maharashtra Employment Guarantee Scheme.' *Economic and Political Weekly*, XXX(41-42): 2663-2676.

Dey, Nikhil, Jean Dreze, and Reetika Khera. 2006. *Employment Guarantee Act: A Primer.* (Delhi: National Book Trust, India)

Do, Quy-Toan and Lakshmi Iyer. 2007. 'Poverty, Social Divisions and Conflict in Nepal.' Unpublished working paper, Harvard Business School.

Drazen, Allen and Marcela Eslava. 2010. 'Electoral Manipulation via Voter-Friendly Spending: Theory and Evidence.' *Journal of Development Economics*, 92: 39-52.

Dreze, Jean and Reetika Khera. 2009. 'The Battle for Employment Guarantee.' *Frontline*, 26(1).

Dube, Oeindrila and Juan F. Vargas. 2012.'Commodity Price Shocks and Civil Conflict: Evidence from Colombia.' Royal Holloway, University of London: Discussion Papers in Economics.

Dunning, Thad and Janhavi Nilekani. 2012. 'Ethnic Quotas and Political Mobilization: Caste, Parties, and Distribution in Indian Village Councils.' Mimeo.

Dutta, Puja, Rinku Murgai, Martin Ravallion, and Dominique van de Walle. 2012. 'Does India's Employment Guarantee Scheme Guarantee Enployment?' World Bank Policy Research Working Paper 6003.

Economic and Political Weekly (EPW). 2009. 'Defeated But Still a Threat.' *Economic and Political Weekly*, XLIV(24): 6.

Fearon, James D. 2007. 'Economic development, insurgency, and civil war,' in Institutions and Economic Performance. Elhanen Helpman ed. Cambridge: Harvard University Press.

Ferraz, Claudio and Frederico Finan. 2011. 'Electoral Accountability and Corruption: Evidence from the Audits of Local Governments.' *American Economic Review*, 101: 1274-1311.

Fetzer, Thiemo. 2013. 'Can Workfare Programs Moderate Violence? Evidence from India,' Mimeo.

Finan, Frederico and Laura Schechter. 2012. 'Vote-Buying and Reciprocity.' *Econometrica*, 80(2): 863-881.

Gilligan, Daniel O., John Hoddinott, Alemayehu Seyoum Taffesse. 2008. 'The Impact of Ethiopia's Productive Safety Net Programme and Its Linkages.' IFPRI Discussion Paper 00839.

Government of India. 2009. 'The National Rural Employment Guarantee Act.'

Grossman, Herschell I. 1991. 'A General Equilibrium Model of Insurrections.' The American Economic Review, 81(4): 912-21.

Grossman, Herschell I. 1999. 'Kleptocracy and revolutions.' Oxford Economic Papers, 51, 267-83.

Gupta, Santanu. 2006. 'Were District Choices for NFFWP Appropriate?' *Journal of the Indian School of Political Economy*, 18(4): 641-648.

Hirshleifer, Jack. 1989. 'Conflict and rent-seeking functions: Ratio versus difference models of relative success.' Public Choice, 63, 101-12.

Humphreys, Macartan and Jeremy M. Weinstein. 2008. 'Who Fights? The Determinants of Participation in Civil War.' *American Journal of Political Science*, 52(2): 436-455.

Imbert, Clement, and John Papp. 2013. 'Labor Market Effects of Social Programs: Evidence of India's Employment Guarantee.' Centre for the Study of African Economies Working Paper WPS/2013-03.

Institute of Applied Manpower Research. 2007. 'All-India Report on Evaluation of NREGA - A Survey of 20 Districts.'

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

Johnson, Doug. 2009a. 'Can Workfare Serve as a Substitute for Weather Insurance? The Case of NREGA in Andhra Pradesh.' Institute for Financial Management and Research, Centre for Micro Finance, Working Paper 32.

Johnson, Doug. 2009b. 'How Do Caste, Gender and Party Affiliation of Locally Elected Leaders Affect Implementation of NREGA?' Institute for Financial Management and Research, Centre for Micro Finance Working Paper 33.

Kaur, Supreet. 2012. 'Nominal Wage Rigidity in Village Labor Markets.' Mimeo.

Khera, Reetika. 2009. 'Group Measurement of NREGA Work: The Jalore Experiment.' Centre for Development Economics Delhi School of Economics Working Paper 180.

Khera, Reetika, and Nandini Nayak. 2009. 'Women Workers and Perceptions of the National Rural Employment Guarantee Act.' *Economic and Political Weekly*, XLIV(43): 49-57.

Khera, Reetika. 2010. 'Wages of Delay.' *Frontline*, 27(10).

Khera, Reetika. 2011. *The Battle for Employment Guarantee*. Oxford University Press.

Klonner, Stefan and Christian Oldiges. 2012. 'Employment Guarantee and Its Welfare Effects in India.' Mimeo.

Kluve, Jochen. 2010. 'The Effectiveness of European Active Labor Market Programs.' *Labour Economics*, 17(6): 904-918.

Labonne, Julien. 2013. 'The Local Electoral Impacts of Conditional Cash Transfers : Evidence from a Field Experiment.' *Journal of Development Economics*, 104: 73-88.

Lal, Radhika, Steve Miller, Maikel Lieuw-Kie-Song, and Daniel Kostzer. 2010. 'Public Works and Employment Programmes: Towards a Long-Term Development Approach.' International Policy Centre for Inclusive Growth Working Paper 66.

Lalwani, Sameer. 2011. 'India's Approach to Counterinsurgency and the Naxalite Problem.' *CTC Sentinel*, 4(10): 5-9.

Lee, David S. 2008. 'Randomized Experiments from Non-Random Selection in U.S. House Elections.' *Journal of Econometrics*, 142(2): 675-697.

Lee, David S., and Thomas Lemieux. 2010. 'Regression Discontinuity Designs in Economics.' *Journal of Economic Literature*, 48(2): 281-355.

Manacorda, Marco, Edward Miguel, and Andrea Vigorito. 2011. 'Government Transfers and Political Support.' *American Economic Journal: Applied Economics*, 3(3).

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. 'Economic Shocks and Civil Conflict: An Instrumental Variables Approach.' *Journal of Political Economy*, 112(4): 725-53.

Miguel, Edward, and Shanker Satyanath. 2011. 'Re-examining Economic Shocks and Civil Conflict.' *American Economic Journal: Applied Economics*, 3(4): 228-232.

Ministry of Home Affairs, Government of India. 2006. 'Annual Report 2006-2007.'

Ministry of Rural Development, Department of Rural Development, Government of India. 2010. 'Mahatma Gandhi National Rural Employment Guarantee Act 2005 - Report to the People 2nd Feb 2006 - 2nd Feb 2010.'

Mukherji, Nirmalangshu. 2012. *The Maoists in India: Tribals under Siege.* (London: Pluto Press, United Kingdom)

Murshed, S. Mansoob and Scott Gates. 2005. 'Spatial-Horizontal Inequality and the Maoist Insurgency in Nepal.' *Review of Development Economics*, 9(1): 121-34.

NCAER-PIF. 2009. 'Evaluating the performance of the National Rural Employment Guarantee Act.'

Niehaus, Paul, and Sandip Sukhtankar. 2013a. 'The Marginal Rate of Corruption in Public Programs.' *Journal of Public Economics*, 104: 52-64.

Niehaus, Paul, and Sandip Sukhtankar. 2013b. 'Corruption Dynamics: The Golden Goose Effect.' *American Economic Journal: Economic Policy*, 5(4): 230-269.

Nupia, O. 2011. 'Anti-Poverty Programs and Presidential Election Outcomes: Familias En Accion in Colombia.' Working Paper, Universidad de los Andes.

Nunn, Nathan and Nancy Qian. 2012. 'Aiding Conflict: The Impact of U.S. Food Aid on Civil War.' NBER Working Paper No. 17794.

oneworld.net. 2011. 'Community MGNREGS Programme for Naxalite Affected Areas.' Report.

Pande, Rohini. 2011. 'Can Informed Voters Enforce Better Governance? Experiments in Low Income Democracies.' *Annual Review of Economics*, 3(1): 215-237.

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

Planning Commission. 2009. 'Report of the Expert Group to Review the Methodology for Estimation of Poverty.'

Planning Commission (MLP Division). 2003. 'Backward Districts Initiative - Rashtriya Sam Vikas Yojana - The Scheme and Guidelines for Preparation of District Plans.'

Pop-Eleches, Cristian and Grigore. 2012. 'Government Spending and Pocketbook Voting: Quasi-Experimental Evidence from Romania.' *Quarterly Journal of Political Science*, 7(30).

Powell, Robert. 2006. 'War as a Commitment Problem.' *International Organization*, 60, 169-203.

Ramana, P.V. 2011. 'India's Maoist Insurgency: Evolution, Current Trends, and Responses.' in *India's Contemporary Security Challenges*, edited by Michael Kugelman, Woodrow Wilson International Center for Scholars Asia Program.

Ramani, Srinivasan. 2009. 'A Decisive Mandate.' *Economic and Political Weekly*, XLIV(21): 11-12.

Ravi, Shamika, Kapoor, Mudit, and Rahul Ahluwalia. 2012. 'The Impact of NREGS on Urbanization in India.' Mimeo.

Samarthan Centre for Development Support. 2007. 'Status of NREGA Implementation: Grassroots Learning and Ways Forward - 1st Monitoring Report.'

Singh, Prakash. undated. 'Naxal Threat and State Response.' Mimeo.

Skaperdas, Stergios. 1992. 'Cooperation, Conflict, and Power in the Absence of Property Rights.' *American Economic Review*, 82(4): 720-39.

Subbarao, K. 1997. 'Public Works as an Anti-Poverty Program: An Overview of Cross-Country Experience.' *American Journal of Agricultural Economics*, 79(2): 678-683.

Subbarao, K., del Ninno, C., Andrews, C., C. and Rodrguez-Alas. 2013. *Public Works as a Safety Net - Design, Evidence and Implementation.* Washington DC: World Bank.

Sundar, Nandini. 2011. 'At War with Oneself: Constructing Naxalism as India's Biggest Security Threat.' in *India's Contemporary Security Challenges*, edited by Michael Kugelman, Woodrow Wilson International Center for Scholars Asia Program.

Times of India. 2006. 'United Progressive Alliance: Partners in Governance.' July 8.

Vanden Eynde, Oliver. 2011. ‘Targets of Violence: Evidence from India’s Naxalite Conflict.’ Mimeo.

Weber, Max. 1919. ‘Politics as a Vocation.’ ‘Politik als Beruf,’ Gesammelte Politische Schriften, pp. 396-450. Duncker and Humblodt, Munich

Yadav, Yogendra. 1999. ‘Electoral Politics in the Time of Change.’ *Economic and Political Weekly*, 34(34/35): 2393-2399.

Zimmermann, Laura. Forthcoming. ‘Public-Works Programs in Developing Countries’, *IZA World of Labor*.

Zucco, Cesar. 2010. ‘Cash Transfers and Voting Behavior: An Assessment of the Political Impacts of the Bolsa Familia Program.’ Mimeo.