ESSAYS IN LABOR AND DEVELOPMENT ECONOMICS IN INDIA

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kalyani Raghunathan

May 2015

ESSAYS IN LABOR AND DEVELOPMENT ECONOMICS IN INDIA

Kalyani Raghunathan, Ph.D.

Cornell University 2015

This dissertation consists of five distinct chapters that answer questions in labor and development economics in the Indian context. These five chapters are divided into two sections - the first deals with empirical questions, and the second with theoretical models.

The first essay studies the impact of the National Rural Employment Guarantee Act (NREGA) on the risk-taking behavior of farmers in India. In the context of a rural economy with very low access to formal credit and insurance, this Act provides a sizable increase in incomes. We argue that the Act does not compete with private employment, and eases either credit or insurance constraints, or both. The introduction of this Act results in an increase in the riskiness of the portfolio of crops grown at the district level, suggesting that the additional income allows farmers to invest in riskier but potentially higher return technologies.

The second essay uses household-level information from rural India to study the impact of changes in the opportunity cost of a mother's time on time-intensive health investments in her children. Negative rainfall shocks decrease the size of the harvest and the need for female labor, thereby increasing the amount of time these mothers have to spend on their children. I show that negative rainfall shocks in households where the mother works increase the probability of the child being vaccinated and of breastfeeding being initiated, and decrease the probability of the child dying at a young age.

The second section of my dissertation consists of theoretical modeling of the Indian labor market. The first essay in this section provides an introduction to the literature on Indian labor markets. The second essay builds a model of the rural labor market, with seasonality in agricultural production and inter-temporal spillovers in productivity, and studies the welfare implications of the introduction of the NREGA. The final essay is a model outlining the linkages between the rural and urban sectors in a stylized model of the Indian labor market. It incorporates the circular nature of rural-urban migration, and studies the welfare impact of the introduction of the NREGA into such an economy.

BIOGRAPHICAL SKETCH

Kalyani Raghunathan was born and raised in Delhi, India. She obtained her Bachelor's degree in Economics from St. Stephen's College, Delhi University, in 2008, and her Master's in Quantitative Economics (M.S.Q.E) from the Indian Statistical Institute, Delhi, in 2010. After completing her Ph.D. from Cornell University she plans to join the International Food Policy Research Institute in Delhi as an Associate Research Fellow.

This document is dedicated to Vinod Raina, whose generosity, kindness, and equanimity have been a constant source of inspiration. I wish you could have been here today.

ACKNOWLEDGEMENTS

I thank Gary Fields, John Abowd, Jim Berry, Arnab Basu and Lars Vilhuber, without whose help this thesis could not have been written. Gary spent many hours discussing large portions of this document with me; painstakingly reading through early versions in order to make my arguments clearer and more precise. John set a high standard for all my work, and encouraged me to reach it. Jim provided invaluable advice and support, especially during the job market period. Arnab gave me many new ideas for my theoretical work and much wisdom regarding life in general. Finally, Lars patiently supervised my work as a research assistant, suffering uncomplainingly through my many errors.

I thank my parents, Raghu and Harriet, who have always emphasized the importance of learning, and of integrity in thought and action. I would not have made it without their constant encouragement and support.

The years at graduate school can seem very long and lonely, and I am lucky to have met people along the way who made it easier to keep going. Many thanks to Hari for being such an easygoing co-author and for his and Venu's hospitality on countless visits to New York; to Tammy, without whose unwavering support and friendship I would have given up many months ago; to Yagna, who appeared late in my years at Cornell, but greatly improved them; to Laura, Andrew and Bella, who reminded me that real life is much richer and more rewarding than graduate school would have you believe; to Janos, Yanlei and Ankita who have borne my complaints since the beginning of this program; and to Alice, Ellen and Jon for uncomplicated friendships that I could count on.

Finally, thanks are also due to Raj, for trying so hard to teach me to keep things in perspective, and for always believing I would make it in the end. My Ph.D. has been harder on you than it was on me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The chapters in my thesis can be divided into two broad sections. The first section consists of two papers where I use household-level and district-level data from India to answer policy-relevant questions. The first of these chapters is co-authored with Siddharth Hari. The second is solo-authored. The second category develops two theoretical labor market models of the Indian labor market. These models are written jointly with Prof. Gary Fields. This is why I use the pronoun "we" in Chapters 2, 5 and 6 chapters of this dissertation, and "I" in Chapter 3.

In Chapter 2, we show empirically that the introduction of a rural employment guarantee act (EGA) in India changes crop choice at a district level. In particular the access to a certain income from the program increases the amount of risk embodied in the district-level crop portfolios. One possible reason is that the income from this employment guarantee eases either credit or insurance constraints and allows farmers to invest in technologies that exhibit greater variance. We provide evidence that this effect is indeed causal, and suggest some mechanisms through which it might be operating.

In Chapter 3, I study health investments in children in rural India. I differentiate districts based on the level of female participation in the labor force, and show that negative rainfall shocks in households where women participate in agriculture can actually improve child health by reducing the opportunity cost of the mother's time. Negative rainfall shocks increase the likelihood of the child being vaccinated and breastfed, and increase the probability of the child surviving the first six months of its life. This finding runs counter to most of

the literature, which discusses the adverse impacts of negative rainfall shocks on household income, and (through this channel) on the early-life outcomes of children.

Chapter 4 serves as an introduction to the features of the Indian rural and urban labor markets that will inform our theoretical models. In Chapter 5, we develop a simple theoretical model of the rural Indian labor market. It models the rural market as one with two seasons, a slack season where the labor market is far from full employment, and a peak season where the labor market is 'tight'. There are inter-temporal spillovers in productivity between the two seasons. We model the introduction of an EGA into this market, and study the welfare implications of the distortions to wages and employment that result. We show that it is possible for recipients of the Act's payments to be made worse off through coordination failures.

Finally in Chapter 6, we extend the previous model to the urban sector, and model linkages between the rural and urban sectors that arise through the movement of labor. In the slack agricultural season rural workers move from the rural to the urban sector in search of work, and in the peak season there is 'reverse migration' of these workers back to the rural areas. We show that an EGA has a negative impact on the peak season agricultural wage, and as a result the welfare effects of the Act are, for the large part, ambiguous. While both the rural and the rural-urban models are simple, they are powerful - delivering the result that the EGA can have negative effects on the welfare of some sections of the population, even if the guarantee is only present in the slack season.

1.1 Providing more than just employment? Evidence from the NREGA in India

This paper studies the effect of the NREGA on the ability of farmers to undertake riskier activities. Risk and income uncertainty are central features of agricultural economies, and in particular of developing economies like India. Only 17% of farmers in India have access to insurance, and only 15-30% of small or marginal farmers have access to formal credit (Mahul et al., 2012; Satyasai, 2012). Irrigation infrastructure is limited, and most agricultural workers rely on the monsoon for the survival of their crops. Fluctuations in weather or idiosyncratic shocks to farmer households can cause large losses in income.

In such an environment it is not surprising that farmers are risk-averse. There is considerable evidence that small economic agents faced with credit constraints and uninsured risk choose low-risk and low-profitability technologies over those that are known to be riskier but also more profitable (Dercon and Christiaensen, 2011; Gine and Klonner, 2006; Karlan et al., 2012; Morduch, 1990; Moser and Barrett, 2006; Rosenzweig and Binswanger, 1993). If this is also the case in India, then the introduction of some additional sure income (for example, the income from a government welfare scheme) could ease some of these constraints and allow farmers to choose riskier production activities. In particular we are interested in the decision to grow portfolios of crops that exhibit greater variance in yields, but also potentially higher average returns.

The NREGA provides a hundred days of employment per year to all households in rural districts of India. Laborers are paid a state-specific minimum wage for working on NREGA sanctioned work, and reports suggest that at the

time of the introduction of the Act the minimum wage was considerably higher than the going agricultural wage (Khera, 2011). For small and marginal farmers, NREGA income could constitute up to 50% of their annual incomes. Thus the NREGA represented a significant easing of income constraints for these households.

In addition to being a large amount of money, the NREGA is designed to be demand-driven, i.e. workers who wanted employment in the program could file an application for work to be opened in their village. In this manner the NREGA differs from other workfare programs across the world, and from other similar programs instituted previously in India. Rural inhabitants who wish to work for the NREGA file an application with their village officials, and work has to be provided within 15 days of receipt of this application, and within 5 kilometers of the workers' place of residence.[1] For this reason, we argue that it is reasonable to think of the NREGA income as being 'sure income', i.e. a source of livelihood that workers could count on receiving.

We use a regression discontinuity design to study the impact of the NREGA on the crop choices at a district level. The NREGA was introduced in a phased manner, with 200 of the poorest districts in the country receiving the program in the first phase in mid-2006, the next 130 poorest in the second phase in mid-2007 and all remaining rural districts in the third phase in mid-2008. We use the phased roll-out to design our empirical methodology. Our outcomes of interest are the standard deviation and coefficient of variation of the yields and the average yield and revenue of the portfolio of crops grown at a district level. We find that districts that received the program show a significant and large increase in these outcome variables relative to those districts that did not have

[1] 5 kilometers is approximately 3 miles.

the NREGA, which suggests that the introduction of the NREGA permits investment in riskier crops. To put the numbers in perspective, we find about a fifth of a standard deviation increase in portfolio variability. We do not find a significant increase in the average returns to a portfolio.

The ranking of the districts in order of poverty was made on the basis of a district's composite score on a backwardness index published by the Planning Commission (Planning Commission, 2003a). This backwardness index uses information on agricultural wages, agricultural productivity and the proportion of the district's population that belonged to a scheduled caste or tribe. The data on these three components comes from the mid to late 1990s, pre-dating the NREGA by almost 10 years, which mitigates our concerns of manipulation of these scores by the districts in order to affect the probability of treatment.

We use the Planning Commission backwardness index to rank the districts within each of the seventeen states for which the information was available. Taking the number of districts in each state that were assigned the program as given, we re-normalize the ranks such that districts with negative ranks are predicted to receive the program, districts with positive ranks are not, and the district with rank 0 is the last district predicted to receive the program in that state. This state-normalized rank then forms the running variable for our regression discontinuity.

The most general form of the equation of interest in our empirical methodology is the following:

$$\begin{aligned} Outcome_{ist} =& \alpha + \beta N\widetilde{REGA}_{it} + \gamma rank_{ij} + \delta rank_{ij}^2 + \eta N\widetilde{REGA}_{it} rank_{ij} \\ &+ \lambda N\widetilde{REGA}_{it} rank_{ij}^2 + \mu_j + \phi_s + \theta Baseline_{is} + \varepsilon_{ist} \end{aligned}$$

where $Outcome_{ist}$ is the dependent variable in district i during season s of year t; $\widetilde{NREGA}_{it}$ is a dummy variable taking the value 1 if the program is predicted to have been introduced in phase 1 in district i in year t; $Rank_i$ is the state normalized rank for district i; $Baseline_{is}$ is the value of the dependent variable in the same season of the previous year; μ_j are state fixed effects, and ϕ_s are season fixed effects.

This equation is of the quadratic flexible form, and is our preferred specification. We also estimate a similar equation in linear, linear flexible and quadratic forms. Across all four specifications, we find a significant increase in the riskiness of the portfolio of crops grown at the district level with the introduction of the NREGA. This increase is robust to the different specifications of the above estimating equation, to changes in sample, and to different measures of riskiness. Our estimates suggest that the NREGA was responsible for an increase in riskiness of about a fifth of a standard deviation, or about Rs. 5000 per hectare for the district as a whole, which is a sizable amount given the incomes discussed above.[2]

Our results are important for several reasons. Firstly, they constitute the first all-India discussion of this topic. Secondly, we provide the first decomposition of the movements in land allocation into those that are risk-increasing and those that are risk-reducing. Thirdly, in light of the current government's decision to scale the program back significantly, it is important to demonstrate that the program might actually have cumulative effects by making farmers more profitable, thereby permitting them to invest more money in their production activities and increase their incomes further.

[2]The current exchange rate is approximately $1=Rs. 60.

1.2 Rainfall Shocks and Cropping Patterns: Implications for Child Well-being

This paper uses household-level data from India to look at the effect of rainfall shocks on the level of health investments the mother makes, including whether or not she initiates breastfeeding, the amount of time spent breastfeeding the child, and whether the child is vaccinated against common diseases; all of which can impact both infant mortality and a surviving child's health in later life. I also look at the likelihood of the child surviving for a range of months post birth.

There is a large literature on the impact of income shocks in early life on the health outcomes of children. In some of these papers, especially those dealing with rural populations, rainfall shocks are used as negative shocks to income. In India greater rainfall is indeed associated with greater production, and in changes in the disease environment, both of which can affect child health. I do not focus on these channels, as they have been studied in detail. Instead I isolate a channel that has received much less attention in the literature - the changes in the mother's labor supply caused by rainfall fluctuations, and the consequent impact on her ability to engage in time-intensive activities that benefit her child. As far as I am aware this is the first paper to look at this question in the context of India, and one of a very small number of such studies in any developing country.

This study exploits the fact that a large number of the rural workers in India are reliant upon rainfall for their livelihood, and as a result shocks in rainfall can cause fluctuations in parental behavior as well as in the child's immediate environment, both of which can have (potentially contradictory) impacts on the

child's well-being. The effect of rainfall on the opportunity cost of the mother's time ist larger in districts where the mother plays an active role in agriculture, as opposed to districts where most of the work in cultivation is performed either by machines or by men. At the same time the effect of rainfall on other factors like household income and disease environments, for example, should be similar across rural agricultural districts regardless of the level of female involvement in agriculture.

There could be concerns that the level of female involvement in agriculture is not exogenous to investments in child health. Districts where cultural norms dictate that women play a larger role in agriculture would tend to see both greater investments in child health through larger control of household resources by the mother, and at the same time a shift towards crops that afford women opportunities to work as well. In this case higher female labor-force participation and the outcome of interest would both be driven by cultural attitudes and norms.

To address this question of endogeneity, I build upon the work of (Carranza, 2012) in using exogenous variations in the type of soil in a particular district to explain differences in female agricultural participation. Women engaged in agriculture in India are disproportionately employed in crop cultivation activities (like transplanting seedlings, and fertilizing and weeding the crops) rather than land preparation activities which are more labor intensive and are generally conducted by men (Basant, 1987; Foster and Rosenzweig, 1996). The texture of the soil in the district in question dictates to a large degree the methods that can be used to till the land. Tillage reduces the need for precisely the tasks that women specialize in and hence also reduces the importance of female labor. In

districts where the soil is loamy and deep tillage using machines is possible, female agricultural labor force participation is lower, as there is a reduced need for the operations that are traditionally undertaken by women. In contrast women still play an important role in crop cultivation in districts with clayey soils.

The texture of soil in a district is exogenously determined, and cannot easily be changed. However the texture of the soil does not in itself determine soil quality, the nature of the crop grown, cropping patterns or agricultural productivity. It determines only the ease with which labor can be substituted for using machines, and hence the relative importance of women in agriculture (Basant, 1987; Burton and Reitz, 1981; Foster and Rosenzweig, 1996). Thus the variation across districts in the texture of the soil can be used as an instrument for the level of female workforce participation. The key assumptions implicit in this strategy are that, firstly, the texture of the soil is indeed correlated with the level of female labor in agriculture, and secondly, that soil texture does not have any direct effect on the health investments made in children, which seems plausible.

The main question of this paper is whether negative rainfall shocks that lead to a reduction in the opportunity cost of the mother's time filter into reduced health investments in her child. Let Y_{idsrt} be the health-investment outcome of interest for child i in district d in state s in region r in time t, which consists of whether or not breastfeeding was initiated, the duration of breastfeeding if initiated, and whether or not the child was vaccinated against common diseases (polio, measles, tuberculosis and diphtheria-pertussis-tetanus (DPT)). α_{0r} is a region-fixed effect that captures cultural norms that are common across districts within the same region. $RAIN_{dsrt}$ is the rainfall shock in the district in question in the 12 months following the child's birth, which is normalized by the long-term

mean and standard deviation of rain in that district.

$FemLFP_{idsrt}$ is a dummy for whether or not the mother of child i was working in the 12 months prior to the month of interview. The vector X_{idsr} contains child-specific covariates like birth order, sex and mother's age at birth as well as household specific variables like the standard of living index, mother's education and father's education. These covariates are not time-varying. δ_{srt} is a state specific time trend, that would capture, for example, changes in the provision of medical services or the dissemination of information regarding the importance of these investments within a state over time. Finally ϵ_{idsrt} is the error term.

With these definitions in place, the main equation of interest can be written as

$$\begin{aligned} Y_{idsrt} =& \alpha_{0r} + \alpha_1 RAIN_{dsrt} + \alpha_2 FemLFP_{idsrt} + \alpha_3 (RAIN_{dsrt})(FemLFP_{idsrt}) \\ &+ \beta X_{idsr} + \delta_{srt} + \epsilon_{idsrt} \end{aligned}$$

The OLS results suggest that relative to households that do not have a working mother, positive rainfall shocks lead (on average) to an increase in child health investments, suggesting that the income effect dominates. This is the opposite of what we would expect if the opportunity cost of time mechanism were more important. We discuss why we think the OLS estimates are likely to be positively biased, and then present the second stage instrumental variable estimates. These are smaller and in the opposite direction, giving us the result that positive rainfall shocks affect households where the mother works by *reducing* investments in young children, and increasing the probability of the child dying in early months. In the most rigorous specifications a one-standard deviation positive rainfall shock decreases the probability that the child receives

certain vaccines by between 18 and 30 percent, and increases early infant mortality by up to 6 percent. In the context of the large literature that studies the positive impacts of increases in rainfall on child health and later outcomes, this is one understudied and interesting effect in the opposite direction.

1.3 For Better or Worse? The Effects of an Employment Guarantee in a Seasonal Agricultural Market (with Gary Fields)

This paper develops a market-level, tractable theoretical model of the Indian rural labor market. It includes certain 'stylized features' - of which seasonality in agricultural production is the most important. We introduce an employment guarantee in order to study the impact of the NREGA on the welfare of daily-wage agricultural workers, and demonstrate the interesting and counter-intuitive result that the introduction of such an Act can potentially have negative consequences for the welfare of workers. At the time of its introduction, the NREGA wage was pegged to be higher than the going slack season agricultural wage. As a result the introduction of the Act causes changes in private sector employment and wages, and these changes drive the negative welfare effect of the Act.

We model the Indian rural labor market as consisting of two types of agents. There are large farmers who hire agricultural workers to work on their land, and casual laborers who work for the large farmers for a daily wage but do not hire in workers for their own land. There are also two seasons in the model - a peak season when the agricultural labor market is close to full employment and agricultural wages are high, and a slack season where a sizable fraction of

the agricultural laborers are unemployed and agricultural wages are low (see (Basu, 2013) and (Johnson, 2009) for a similar segregation).

Most importantly for our model, there are inter-temporal spillovers in productivity from one season to the next. The amount of slack season labor hired to tend to the growing crop affects the size and quality of the harvest in the peak season, and hence the marginal productivity of the peak season worker. The greater the amount of slack season labor hired, the more care the crops receive and hence the more bountiful the harvest. In addition, we assume that slack season labor can be completely substituted for by natural factors. For example, rain could replace the need for irrigation, and some plants could survive even if no workers were engaged in weeding. Peak season labor, however, is essential to the production of the crop - if no-one is hired in the peak season then the farmer receives no income. This reflects the fact that peak season activities are time-sensitive - if the farmer cannot find laborers to harvest the crop at the right moment then the plants will die in the fields.

In this simple setup we demonstrate that it is possible that the NREGA actually makes agricultural workers worse off, especially when the inter-temporal spillovers are large. In the presence of inter-temporal spillovers, the NREGA causes a fall in the peak season wage by attracting workers away from agriculture in the slack season, thereby reducing the marginal product of the peak season agricultural workers. We illustrate this counter-intuitive result with a numerical example, and provide some simple comparative statics with respect to the efficiency of slack season labor and the size of the guarantee. As the size of the guarantee increases, the negative effects on worker utility increase, because the guarantee attracts more workers away from agriculture. Similarly,

increasing the efficiency of the slack season workers increases the negative welfare impact of the NREGA, because it strengthens the link between the amount of slack season labor hired and the amount of peak season labor demanded.

1.4 A Multi-sector Labor Market Model for India

The last chapter in this dissertation links the rural model of the previous chapter with an urban sector. In this chapter we build a stylized model of the urban and rural sectors and the linkages between them that arise as a result of the movement of labor from one sector to another. A large proportion of the internal migration in India consists of workers moving from the rural to the urban sector in search of work in the slack season, and moving back to the rural areas in the agricultural peak season. Our model captures both the outward and the reverse migration as equilibrium features. While we retain the assumption of seasonality from the previous model, we do not have any inter-temporal spillovers in this model, so this can be thought of as a special case of a more general model that would also have spillovers in productivity from the slack to the peak season.

Workers in our model are differentiated depending on where they are from (rural-born or urban-born), and on the amount of education they have (educated or uneducated). We assume that all workers from the rural areas are uneducated, but that urban workers can be either educated or uneducated. We model three kinds of jobs in the urban sector, which have a clear hierarchy in terms of their wages and status values. At the top of the job ladder are the managers' sector jobs. These are formal sector jobs that are only available to those work-

ers who have education. The next rung below this are the office-workers' jobs, which hire educated workers preferentially, but also hire uneducated workers if there are jobs left over after all the educated workers who want jobs in this sector have one. Finally the bottom-most rung in the urban sector consists of the free-entry sector jobs that are available to anyone who wants them. These are jobs like selling tomatoes on the side of the street, working as informal construction workers and so on There is a fixed total product in this sector, and workers working in this sector earn the average product. We assume that those workers with education do not work in this sector because there is a status cost involved with working here that is greater than the wage they could receive.

In the rural sector there is only one type of work, agriculture, but there are two seasons that differ in the wages that workers can receive and in their productivities. The seasons are similarly defined as in the model in Chapter 5. In addition to working in the agricultural sector, workers can also migrate to the urban sector in search of work in either an office-worker job or in the free-entry sector. Since this is a single-year model, we assume that urban free-entry sector workers have an advantage in the amount they can earn, say for example they know the best neighborhoods to sell tomatoes, or the busiest intersections to sell magazines to passing cars etc, and that rural free-entry sector workers do not learn from their urban counterparts by the time the world ends.

The model solves for the equilibrium search strategies for all types of workers, their ex-ante choices of which sector to search in, and the ex-post outcomes of how many workers get a job in each sector and how much they earn. We then introduce an employment guarantee into this pre-EGA model, and study what happens to the ex-ante search choices, as well as the ex-post outcomes for all

types of workers. We show that the introduction of the employment guarantee causes a reduction in the peak season wage, though this time the mechanism is different from the mechanism in Chapter 5. In this model, the EGA provides a viable alternative to the agricultural work and to migration for the rural population, thereby reducing the number of workers who migrate in search of the office-workers' jobs. In doing so, it also reduces the number of rural workers who actually get jobs in that sector. This increases the number of workers who want to find a job in agriculture in the peak season, and thus depresses the wage that each of these workers receives.

We conclude this chapter with welfare analyses conducted using the cumulative income distributions for the economies with and without the employment guarantee. We show that because of the fall in the peak season agricultural wage, the impact of the employment guarantee is, for the large part, ambiguous.

The rest of the dissertation presents the papers discussed above in their full detail.

CHAPTER 2

PROVIDING MORE THAN JUST EMPLOYMENT? EVIDENCE FROM THE NREGA IN INDIA (WITH SIDDHARTH HARI)

2.1 Introduction

Risk is a central feature of agricultural economies, particularly in a developing country context, and arises as a result of shocks that are either covariate or idiosyncratic in nature. In countries where irrigation infrastructure is poor, like India, weather-related shocks – for example, extremes of rainfall in either direction – significantly and unpredictably affect agricultural production. In addition, fluctuations in prices add to uncertainty about future agricultural profits and farmer incomes. Apart from these aggregate events which affect entire villages or regions at a time, individual farmers are often also subject to idiosyncratic shocks, such as a death in the family or individual health shocks, which can significantly hamper work and require large unexpected expenditures.[1]

A combination of low savings and lack of access to formal credit means that farmers in these countries are unable to smooth consumption in bad times, further blunting their ability to cope with risks. Mutual insurance networks – such as borrowing from friends and family, or within one's caste group – are widely prevalent (see (Ligon et al., 2002; Munshi and Rosenzweig, 2013; Rosenzweig, 1986; Rosenzweig and Stark, 1989; Townsend, 1994) for evidence on the importance of these networks in the Indian context). However, these networks are able to provide insurance only against idiosyncratic shocks, leaving farmers vulner-

[1]Evidence from South Africa suggests that spending on funerals and weddings can amount to as much as an entire year's income, see (Case et al., 2013).

able to covariate shocks affecting production in agriculture. In addition, the informal sharing norms in these networks can hinder long-term growth through adverse incentive effects and restrictions on savings ((Grimm et al., 2011)).

In this paper, we use a regression-discontinuity design to study changes in risk taking behavior in response to a large-scale government-sponsored employment guarantee program, the National Rural Employment Guarantee Act (NREGA), in India. Though the program was not assigned to districts in a random manner, assignment was based on a poverty index (created almost ten years prior) that we argue was exogenous to program placement. We recover the information used in the index and use that to predict the assignment of districts to the NREGA. This empirical methodology allows us to isolate the causal effect of the program on our outcomes of interest - the riskiness and the average returns of the portfolio of crops grown.

There are several potential measures of risk-taking behavior that we could study. These include spending more time on one's own farm as opposed to seeking wage labor in the private sector, shifting to non-farm self-employment, and growing riskier crops, among others. Here we focus on crop choices made by farmers, and ask the following question: "What is the impact of the presence of some form of 'sure income' on the ability of farmers in India to grow a portfolio of crops that is characterized by higher risk but also potentially higher returns?" Our hypothesis is that when farmers are given access to an additional source of income they may shift land and other resources away from crops that are low return but also low risk to other crops that are potentially more profitable but pose greater risk.

To preview our results, we find evidence of an increase in the riskiness of

the overall portfolio of crops at a district level when the employment guarantee program is introduced. We do not have cost of cultivation data, but we use information on both crop revenues and crop yields to calculate our measures of portfolio risk. The results are consistent across the two sets of measures and various specifications of the econometric model. Our results suggest that the mix of crops being chosen by farmers is affected by the additional income they may receive from participation in the employment guarantee program. We do not find a significant increase in the mean yields or mean revenues of the portfolios of crops, but show that riskier portfolios on average do generate higher returns. Finally, we decompose the shifts in land allocated to various crops into those movements that increase portfolio risk and those that lower it. We do not find evidence of an increase in the total amount of land reallocation, or in its composition. We discuss how our findings of an increase in portfolio risk but of no significant change in the volume of land allocation can be reconciled, and why the data we have may not be adequate to answer the question of changes in land allocation.

India provides an ideal setting for the study of the impact of a welfare program on the production decisions of farmers for a number of reasons. First, agriculture and allied activities still employ a large fraction of the population. As of 2001, agricultural workers constituted a sizable 56.6% of the working population.[2] Second, access to credit and insurance is very low among the rural population. Only about 17% of farmers have access to crop insurance (see (Mahul et al., 2012)) and between 70 to 85% of small and marginal farmers do not have access to any formal credit.[3] Third, agriculture is heavily rainfall-dependent.

[2]Census of India. Agricultural workers include self-employed as well as those daily wage laborers.

[3](Satyasai, 2012) finds that as of 2006-07, 80% of farmers with less than one hectare of land and almost 70% of farmers with between 1 and 2 hectares of land in India did not have access

The seasonality of Indian agriculture translates into extended slack agricultural periods where incomes are low and work is hard to come by, in between periods of planting and harvesting when agricultural wages are high and the village economy is close to full employment. There is a large body of literature that suggests that the seasonal fluctuation in incomes together with the inability to guarantee future incomes can have an effect on production decisions, encouraging farmers to engage in activities that are less risky. In this context of credit constraints and uninsured risk it is hardly surprising that risk-averse agents choose low-risk and low-profitability technologies over technologies that require higher fixed costs or embody greater risk (Dercon and Christiaensen, 2011; Fafchamps and Pender, 1997; Gine and Klonner, 2006; Rosenzweig and Binswanger, 1993). Access to sure income from a government welfare program could have large effects on planting decisions. This paper evaluates these effects.

In recognition of the high incidence of poverty and the absence of social safety nets in rural India, the Indian government launched the NREGA in 2006. Under the provisions of this Act rural households were entitled to one hundred days of employment per year in the form of unskilled manual labor, at a state-specific minimum wage. Most of the labor supplied under the program is used for the construction of canals and ponds for irrigation, and for improving rural connectivity through the construction of mud roads. Though not specifically mandated, NREGA work in many districts occurs during the slack season (see (Imbert and Papp, 2011)).[4] Thus the Act not only increases overall household

to formal bank credit which could help smooth consumption over the course of the year.

[4]Some of this is driven by the timing of workers' demand for work, some by lobbying by farmers who need laborers during the peak season, and some imposed by the state through 'work calendars' (see (Johnson, 2009)). The exact timing of the slack season varies by state and district because the monsoon reaches different parts of the country at different times.

incomes, it also tends to do so in precisely the season where incomes are low.

To the extent that the introduction of the NREGA affects production technologies, the additional income associated with higher profitability is an important spillover effect of the program. Since, as discussed above, low incomes and lack of insurance opportunities mean that farmers often choose to forego potentially higher incomes by playing it safe, the NREGA could act as an implicit insurance mechanism by raising incomes in the bad states of nature and thereby helping low-income farmers increase their average incomes. Identifying this unintended positive spillover of the Act is particularly important at this time when significant scaling back of the program is being discussed by the current government.[5]

In addition to acting as a substitute for formal insurance, the program could be relaxing other constraints farmers face, for example credit constraints. Increasing annual incomes could help farmers incur the fixed costs associated with switching production away from crops they have been growing in the past, and into crops which they perceive will provide them with the highest returns. If that were the case, a relaxation of credit constraints by the NREGA would affect the total amount of land reallocation among crops from one year to the next. In order to test this hypothesis we develop a measure of overall mobility at a district level, and also decompose it into mobility in the direction of increased risk taking and decreased risk taking.

The rest of this paper is organized as follows: Section 2.2 describes the empirical literature on risk-aversion and production decisions and the NREGA program details, Section 2.3 presents a simple theoretical model and Section 2.4

[5] http://www.thehindu.com/news/national/mnrega-may-be-restricted-only-to-backward-tribal-districts/article6413021.ece.

discusses the measures of risk. Section 2.5 lays out the empirical methodology we employ and Section 2.6 presents the results. Section 2.7 concludes.

2.2 Empirical context and program details

In this section we provide some evidence from developing countries regarding the impact of risk and credit constraints on production decisions, and discuss the findings from those papers that deal specifically with the Indian context. We then discuss the main features of the NREGA and why it can be considered a sure source of income for farmers. We also briefly review the existing literature on the impact of the NREGA on labor market and other outcomes. Since most of the focus of the literature on NREGA in India has been on wage and employment and not on risk or production decisions, the purpose of discussing these papers is largely to highlight the methods that most authors have employed in attempting to establish causality.

2.2.1 Credit constraints, uninsured risk and production decisions

Lack of access to credit and uncertainty about future incomes are two of the most important reasons why poor farmers might be unable or unwilling to invest in risky technologies, even when they know the risky technology to be more profitable than the safe one. These two reasons are closely related, however - being able to provide collateral or guarantee a steady income is often a prerequisite for receiving credit, while receiving credit and being able to invest

in a profitable venture might often be the only way that an individual can increase her earning prospects.

There are several reasons why new agricultural technologies might be riskier than existing ones ((Foster and Rosenzweig, 2010)). First, new varieties of seeds might be more susceptible to weather shocks, or require a steady irrigation source, which is often absent in developing countries. Second, lack of knowledge about input management may increase the variability of yields from such technologies. Lastly, new technologies often require greater up-front investments before the resolution of the uncertainty, e.g. high yield variety (HYV) seeds need fertilizer investments before the weather shock or failure of the crop is fully known, and those investments (unlike the quantity of labor hired) cannot be adjusted ex-post.

For the most part it is difficult to disentangle the effect of credit constraints from the effect of insurance convincingly. In one of the few studies to do so, (Karlan et al., 2012) conduct experiments in Ghana where poor farmers are randomly assigned opportunities to purchase insurance or cash grants, or a combination of the two. They report a high demand for insurance. The access to insurance significantly increases the investments made by the farmers, including amount spent on land preparation and on chemicals used on the land, but the effects of the cash grant are relatively small, suggesting that uncertainty plays a larger role than credit constraints in this context.

(Dercon and Christiaensen, 2011) study the investment in fertilizer by agricultural households in Ethiopia. Fertilizer use improves mean yields from the land, but involves a substantial sunk cost. Farmers are faced with uncertainty about the weather and hence the success or failure of their harvest. In the ab-

sence of means to smooth their consumption they under-invest in fertilizer.

The divisibility of the investment also seems to matter. (Fafchamps and Pender, 1997) study investments in non-divisible assets, in particular irrigation wells in India. These wells are profitable, but because they are non-divisible and irreversible in nature poor farmers are unwilling to invest in them, partly for fear of being unable to buffer against short-term consumption shortfalls caused by weather and other shocks during the investment period. Simulation results demonstrate that providing credit to these farmers can have large impacts on the level of investment and hence potentially also impact the profitability of the land.

(Gine and Klonner, 2006) study the barriers to the adoption of a profitable technology by members of a fishing community in Tamil Nadu in India. They find that lack of asset wealth is a strong indicator of delays in the take-up of the technology, and interviews with the fishermen corroborate the hypothesis that credit constraints are the main reason, followed by higher risk aversion among poorer households. (Moser and Barrett, 2006) look at the adoption of a new rice technology in Madagascar and find that (among other things) the household having a stable source of income is a strong predictor of the decision to adopt, the decision of how much land to allocate to the new technology, and continued use of the technology.

There is also evidence that farmers are well aware of risk considerations and adjust their asset portfolios in response to the expectation of shocks. (Rosenzweig and Binswanger, 1993) show that poor farmers in India adjust their portfolios of crops to reflect the variability of rainfall they face. When rainfall variability increases, they tend to choose a portfolio that is less influenced by rain-

fall, but which also generates lower average returns. The same effects are not present among richer farmers. Morduch (1990) also shows that poorer farmers who are exposed to risk plant less risky crops than wealthier farmers. (Wadood and Lamb, 2006) use ICRISAT data from the semi-arid tropics in India to study the question of crop choice. One of the ways in which households in their dataset mitigate risk is by changing their demand for or supply of labor from the off-farm market. They find that when households are faced with greater employment risks on the off-farm market, they compensate for this by choosing a portfolio of crops that has less risk.

The main lesson to be learned from these papers is that increasing the incomes of these farmers or reducing the amount of uninsured risk they have to face can indeed increase the risks they are willing to bear, potentially making their farms more profitable.

2.2.2 Details of the NREGA program and other studies of its effects

There are three features of the NREGA that make it both interesting and novel to study. The first feature concerns the scope of the program. NREGA incomes are large and significant when compared to the incomes of the individuals who take up the Act's benefits.[6] Small and marginal farmer annual incomes are estimated to be between Rs. 20,000-Rs. 30,000 (see table 14 in (Dev, 2012)).[7] Assuming they were able to work their full entitlement under the NREGA at a wage rate

[6]Survey data from 2007 finds that the two largest groups of NREGA workers are self-employed agricultural workers and landless laborers, not large farmers who hire in laborers.

[7]As of October 2014, the exchange rate is approximately $1=Rs. 60.

of about Rs. 100 a day they would have the ability to increase their incomes by as much as 30% - 50%. This is a large amount of money, with potentially large effects on behavior. Secondly, the annual cost of the NREGA is close to 1% of India's GDP, making it the world's largest public works scheme, and it currently benefits close to 50 million households in the country. This program clearly represents an important topic of study.

The second feature of the NREGA that makes it attractive to analyze from an econometric point of view is the fact that it was rolled out across districts in India in a phased manner. The Act was first introduced in 2006 in 200 of the country's poorest districts, then rolled out to an additional 130 districts in 2007, and was finally made available in every rural district of the country by mid-2008. These three waves will be referred to in this paper as Phases 1, 2 and 3. While the roll-out was not conducted randomly, other aspects of the program design allow us to identify exogenous factors that determined treatment, which we will discuss later.

The third feature, perhaps most important for the purpose of this paper, is that the Act was designed to be 'demand-driven', i.e., that work would be provided within 5 kilometers (approximately 3 miles) of the worker's place of residence within 15 days of the worker filing for work with the village authorities. This feature means that unlike other workfare programs that are initiated and conducted in a 'top-down' or supply-driven manner, the NREGA actually guarantees that work is available to those who need it, when they need it. In addition to this feature, much of the NREGA work is provided in the slack season, precisely when workers' agricultural incomes are low and agricultural work is hard to come by. This allocation of work to the agricultural slack season means that

there was very little overlap in time spent working on harvesting or planting and time spent working on the NREGA. NREGA income can thus be thought of as being separate from income generated from land.

The NREGA guidelines provide for up to 100 days of manual unskilled work per rural household. Each household is provided with a job card which contains the names of all adults who are eligible to work on NREGA work sites, and which is used to record the total number of days the household has used in that particular year. A "household" is informally defined as the set of individuals who cook around one common stove, or *chulha*. Workers apply to their village-heads for work, and work is sanctioned at the block level[8].

According to the official guidelines, sanctioned projects are to be chosen from a schedule of works, which includes the building of canals, roads and ponds, afforestation, leveling of fields, and even the building of toilets on the lands of disadvantaged communities - the focus being on the "creation of durable assets and strengthening of the livelihood resource base of the rural poor" (see (Chakraborty, 2007)). The prioritization of certain works over others is democratically decided in a *gram sabha*, or village meeting. Workers are paid for work completed on either a piece-rate or a daily-wage basis, subject to the worker receiving at least the minimum wage per day. Minimum wages are state-specific and subject to a national minimum. Most of the state-specific minimum wages are set higher than the national minimum, however, so the national minimum is typically not binding. Piece rate wages paid to workers differ depending on the nature of the work and of the ground - for example hard stony earth fetches more per cubic meter dug than soft earth.

[8] A block is a collection of villages within a district.

In the first few months of implementation many of the above guidelines were repeatedly flouted. Many of the problems arose from the fact that workers were not aware of their rights under this new Act - they did not know that they could apply for work, nor were they aware that they were entitled to unemployment insurance if their payment was delayed for more than two weeks. Other problems with implementation included the siphoning off of funds intended for the workers or for materials, the illegal use of machinery to complete work which was intended for laborer and the fudging of the attendance sheets (the "muster-rolls") by the work site supervisors in order to add the names of cronies or to over-report the number of days worked by laborers so that the extra funds could be pocketed (see (Dutta et al., 2012; Niehaus and Sukhtankar, 2013)).

Despite these initial adjustment problems, there is some evidence that the program has indeed had a tangible effect on both labor market as well as peripheral outcomes. A large focus of the literature has been on determining the direction and magnitude of labor-market effects of the program. Studies have found that the introduction of the NREGA lead to an increase in private and public sector wages and employment ((Azam, 2012; Berg et al., 2012; Imbert and Papp, 2011), and an increased investment by farmers in labor-saving technologies ((Bhargava, 2014)). In addition, (Zimmermann, 2013) showed that small farmers are substituting away from private sector work and allocating more time to working on their own farms - a riskier activity. Both (Zimmermann, 2013) and (Johnson, 2009) find that the take up of work under the program increases following a bad weather shock, suggesting that the NREGA could be substituting for weather insurance. Together, these last two papers can be taken as some evidence that the program might indeed be acting as a social safety net.

Finally, the paper closest to ours is (Gehrke, 2014). She studies shifts in crop choice as a result of the NREGA using data from one state in India (Andhra Pradesh), and finds evidence for increased risk taking. She uses difference-in-difference (DID) methods combined with matching and fixed-effect models, and the Young Lives Survey (YLS) to study this question. Her paper provides valuable insight into risk-taking behavior among farmers. We draw inspiration for our question from her study.

The contribution of this paper is three-fold. First, we analyze the question of the effect of the NREGA on crop choice at an all-India level, a broader scope than previous work on this question, and one which is of considerable import. Given that Andhra Pradesh is among the best performing states with regard to NREGA implementation, it is interesting to see if Gehrke's results extend across the country. Second, we use a regression-discontinuity design to obtain our estimates. Since the program roll-out was based on a poverty index, we are concerned that program assignment was not exogenous to other district characteristics that could influence crop choice, like farmer incomes and the number of small or marginal farmers. DID methods, while informative, would provide biased estimates of the effect of the program. Finally, we offer a theoretical decomposition of crop-change behavior into risk-increasing and risk-reducing land reallocations, which informs our results further. We find evidence of increased risk taking as measured by both the standard deviation and coefficient of variation of the portfolio of crops, and that these results are reasonably robust to changes in specification and in sample-selection.

2.3 Theoretical Model

Consider the problem of a farmer with log utility who has to allocate land to growing 2 crops, R and S. Crop S is 'safe' and yields a return of R_s with probability 1. Crop R is risky, and yields a return of R_r with probability p_r and 0 with probability $1 - p_r$. We assume that $p_r R_r \geq R_s$, i.e., that the riskier crop yields an expected return that is at least as large as the return from the safe crop.

The farmer does not have access to any borrowing or saving technology, and therefore consumes her income each period. This assumption is reasonable given the evidence on the lack of credit market instruments available to small farmers in rural India.[9] Apart from income from agriculture, the farmer earns income from non-agricultural sources, denoted by $\bar{y}$. This $\bar{y}$ includes transfers and subsidies from the government and does not depend on the amount of land allocated to either crop. In our model, the introduction of the NREGA can be thought of as an increase in $\bar{y}$.

Let λ denote the proportion of land allocated to the risky crop, R. The farmer's problem is then to allocate land among the two crops in order to maximize her expected utility, which can be written as:

$$Max_{\lambda}\{p_r[Log[(1-\lambda)R_s + \lambda R_r + \bar{y}]] + (1-p_r)[Log[(1-\lambda)R_s + \bar{y}]]\}.$$

Taking first order conditions and simplifying, we get:

$$\frac{(1-\lambda)R_s + \lambda R_r + \bar{y}}{(1-\lambda)R_s + \bar{y}} = \frac{p_r(R_r - R_s)}{(1-p_r)R_s},$$

which yields

$$\lambda = \frac{(R_s + \bar{y})(p_r R_r - R_s)}{R_s(R_r - R_s)}.$$

[9] See (Satyasai, 2012).

Given our assumption that $p_r R_r \geq R_s$ we are guaranteed that $\lambda \geq 0$. However, if $p_r R_r > R_s$, then for high enough values of $\bar{y}$ we might reach a corner solution with $\lambda = 1$, i.e., with the farmer allocating all her land to the risky crop.

Now, the introduction of the NREGA can be thought of as an increase in $\bar{y}$. As is clear from the solution for λ, such an increase will increase the proportion of land allocated to the risky crop. Therefore we would expect to see an increase in land allocated to risky crops in response to the program.

2.4 Measures of Risk

In this paper we focus on two broad sets of measures for risk and changes in cropping patterns. The first set should be familiar to most, and includes the standard deviation and coefficient of variation of the yields and revenues of the overall portfolio of crops grown by farmers in a district. The second is a measure we develop to measure overall changes in cropping patterns and then decompose these changes into risk-increasing changes and risk-reducing changes. We defer the discussion of the measure for changes in cropping patterns to the Appendix, and here provide the formulation of the risk measures on which our main results will be based.

2.4.1 Using yields

The first measure of risk we use is based on the variance in crop yield, which is measured simply as units of output per unit of land (standardized to a common set of units). Let $\lambda_{ist} = (l_{1ist}, l_{2ist}..., l_{n_{is}ist})$ be the allocation of land to the n_{is} crops

grown in district i in season s. n_{is} is the complete list of crops grown in any year in that district-season combination, and if the crop if not grown in any one year the yield is 0 for that year. We distinguish between seasons here because the types of crops grown and the weather conditions vary across seasons as well as across districts. Let $\upsilon_{ist} = (y_{1ist}, y_{2ist}..., y_{n_{ist}ist})$ be the yields for those n_{is} crops in year t. Let the time-series season-specific mean yield of a crop k be given by

$$\bar{y}_{kis} = \sum_t \frac{y_{kist}}{T}$$

where T is the total number of time periods for which we have data. The co-variance in yields of two different crops k and m is given by

$$cov(y_{kis}, y_{mis}) = \sum_t \frac{(y_{kist} - \bar{y}_{kis})(y_{mist} - \bar{y}_{mis})}{T-1}.$$

Then we define the risk of the portfolio of crops grown in that district-state-season-year combination as

$$\rho_{ist} = \sqrt{\lambda'_{ist} \times \Upsilon_{is} \times \lambda_{ist}}.$$

where

$$\Upsilon_{is} = \begin{bmatrix} var(y_{1is}) & cov(y_{1is}, y_{2is}) & . & . & . & cov(y_{1is}, y_{n_{is}is}) \\ cov(y_{2is}, y_{1is}) & var(y_{2is}) & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ cov(y_{n_{is}is}, y_{1is}) & . & . & . & . & var(y_{n_{is}is}) \end{bmatrix},$$

is simply the variance-covariance matrix of the yields in a given district-season combination.

In this calculation crop riskiness is specific to the state-district-season we are studying, and a crop that is the riskiest in one particular state and district need not be risky at all from the point of view of another state and district.

Let the district-state-season-year specific weighted mean yield be given by

$$\bar{y}_{ist} = \lambda_{ist} \cdot \upsilon_{ist}.$$

Then we define the coefficient of variation as

$$cv_{ist} = \frac{\rho_{ist}}{\bar{y}_{ist}}.$$

The standard deviation and the coefficient of variation are the two measures of risk we use in our analysis.

2.4.2 Using revenues

Fluctuations in prices are a large source of risk for farmers, and so we also use revenues from the portfolio of crops (evaluated at retail prices) in order to construct measures of risk. The setup is analogous to the previous section, with revenues instead of yields.

We create the same two measures - the standard deviation of the revenues evaluated at retail prices, and the coefficient of variation. The retail prices we use are at a state-district-year level. Unfortunately these prices are not available at a seasonal level within districts.[10]

One possible criticism of using the retail measures described above is that fluctuations in yield could be causing fluctuations in prices, and hence the two

[10]We also tried using government announced minimum support prices (MSPs) but did not find significant results. The advantage of using the MSPs is that they are unaffected by fluctuations in yields, and are only adjusted for inflation each year. The disadvantage is that MSPs are country-wide and so we lose a lot of within-state and across-state variation, as well as the fact that the basket of crops that the government chooses to provide support for are likely to differ in some way, either in their inherent riskiness or in the number of farmers that grow them.

cannot be neatly separated. With the data we have we cannot easily address this question. The regression of prices on yields in the same year (including fixed state, season and crop effects) does not yield a significant coefficient, though replacing contemporaneous yields with yields in the previous season does yield a coefficient on yields that is significant at the 5% level. This suggests that while prices and yields are correlated in a lagged fashion, contemporaneous correlation might not be too much of a concern.

2.5 Empirical Methodology

In this section we discuss the manner in which the NREGA was rolled out across districts in India, and how we use this information to develop our regression discontinuity design. We outline in detail the algorithm we use, and the construction of our running variable. We then discuss the main assumption of the RD design, the implicit control assumption, and provide evidence in support of this assumption being valid for our setting. We then discuss the data we use in this paper, and finally we present the estimating equations for the OLS and RD specifications.

2.5.1 Program Roll-out and the Regression Discontinuity Design

The phased roll-out of the NREGA allows for the analysis of the causal effect of the program on a wide range of outcomes. As mentioned earlier, many of the studies of the NREGA have used a difference-in-difference (DID) approach to

measure the impact of the Act on labor market outcomes. Such an approach is problematic as the program was not introduced in a random manner and was in fact deliberately targeted at the poorest districts first. Many of the earlier studies also use data which were collected after Phase 2 of the program was rolled out, meaning that they can only compare Phase 2 and Phase 3 districts. Since many of the Phase 3 districts are some of the richest in the country (and are from the richest or most developed states like Haryana, Punjab, Kerala, Tamil Nadu and Gujarat) the common-trend assumption can be hard to defend. Indeed, (Zimmermann, 2013) provides evidence that it does not hold for some outcomes.

Instead we use the fact that the NREGA was assigned to districts based on a poverty measure in order to develop a regression discontinuity design that we believe is the most robust methodology for analyzing this program. The poverty measure was based on data that pre-dated the NREGA by almost ten years, and was itself published two years before the NREGA was announced ((Planning Commission, 2003a)).

Regression discontinuity essentially relies on a "jump" in treatment probabilities around a certain cut-off to identify program impact. If treatment status is based on a certain cut-off, then units within a small range on either side of the cut-off are likely to be similar on both observable and unobservable characteristics. For this small range around the cut-off therefore, program assignment can be thought of as random.

The NREGA program was first introduced in April 2006 in the poorest 200 districts in the country. In April 2007, the next 130 poorest districts got the program. Finally, in April 2008 the program was introduced in all remaining

districts[11]. Figure 2.1 shows the phased roll-out of the program. Since the allocation was based on some poverty measure (or measures) it lends itself to a regression discontinuity design, provided we can recover the ranking on which this decision was based.

We use the information in (Zimmermann, 2013) to recover the algorithm used to rank the districts in order of poverty. The actual treatment decision consisted of two stages. In the first stage the number of districts in each state to be allocated the NREGA in the phase in question was determined on the basis of the proportion of people in that state who were poor. In the second stage the exact identity of the districts to receive the program from each state was determined on the basis of a 'backwardness' ranking, with the most backward districts within a state receiving the program first.

The choice of the particular poverty measure employed for the first stage and the data on which the poverty calculation was based are not publicly known. (Zimmermann, 2013) makes an educated guess at the criterion used. She uses the poverty headcount ratio calculated from the 1993-1994 National Sample Survey (NSS) to calculate the "incidence of poverty" for each state. Since the poverty measure is not known with certainty, we do not follow (Zimmermann, 2013) in this first stage. Instead, we simply take the total number of districts allocated the program in a particular state as given, and then proceed to the second stage.

In the second stage, districts within a state were ranked in terms of their backwardness, and the districts with the highest rankings within each state received the program first. The ranking used for this purpose is publicly available

[11]India has a total of 655 districts, of which 30 are urban districts and hence did not receive the program in any wave of the roll-out.

from (Planning Commission, 2003a), a report which provides a list of districts along with the calculation of an "Index of Backwardness". This is a composite index which assigns a score to each district along three dimensions - the percentage of Scheduled Caste and Scheduled Tribe individuals in its population[12] (from the 1991 census), agricultural wages (from 1996-97), and the output per agricultural worker (from 1990-93). These three dimensions are then aggregated to form a composite index for the district.

The Planning Commission ranking is available only for 447 districts in 17 states.[13] (Zimmermann, 2013) suggests that part of the reason for the omission of entire states might be because of "internal stability and security issues" during the time that the data used in constructing the index was collected. As a result of this it is plausible that these states played a larger role in determining which of their districts received the program, which would violate the implicit control assumption of the regression discontinuity design. We simply omit those districts for which we do not have information on the ranking.

One complication is that program allocation was not based entirely on these backwardness rankings. Broadly speaking, two additional factors determined treatment. First, given the high variance in the backwardness index, some states, e.g. Kerala, had no districts among the 200 most backward. Owing to political pressures, however, it was decided that treatment would be made such that every state in India had at least one district which received the program in each phase. Second, many districts in India are affected by the Maoist (an extremist left wing) movement. The government viewed the NREGA as a tool

[12]These are two groups of historically disadvantaged individuals that are formally recognized by the Constitution of India.

[13]India has a total of 29 states and 7 union territories. The 17 states for which the ranking is available are the most populous states, however, and according to 2011 Census data they make up 94.66% of India's population.

to combat the influence of the Maoists and decided that all districts affected by the movement would be assigned the NREGA in phase 1. These districts are in the states of Andhra Pradesh, Bihar, Jharkhand, Madhya Pradesh, Chhattisgarh, Orissa and Uttar Pradesh. The list of 32 'extremist affected districts' is available in (Planning Commission, 2003b).

Using these elements of the NREGA program design, we construct an algorithm that predicts program assignment and use this to implement our RD strategy. The program prediction algorithm works as follows:

- **Step 1:** Rank districts within a state s using the Backwardness Index published by the Planning Commission.
- **Step 2:** Take the total number of NREGA districts *actually* assigned to a state s, n_s, as given. Allocate the available n_s slots starting with the most backward district in the state receiving the program first.
- **Step 3:** Create state-normalized ranks by assigning the most backward district in a state a rank of $-n_s$, the second-last district the rank of $-n_s + 1$ and so on, such that the district with state-normalized rank of 0 is the last district predicted to receive the NREGA in each state.

The construction of the state-normalized rank ensures that the discontinuity is evaluated at a common point across states, in this case at the value 0. The state-normalized rank is the 'running variable' in our regression discontinuity design, i.e., the numerical variable which predicts the probability of treatment.

Clearly this process can be performed so as to generate cut-offs for both phases, but we focus for now on the Phase 1 cut-off. This is because we believe that districts in Phases 1 and 2 were more similar to each other than Phase

1 and 2 districts were to Phase 3 districts, so the comparison of cropping choices along this boundary is more justifiable. By construction, districts with negative state-normalized ranks are predicted to have received the program, and districts with strictly positive state-normalized ranks are predicted to not have got it. If we restrict attention to a small enough bandwidth around the cut-off of 0, treatment prediction can be thought of as being random since districts have similar backwardness index scores, but differ in treatment status. Alternatively, we can expand our bandwidth around the cut-off and use a flexible polynomial specification. Since the number of districts and state-normalized ranks is small, we prefer to use all districts without restricting our attention to a bandwidth. We present robustness checks which exclude Phase 3 districts, and employ a number of different specifications.

How well does the algorithm do in predicting program allocation? Table 2.1 gives a state-wise break-up of the actual number of districts assigned the NREGA in Phase 1 and the algorithm prediction success. As is clear from the table, the algorithm performs better in some states than in others. The overall probability of success in prediction is 80%. The table also reports the number of false negatives - i.e. the number of districts our algorithm predicts as not having received the program that did in fact actually receive the program in Phase 1.

Given the imperfect prediction of program assignment - the cut-off of 0 does not predict actual treatment deterministically - we use a fuzzy RD design. Figure 2.2 plots the mean probability of receiving the program for each state-normalized rank. The fitted curves are quadratic, with 95% confidence intervals plotted as well. The 95% confidence intervals do not overlap, though the discontinuity is not very clear.

One of the steps that was followed in the actual assignment of the program was to allocate the program to all 32 'extremist' districts regardless of their ranking in the Planning Commission Backwardness index. The 32 extremist districts were *not* always among the lowest ranked districts in their states. Often these districts were close to the state-specific cut-off, but sometimes they were significantly more 'advanced'. Not taking this into account while reconstructing the algorithm results in a greater probability of incorrect assignment around the cut-off of 0 that visibly dampens the discontinuity at this point. To demonstrate this we also present the graph of the fuzzy RD without the inclusion of the extremist districts (see Figure 2.3). As can be seen from this figure, the discontinuity is much stronger at the cut-off point. We thus also present results for the sample without the extremist districts.

2.5.2 Possible Concerns

One of the concerns with any RD design is potential manipulation of the assignment variable by the beneficiaries of the program. Since treatment is on the basis of a cut-off, if units could slightly misreport or manipulate their scores then the "quasi-randomness" of allocation around the cut-off would no longer hold. In our case, this means that if it were possible for districts to misreport their scores on the backwardness index, then we would be worried that districts on either side of the cut-off systematically differ on unobservable characteristics such as potential benefits from the program or political influence.

This is unlikely to be the case for the NREGA. Firstly, the data used for construction of each of the dimensions of the backwardness index comes from the

mid-1990s, so it pre-dates the introduction of the NREGA by approximately 10 years. Secondly, the same ranking had been used previously for the allocation of other welfare programs, but with different cut-offs of 100 or 150 districts. So even if districts could manipulate their ranks, there is no reason to think that they would know to aim for the NREGA cut-off of 200 districts in Phase 1.

To further demonstrate that the districts did not have control over their ranking, Figure 2.4 depicts the relationship between the backwardness index score and the overall district ranks (which range from 1 to 447). If there was strategic misreporting we would expect to see clustering just below the cut-offs of 200 (Phase 1) and 330 (Phase 1 and 2 together), which does not seem to be the case. As can be seen from the graph, the relationship is smooth, with a couple of inflection points where the curve becomes flatter or steeper. These inflection points are at around 100 and 400 districts respectively. The backwardness index score ranges from 0.08 to 2.3.

Finally, Figure 2.5 depicts state-wise graphs of the state-normalized rank versus the composite ranking of districts. In this graph, discontinuities around the normalized cut-off of 0 would be suggestive of manipulation on the part of individual states. We do not see any evidence of such discontinuities.

The evidence presented above gives us confidence in the fact that the individual districts did not have control over their scores in the Backwardness Index.

2.5.3 Data

The information needed to recreate the algorithm for the district assignment of NREGA comes from the (Planning Commission, 2003a). This document provides the score of each of the 447 districts on each of the three indicators - agricultural wages, percentage of the population that is made up of SC/ST individuals and output per agricultural worker, the composite score which is a combination of these three scores, and then the ranking of the districts according to the composite score. As described above we use the ranking of districts in this document to create the running variable of our RD, the state-normalized rank.

Information on crop yields, land use and season comes from a dataset published by the Ministry of Agriculture, Government of India. This dataset is a district level panel for the years 1998 to 2010, and provides information on area under cultivation and total production for each crop grown in all districts in India, across *rabi*, *kharif*, autumn, winter and summer seasons of the year. The number of crops varies across districts in a given year and season, across years within a particular district and season, and across seasons within a particular district-year combination. Data on crop production is in metric tons and on land use is in hectares (ha).

The price data used is retail prices for the period 2001-2010 from the Retail Prices Information System, Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, Government of India[14]. The price data is at a sub-state regional level. More accurate measures of district level prices might match districts with the closest available regional price value, but without any way of analyzing the distance between individ-

[14]This data is available at http://rpms.dacnet.nic.in/Bulletin.aspx

ual districts and the regional focal point we simply aggregate the information to a state level. All prices have been deflated using the Consumer Price Index (CPI), and are evaluated at a 1986-87 base. The CPI is available at a state-year level. We calculate the simple mean of the prices for a particular state-crop-year combination and use that number for every district within the state. We should mention here that the price data is limited and covers only about a third of the crops in our sample. We have been unable to find information on other crops from other sources.

We have also tried using government minimum support prices (MSPs) instead of retail prices. The MSPs are available for an even smaller number of crops than the retail price information, and are also centrally-announced prices and so do not vary by state. Using the MSPs does not give us any significant results for revenues. The results with MSPs are not reported here.

2.5.4 Empirical Methodology

OLS estimates

We start with the specification for the OLS estimates of the effect of being a district with the NREGA on the risk-taking behavior of farmers. The main equation we use is as follows. For a district i in state j in season s and year t:

$$Outcome_{ist} = \alpha + \beta NREGA_{it} + \gamma t + \delta_s + \chi_j + \varepsilon_{ist} \tag{2.1}$$

$Outcome_{ist}$ is the outcome of interest - the standard deviation or coefficient of variation of revenue or yield for district i in state j in season s in time t,

$NREGA_{it}$ is a dummy which takes the value 1 if the district was assigned to the NREGA program in year t. The equation includes a linear time trend and state and season fixed effects. Different versions of the results are presented with different combinations of the above covariates, and often include as well the level of previous risk from the year prior as a control.

Regression discontinuity estimates

The most flexible form of the main equation of interest is:

$$\begin{aligned} Outcome_{ist} = \quad & \alpha + \beta \widetilde{NREGA}_{it} + \gamma rank_{ij} + \delta rank_{ij}^2 + \eta \widetilde{NREGA}_{it} rank_{ij} \\ & + \lambda \widetilde{NREGA}_{it} rank_{ij}^2 + \mu_j + \phi_s + \theta Baseline_{is} + \varepsilon_{ist} \end{aligned} \tag{2.2}$$

where $Outcome_{ist}$ is the dependent variable in district i during season s ; $\widetilde{NREGA}_{it}$ is a dummy variable taking the value 1 if the program is predicted to have been introduced in phase 1 in district i in year t; $Rank_i$ is the state normalized rank for district i; $Baseline_{is}$ is the value of the dependent variable in the same season of the previous year; μ_j are state fixed effects, and ϕ_s are season fixed effects.

In this specification, we allow for the regression slopes and intercepts to vary on either side of the treatment cut-offs. The coefficient of interest is β, which is the intent-to-treat (ITT) estimate of the program. We vary this specification and present results for four variants - a linear form on either side of the cut-off, a linear flexible form with slopes allowed to vary, a quadratic form and then finally the quadratic flexible form of the equation above. Each of the other specifications is simply a special case of the quadratic flexible form, and so the

latter is our preferred specification, even though we present results from all four equations.

The dependent variable is the chosen measure of risk. In the above specification the main variable of interest is the dummy for whether or not the district got the program in phase1, $NREGA_{it}$. This captures the jump in risk-taking behavior at the normalized cut-off of 0.

Theoretical decomposition of land reallocations

The last set of tests we conduct is to check whether the manner in which districts reallocate their land changes after the introduction of the NREGA. The basic specification we use for this is the same as in equation 2.2. The outcomes of interest are the total amount of 'churning' or changes in crop land allocation, the quantity of land that was reallocated in order to increase risk, and the quantity of land that was reallocated in order to decrease risk. The theoretical justification for using this measure and the results of the decomposition are in the Appendix.

2.6 Results

In this section we present results for both the OLS and the RD specifications, with various different outcome variables. All the results we present here exclude Phase 3 districts entirely, and focus only on Phase 1 and 2 districts. The Phase 3 districts are the richest in the country and received the program last. It is harder to argue, therefore, that these districts are comparable to districts that received the program in the first or second phases. Phase 3 districts will be in-

cluded in the robustness checks, and we will show that their inclusion does not alter our results significantly.

Average yields and revenues

Tables 2.2 and 2.3 present regression discontinuity results for average portfolio yields and average portfolio revenues in the post-program period in 2007. As a reminder, the average portfolio yield is given by:

$$\overline{y}_{ist} = \lambda_{ist} \cdot \upsilon_{ist}.$$

The average portfolio revenue is calculated analogously.

As can be seen from Tables 2.2 and 2.3, there is no significant difference between districts that were predicted to receive the program and those that were not, in terms of the average returns to the portfolio of crops. A priori we would expect that riskier crops should also be higher return crops, so that if we see an increase in the riskiness of the portfolio then we should also see an increase in the mean return. That hypothesis cannot be rejected with the evidence presented here. The portfolio of crops in districts that were predicted to receive the program do not seem to be generating significantly higher average returns, at least in 2007.

Though we cannot show that districts do indeed earn higher average returns from their portfolios, we can show in a very simple manner that portfolios that are riskier are associated with higher returns. Table 2.4 presents the simple correlations between the district crop portfolio risk (measured by the standard deviation) and its return, as calculated by using the yields and the revenues at retail prices for the two major agricultural seasons. The returns and risk are

indeed positively correlated, suggesting that portfolios with greater risk on average have greater mean returns as well.

OLS results

To begin with, we present the results for the OLS specification from equation 2.1. Tables 2.5 and 2.6 present the results for the standard deviations of both yields and revenues. In these specifications, $NREGA_{it}$ is a dummy that takes on value 1 if the district i in state j received the NREGA in time t. Since the districts that received the program were poorer to begin with, we should expect that the estimates on the receipt of the NREGA should be biased downward, giving us smaller results than we would get in the case of the regression discontinuity.

We restrict the sample to the years 1998 to 2007, and remove all Phase 3 districts from consideration. As can be seen from the tables, there is a strong significant effect of the NREGA dummy on the measures of risk employed, in the case of both revenues and yields. In fact the OLS results are indeed smaller than the RD estimates for the sample without Phase 3 districts, as will be seen in the next section.

Regression Discontinuity Results

In this section we present the results from the regression discontinuity estimating equation presented in Section 2.5 above. We begin with the sample that includes the extremist districts, and present results that use all districts without restricting attention to a specific bandwidth around the cut-off. All tables include lagged values of the relevant dependent variable, and fixed state and

season effects. Since we are using the entire sample, we employ different specifications and allow for flexibility in slope and intercept in order to try to capture as best as possible the behavior of the data.

Tables 2.7 and 2.8 present the baseline results for the year 2005 before the program was introduced in any of the districts. These results include extremist districts, but exclude the richest districts that got the program in Phase 3. We should expect that there be no significant differences between districts on the yield or revenue risk measures prior to the introduction of the program. For both yields and revenues, the coefficient on the predicted assignment of NREGA variable is positive, however it is not statistically significant.

Thus there does seem to be evidence that the Phase 1 and 2 districts were *not* significantly different from one another in terms of risk-taking behavior prior to the introduction of the NREGA.

The next set of tables, Tables 2.9 and 2.10, present the results for the year 2007 for both yields and revenues. We see an increase in the standard deviation of yields and revenues that is large, and persists across all specifications. Are the numbers for revenue increases reasonable? When we restrict our attention to the crops for which we have information on prices, the coefficients on NREGA receipt for increases in the standard deviation of yields range from .726 to 1.043 tons per hectare. These evaluated at the mean price of Rs. 5096.25 per ton (1000 kg) yield increases in the standard deviation of revenues in the range of Rs. 3700 to Rs. 5300 per hectare, which is comparable to the coefficients presented here. The numbers for the increase in the standard deviations of revenues are larger than this range, but given the dispersion in the price distribution the numbers are not unreasonable. Also, the numbers for the increase in the standard devia-

tion of revenues seem large, but one should remember that these are per 10,000 square meters, and are not the measures of increases in the standard deviations of profits since we do not have cost of cultivation information.

Figures 2.6 to 2.9 plot graphically the coefficient of variation of revenues and yields (with extremist districts included) for the baseline and for the post-program periods. The graphs look very similar in shape, though not in scale. The vertical line is at the cut-off point for the allocation of the program. The baseline pictures do not show any discontinuities at the cut-off. The graphs for the year 2007 do, but the discontinuity is not as clear as one would have hoped for.

Tables 2.13 and 2.14 present the same RD baseline results but for the sample where the set of extremist districts has been excluded, and the state-normalized rank calculated accordingly. Again we see no differences in the districts prior to the introduction of the program. However, the results for the year 2007 presented in Tables 2.15 and 2.16 show no significant results. The magnitudes of the coefficients are very similar to the results with extremist districts, and there is a positive effect on risk-taking in all four tables, though the standard errors have increased so that none of the changes appear significant.

The change in the results when the extremist districts are excluded might simply be a result of the fall in the size of the sample, or it might be because of something specific to the extremist districts. If these are districts where the government is already investing more resources in order to appease the rebel groups, then there may be other programs going on at the same time that are driving greater increases in the risk of the crop portfolios than in other districts.

2.6.1 Robustness

All of the above results present estimates for the sample that excludes Phase 3 districts entirely. Just to show that the inclusion of these districts does not have any impact on the results, in the Appendix we present similar tables with Phase 3 districts included. These tables are in the same format as the ones in the main results section.

In addition, we have tables for the other two measures of risk that we discussed above, and also results for the same specifications but with the sample restricted to the set of crops for which we have price information.

As can be seen from the tables in the Appendix (tables 2.17 to 2.22) the quantitative results are mainly unchanged. The introduction of the NREGA still seems to have increased risk-taking behavior, though statistical significance has declined somewhat. Most of the changes in the results for the year 2007 show the same pattern as before, with the results on revenues being significant in the majority of the specifications. In addition, when we restrict our attention only to the yields of those crops for which we have price information (Tables 2.21 and 2.22) we see that the standard deviation of the yield has increased. This result holds whether or not we include the Phase 3 districts.

Our results are not completely robust, but within the limitations of the dataset we have we show that at least some of the main conclusions hold even when the sample is changed and the estimation methods are altered.

2.6.2 Impact of the NREGA on formal credit

While it seems fairly clear that the NREGA has had an effect on the riskiness of the portfolios of crops being grown at a district level, we have as yet been unable to pinpoint the exact channel through which this operates. One possibility is the relaxation of credit constraints. Using data on formal credit, we estimated the impact of NREGA on loan take up, and found no impact. This may not be very surprising, given that formal credit is a small share of the total credit small and marginal farmers have access to. On the other hand, data on informal credit is very hard to obtain. Using data from a field survey in West Bengal, (Dey, 2010) finds that informal loans (as measured by credit given by small grocers and shop keepers) has gone up as in response to the program. It would be interesting to see if this finding holds for other parts of India.

2.7 Conclusion

The NREGA is an ambitious, large scale employment guarantee program. While the focus of the program from a policy maker's point of view has been on the higher incomes and employment generated, as well as the public works built using labor supplied under the program, the NREGA can affect the rural economy in several other ways. In this paper we studied the role of the program in providing insurance or reducing credit constraints.

Our main contribution has been to look at changes in crop-choice across the country as a result of the introduction of this program. We use a regression-discontinuity design, which we believe is one of the more robust ways of

causally estimating the causal effect of the NREGA. We also provide a theoretical decomposition measure that divides total changes in land allocation into those that increase risk and those that reduce it. The data on crop yields is limited, and our methodology cannot identify the exact mechanism that is driving the increase in risk-taking. In particular we do not know if the main impact of the NREGA operates through the relaxation of credit constraints, or through the provision of insurance. We find that the amount of risk in the district-level crop portfolios has increased as a result of the introduction of the program, suggesting that the program might well be easing one or both of these constraints.

In order to try to disentangle credit constraints from the role of uninsured risk we looked at the effect of the NREGA on the total changes in land allocation. We find no increase in the total number of changes in land allocation, which suggests that the reason farmers are not changing their cropping patterns might have less to do with there being a fixed cost to adoption, and more to do with uninsured risk.

2.8 Figures and tables

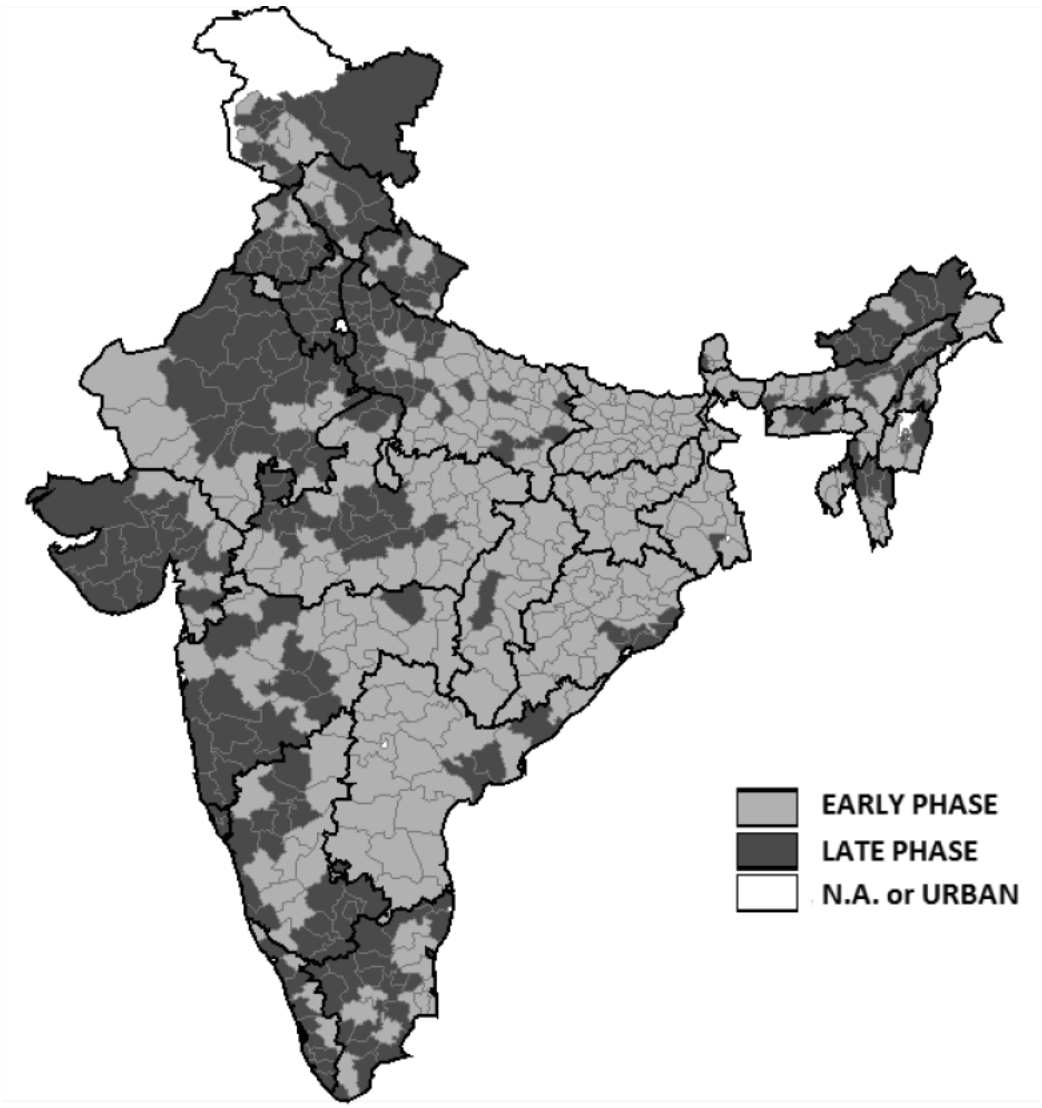

Figure 2.1: Country-wide roll-out of the NREGA: (Source (Imbert and Papp, 2013)). Early phase refers to Phases 1 and 2 in our paper, and the late phase refers to Phase 3 districts.

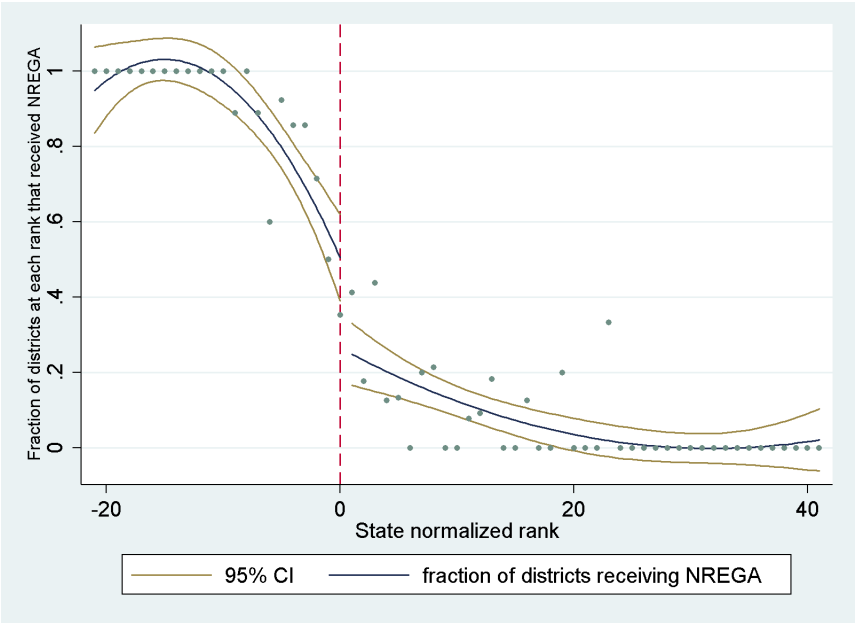

Figure 2.2: Discontinuity of treatment at the Phase 1 cut-off - Extremist districts included

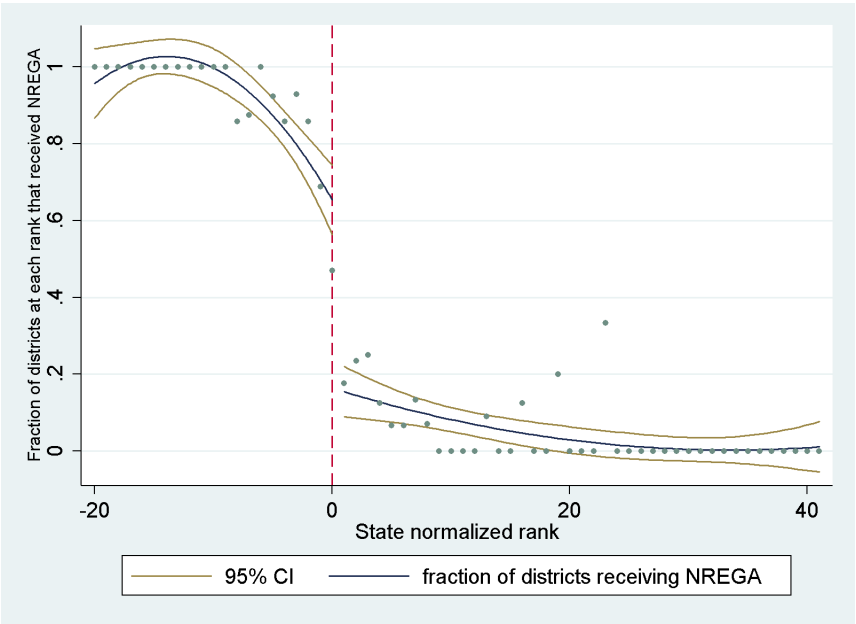

Figure 2.3: Discontinuity of treatment at the Phase 1 cut-off - Extremist districts excluded

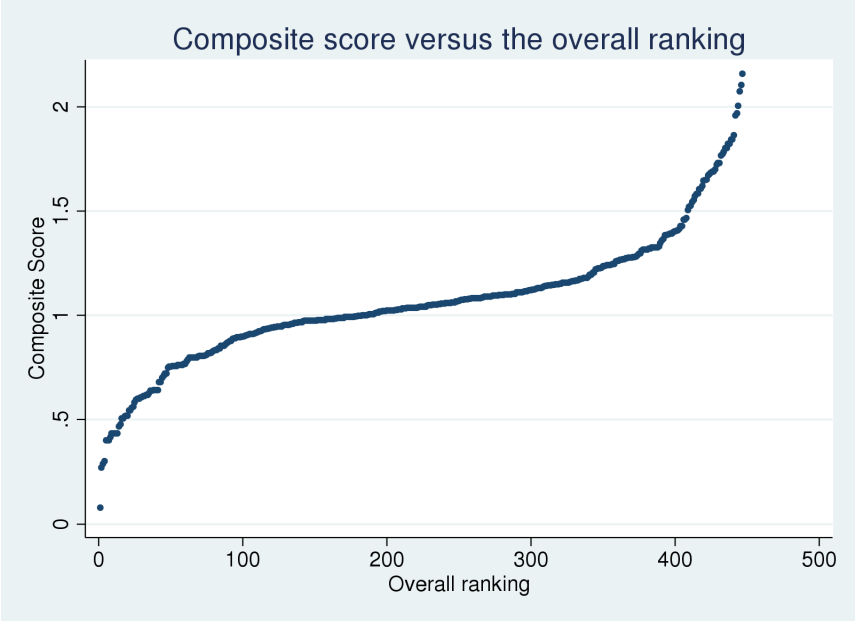

Figure 2.4: The composite score on the Backwardness Index vs. the overall district ranking (Source: (Planning Commission, 2003a))

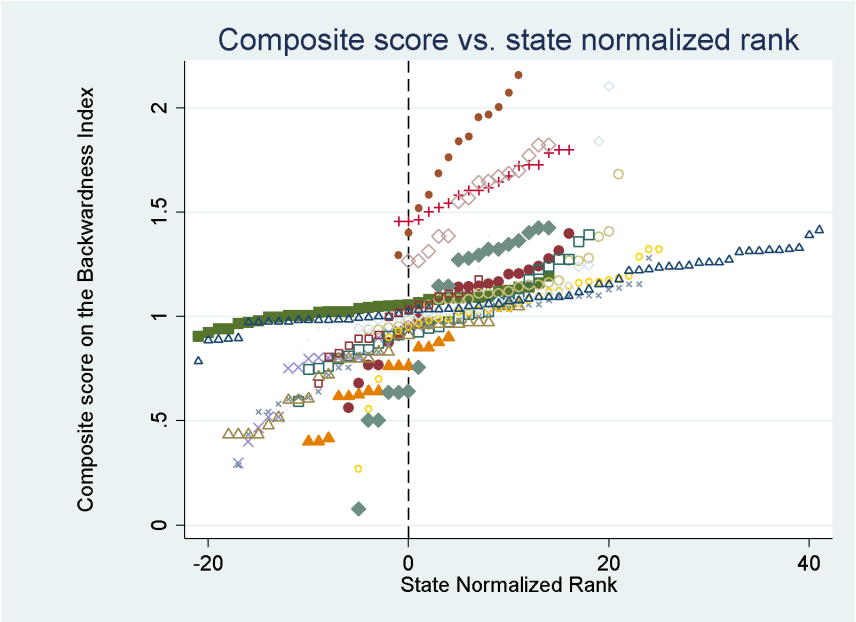

Figure 2.5: The composite score vs. the state-normalized rank. Each symbol refers to a particular state.

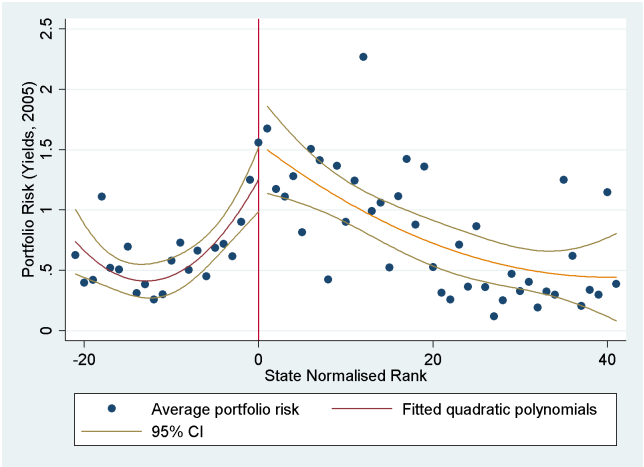

Figure 2.6: Standard deviation of yield - baseline (2005)

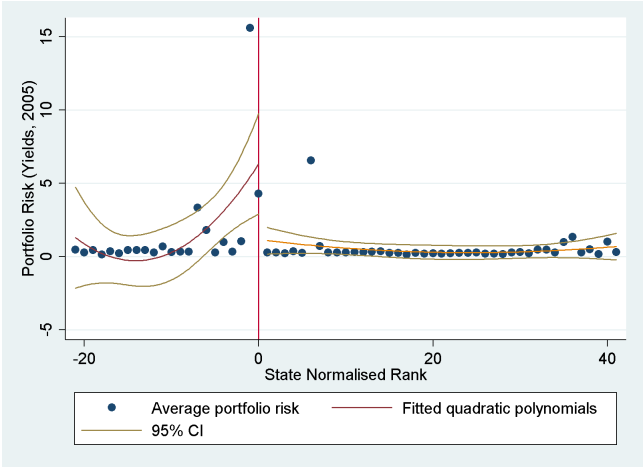

Figure 2.7: Standard deviation of yield - endline (2007)

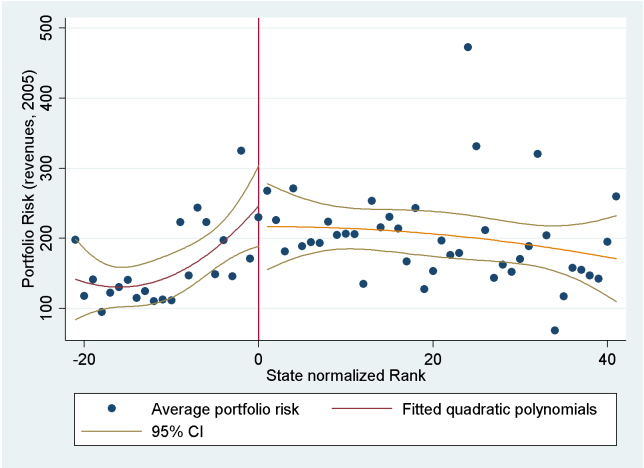

Figure 2.8: Standard deviation of revenues - baseline (2005)

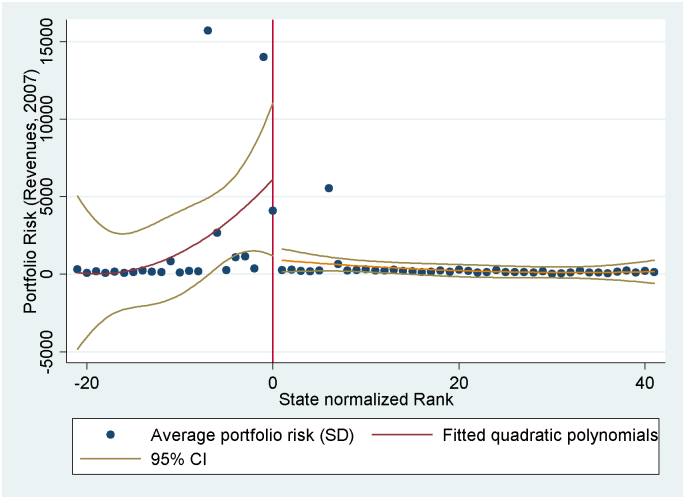

Figure 2.9: Standard deviation of revenues - endline (2007)

Table 2.1: Algorithm Success Rate for Phase 1

State	Phase 1 districts	Correctly predicted	False negatives	Non-phase 1 districts	False neg. rate (%)	Success (%)
Andhra Pradesh	13	10	3	8	37.5	76.9
Assam	7	7	0	16	0	100.0
Bihar	22	16	6	14	42.8	72.7
Chhattisgarh	11	9	2	4	50	81.8
Gujarat	6	4	2	14	14.3	66.7
Haryana	2	0	2	16	12.5	0.0
Jharkhand	18	17	1	2	50	94.4
Karnataka	5	4	1	21	4.7	80.0
Kerala	2	1	1	11	9	50.0
Madhya Pradesh	18	13	5	24	20.8	72.2
Maharashtra	12	11	1	18	5.6	91.7
Orissa	19	17	2	11	18.1	89.5
Punjab	1	1	0	14	0	100.0
Rajasthan	6	5	1	25	4	83.3
Tamil Nadu	6	4	2	20	10	66.7
Uttar Pradesh	22	19	3	41	7.3	86.4
West Bengal	10	8	2	7	28.6	80.0
Total	180	146	34	266	12.8	80.0

Note: Because of the design of our prediction model, which always has exactly as many predicted phase 1 districts as were allocated to the state, the number of false negatives always equals the number of false positives, although they are different districts. The false positive rate, which is not presented in the table, equals False negatives/Phase 1 districts.

Table 2.2: Effect of the NREGA on average portfolio returns in 2007 (yields)

Dependent variable: Average portfolio yields				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.652	0.618	0.655	0.905
	(0.52)	(0.53)	(0.52)	(0.71)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	949	949	949	949

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.3: Effect of the NREGA on average portfolio returns in 2007 (revenues)

Dependent variable: Average portfolio revenues				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	203.556	178.448	202.918	850.177
	(482.00)	(483.72)	(482.07)	(655.98)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	873	873	873	873

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.4: Are riskier portfolios also higher return on average?

	Correlation between the average returns of the portfolio and its risk, as measured by	
Season	Std. dev. of yields	Std. dev. of revenues
Kharif	0.6919	0.6872
Rabi	0.2483	0.3234

2.8.1 Results with extremist districts included

OLS results

Table 2.5: Effect of NREGA on riskiness (std. dev. of yields) - OLS results

Dependent Variable: Standard deviation of yields						
NREGA	0.443*	0.519**	0.594**	0.632**	0.274	0.299
	(0.242)	(0.242)	(0.257)	(0.258)	(0.323)	(0.325)
Previous risk	No	Yes	No	Yes	No	Yes
Season FE	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes	Yes
Time trend	No	No	No	No	Yes	Yes
Observations	8928	8928	8928	8928	8928	8928

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.6: Effect of NREGA on riskiness (std. dev. of revenues) - OLS results

Dependent Variable: Standard deviation of revenues (in Rs. '000s)						
NREGA	4.331	6.712**	6.716*	8.319**	3.684	2.896
	(3.310)	(3.236)	(3.491)	(3.467)	(5.679)	(5.728)
Previous risk	No	Yes	No	Yes	No	Yes
Season FE	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes	Yes
Time trend	No	No	No	No	Yes	Yes
Observations	5434	5434	5434	5434	5434	5434

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Baseline RD Results for year 2005

Table 2.7: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2005

Dependent variable: Standard deviation of yields				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.003	0.004	0.003	0.000
	(0.01)	(0.01)	(0.01)	(0.01)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1120	1120	1120	1120

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.8: Effect of NREGA on riskiness (std. dev. of revenues)

Dependent variable: Standard deviation of revenues (in Rs. '000s)				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	.031	.030	.031	.005
	(.029)	(.029)	(.029)	(.039)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	711	711	711	711

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Endline RD Results for year 2007

Table 2.9: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2007

Dependent variable: Standard deviation of yields				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.363*	0.349*	0.363*	0.580**
	(0.20)	(0.20)	(0.20)	(0.27)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	949	949	949	949

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.10: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2007

Dependent variable: Standard deviation of revenues (in Rs. '000s)				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	5.561*	5.411*	5.569*	7.920*
	(3078.16)	(3087.26)	(3079.81)	(4167.31)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	865	865	865	865

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

2.8.2 Results without extremist districts

OLS results

Table 2.11: Effect of NREGA on riskiness (std. dev. of yields) - OLS results, no extremist districts

Dependent Variable: Standard deviation of yields						
NREGA	0.364	0.439*	0.503**	0.540**	0.196	0.220
	(0.230)	(0.230)	(0.240)	(0.241)	(0.309)	(0.310)
Previous risk	No	Yes	No	Yes	No	Yes
Season FE	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes	Yes
Linear time trend	No	No	No	No	Yes	Yes
Observations	8137	8137	8137	8137	8137	8137

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.12: Effect of NREGA on riskiness (std. dev. of revenues) - OLS results, no extremist districts

Dependent Variable: Standard deviation of revenues (in Rs. '000s)						
NREGA	1.517*	1.516*	1.603*	1.588*	1.294*	1.281*
	(.854)	(.851)	(.895)	(.876)	(.715)	(.698)
Previous risk	No	Yes	No	Yes	No	Yes
Season FE	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes	Yes
Lin. time trend	No	No	No	No	Yes	Yes
Observations	3957	3957	3957	3957	3957	3957

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Baseline RD results for year 2005

Table 2.13: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2005, no extremist districts

	Dependent Variable: Standard deviation of yields			
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	-0.001	-0.000	-0.002	-0.010
	(0.01)	(0.01)	(0.01)	(0.02)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1013	1013	1013	1013

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.14: Effect of NREGA on riskiness (std. dev. of revenues) - RD results in 2005, no extremist districts

Dependent Variable: Standard deviation of revenue (in Rs. '000s)				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	-.013	-.013	-.013	-.012
	(.010)	(.010)	(.010)	(.014)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	636	636	636	636

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Endline RD results for year 2007

Table 2.15: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2007, no extremist districts

	Dependent Variable: Standard deviation of yields			
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.319	0.287	0.327	0.376
	(0.20)	(0.20)	(0.20)	(0.27)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	857	857	857	857

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.16: Effect of NREGA on riskiness (std. dev. of revenues) - RD results in 2007, no extremist districts

	Dependent Variable: Standard deviation of revenue			
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	6.124*	5.565	6.210*	6.709
	(3.707)	(3.775)	(3.716)	(5.007)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	708	708	708	708

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

2.9 Appendix

2.9.1 Phase3 districts included in the sample

Table 2.17: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2007, Phase 3 districts incl.

Dependent Variable: Standard deviation of yield)				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.221	0.258*	0.285*	0.527**
	(0.15)	(0.16)	(0.16)	(0.21)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1361	1361	1361	1361

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.18: Effect of NREGA on riskiness (std. dev. of revenues) - RD results in 2007, Phase 3 districts incl.

Dependent Variable: Standard deviation of revenue (in Rs. '000s)				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	3.696	4.133*	4.524*	7.109**
	(2.301)	(2.338)	(2.461)	(3.155)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1238	1238	1238	1238

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

2.9.2 Results with other measures of risk

Table 2.19: Effect of NREGA on riskiness (coeff. of var. of yields) - RD results in 2007, Phase 3 districts excl.

Dependent Variable: Coefficient of variation of yield				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.037	0.035	0.037	0.068**
	(0.02)	(0.02)	(0.02)	(0.03)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	949	949	949	949

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 2.20: Effect of NREGA on riskiness (coeff. of var. of revenues) - RD results in 2007, Phase 3 districts excl.

Dependent Variable: Coefficient of variation of revenues at retail prices				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.258*	0.253*	0.258*	0.342*
	(0.14)	(0.14)	(0.14)	(0.19)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	865	865	865	865

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

2.9.3 Yield results with the restricted sample of crops

Table 2.21: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2007, Phase 3 districts excl., restricted sample

Dependent Variable: Standard Deviation of Yields				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.726*	0.705*	0.725*	1.043*
	(0.39)	(0.39)	(0.39)	(0.53)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	785	785	785	785

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

The sample for this table is restricted to those crops for which we have information on both yields and prices.

Table 2.22: Effect of NREGA on riskiness (std. dev. of yields) - RD results in 2007, Phase 3 districts incl., restricted sample

Dependent Variable: Standard deviation of yields				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.483	0.542*	0.592*	0.941**
	(0.30)	(0.30)	(0.32)	(0.41)
Lagged risk	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1096	1096	1096	1096

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

The sample for this table is restricted to those crops for which we have information on both yields and prices.

2.9.4 A Measure of Mobility

Apart from the "implicit insurance" channel, the NREGA could affect individual decision making by relaxing other constraints as well. For example, by relaxing credit constraints the Act could permit farmers to incur the fixed costs associated with switching production away from crops they have been growing in the past, and into crops that they perceive will provide them with the highest returns. If that is the case, we would observe greater mobility in crop choice as a result of the program. In order to test whether or not this is really the case we develop a measure of crop mobility and describe how to decompose it into movements that are risk-increasing and those that are risk-reducing.[15]

Consider an agent who enters a period with some assets, a and a "crop inheritance", q. This crop inheritance should be thought of as the crop which was being grown on his land in the previous period. At the beginning of this period, the agent receives a preference shock, which provides a preference ordering over the crops. This preference shock should be interpreted as a reduced form for information that the analyst does not have. For example, the agent might learn of a new technology that he wants to implement, or learn that his parcel of land is more suited to growing certain kinds of crops.

Once this preference shock is realized, the agent might want to switch away from his inherited crop, and towards some other crop. Now, to make this switch, the agent must incur a fixed cost F. This can represent the cost of new seeds, of equipment that needs to be purchased for the new crop, or simply of the time and effort required to acquire information about how to tend for the new crop. If the agent's asset levels are lower than F, he would need to borrow. If

[15]We would like to thank Debraj Ray for suggesting this conceptual framework.

access to credit is limited and borrowing constraints bind, then the agent may be forced to continue production of q. In such a situation, if the introduction of the NREGA relaxes borrowing constraints, then we would expect to see increased mobility across crops. Some of this increased mobility would be in the direction of increased risk-taking, and some of it would just reflect the idiosyncratic preference shock. The natural question then is, is it possible to decompose the overall mobility in crop choice into movement towards *in*creased risk, and movement towards *de*creased risk?

To fix ideas, consider a district growing n crops. We arrange the crops in decreasing order of riskiness as measured by the standard deviation of yields or revenues, so that crop 1 is the riskiest and crop n is the least risky. Let $Y_1 = (a_1, a_2..., a_n)$ represent the land allocation vector in this district in year 1 and $Y_2 = (b_1, b_2..., b_n)$ be the allocation in year 2.

Define $\triangle = Y_2 - Y_1$ to be the "switch vector", which records all the shifts in land allocation across the years. $|\triangle| = \sum_{i=1}^{n} |(a_i - b_i)|$ is one possible measure of the overall mobility in crop choice across the year. A larger value for $|\triangle|$ indicates more "churning" or movement within crops, and a smaller value indicates more persistence in land allocations among crops. From the definition, it follows that $|\triangle| \in [0, 2]$.[16] Now we develop a measure which decomposes $|\triangle|$ into risk-increasing moves and risk-reducing moves.

What is a basic property would we want such a decomposition to satisfy? Consider an illustrative example where the change in land distribution vector is given by

$$\triangle = (0.1, 0.4, -0.2, -0.3) \tag{2.3}$$

[16] $\sum_{i=1}^{n} |a_i| - \sum_{i=1}^{n} |b_i| \leq \sum_{i=1}^{n} |(a_i - b_i)| \leq \sum_{i=1}^{n} |a_i| + \sum_{i=1}^{n} |b_i|$, and $\sum_{i=1}^{n} |a_i| = \sum_{i=1}^{n} |b_i| = 1$.

Looking at this vector we can see that district reduced production of crops 3 and 4, and increased its production of crops 1 and 2. Since we have arranged crops in order of decreasing risk, we would say that all the shifts that have taken place in this district have been in the direction of greater risk.

More generally, if the $\triangle$ vector satisfies "single crossing at 0", then our decomposition should be able to unambiguously label the shifts as being risk-taking or risk-reducing. The "single crossing" property can be satisfied in two ways. One possibility is that the initial few entries of the vector are entirely positive, and the latter entries are all negative values. This means there has been a move from the less risky crops to the more risky crops, and no move in the reverse direction. An example of such a vector is the one depicted above in equation 2.3. In this case our decomposition should assign all the moves as risky. Conversely, if the initial components of $\triangle$ are negative and the latter components are positive, we should be able to say that all moves have been towards lower risk. An example of such a vector is $\triangle = (-0.1, -0.4, 0.2, 0.3)$.

Of course, the $\triangle$ vectors in the data do not satisfy this single crossing at zero property, and we therefore need to define a measure of decomposition which can be used for general vectors, but which also satisfies the above criterion when $\triangle$ does indeed "cross" zero just once. We explain this measure using a simple example of a vector that does not satisfy the single-crossing property. Let

$$\triangle = (-0.1, 0.3, -0.4, 0.2) \tag{2.4}$$

In this example we are moving out of crops 1 and 3, and into crops 2 and 4. More generally say we are moving out of crops i and k and into crops j and l. Take crop i which we are "moving out" of and transfer the land lost into crops j and l in a manner proportional to the *final* land allocated to the crops which

are "moving into". Precisely, let T_{ij} represent the transfer from crop i to crop j. Then in the above example

$$T_{12} = \frac{0.3}{0.3 + 0.2}(0.1)$$
$$T_{14} = \frac{0.2}{0.3 + 0.2}(0.1).$$

Similarly, for crop 3 we have

$$T_{32} = \frac{0.3}{0.3 + 0.2}(0.4)$$
$$T_{34} = \frac{0.2}{0.3 + 0.2}(0.4).$$

Then set all other transfers equal to 0. So $T_{21} = T_{23} = T_{41} = T_{42} = 0$.

Now since we have arranged crops in decreasing order of crop riskiness, we define a transfer T_{ij} to be *risk-increasing* if $i > j$ and *risk-reducing* if $i < j$. In our example, therefore, T_{12}, T_{14} and T_{34} are risk-reducing moves, and T_{32} is a risk-increasing move. Finally, define:

$$R_+ = \left\{ \sum T_{ij} | i > j \right\}$$

and

$$R_- = \left\{ \sum T_{ij} | T_{ij} i < j \right\}$$

R_+ is a measure of the value of all risk taking moves, and R_- is a measure of the value of all risk reducing moves. It is then trivial to see that

$$R_+ + R_- = \frac{|\triangle|}{2}$$

We therefore have a valid decomposition. This decomposes all the movement in land allocation in the economy into those moves that are in the direction of increased risk and those that are not. This is the third measure of risk and mobility which we use to estimate the impact of the program.

In order to calculate the total changes and divide them into risk-taking and reducing, we rank crops in a particular district-state-season combination based on the standard deviation of the revenues at retail prices over the period for which we have this information. Thus a particular crop could be the riskiest in one district and have a completely different ranking in another district. Having ranked crops in this manner, we can look at changes in the allocation of total land across these crops and how that changes from one year to the next, in the manner described above.

Tables 2.23, 2.24 and 2.25 present the results for the endline for the total number of changes of crop land allocation from 2006 to 2007, and for how many of those changes are risk-increasing and how many are risk-reducing. In order to perform this calculation we rank crops based on the standard deviation over time in the revenues generated from them.

The results show that there do not seem to be significant differences across districts that received the NREGA and those that did not. Of course the calculations here and in the previous results are not necessarily capturing the same measure, and so this is not in contradiction to the increase in the coefficient of variation results we presented earlier. However it is interesting to see that (though insignificant) the coefficients are in line with our hypothesis - districts receiving NREGA seem to be making fewer risk-reducing moves, more risk-increasing moves, and more moves overall.

One of the ways to reconcile the finding of higher risk in yields and revenues through higher variances in portfolios with the fact that we do not find churning in the land allocation is the following: Many farmers do not necessarily take on greater risk by choosing to cultivate a different crop entirely, but simply by sub-

stituting one low-risk variety with another higher-risk higher-yield variety of the same crop (see (Mobarak and Rosenzweig, 2014)). Since our dataset does not distinguish between the different varieties of crops, we might simply be under-equipped to answer the question of whether the number of risk-increasing or risk-reducing moves has changed since the introduction of this program. It is possible that crop changes occur within a crop but across varieties with greater frequency than they occur across crops.

Total Changes in Crop Land Allocation

Table 2.23: Effect of NREGA on total land allocation changes - RD results in 2007, Phase 3 districts excluded

Dependent Variable: Total amount of land reallocated				
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.062	0.020	0.066	0.116
	(0.41)	(0.42)	(0.41)	(0.56)
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1044	1044	1044	1044

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Risk-reducing moves

Table 2.24: Effect of NREGA on risk-reducing land reallocations - RD results in 2007, Phase 3 districts excluded

	Dependent Variable: Total amount of land reallocated in a risk-reducing fashion			
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	-0.355	-0.372	-0.353	-0.155
	(0.26)	(0.26)	(0.26)	(0.35)
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1044	1044	1044	1044

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Risk-increasing moves

Table 2.25: Effect of NREGA on risk-increasing land reallocations - RD results in 2007, Phase 3 districts excluded

	Dependent Variable: Total amount of land reallocated in a risk-increasing fashion			
Specification	Linear	Lin. Flex.	Quadratic	Quad. Flex.
$\widetilde{NREGA}_{it}$	0.378	0.357	0.380	0.284
	(0.31)	(0.31)	(0.31)	(0.42)
State FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	1044	1044	1044	1044

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

CHAPTER 3

RAINFALL SHOCKS AND CROPPING PATTERNS: IMPLICATIONS FOR CHILD WELL-BEING

3.1 Introduction

Recently the question of the effect of early-life conditions on mortality and later-life outcomes like health, education and wages has received significant attention within economics. The central theory behind this link is the fetal-origins hypothesis, attributed to a series of papers in the early 1990s by epidemiologist David J. Barker, which posits that fetal conditions such as nutrient intake can have significant long-run impacts on health conditions. This, coupled with other evidence that the health of a young child has a role to play in cognitive development and consequently in the amount of schooling obtained and wages earned, means that early-life events can potentially have important long-lasting economic impacts[1].

This paper uses data from India to look at the effect of rainfall shocks on the level of health investments the mother makes in her child, including whether or not she initiates breastfeeding, the amount of time spent breastfeeding the child and vaccinations against common diseases, all of which can impact both infant mortality and a surviving child's health in later life. I also look at the likelihood of the child surviving for a range of months post birth. While there are a number of mechanisms through which rainfall might be operating, I isolate one channel in particular - the changes in the mother's labor supply caused by

[1]See (Almond and Currie, 2011) for a comprehensive overview of the fetal origins hypothesis literature in economics.

rainfall fluctuations, and the consequent impact on her ability to engage in time-intensive activities that benefit her child. As far as I am aware this is the first paper to look at this question in the context of India, and one of a very small number of such studies on any developing country.

This study exploits the fact that a large number of the rural workers in India are reliant upon rainfall for their livelihood, and as a result shocks in rainfall can cause fluctuations in parental behavior as well as the child's immediate environment, both of which can have (potentially contradictory) impacts on the child's well-being. The effects of rainfall on the opportunity cost of the mother's time are larger in districts where the mother actually plays an active role in agriculture, as opposed to districts where most of the work in cultivation is performed either by machines or by men. At the same time the effect of rainfall on other factors like household income and disease environments, for example, should be similar across rural agricultural districts regardless of the level of female involvement in agriculture.

There could be concerns that the level of female involvement in agriculture is not exogenous to investments in child health. Districts where cultural norms dictate that women play a larger role would tend to see both greater investments in child health through larger control of household resources by the mother, and at the same time a shift towards crops that afford women opportunities to work as well. In this case higher female labor-force participation and the outcome of interest would both be driven by cultural attitudes and norms. I build upon the work of (Carranza, 2012) in using exogenous variations in the type of soil in a particular district to explain differences in female agricultural participation.

Women engaged in agriculture in India are disproportionately employed

in crop cultivation activities (like transplanting seedlings, and fertilizing and weeding the crops) rather than land preparation activities which are more labor intensive and are generally conducted by men ((Basant, 1987; Foster and Rosenzweig, 1996)). The texture of the soil in the district in question dictates to a large degree the methods that can be used to till the land. Tillage reduces the need for precisely the tasks that women specialize in and hence also reduces the importance of female labor. In districts where the soil is loamy and deep tillage using machines is possible, female agricultural labor force participation is lower, as there is a reduced need for the operations that are traditionally undertaken by women. In contrast women still play an important role in crop cultivation in districts with clayey soils.

The texture of soil in a district is exogenously determined, and cannot easily be changed. However the texture of the soil does not in itself determine soil quality, the nature of the crop grown, cropping patterns or agricultural productivity. It determines only the ease with which labor can be substituted for using machines, and hence the relative importance of women in agriculture (Basant, 1987; Burton and Reitz, 1981; Foster and Rosenzweig, 1996). Thus the variation across districts in the texture of the soil can be used as an instrument for the level of female workforce participation. The key assumptions implicit in this strategy are that, firstly, the texture of the soil is indeed correlated with the level of female labor in agriculture, and secondly, that soil texture does not have any direct effect on the health investments made in children, which seems plausible.

It is well known that states in India differ considerably in terms of the treatment of women, with the more progressive and better developed southern states having more equal societies, while the large northern and central states

perform poorly in terms of sex ratios and women's agency. These regional differences in son-preference have been attributed to economic, social and religious factors (see (Arnold et al., 1998)). If the analysis were conducted at the state or regional level, a valid concern would be that maybe the differences in child investments was driven by cultural norms and the access of women to household resources. However there is not much evidence of differences in these cultural norms across districts within the same state or the same region. As a result comparing districts within regions that differ only in terms of their soil texture is a valid comparison of the relative importance of women.

The two well-known and well-studied mechanisms via which child health and mortality could be affected by weather are fluctuations in household income and changes in disease environments. In India, a failed monsoon, or one that arrives too late, could result in significant reductions in farmers' profits, and this income effect could reduce investments in child health.[2] In addition, in a tropical country like India where diseases like malaria still claim many lives every year, a slight alteration in weather conditions could drastically increase chances of infection. Illness of the mother during childbearing and of an infant in its first few months can both result in increased infant mortality or poor health of the young child.

These two mechanisms have received a lot of attention in the literature. However, holding other determinants of cropping patterns and agricultural yields constant there is no reason to believe that the effects of rainfall shocks on agricultural yield or disease environments should differ across regions with

[2]While it has been posited in the context of other countries that extremely heavy rains could lead to crops getting washed away or render the use of fertilizer ineffective, thus also resulting in a lower yield, this has been shown to be less of an issue in the Indian context (barring perhaps the detrimental effects of heavy flooding). See (Adhvaryu et al., 2013; Jayachandran, 2006).

different levels of female labor force participation.

The additional mechanism which has perhaps not been so well-studied, and which *would* vary across districts depending on the level of female involvement in agriculture, is the change in maternal labor supply as a function of rainfall patterns, which operates through the opportunity cost of the mother's time.

The amount of time spent breastfeeding the child or taking him or her for vaccinations at the clinic is a function of the amount of time the parents have, and is likely to vary depending on the size and quality of the harvest in that particular year. In years when the harvest is bountiful and there is much farming work to be done, working mothers with infants do not have as much time to breastfeed their children or make health investments as they would in times when the rains are poor and agricultural labor is less in demand. However the magnitude of the effects of rainfall on the mother's opportunity cost of time would differ depending on how much work there was for her in agriculture. Comparing the effects of rainfall shocks on child health investments in districts where women participate actively in agriculture to those districts where women do not play an active role in agriculture gives us an idea of the importance of the opportunity cost of the mother's time.

I note here that it is plausible that greater household income from a good harvest during the year of the child's birth might mean that the mother needs to work less if the household can hire an additional unit of labor to help with the farm work. Given the gender-specific task divisions, however, this would mean hiring an additional female unit of labor. For various cultural reasons it is unusual for women in rural India to work on farms other than their own (indeed, to engage in any form of paid activity, see (Olsen and Mehta, 2006)), so finding

a woman to replace a nursing mother is not likely to be easy. This means that in areas where the woman's labor is essential to cultivation we should expect negative rainfall shocks to *increase* the amount of time the mother spends with her child.

The remaining sections of this paper are organized as follows. Section 3.2 reviews the literature on parental behavior, soil, and agriculture in India, Section 3.3 describes the data set and provides some basic summary statistics. Then Section 3.4 lays out the empirical strategy, and sections 3.5 and 3.6 presents the results. Section 3.7 concludes.

3.2 Literature Review

I review three strands of the literature that are relevant to our study. The first is the effect of economic conditions on parental behavior and child outcomes, the second is the relationship between soil texture and female labor force participation, and the third is the relationship between rainfall, agricultural labor and crop yields in India.

3.2.1 The effect of economic conditions on parental behavior and child outcomes

There are several ways in which parents could respond to changes in the economic conditions. Risky behavior like smoking and drinking and household expenditure on food versus non-health items could change in ways that either

reinforce or counter the effects of the economic changes. Parents could target the timing of the births to coincide with high-income periods, or with low labor-demand periods, or with periods when infant mortality risk is the lowest. Parents could choose the sex of the child depending on whether incomes are high or low - for example if women are considered to be economic liabilities then parents might choose to selectively abort girl children during drought years and years of low income, or to neglect their health and nutrition. Finally families can adjust their composition through migration and marriage.

There is a considerable body of evidence (based mostly in developed countries) that suggests that mortality and health-related behavior move in a countercyclical fashion with the state of the economy (Chatterji and Frick, 2005; Dehejia and Lleras Muney, 2004; Ruhm and Black, 2002; Ruhm, 2000, 2003)). Chatterji and Frick (2005) find that returning to work soon after childbirth is associated with a reduction in breastfeeding duration among mothers that initiate breastfeeding, and a reduced probability of initiating breastfeeding at all. Dehejia and Lleras Muney (2004) find that babies conceived during recessions have lower incidence of low or very low birth weight, a reduced rate of neo- or post-natal mortality and fewer congenital malformations. Ruhm (2000) studies total mortality and ten other sources of fatalities, and finds that almost all of these exhibit pro-cyclical fluctuations, with suicide rates being one notable exception. He also finds that smoking and obesity increase when the economy is stronger. Ruhm and Black (2002) investigate the relationship between drinking behavior and macroeconomic conditions and find that the lower income during recessions leads to lower consumption of alcohol. Finally, following on his earlier work, Ruhm (2003) finds that when economic conditions worsen, there is a decline in smoking, body weight and in physical inactivity, with the largest effects

seen among those who were most at risk. The author suggests that the main reason for this change is that a fall in the opportunity cost of time for workers during an economic downturn makes it less costly to invest time in healthy behavior.

In developing countries, however, particularly in rural sectors, the access to insurance is limited and government spending follows a pro-cyclical pattern, indicating that individuals are considerably less insulated from economic downturn. In such situations it is plausible that mothers in poor families actually *increase* their labor supply in response to such shocks, rather than decrease it, and hence have less time to care for their infants than in times of positive income shocks. There is evidence going both ways on this question. (Bhalotra, 2010) finds that in India infant mortality increases in the wake of an economic shock, and that in good economic times ante-natal care, vaccinations and treatment for illnesses improve. Using data on a number of African countries, (Artadi, 2005) demonstrates that in countries where low infant-mortality months coincide with the high labor demand months in agriculture and the costs of reduction in income from the mother's inability to work outweigh the benefits from reduced infant mortality, parents could plan their pregnancies for periods when infant mortality is highest, not lowest.

There are only a couple of papers I have found that deal with the relationship between weather events and breastfeeding behavior in developing countries. (Thai and Myrskylä, 2012) conclude that positive rainfall shocks in the year following the child's birth have negative effects on the amount of time spent breastfeeding the child, and that this effect is larger for households involved in farm work (work that is not conducive to being able to breastfeed the

child) compared to those not involved in farm work or those employed in urban areas. (Rosales, 2014) finds that exposure to El Niño decreased the duration of exclusive breastfeeding and increased the duration of non-exclusive breastfeeding.

In India, where son-preference is well-documented, there could also be changes in sex-selection and infant care behavior. (Millett and Shah, 2012) do not find any evidence of sex-selection among Indian households based on drought or poor rainfall conditions, though there is some evidence supporting the hypothesis that boys are favored over girls in periods of low income. (Rose, 1999) finds that among Indian households with limited means of smoothing consumption over time, positive rainfall shocks in the first year of birth significantly increase the probability that the surviving child of a couple is a girl. This suggests that while girls benefit disproportionately from positive income shocks, they also suffer disproportionately from adverse income shocks. (Behrman and Deolalikar, 1990) find that when food prices rise there is some evidence of asymmetric adjustment within Indian households, with women's food intake being adjusted downward by more than the household average. (Thomas et al., 2004) present evidence that the crisis in the late 1990s in Indonesia raised household investments in the education of older male children at the cost of girl children aged 10-14.

Since one of the datasets I use does not present the entire birth history of the woman but only of those children born in the last four years I cannot check adequately whether sex-selection might be a concern in my sample. I will later present some suggestive evidence that the sex of the child whose health history I am studying does not appear to depend on the weather conditions in the year

prior, or by the soil texture of the district (see Section 3.5).

Lastly, marriage and migration are two other methods families can use to make short or long-term adjustments to their overall sex-composition in response to changing economic conditions. (Mbiti, 2006) shows that rainfall shocks in female-labor-intensive rice-producing areas in India lead to greater increases in the demand for adult women than similar shocks in male-labor-intensive wheat-producing areas, which changes the rates of marriage and migration. Similar results are to be found among tea and orchard plantations in China, where an increase in household income from an increase in the value of tea (a female-intensive crop) has a greater impact of improving sex ratios than from an increase in the value of orchard plantations ((Qian, 2005)). Out-migration might be a concern if it differed significantly across various soil types. (Carranza, 2012) shows that this is not likely to be an issue.

3.2.2 Long-term effects of early-life investments

I presented some evidence in the previous section that the investments parents make in their children and in their own health behavior can be affected by the economic conditions they face, both through an income effect and an opportunity cost of time effect. The health of the mother can translate into the health of the child. Low investments in breastfeeding and in vaccinations can lead to childhood illnesses, and these illnesses can have significant effects on the child's later life outcomes.

There is a large body of evidence pointing to the fact that early-life events can affect physical growth, cognitive development and later life earning poten-

tial. Extreme events like drought, civil war and pandemics are found to have significant retarding effects on the rate of growth of children and height-for-age scores in later life, and to be associated with higher rates of physical disability (Alderman et al., 2006; Almond, 2006; Hoddinott and Kinsey, 2001). Rainfall shocks in the year of birth have also been found to have significant effects on adult heights (Maccini and Yang, 2009; Skoufias et al., 2011).

Sibling or twin comparisons provide even greater validity by comparing children with the same parental characteristics and home environments. It has been found that being the heavier twin or sibling positively affects adult height and health, increases the weight of offspring many years later, reduces the likelihood of experiencing complications during delivery, increases test scores and labor market earnings and is associated with a reduced probability of being a high school dropout (Behrman and Rosenzweig, 2004; Black et al., 2007; Johnson and Schoeni, 2007; Oreopoulos et al., 2006; Royer, 2009).

Early-life rainfall shocks have also been found to have effects on a number of educational measures, including grade progression, test scores and the probability of graduating (Maccini and Yang, 2009; Millett and Shah, 2012; Szott, 2012; Thai and Falaris, 2011). Changes in the opportunity cost of the child's time also translate into better educational outcomes in rural economies where children participate in the labor market (Millett and Shah, 2012). Other factors like malnutrition at a young age (Glewwe and King, 2001) and exogenous early-life events like epidemics (Almond, 2006), civil war and droughts (Alderman et al., 2006) have all been found to have significant negative effects on the cognitive development and school performance of children, and on labor market outcomes.

The amount of time spent breastfeeding the child can also have long-term impacts on the child's health and development. Breastfeeding is widely regarded as one of the most important infant protection mechanisms, and also has important implications for the health of the mothers. The World Health Organization (WHO) recommends that mothers exclusively breastfeed their children till at least 6 months of age in order to boost their immunity and improve their growth and development, and then continue to breastfeed in combination with other food for up to two years.[3] There are a number of papers that study the impact of breastfeeding on cognitive development, with the finding that breastfeeding is associated with improved cognitive outcomes in the short- and long-term (Anderson et al., 1999; Belfield and Kelly, 2010; Kramer et al., 2008; Michaelsen et al., 2003; Rees and Sabia, 2009). There is also evidence that breastfeeding improves the probability of being in good health later in life, reduces risk of infection during infancy, increases motor scores, decreases the incidence of obesity and diabetes and improves visual acuity (Belfield and Kelly, 2010; Dieterich et al., 2013; Michaelsen et al., 2003).

3.2.3 Soil texture and its effect on the demand for female labor

The texture of the soil is a physical property which depends on the nature of the components it is made up of and their size. The percentage by weight of three mineral fractions - sand, silt and clay - determine the classification of the soil texture. Larger looser particles make up sandy soil, finer more densely packed particles comprise clayey soil, and loamy soils lie in the middle of the two. 'Loamy' soil is defined as 'a mixture of sand, silt and clay that exhibits the properties of

[3] http://www.who.int/nutrition/topics/infantfeeding_recommendation/en/.

each fraction about equally' ((Thompson and Troeh, 1978)).

The soil texture in a region is the result of geological and meteorological factors such as wind, rain and the parent rock materials, and cannot be affected by cropping practices. Thus it is valid to think of it as being exogenous to the choices people might make in deciding what to grow on the soil. The texture of the soil dictates the degree to which the soil retains water and nutrients, the ease with which it can be worked, its compressibility and so on.

Soils of different textures have different advantages and disadvantages for plant growth, and it is not possible to classify soil types into those that are *inherently* better suited to cultivation, or are more productive. The productivity of a soil type depends on the organic matter it contains, not just on the size of the soil particles ((Thompson and Troeh, 1978)). For example, sandy soils exhibit good aeration, but have low water and nutrient-retention capacities. Clayey soils have high water and nutrient retention, and are also electrically charged, which helps them hold plant nutrients to their surface, unlike sand. However such soils also have low levels of aeration, which is detrimental to plant growth. Loamy soils lie somewhere in between the above two types of soil - they have better aeration than clayey soils, and better water-retention than sandy soils. The presence of organic matter helps compensate for the low clay content in sandy soils, and improves aeration in clayey soils.

One important role soil texture plays is in determining the type of tilling operations that can be undertaken in preparing the land. Plowing and other tillage operations require much more power in the case of clayey soils, while loamy soils are easier to till at all levels of moisture content ((Müller and Schindler, 1999)). Deep tillage using machines can only be done where the soil is not too

clayey. Deep tillage reduces the need for precisely those tasks that are typically performed by women, so districts with soil texture that permit deep tillage see lower labor force participation by women.

Deep tillage "consists in breaking and turning over the soil, which facilitates root development, reintegrates moisture, nutrients and organic matter, and uproots weeds" ((Carranza, 2012)). Deep tillage using machines is a labor saving technology, and so it reduces total employment. In addition, it increases the duration of land preparation, while reducing the need for weeding, fertilizing and transplanting activities. Interestingly, this has gender implications. Since the latter set of tasks are typically conducted by women while the former predominantly by men, tillage also has the *gender-biased* effect of reducing the demand for female labor relative to that of male labor ((Alesina et al., 2013; Basant, 1987; Burton and Reitz, 1981; Nyangal et al., 2012)).

Aside from the effects of tillage on the importance of female labor, there is no evidence that deep tillage alone has any significant effects on the productivity of the soil, the types of crops grown or the cultivated area (see Section 3.5). I argue that the only effect deep tillage has is to reduce the importance of and demand for female labor in agriculture. We would expect, therefore, a negative relationship between the presence of loamy soils and both total agricultural employment as well the share of female employment in the total.

3.2.4 Agricultural labor, crop yields, and the monsoon in India

In India, the level of activity in agriculture is largely a function of the amount of rainfall a region receives. Agriculture is predominantly rain-dependent, and

yields fluctuate substantially with changes in rainfall. In the absence of systematic irrigation systems a failed or delayed monsoon can wreak havoc on the lives of those engaged in farming and related pursuits. (Jayachandran, 2006) estimates that annual rainfall below the 20th percentile for a district causes a 7% fall in crop yield.[4] (Rosenzweig and Binswanger, 1993) find that in areas heavily dependent on rainfall a delay of just 16 days in the start of the monsoon can cause a 6% reduction in crop profits. As of 2010 only 35% of agricultural land was irrigated (meaning that it was 'purposely provided with water'), even though agricultural land accounted for more than 60% of total land area.[5] In addition, the lack of assured irrigation disproportionately affects small and marginal farmers, who are unlikely to have access to credit or other consumption smoothing mechanisms. As of 2005-06, small and marginal farmers formed 69.6% of the partly-irrigated population, and 80.5% of the wholly-unirrigated population.[6]

Previous work on India has shown that while low rainfall is a bad shock to agriculture, high rainfall is a good shock ((Adhvaryu et al., 2013; Jayachandran, 2006)). Figures 3.2 and 3.3 show the relationship of the yield and production of major crops in India to rainfall. The yield, production and rainfall have been converted to standardized variables using the long-term average yield and standard deviations for a particular state-district combination. The standardized rainfall variable was then divided into 100 bins, and the average yield and production deviations were calculated for each of these bins. This average deviation is what has been plotted here. As can be seen from these figures, in most

[4]She defines a rainfall shock variable that takes the value 1 if the annual rainfall is above the 80th percentile for that district, 0 if it is between the 20th and 80th percentiles, and -1 if it is below the 20th percentile. This means that annual rainfall above the 80th percentile results in a 7% increase in crop yield.

[5]Figures from http://data.worldbank.org.

[6]Figures from the Department of Agriculture and Cooperation, Agricultural Census division

cases the relationship between rainfall and yield and production is positive - so that higher rainfall is on average better for agriculture in India.

(Jayachandran, 2006) provides some evidence that workers substitute working for private employers (farmers who hire casual labor) with working on their own landholdings when there are shocks to wages. However weather shocks are correlated across farms within districts (and indeed across neighboring districts) which means that such substitution is not likely to be possible when private employment falls due to poor rainfall. Thus a fall in agricultural activity caused by low rainfall is indeed likely to reduce the amount of time women spend working in the fields.

I provide some evidence of the relationship between rainfall deviations and the probability that women are working in Figures 3.4 and 3.5, which plot these relationships for the whole of our sample as well as the subsample of rural women (the construction of these figures will be discussed later in the paper). The relationship is increasing and concave up to about the 60th percentile of rainfall shocks, and declines thereafter.

Rice and wheat are the dominant staple crops in India, with 80% of the rural households in a representative survey of India growing at least one of these crops ((Mbiti, 2006)). The majority of rice in India is grown using the method of transplantation. The task of transplanting the seedlings is a delicate business, and is also the most labor-intensive part of the whole production process. Women are considered to have a comparative advantage in transplantation as they have a gentler, more nimble touch. As a result this stage of cultivation is handled almost exclusively by women. Women also help in the weeding, harvesting and threshing of the crop. By contrast, wheat requires more hard labor

and men play the predominant role in its cultivation. Across Asia, women play a more important role than men in rice cultivation, with their labor contributing as much as 80% of the work in India ((Mowbray, 1995; Romero-Paris, 2009; Unnevehr and Stanford, 1983)).

This difference in the importance of women's labor has been advanced as a possible reason for the differences in the status of women between what were once primarily rice-producing (southern) states and primarily wheat-producing (northern) states of India (see (Bardhan, 1974)). Now however the distinction between states on the basis of crops has become more blurred, as most big states produce some wheat *and* some rice at the same time. Without a clear demarcation of cropping patterns, I cannot simply use the nature of the crops grown in assessing the importance of female labor in the agricultural process; also as mentioned earlier the decision of what to plant in a particular district might not be completely exogenously determined by environmental factors. For example, districts where women play a more dominant role might choose to plant crops that allow for more labor market opportunities for female agricultural laborers, and vice versa. Instead, I propose to build upon the work of (Carranza, 2012) and use variation across districts in the nature of the soil as exogenous variation in the importance of female labor in agriculture.

3.3 Data

Information on child health investments and other household information comes from the household and individual level files from various demographic and health surveys. These contain information on state and district. The state-

district information was used to match these files to information on geographic coordinates and rainfall, so that one large file with information on child outcomes, duration of breastfeeding, household and parent characteristics, type of place of residence (rural/urban) and district-level rainfall was created. This was then used for the analysis.

3.3.1 District-level agricultural and other data

Information on district level variables was obtained from the International Crops Institute for the Semi-Arid Tropics (ICRISAT) District Level Database Documentation, which compiles information from a number of sources on all districts from the 19 major states for the period 1966-67 to 2007-08. This is an unusually rich and detailed dataset, with information on production, irrigated area, fertilizer consumption, length of roads, population, soil type, length of growing period, rainfall and livestock holdings, among other variables. The total number of districts covered in this dataset is 331.

I also use information from the India Agriculture and Climate dataset, the collection of which was jointly funded by the World Bank and the Electric Power Research Institute. This dataset contains information on 271 districts over the period of 1957-1987. I use this dataset to obtain selected soil characteristics like soil pH levels and topsoil depth, and for calculating the long-term average temperatures.

Information on crop yields, land use and season comes from a dataset published by the Ministry of Agriculture, Government of India. This dataset is a district level panel for the years 1996 to 2009, and provides information on area

under cultivation and total production for each crop grown in all districts in India, across *rabi*, *kharif*, autumn, winter and summer seasons of the year. The number of crops varies across districts in a given year and season, across years within a particular district and season, and across seasons within a particular district-year combination. Data on crop production is in metric tonnes and on land use is in hectares. I calculate the average yield per year (all seasons) across this time period, and this is the measure of yield that I employ.

3.3.2 Soil Characteristics

Information on soil texture was obtained from the Food and Agriculture Organization (FAO) Digital Soil Map of the World and Derived Soil Properties (CDROM). The Digital Soil Map provides information on soil texture by geographical coordinates, with the scale of the original map being 1:5,000,000. The scale of the map allows for sub-district soil texture mapping. The map also comes with shape files that provide district boundaries and centroid points for all districts of India.

I calculate the proportion of each sub-district covered in fine clayey soil, medium loamy soil and coarse sandy soils. We then use the fraction of the district's area covered by that sub-district to weight the soil texture composition for the entire district. I end up with a measure of the fraction of a district covered by fine, medium and coarse soils, which I then use for my analysis.

3.3.3 Rainfall data

The monthly data on rainfall was obtained from the Center for Climatic Research, University of Delaware (version 3.02). This dataset was compiled using station records of monthly precipitation obtained from several up-to-date sources, including the National Center for Atmospheric Research's (NCAR) daily rainfall data for India. Station values were interpolated to a 0.5 degree by 0.5 degree latitude-longitude grid, with the grid nodes centered on .25 degree. The readings at stations closer to the grid node were assigned a higher weight. In this version of the data the number of nearby stations that influenced a grid-estimate was increased to twenty from the previous seven.[7]

The coordinates of the (approximate) center of each district were obtained from the ICRISAT database and were then linked to the closest node in the precipitation dataset. The latitude and longitude were then rounded up or down to the closest .25 degree increment based on inspection. Data on precipitation was collected from 1940 till 1999, and rolling 35-year averages of rainfall were calculated for each calendar year and district combination.

3.3.4 Household level information

Household level information on a representative sample of Indian households was obtained from the District-Level Household Survey (DLHS) and the National Family Health Survey (NFHS). The NFHS is the DHS conducted in India, and is representative at a state level. The rounds of the NFHS that I use were

[7]See http://climate.geog.udel.edu/~climate/html_pages/archive.html for more information on the dataset and the method of collection.

conducted in 1992-93 and 1998-99. We also use the third round of the DLHS from 2007-08. The DLHS was modeled after the NFHS, but intended to be a much larger district-representative survey. The scope of the two surveys is similar, and the modules that deal with health and reproduction exhibit considerable overlap in design and content. As a result these two surveys can be used together.

The DLHS and NFHS surveys collect detailed information on the reproductive behavior of women aged 15-49 at the time of the survey, as well as some basic information on educational attainment, religion and ethnicity for both parents, child outcomes like height and weight, and household-level information on consumer durables ownership, place of residence and so on. The second round of the DLHS does not have any information on whether the respondent works or has worked in the past, and hence could not be used.

I use the women level dataset which contains information for multiple children per mother on the date of birth, sex, birth order, preceding and succeeding birth intervals, whether or not the child is alive, and if not, then the age at which the child died. Information on the duration of breastfeeding, on vaccinations, ante-natal checkups and treatment during illnesses is also collected. The only information I have on the mother's employment is whether or not she worked sometime in the last 12 months. As a result I restrict the sample to those children who were born in the year before the time of interview. This considerably reduces the sample size, especially in the case of the NFHS, which records this information for every child ever born.

I dropped all observations where the mother had twins, as these are generally unexpected and unplanned and could bias the results. In order to make

sure the rainfall shocks do indeed correspond to the shocks experienced by the mothers at the time of the child's birth, I also restrict the sample to those mothers who are living in the same place since the time of the child's birth. I also drop mothers who had their first birth before the age of 13 or those that had had more than 10 children at time of interview as these mothers are likely to be substantially different from others.

The end result is a mother micro-panel nested within a district level panel, where each row of the data matrix corresponds to a birth of that mother. For each birth we have the month and year of birth, and the district of residence. These identifiers are then used to link this dataset to the rainfall and soil information.

3.4 Empirical Strategy

The main question of this paper is whether rainfall shocks that lead to changes in the opportunity cost of the mother's time filter into reduced health investments in her child. Let Y_{idsrt} be the health-investment outcome of interest for child i in district d in state s in region r in time t, which consists of whether or not breastfeeding was initiated, the duration of breastfeeding if initiated, and whether or not the child was vaccinated against common diseases (polio, measles, tuberculosis and diphtheria-pertussis-tetanus (DPT)). α_{0r} is a region-fixed effect that captures cultural norms that are common across districts within the same region. $RAIN_{dsrt}$ is the rainfall shock in the district in question in the 12 months following the child's birth, which is normalized by the long-term mean and standard deviation of rain in that district.

$FemLFP_{idsrt}$ is a dummy for whether or not the mother of child i was working in the 12 months prior to the month of interview. The vector X_{idsr} contains child-specific covariates like birth order, sex and mother's age at birth as well as household specific variables like the standard of living index, mother's education and father's education. These covariates are not time-varying. δ_{srt} is a state specific time trend, that would capture, for example, changes in the provision of medical services or the dissemination of information regarding the importance of these investments within a state over time. Finally ϵ_{idsrt} is the error term.

With these definitions in place, the main equation of interest can be written as

$$\begin{aligned} Y_{idsrt} =& \alpha_{0r} + \alpha_1 RAIN_{dsrt} + \alpha_2 FemLFP_{idsrt} + \alpha_3 (RAIN_{dsrt})(FemLFP_{idsrt}) \\ &+ \beta X_{idsr} + \delta_{srt} + \epsilon_{idsrt} \end{aligned}$$

Due to concerns that the measure of female labor force participation could be endogenous, I instrument for female labor force participation using the fractions of loamy and clayey soil in the district. I claim that soil texture affects the outcome variable only through female labor supply, and no other channel. In Section 3.5 I present some evidence that districts with differing soil textures do not differ on other important dimensions like agricultural productivity. I also present results regarding the relationship between female labor force participation, child health investments and rainfall shocks. Section 3.6 then presents the main results of this paper based on the above empirical specification.

In all the results below I employ regional fixed effects. As mentioned earlier, cultural differences in cropping patterns and the evaluation of women tend to persist over regions larger than states. By controlling for these differences, I iso-

late the effect of female labor force participation. Figure 3.1 provides a graphical depiction of the six regions we have divided the country into.

3.5 Descriptive statistics and results

This section will present some descriptive statistics as well as the OLS and instrumental-variable results for the main specification.

3.5.1 Characteristics by Soil Texture

I first want to check that various district level characteristics do not differ substantially across differing soil textures. In particular I am interested in yields of major crops being more or less similar across districts with differing soil textures, in order to demonstrate that soil texture affects female labor only through deep tillage and not through other channels like income or cropping patterns. Table 3.3 presents some of these results. The coefficients reported are from the regression of the variable in question on the fractions of the district's soil that is clayey or loamy, conditional on state fixed effects.

The first section of the table presents the results for some district level characteristics like the length of the agricultural growing period, annual average rainfall and temperature, and irrigated area. The second section presents results for some select soil characteristics like pH levels and topsoil depth. The third section presents results on the yields of various crops over the period 1996-2010 for the two main planting seasons, the *rabi* and *kharif* seasons. As can be seen from the table there are some differences in these characteristics by soil texture,

however overall it seems that not a large fraction of these differences (particularly with respect to yields) are significantly different across districts.

3.5.2 Sex-selection, soil, and weather shocks

In tables 3.1 and 3.2 I present some results indicating that there is no evidence that the probability of the last child being male depends either on previous rainfall shocks or on the soil composition of districts. If sex-selection is indeed a concern, then we would expect that a girl born into a family with boys would be less likely to be aborted than a girl born into a family with a large number of daughters. Unfortunately the data I have do not enable me to construct the entire reproductive history for each woman, and so I cannot control for sex composition of the family at the time of birth of the child. I continue to control for the sex of the child in all specifications.

3.5.3 Seasonality of births

Another concern for our study would be if there was a seasonality in birth patterns that varied with rainfall shocks. Households could deliberately choose to have children at a certain time of the year because the agricultural seasonality would mean that there would be less agricultural work, or greater incomes, for example. We conduct two different tests to show that this does not seem to be the case in our sample.

In figure 3.6, we depict the fractions of total births that occur in each month for both the rural and urban populations. If it were agricultural seasons that

was driving this seasonality, we would expect to see differences between the two populations. However, the figure shows that the births show very similar patterns in both rural and urban areas.

In figure 3.7 we plot the fractions of births that occur in each month for rural respondents, with separate lines for contraceptive users and those that report never using contraceptives. We restrict ourselves to modern methods of contraceptives (IUD, pills, injectables, male and female condoms) as these are proven to be much more effective at preventing unwanted pregnancies. Again, if households were using contraceptive methods to alter their fertility patterns then we would expect a divergence between the birth patterns for contraceptive users and non-users. The figure shows that the two lines actually move quite closely with one another.[8]

3.5.4 How does female labor force participation respond to rainfall shocks?

Now I want to check how female labor force participation responds to rainfall shocks in the same year. My initial hypothesis is that in rural areas, years with negative rainfall shocks result in a smaller harvest, and this in turn means less work for women. Following a year with a bad harvest due to a negative rain shock it is possible that the household adjusts by actually *increasing* its labor supply in order to increase its income, but it is also possible that in the year of the poor shock the household is unable to find other income-generating avenues

[8]The test of significance for the coefficient on contraceptive use in the regression of fraction of births on the contraceptive use dummy and the month of the year also shows no significant differences between the two populations.

and instead is forced into involuntary unemployment.

I plot the relationship between the two variables of interest for the entire sample in Figure 3.4 and for the rural sample of women in Figure 3.5. The x-axis is rainfall deviations grouped into 100 equally sized bins. The y-axis plots the mean predicted values from the regression of the labor force participation dummy on state fixed effects and other covariates. The overall relationship is not very strong, but is positive and weakly concave.

3.5.5 Does soil texture affect female labor force participation?

Table 3.4 reports the results of the first stage regressions of female labor force participation on soil texture in both rural and urban areas. All columns include regional fixed effects. The first two columns do not include time trends, while the latter two have state time trends as well. I control in these regressions for some demographic characteristics like religious dummies, caste dummies, years of schooling of the mother and household size and standard of living.

In the two columns without time trends we see that rural districts with a larger fraction of clayey or fine soil see a significantly higher rate of female labor force participation. A larger fraction of loamy soil is associated with lower female labor force participation relative to sandy soils, though the coefficient is not significant. The F-statistic on the fine and loamy soils variables is 22.00, strongly rejecting the null hypothesis that both coefficients are jointly equal to 0. In the case of urban areas, neither coefficient is significant, though the signs are still in the correct direction. The F-statistic is also considerably lower at only 5.21.

When state time trends are added the significance of the soil instrument declines somewhat. However as can be seen from Table 3.4, the clayey soil coefficient is still significantly positive, and the F-statistic on the fine and loamy soils variables for rural areas is significant at 9.15. For the urban areas neither coefficient is significant, which is reassuring as well.

In all specifications it can be seen that a higher standard of living index is associated with a lower labor force participation by women. Mother's education has a positive correlation with her working only in urban areas. Being Muslim is associated with a lower female labor force participation in both rural and urban areas, and being of a scheduled caste or tribe is positively associated with the mother working in the 12 months preceding.

For the remainder of the results I restrict my attention to rural areas only, and employ regional fixed effects in all specifications. Standard errors are clustered at the district level in all tables depicted here, and sampling weights are employed.

3.5.6 OLS results for the main specification

We start by describing the OLS results for the effect of changes in the opportunity cost of women's time on their investments in their children. Tables 3.5 and 3.6 show the main results of the interaction between rainfall shocks and female labor force participation on child health investments. The rainfall shock is defined as deviations from a district-month specific long-term mean. Keeping this in mind the coefficients can be interpreted in the following way. In households where the mother is working relative to households where she is not, a

one standard deviation increase in rainfall in a district leads to no increase in the probability of breastfeeding, and no significant change in the months spent breastfeeding. The overall effect of rainfall is positive and insignificant on the probability of receiving the BCG and Polio vaccines, positive and significant in the case of DPT and negative and significant in the case of measles. In general, therefore, the OLS results suggest that the overall effect of the interaction of rain and the mother working is mostly positive but not very strong.

On the other hand, tables 3.7 and 3.8 show the OLS results for the effect of the interaction between rainfall shocks and female labor force participation on child mortality. The effect of the interaction term is insignificant, and the overall effect of the mother working is to increase the probability that the child will die within 1 month by about 2.7 percent, and to decrease the probability that the child is alive at up to 12 months post birth by about 3.1 percent.

Overall, therefore, the OLS results are mixed - while a positive rainfall shock in a household where the mother is working serves to increase the probability the child is vaccinated, it also increases the probability that the child dies at a young age.

3.5.7 Discussion

I have presented preliminary results so far that have nonetheless documented some interesting findings. Female labor force participation appears to be positively related to rainfall shocks in the same year, which is in line with our hypothesis that good harvest years mean plentiful employment, and bad harvest years mean unemployment (plausibly involuntary) for rural workers. My hy-

pothesis is that rainfall shocks in the early years of a child's life could actually have some positive effects, if they permit the mother to invest more in child health because she has more time to do so. I will now present the results based on the empirical specification described above.

3.6 Instrumental Variable results for the main specification

We present here again the main specification for which we will present results. It is:

$$\begin{aligned} Y_{ihdsrt} = \quad & \alpha_{0r} + \alpha_1 RAIN_{dsrt} + \alpha_2 FemLFP_{hdsrt} + \alpha_3 (RAIN_{dsrt})(FemLFP_{hdsrt}) \\ & + \beta X_{ihdsr} + \delta_{srt} + \epsilon_{ihdsrt} \end{aligned}$$

where the details regarding the variables can be found in Section 3.4. I instrument for the endogenous variables $FemLFP_{hdsrt}$ using the fractions of clayey and loamy soils in the district and for $(RAIN_{dsrt})(FemLFP_{hdsrt})$ using the interaction between the rainfall shock and the fractions of clayey and loamy soils in the district.

The second stage of the instrumental variables estimation gives different results from the OLS. Tables 3.9 and 3.10 depict the second stage IV results with and without state time trends.

Let us look first at Table 3.9. Comparing households where the mother is working to households where the mother is not working, we can see that a one standard deviation increase in rainfall leads to a 17.9 percent fall in the probability that the child is breastfed, a 19.8 percent increase in the child receiving the BCG vaccine, a 34 percent decline in the probability of receiving the polio vac-

cine, a 15.6 percent fall in the probability of getting the DPT vaccine, and about a 1.8-2.3 month increase in the number of months spent breastfeeding. The direct effect of the mother working on the number of months the child is breastfed is negative and significant.

The results are similar when state time trends are added, as in Table 3.10. Here we see that comparing households where the mother is working to households where the mother is not working, a one standard deviation increase in rainfall leads to a 31 percent fall in the probability of receiving the BCG vaccine, a 37 percent fall in the probability of receiving the polio vaccine, a 24 percent decline in chances of getting the measles vaccine, and a 42 percent fall in the likelihood of receiving the DPT vaccine. The results in this case unambiguously favor the opportunity cost of time theory - rainfall shocks that increase the amount of labor the mother provides to the market decrease her investments in her children. Even the effect on the number of months the child is breastfed is negative, with a positive rainfall shock and the mother working being associated with between a 2.9 and 3.8 month decline in duration of breastfeeding.

It is interesting also to note the directions of the direct effect of the rainfall deviations in the twelve months prior to the child's birth. Let us look at table 3.9, without state time trends. If we compare two households in which the mother does not work, then a one standard deviation increase in rainfall (versus no change in rainfall from the long-term average) has the effect of increasing the probability that the child will receive a vaccination by 18.2 percent in the case of BCG, 20.7 percent in the case of polio, 8 percent in the case of measles and 13.1 percent in the case of DPT. This seems to be the income effect at work - all else equal, an increase in the amount of rainfall increases household income

and hence the investments parents can make in their children. Including state time trends dampens the effect of the rainfall shocks somewhat, but the positive association remains in the case of all vaccinations.

If instead we compare two households where the mother *did* work but which received different rainfall shocks, then the direct income effect of a positive rainfall shock is offset by the substitution effect (from the interaction term), rendering the overall impact on the probability that the child will receive a vaccination negative. Looking at the table 3.10, a one standard deviation increase in rainfall is associated with a 22.5 percent decline in the probability of receiving a BCG vaccine, and similarly a 27.8 fall in the case of polio, an 18.9 percent fall in the case of measles and a 31.4 percent decline in the probability of receiving the DPT vaccine. The same effects can be seen in table 3.9, but are larger when the state time trends are not included.

The IV results for child mortality vary depending on the specification employed. With regional fixed effects but no time trends, Table 3.11 shows that a one standard deviation increase in rainfall increases the probability that the child will die before the age of 1 month by about 9 percent, and increases the probability that the child will die before the age of 6 months by 26 percent. However including state time trends renders these estimates insignificant, and reduces their magnitude considerably.

3.6.1 Discussion

The reason why we might have thought the OLS results might be biased is that districts could vary in terms of cultural differences that would drive both fe-

male labor force participation as well as investments in children. For example districts where women are valued more would see a larger control of resources by the mother, but could also see a shift in the choice of crops to be cultivated such that women work more. If this is the direction of the confounding, then we would expect higher labor force participation to be associated with greater child health investments. As a result we would expect the OLS estimates to be biased upward.

I use an instrument that I argue is exogenous to the choice of health investments in children apart from the impact it has through mother's labor force participation. Districts with greater proportions of clayey soil to loamy soil see greater female labor force participation, regardless of regional cultural differences. My instrumental variable results suggest that once we control for these regional cultural differences, the effect of the mother's participation in the labor force is actually smaller than and in the opposite direction to that of the OLS results. Women who work have less time to spend with their children, and especially in the wake of positive rainfall shocks which demand more of their labor, this has a detrimental effect on child health investments and on the child's probability of survival.

3.7 Conclusion

I presented evidence that among working mothers, positive rainfall shocks lead to decreased investments in child health as measured by breastfeeding and vaccinations. The two competing theories are that an increase in labor force participation by women leads to an increase in household income and hence in

spending on 'normal' goods like child health. This is the income effect. The opportunity cost effect is that with the mother working, she has less time to invest in her child, in particular vaccinations and time spent breastfeeding, both of which take time to complete.

I used an instrument for female labor force participation that allowed me to plausibly argue that an explanation for this reduced investment is that the mother's opportunity cost of time is higher during positive rainfall shocks, and thus she spends more time working than attending to her child. The instrument is the texture of the soil in the district in question. I show that districts do not differ significantly in crop yields and other descriptive statistics based on their soil texture. I also argue that the texture of a district's soil and the fraction of the soil that is fine or loamy does not impact child health investments except through the mother's labor force participation.

The OLS results suggest that relative to households that do not have a working mother, positive rainfall shocks lead (on average) to an increase in child health investments, suggesting that the income effect dominates. We discuss why we think the OLS estimates are likely to be positively biased, and then present the second stage instrumental variable estimates. These are smaller and in the opposite direction, giving us the result that positive rainfall shocks affect households where the mother works by reducing investments in young children, also plausibly causing an increase in the probability of the child dying in early months.

3.8 Figures and Tables

Figure 3.1: Division of India into Regions

Figure 3.2: Yield vs. Rainfall

Figure 3.3: Production vs. Rainfall

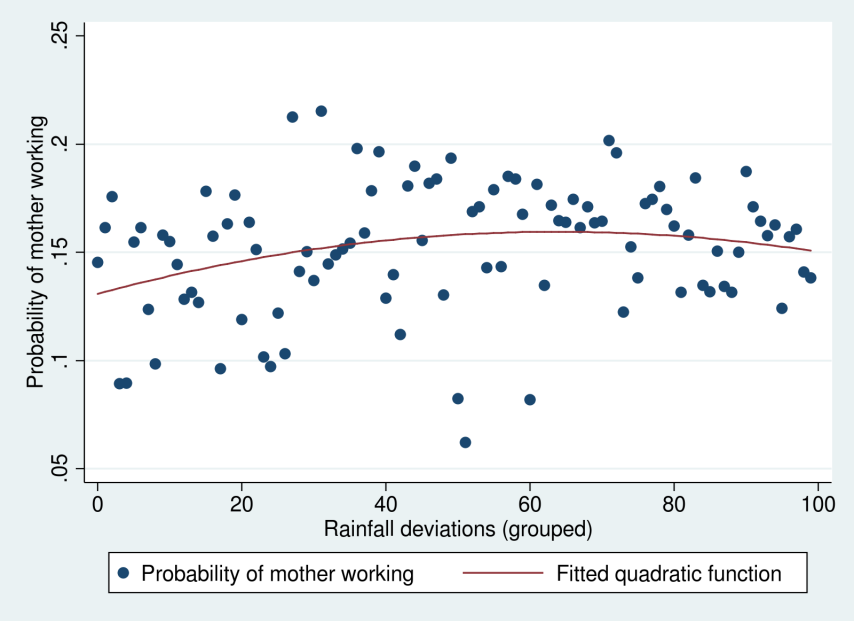

Figure 3.4: Relationship of mother's work to rainfall deviations for the whole sample- regional FE

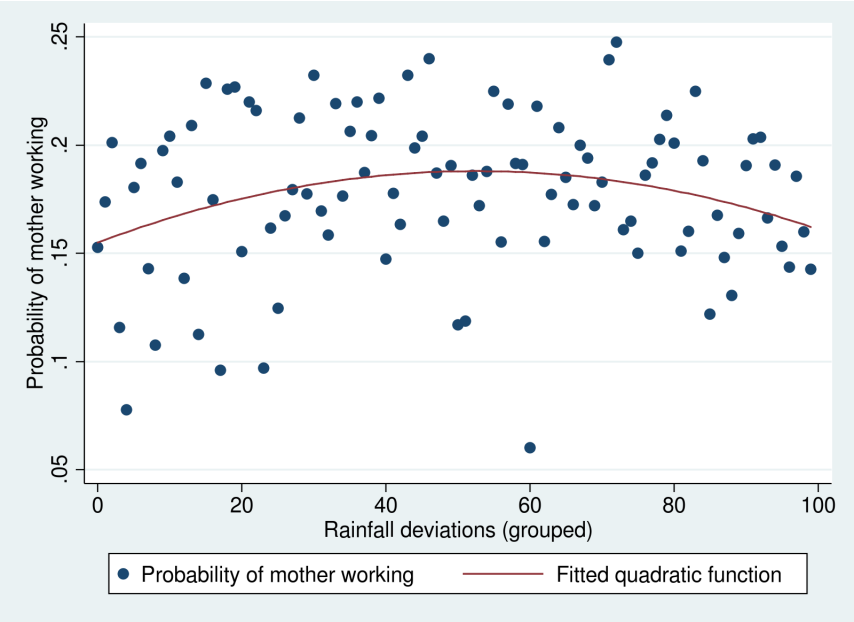

Figure 3.5: Relationship of mother's work to rainfall deviations for the rural sample- regional FE

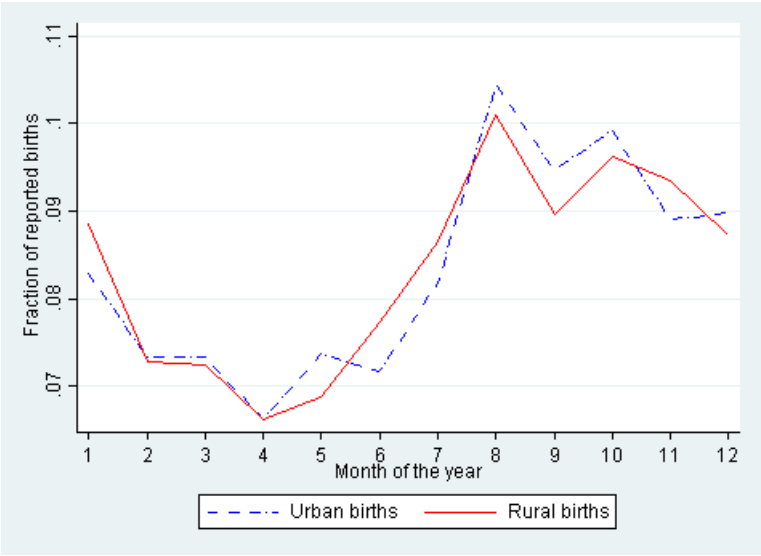

Figure 3.6: Fractions of births in each month for rural and urban areas

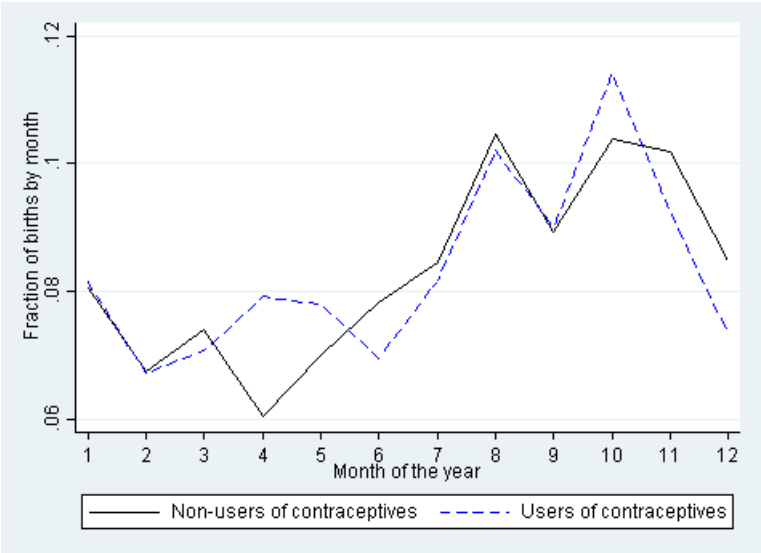

Figure 3.7: Fractions of births in each month for contraceptive users and non-users in rural areas

Table 3.1: Relationship of the probability that the last child is male to rainfall deviations for the rural sample

	Prob. last child was male		
Rainfall in the year prior to child's birth	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Age of the mother	0.003	0.004	0.003
	(0.002)	(0.002)	(0.002)
Birth Order	0.013	0.013	0.014
	(0.016)	(0.016)	(0.016)
Mother's education	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
Father's education	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
Size of the household	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
Hindu	0.020	0.019	0.021
	(0.025)	(0.029)	(0.025)
Muslim	0.056*	0.058*	0.057*
	(0.032)	(0.035)	(0.032)
Caste	0.022	0.023	0.022
	(0.017)	(0.018)	(0.018)
Household Standard of living index	0.003*	0.003*	0.003*
	(0.001)	(0.002)	(0.001)
Regional fixed effects	Yes	Yes	Yes
State time trends	No	Yes	No
Regional time trends	No	No	Yes
Observations	19089	19089	19089

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

Standard errors clustered at the district level.

Table 3.2: Relationship of the probability that the last child is male to soil types for the rural sample

	Prob. last child was male		
Fraction of district covered in fine soil	0.036	-0.001	0.035
	(0.058)	(0.057)	(0.059)
Fraction of district covered in loamy soil	0.028	0.016	0.027
	(0.053)	(0.053)	(0.054)
Mother's age	0.003	0.004	0.003
	(0.002)	(0.002)	(0.002)
Birth Order	0.013	0.014	0.014
	(0.016)	(0.016)	(0.016)
Mother's years of education	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
Father's years of education	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
Size of household	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
HINDU	0.020	0.019	0.021
	(0.025)	(0.029)	(0.025)
MUSLIM	0.055*	0.057	0.056*
	(0.032)	(0.035)	(0.033)
CASTE	0.022	0.023	0.022
	(0.017)	(0.018)	(0.018)
Household standard of living index	0.003**	0.003*	0.003*
	(0.001)	(0.002)	(0.001)
Regional fixed effects	Yes	Yes	Yes
State time trends	No	Yes	No
Regional time trends	No	No	Yes
Observations	19089	19089	19089

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

Standard errors clustered at the district level.

Table 3.3: Various Characteristics by Soil Texture

Variable	**Clayey Soil** (1)	**Loamy Soil** (2)	**Difference** (3) = (1) - (2)	**Obs.**
Growing period	21.805	34.4**	-12.605	341
	15.237	13.798	9.835	
Average rainfall	-44.756	158.678	-203.434	573
	226.454	193.801	145.774	
Average temperature	-0.448	−1.302***	−0.854***	271
	0.379	0.359	0.256	
Net Irrigated Area	21.358	39.826	18.468	271
	39.85	37.762	26.999	
Gross Irrigated Area	39.106	57.89	18.784	271
	52.405	49.66	35.505	
Literacy	-0.01	0.011	0.02	271
	0.03	0.029	0.02	
Soil Characteristics				
Topsoil 0-25 cm	0.03	0.076	0.046	271
	0.09	0.085	0.061	
Topsoil 25-50 cm	-0.23	-0.209	0.021	271
	0.145	0.137	0.098	
Topsoil 50-100 cm	0.266	-0.11	−0.377***	271
	0.18	0.17	0.122	
Topsoil 100-300 cm	-0.155	−0.43***	−0.275***	271

Continued on next page

Table 3.3 – *Continued from previous page*

Variable	**Clayey Soil** (1)	**Loamy Soil** (2)	**Difference** (3) = (1) - (2)	**Obs.**
	0.134	0.127	0.091	
Topsoil > 300 cm	0.088	0.673***	0.584***	271
	0.159	0.151	0.108	
$4.5 < pH < 5.5$	-0.135	0.143	0.278***	271
	0.162	0.153	0.11	
$5.5 < pH < 6.5$	−0.789***	−0.606***	0.183	271
	0.18	0.17	0.122	
$6.5 < pH < 7.5$	0.175	0.364***	0.189	271
	0.147	0.14	0.1	
$7.5 < pH < 8.5$	0.639***	-0.025	−0.665***	271
	0.179	0.169	0.121	
$8.5 < pH < 9.5$	0.104	0.112	0.008	271
	0.137	0.13	0.093	
Agricultural Yields				
Bajra	0.093	0.363	−0.270***	574
	0.144	0.121	0.096	
Barley	0.992	0.780	0.212	574
	0.277	0.222	0.189	
Cotton	-0.513	-0.884	0.372**	574
	0.310	0.300	0.175	
Gram	0.080	0.037	0.043	574

Continued on next page

Table 3.3 – *Continued from previous page*

Variable	**Clayey Soil** (1)	**Loamy Soil** (2)	**Difference** (3) = (1) - (2)	**Obs.**
	0.082	0.070	0.054	
Groundnut	0.053	0.186	-0.132^{**}	574
	0.105	0.092	0.067	
Jowar	0.419	0.202	0.217^{***}	574
	0.108	0.092	0.072	
Jute	3.053	3.203	-0.150	574
	2.117	1.661	1.722	
Maize	0.340	0.631	-0.291^{***}	574
	0.166	0.144	0.106	
Moong	-0.024	0.035	-0.059^{**}	574
	0.038	0.033	0.024	
Potato	6.137	26.793	-20.656	574
	127.772	75.821	93.123	
Rice	-0.008	0.736	-0.744^{***}	574
	0.182	0.155	0.114	
Sesamum	0.069	0.093	-0.024	574
	0.045	0.039	0.028	
Sugarcane	20.391	23.578	-3.186	574
	8.960	5.533	5.948	
Sunflower	-0.231	0.050	-0.282	574
	0.320	0.254	0.214	

Continued on next page

Table 3.3 – *Continued from previous page*

Variable	**Clayey Soil** (1)	**Loamy Soil** (2)	**Difference** (3) = (1) - (2)	**Obs.**
Tobacco	-0.979	-0.506	-0.473	574
	0.562	0.394	0.386	
Wheat	0.700	0.806	-0.106	574

Table 3.4: First stage with regional fixed effects

	Dep. var.: Mother worked in last 12 months			
	Rural	Urban	Rural	Urban
Clayey soil	0.151**	0.035	0.143**	0.041
	(0.067)	(0.046)	(0.067)	(0.050)
Loamy soil	-0.084	-0.057	-0.025	-0.053
	(0.058)	(0.037)	(0.056)	(0.040)
Mother's education	-0.002	0.004*	-0.002	0.005*
	(0.002)	(0.002)	(0.001)	(0.002)
Father's education	-0.008***	-0.001	-0.008***	-0.000
	(0.002)	(0.002)	(0.001)	(0.002)
Household size	-0.001	-0.002	-0.001	-0.002
	(0.002)	(0.001)	(0.002)	(0.001)
Hindu	0.013	-0.036	0.010	-0.041
	(0.024)	(0.034)	(0.027)	(0.034)
Muslim	-0.100***	-0.067*	-0.100***	-0.070**
	(0.029)	(0.035)	(0.031)	(0.035)
Scheduled caste or tribe	0.051***	-0.005	0.043**	-0.007
	(0.017)	(0.022)	(0.018)	(0.022)
HH standard of living	-0.007***	-0.004***	-0.008***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.002)
State time trends	No	No	Yes	Yes
Observations	19089	6826	19089	6826
F-statistic	22.00	5.21	9.15	3.96

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

The reported F-statistic is for the test of joint significance of the coefficients on the clayey and loamy soil variables.

Table 3.5: OLS results: Effect of rainfall and female LFP on child health (regional FE)

Dependent variables:		Months of breastfeeding		Dummy for child vaccinated against			
	Bfding dummy	w/ zeros	w/o zeros	BCG	Polio	Measles	DPT
$RAIN_{dst} * FemLFP_{hdst}$	0.003	-0.074	-0.174	0.023	0.026	-0.007	0.027*
	(0.013)	(0.228)	(0.218)	(0.015)	(0.016)	(0.014)	(0.015)
Mother worked in last 12 months	-0.004	-0.014	0.080	-0.019	-0.008	-0.040***	-0.014
	(0.011)	(0.212)	(0.198)	(0.017)	(0.017)	(0.015)	(0.017)
Rainfall deviation in last 12 months	0.008	-0.092	-0.111	0.059***	0.078***	0.023***	0.056***
	(0.006)	(0.107)	(0.097)	(0.008)	(0.009)	(0.007)	(0.008)
Mother's age at time of birth	0.002	0.121***	0.098***	0.007***	0.004	0.004**	0.006**
	(0.002)	(0.029)	(0.027)	(0.002)	(0.002)	(0.002)	(0.002)
Birth order	-0.012	6.809***	8.388***	0.100***	0.122***	0.342***	0.124***
	(0.011)	(0.269)	(0.257)	(0.014)	(0.014)	(0.018)	(0.014)
Child is male	-0.013	-0.013	0.061	-0.004	0.003	0.023*	0.016
	(0.011)	(0.184)	(0.157)	(0.014)	(0.014)	(0.013)	(0.014)
Mother's years of schooling	-0.000	-0.029	-0.043	0.008***	0.006***	0.005***	0.007***
	(0.001)	(0.031)	(0.036)	(0.002)	(0.002)	(0.002)	(0.002)
Father's years of schooling	0.002*	0.016	-0.012	0.011***	0.010***	0.007***	0.010***
	(0.001)	(0.024)	(0.024)	(0.002)	(0.002)	(0.002)	(0.002)
Household size	0.005***	0.113***	0.074***	-0.003	-0.003*	0.001	-0.003
	(0.001)	(0.025)	(0.022)	(0.002)	(0.002)	(0.002)	(0.002)
Household standard of living index	-0.000	-0.043***	-0.039**	0.010***	0.008***	0.003***	0.010***
	(0.001)	(0.017)	(0.016)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	19088	18764	17633	11260	11270	11083	11145

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Other covariates include dummies for Hindu, Muslim and scheduled caste or tribe households. We also control for the number of months since the child's birth.

Table 3.6: OLS results: Effect of rainfall and female LFP on child health (regional FE and state time trends)

Dependent variables:		Months of breastfeeding		Dummy for child vaccinated against			
	Brstfding dummy	w/ zeros	w/o zeros	BCG	Polio	Measles	DPT
$RAIN_{dst} * FemLFP_{hdst}$	0.003	-0.106	-0.230	0.022	0.026*	-0.006	0.026*
	(0.013)	(0.236)	(0.231)	(0.016)	(0.015)	(0.014)	(0.015)
Mother worked in last 12 months	-0.002	-0.017	0.059	-0.029*	-0.027*	-0.043***	-0.024
	(0.012)	(0.216)	(0.200)	(0.017)	(0.016)	(0.015)	(0.017)
Rainfall deviation in last 12 months	0.001	-0.158	-0.085	0.032***	0.044***	0.015*	0.041***
	(0.007)	(0.141)	(0.128)	(0.010)	(0.010)	(0.008)	(0.009)
Mother's age at time of birth	0.001	0.118***	0.094***	0.009***	0.006***	0.005***	0.008***
	(0.002)	(0.029)	(0.027)	(0.002)	(0.002)	(0.002)	(0.002)
Birth order	-0.008	6.834***	8.384***	0.104***	0.136***	0.340***	0.124***
	(0.011)	(0.271)	(0.259)	(0.013)	(0.014)	(0.018)	(0.014)
Child is male	-0.012	-0.008	0.072	-0.008	0.002	0.020	0.013
	(0.011)	(0.184)	(0.157)	(0.014)	(0.013)	(0.013)	(0.014)
Mother's years of schooling	-0.001	-0.030	-0.043	0.007***	0.005***	0.005***	0.006***
	(0.001)	(0.031)	(0.037)	(0.002)	(0.002)	(0.002)	(0.002)
Father's years of schooling	0.002*	0.015	-0.013	0.011***	0.011***	0.008***	0.011***
	(0.001)	(0.025)	(0.024)	(0.002)	(0.002)	(0.002)	(0.002)
Household size	0.005***	0.116***	0.074***	-0.001	-0.001	0.002	-0.001
	(0.001)	(0.025)	(0.022)	(0.002)	(0.002)	(0.002)	(0.002)
Household standard of living index	-0.000	-0.046***	-0.039**	0.007***	0.005***	0.003**	0.008***
	(0.001)	(0.017)	(0.017)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	19088	18764	17633	11260	11270	11083	11145

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Other covariates include dummies for Hindu, Muslim and scheduled caste or tribe households. We also control for the number of months since the child's birth.

Table 3.7: OLS results: Effect of rainfall and female LFP on child mortality (regional FE)

Dependent variables:	Child is alive	Child died in	
		< 1 month	< 6 months
$RAIN_{dst} * FemLFP_{hdst}$	0.006	-0.010	-0.021
	(0.011)	(0.012)	(0.022)
Mother worked in last 12 months	-0.031***	0.027**	0.024
	(0.011)	(0.011)	(0.020)
Rainfall deviation in last 12 months	0.011**	-0.009*	-0.019*
	(0.005)	(0.005)	(0.010)
Mother's age at time of birth	0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.003)
Birth order	0.020***	-0.016**	-0.021
	(0.007)	(0.008)	(0.016)
Child is male	-0.024***	0.029***	0.046**
	(0.009)	(0.009)	(0.018)
Mother's years of schooling	0.001*	-0.001	-0.003*
	(0.001)	(0.001)	(0.002)
Father's years of schooling	0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.002)
Household size	0.007***	-0.006***	-0.012***
	(0.001)	(0.001)	(0.002)
Hindu	-0.024*	0.029**	0.061**
	(0.013)	(0.014)	(0.031)
Muslim	-0.008	0.024	0.059
	(0.017)	(0.018)	(0.040)
Scheduled caste of tribe	-0.011	0.016	0.041*
	(0.010)	(0.012)	(0.022)
Household standard of living index	-0.000	0.000	-0.001
	(0.001)	(0.001)	(0.002)
Months since child's birth	-0.001	-	-
	(0.001)		
Observations	19088	16492	7812

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Sample for last two columns restricted to those children who are being observed at least x months after birth, or those that are being observed less than x months after birth but have already died, $x = 1, 6$.

Table 3.8: OLS results: Effect of rainfall and female LFP on child mortality (regional FE and state time trends)

Dependent variables:	Child is alive	Child died in	
		< 1 month	< 6 months
$RAIN_{dst} * FemLFP_{hdst}$	0.006	-0.010	-0.021
	(0.012)	(0.013)	(0.023)
Mother worked in last 12 months	-0.031***	0.029***	0.025
	(0.011)	(0.011)	(0.021)
Rainfall deviation in last 12 months	0.012**	-0.010*	-0.023**
	(0.006)	(0.006)	(0.011)
Mother's age at time of birth	0.001	-0.001	0.000
	(0.001)	(0.001)	(0.003)
Birth order	0.021***	-0.016**	-0.021
	(0.007)	(0.008)	(0.016)
Child is male	-0.024***	0.029***	0.048***
	(0.009)	(0.009)	(0.018)
Mother's years of schooling	0.001*	-0.001	-0.003*
	(0.001)	(0.001)	(0.002)
Father's years of schooling	0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.002)
Household size	0.007***	-0.007***	-0.012***
	(0.001)	(0.001)	(0.002)
Hindu	-0.019	0.021	0.050
	(0.015)	(0.015)	(0.032)
Muslim	-0.004	0.016	0.047
	(0.018)	(0.019)	(0.042)
Scheduled caste of tribe	-0.012	0.018	0.041*
	(0.011)	(0.012)	(0.022)
Household standard of living index	-0.000	0.001	-0.001
	(0.001)	(0.001)	(0.002)
Months since child's birth	-0.001	-	-
	(0.001)		
Observations	19088	16492	7812

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Sample for last two columns restricted to those children who are being observed at least x months after birth, or those that are being observed less than x months after birth but have already died, $x = 1, 6$.

Table 3.9: Second stage results with the IV: regional FE

Dependent variables:		Months of breastfeeding		Dummy for child vaccinated against			
	Brstfding dummy	w/ zeros	w/o zeros	BCG	Polio	Measles	DPT
$RAIN_{dst} * FemLFP_{hdst}$	-0.179**	4.698**	5.676**	-0.575*	-0.719**	-0.206	-0.444**
	(0.070)	(2.344)	(2.500)	(0.311)	(0.289)	(0.166)	(0.212)
Mother worked in last 12 months	-0.009	-2.889*	-3.315*	0.773***	0.379**	0.168	0.288**
	(0.056)	(1.748)	(1.929)	(0.184)	(0.151)	(0.121)	(0.144)
Rainfall deviation in last 12 months	0.039***	-1.488***	-1.775***	0.182***	0.207***	0.080***	0.131***
	(0.011)	(0.367)	(0.390)	(0.053)	(0.050)	(0.028)	(0.036)
Mother's age at time of birth	0.001**	0.018	0.011	0.006***	0.003	0.007***	0.004**
	(0.001)	(0.013)	(0.015)	(0.002)	(0.002)	(0.002)	(0.002)
Birth order	-0.004	2.623***	2.878***	0.077***	0.072***	0.458***	0.161***
	(0.004)	(0.125)	(0.135)	(0.012)	(0.011)	(0.012)	(0.010)
Child is male	-0.006*	-0.009	0.005	0.006	0.007	0.017**	-0.001
	(0.004)	(0.073)	(0.077)	(0.011)	(0.009)	(0.008)	(0.009)
Mother's years of schooling	0.003***	-0.205***	-0.247***	0.020***	0.016***	0.010***	0.011***
	(0.001)	(0.030)	(0.035)	(0.003)	(0.003)	(0.002)	(0.002)
Father's years of schooling	0.001*	-0.081***	-0.098***	0.018***	0.016***	0.010***	0.013***
	(0.001)	(0.023)	(0.026)	(0.003)	(0.003)	(0.002)	(0.002)
Household size	0.001	0.083***	0.081***	-0.006***	-0.005***	-0.002	-0.003*
	(0.001)	(0.013)	(0.013)	(0.002)	(0.002)	(0.001)	(0.002)
Household standard of living index	0.000	-0.051***	-0.053***	0.011***	0.007***	0.004***	0.009***
	(0.000)	(0.010)	(0.011)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	19088	18764	17633	11260	11270	11083	11145

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Other covariates include dummies for Hindu, Muslim and scheduled caste or tribe households. We also control for the number of months since the child's birth.

Table 3.10: Second stage with the IV: regional FE and state time trends

Dependent variables:		Months of breastfeeding		Dummy for child vaccinated against			
	Brstfding dummy	w/ zeros	w/o zeros	BCG	Polio	Measles	DPT
$RAIN_{dst} * FemLFP_{hdst}$	-0.089	0.798	1.333	-0.315**	-0.377**	-0.240*	-0.425**
	(0.059)	(1.439)	(1.301)	(0.145)	(0.151)	(0.136)	(0.169)
Mother worked in last 12 months	-0.144	-3.852**	-2.978*	0.226	0.071	-0.242	-0.096
	(0.092)	(1.958)	(1.743)	(0.186)	(0.175)	(0.170)	(0.192)
Rainfall deviation in last 12 months	0.011	-0.122	-0.197	0.090***	0.099***	0.051**	0.111***
	(0.010)	(0.226)	(0.206)	(0.027)	(0.028)	(0.024)	(0.030)
Mother's age at time of birth	0.002***	0.032***	0.023**	0.003**	0.002	0.006***	0.003**
	(0.001)	(0.012)	(0.011)	(0.002)	(0.001)	(0.001)	(0.002)
Birth order	-0.002	2.470***	2.748***	0.089***	0.073***	0.458***	0.165***
	(0.004)	(0.109)	(0.114)	(0.008)	(0.008)	(0.011)	(0.009)
Child is male	-0.007*	-0.018	0.018	0.000	0.002	0.016**	-0.003
	(0.004)	(0.066)	(0.062)	(0.008)	(0.007)	(0.008)	(0.009)
Mother's years of schooling	-0.000	-0.036*	-0.036*	0.010***	0.006***	0.004***	0.007***
	(0.001)	(0.019)	(0.018)	(0.002)	(0.001)	(0.002)	(0.002)
Father's years of schooling	-0.001	-0.027	-0.031**	0.011***	0.009***	0.006***	0.010***
	(0.001)	(0.017)	(0.016)	(0.002)	(0.002)	(0.002)	(0.002)
Household size	0.002***	0.019*	0.003	-0.002	-0.001	0.001	-0.001
	(0.001)	(0.010)	(0.010)	(0.001)	(0.001)	(0.001)	(0.001)
Household standard of living index	-0.001	-0.013	-0.006	0.005***	0.003***	0.000	0.004***
	(0.000)	(0.008)	(0.008)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	19088	18764	17633	11260	11270	11083	11145

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Other covariates include dummies for Hindu, Muslim and scheduled caste or tribe households. We also control for the number of months since the child's birth.

Table 3.11: Second stage mortality results (regional FE)

Dependent variables:	Child is alive	Child died in	
		< 1 month	< 6 months
$RAIN_{dst} * FemLFP_{hdst}$	-0.094**	0.095**	0.264*
	(0.047)	(0.046)	(0.142)
Mother worked in last 12 months	0.062	-0.050	-0.158
	(0.041)	(0.043)	(0.106)
Rainfall deviation in last 12 months	0.025***	-0.024***	-0.066***
	(0.008)	(0.008)	(0.024)
Mother's age at time of birth	0.001***	-0.001**	-0.002**
	(0.000)	(0.000)	(0.001)
Birth order	-0.016***	0.015***	0.044***
	(0.003)	(0.004)	(0.008)
Child is male	-0.008***	0.010***	0.017***
	(0.003)	(0.003)	(0.006)
Mother's years of schooling	0.004***	-0.004***	-0.010***
	(0.001)	(0.001)	(0.002)
Father's years of schooling	0.002***	-0.001	-0.003*
	(0.001)	(0.001)	(0.001)
Household size	0.002***	-0.002***	-0.003**
	(0.000)	(0.000)	(0.001)
Hindu	0.000	0.002	-0.010
	(0.004)	(0.004)	(0.014)
Muslim	0.006	0.001	-0.012
	(0.007)	(0.007)	(0.018)
Scheduled caste of tribe	-0.003	0.004	0.015
	(0.003)	(0.004)	(0.009)
Household standard of living index	0.000**	-0.000*	-0.001**
	(0.000)	(0.000)	(0.001)
Months since child's birth	-0.001*	-	-
	(0.000)		
Observations	19088	16492	7812

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Sample for last two columns restricted to those children who are being observed at least x months after birth, or those that are being observed less than x months after birth but have already died, $x = 1, 6$.

Table 3.12: Second stage mortality results (regional FE and state time trends)

Dependent variables:	Child is alive	Child died in	
		< 1 month	< 6 months
$RAIN_{dst} * FemLFP_{hdst}$	-0.049	0.057	0.126
	(0.039)	(0.037)	(0.091)
Mother worked in last 12 months	0.023	-0.007	-0.079
	(0.057)	(0.057)	(0.108)
Rainfall deviation in last 12 months	0.010	-0.009	-0.024
	(0.007)	(0.007)	(0.017)
Mother's age at time of birth	0.001**	-0.001**	-0.002*
	(0.000)	(0.000)	(0.001)
Birth order	-0.014***	0.013***	0.041***
	(0.003)	(0.003)	(0.008)
Child is male	-0.008***	0.010***	0.016***
	(0.003)	(0.003)	(0.006)
Mother's years of schooling	0.001***	-0.001**	-0.004***
	(0.000)	(0.001)	(0.001)
Father's years of schooling	0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.001)
Household size	0.003***	-0.003***	-0.005***
	(0.000)	(0.000)	(0.001)
Hindu	-0.006	0.009*	0.012
	(0.005)	(0.005)	(0.014)
Muslim	-0.005	0.013*	0.017
	(0.007)	(0.007)	(0.017)
Scheduled caste of tribe	-0.007*	0.008**	0.022**
	(0.004)	(0.004)	(0.009)
Household standard of living index	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)
Months since child's birth	-0.001	-	-
	(0.000)		
Observations	19088	16492	7812

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district level.

Sample for last two columns restricted to those children who are being observed at least x months after birth, or those that are being observed less than x months after birth but have already died, $x = 1, 6$.

CHAPTER 4

INTRODUCTION TO THE THEORETICAL MODELS OF THE INDIAN LABOR MARKET

The next two chapters of my dissertation construct theoretical models of the Indian labor market and derive the welfare implications of the introduction of an employment guarantee act along the lines of the National Rural Employment Guarantee Act (NREGA). Since the subject matter of both of these chapters is similar, this chapter will present the main features of Indian rural and urban labor markets, details of the NREGA, and a review of the relevant literature. Some of the discussion of the features of the NREGA will be familiar from Chapter 2, and so will be presented in less detail. A larger part of the discussion will be reserved for theoretical models and empirical studies of the Indian labor market.

The final section of this chapter will outline the basic features of the rural-only and the rural-urban models of the two chapters to follow. These features will draw heavily on the literature on the Indian rural and urban labor markets and their linkages.

4.1 Features of the Indian rural and urban labor markets

In this section we will describe the important features of the NREGA, certain stylized facts about Indian rural labor markets, and the nature of the linkages between the rural and the urban markets that operate through the migration of workers in search of jobs. These stylized facts will aid us in the development of our models in the next two chapters. We start by describing in some detail the

NREGA and the rural labor market.

4.1.1 The Indian rural labor market and the NREGA

The National Rural Employment Guarantee Act (NREGA) - or the Mahatma Gandhi National Rural Employment Guarantee Act, as it is now known - was introduced in 200 of India's poorest districts in early 2006, and extended to cover all rural areas of the country by mid-2008. The NREGA provides for a hundred days of unskilled manual work per rural household at a state-wise minimum wage (subject to a national minimum) and within 5 kilometers of the worker's place of residence.

Poverty reduction schemes have historically been a significant part of India's public works programs.[1] However as an act rather than a scheme the NREGA is the first program to provide the *right* to employment to the rural poor. It is more durable, provides more security to the workers, and cannot be canceled or changed without an amendment in Parliament - all features that are desirable from the point of view of poverty reduction and improvement of the livelihoods of rural agricultural laborers (see (Dey and Drèze, 2007) for more details on the provisions of the Act).

The benefits of such an Act have been widely discussed. At the time of implementation its main goals were stated as the "creation of durable assets and strengthening of the livelihood resource base of the rural poor" ((Chakraborty, 2007)). The social benefits, as envisioned by those who fought for its enactment,

[1]Most notably the Maharashtra Employment Guarantee Scheme, the Employment Assurance Scheme, National Rural Employment Programme, Jawahar Rozgar Yojana and the Sampoorna Grameen Rozgar Yojana.

are many and far-reaching. The Act would help protect rural households from poverty and hunger by providing work in the slack season when money was hard to come by, it would help reduce rural-urban migration by providing employment to families within the village, it would create assets of value for village farmers, it would aid in the empowerment of women and the increase in their agency within the household by providing them with income security, and so on (for more details, see (Khera, 2011)).

An employment guarantee scheme like the NREGA provides two types of benefits - transfer benefits, which come from increasing a worker's income, and stabilization benefits, which come from smoothing the income flow over time, especially when the markets exhibit considerable seasonality ((Ravallion, 1990)).

(Jha et al., 2012) calculate the direct transfer benefits as the shares of NREGA wages net the opportunity cost of time in household income. They find that these range from about 4% to about 20% depending on the type of household studied. However both (Jha et al., 2012) and (Imbert and Papp, 2011) show that these benefits are higher for the poor and for those with small land-holdings and help redistribute income from the larger households (the land-owners and net buyers of labor) to the smaller households (net suppliers of labor).

Indian agricultural markets do indeed exhibit considerable seasonality, making the stabilization benefits of an act like the NREGA all the more important. The timing of the busy periods of sowing and harvesting depends on the arrival of the monsoons in different parts of the country. The dependence of cultivation practices on a weather phenomenon like the monsoon results in the fluctuation of periods of intense activity with periods of inactivity. The sowing and harvest periods are the busiest times, with farmers who own land finding it necessary

to hire in as much labor as they can find in order to bring the harvest in in a timely fashion. The slack season consists of some work on the land in the form of weeding, irrigating, and preparing the land for the next season, but by and large the period between the planting of the crop and its maturity is when workers struggle to find employment. We incorporate this seasonality by allowing for two seasons in our model, a slack season followed by a peak season, both of equal length.

There is nothing in the operational guidelines that dictates that NREGA work only be undertaken at certain times of the year. However many districts implement these programs only in the slack season (see (Imbert and Papp, 2011)). In some cases this is driven by demand for work by laborers, in some by lobbying by farmers, who need laborers during the peak season. In other cases, as in Andhra Pradesh, it is imposed by the state through 'work calendars', which are in violation of the NREGA guidelines ((Association for Indian Development, 2009; Johnson, 2009)).[2] Whatever the reason for restricting NREGA works to the slack agricultural season, it seems to be a fairly widespread phenomenon. In our models in Chapters 5 and 6 the rural employment guarantee is only offered in the slack season.

The exact timing of the slack season differs from region to region. However studies of the Indian labor markets typically assume that the dry period of January to June is the slack agricultural season, while July to December is the peak season (see (Klonner and Oldiges, 2014)). With this characterization it would appear that the number of days provided under the NREGA, one hundred days, is less than the length of the slack season. When one takes into account that these

[2]The guidelines say that work should be provided whenever it is demanded, so work is not supposed to be restricted to certain seasons of the year under the provisions of the Act.

hundred days are not an individual entitlement but a household level allocation, it is evident that workers in NREGA work sites must find other sources of income for a large part of the slack season.

The seasonality of crop cycles naturally leads to intra-year fluctuations in labor demand and in agricultural wages as well. Many studies of Indian labor markets remark on the high levels of involuntary unemployment in the slack season, suggesting that labor markets do not clear and that on the whole wages are rigid downward in the slack season ((Bardhan, 1979; Drèze and Mukherjee, 1989; Ghose, 1980; Mukherjee and Ray, 1991; Swamy, 1997)). Several explanations have been forwarded for this downward rigidity of wages, for example, theories of peasant resistance to wage cuts or implicit collusion on the part of the workers, nutrition-based efficiency wages, and the pegging of casual labor wages to those of permanent (tied) laborers. (Swamy, 1997) and (Ghose, 1980) provide empirical tests of the latter two explanations, while (Mukherjee and Ray, 1991) build a theoretical framework to model the phenomenon of worker collusion and the resistance to wage cuts in the slack season. In the models that follow in Chapters 5 and 6 of this dissertation we model the agricultural market as being competitive in nature in order to maintain simplicity. Given the literature just alluded to, however, we acknowledge the fact that a model with downward wage rigidity in the slack season might be a more accurate representation of the Indian labor market.

In their study of the nature of the rural market in Palanpur, as well as their extensive review of the literature on Indian rural markets, (Drèze and Mukherjee, 1989) also present several stylized facts. They outline evidence suggesting that most laborers search for work within their own villages or in neighboring

villages. Rural-rural permanent migration is uncommon, but since many villages are within a few kilometers of each other, a worker can easily move from one to the other in search of employment. Clearly in the slack season when employment is hard to come by in all villages workers will not be able to find employment simply by moving from one village to another. However in the peak season the labor market is tighter, and a worker who doesn't find work in his own village can potentially look for work with a landowner in a close-by village. As a result of this easy movement from one village to another, (Drèze and Mukherjee, 1989) highlight the large number of laborers and farmers in these village markets. This is how we model our rural sector as well - with a large number of laborers and farmers.

(Drèze and Mukherjee, 1989) also report that the daily wage is uniform across workers within a particular season, regardless of their skill level. Since skill differences don't seem to play a role in wage determination, we simplify this by modeling all workers as being identical.

Finally (Drèze and Mukherjee, 1989) point out that much of the agricultural labor hired in India is casual in nature. A casual labor market means that workers are not bound to specific employers, but sell their labor to the highest bidder. Farmers do not hire workers for several periods (or years) at a time, but hire as many workers as they need for the specific task at hand. In contrast to these casual markets, labor tying means that at the start of the year the farmer hires a certain number of workers to work with him for both seasons (or longer), perhaps supplementing these permanent workers with casual labor in the peak season if his labor needs increase. Such a contract often arises out of the farmer acting as a creditor for a worker in the slack season, with the worker in return

promising his labor services in the peak season when labor is hard to come by. These tied labor contracts, either annual or of longer duration, are common in Andhra Pradesh, Rajasthan, Madhya Pradesh and Punjab.[3]

Tied labor contracts do exist in certain parts of the country, but these contracts have been declining over time ((Mukherjee and Ray, 1995)). Instead, the large majority of agricultural labor is employed on a daily basis by farmers, with the farmers' need for labor in the season in question dictating how many workers they hire. (Basu, 2013) suggests that the main beneficiaries of an employment guarantee like the NREGA would be those laborers who are involuntarily unemployed in the slack season or are indebted to a landlord, and hence would be more likely to be workers in areas where tied labor contracts are prevalent, as these contractual arrangements by design generate slack season involuntary unemployment. As far as we aware there has been no comprehensive study of the relationship between the presence of tied labor contracts and the take-up of NREGA work. However our field experience, and preliminary figures from field surveys, suggests that many of the beneficiaries of the NREGA work-sites are in fact casual laborers.

Table 4.1 shows the distribution of primary occupations for a sample of villagers surveyed in ten districts in the large north and central Indian states of Jharkhand, Chhattisgarh, Rajasthan, Bihar, Madhya Pradesh and Uttar Pradesh. These are some of the poorest regions in the country and the people interviewed were selected only from among those working for the NREGA (drawn at random from the work-site attendance sheets), so this sample is not representative. With this caveat, however, it is interesting to look at the figures. The

[3]See (Basu, 2013) for a longer discussion of why these contracts arise, and more references to studies of their development.

Table 4.1: Primary Occupations of NREGA workers (survey data)

Occupation	Freq.	Percent	Cum.
Casual Labor	560	53.03	53.03
Self-employment (agriculture)	411	38.92	91.95
Self-employment (non-agriculture)	30	2.84	94.79
Regular Employment	8	0.76	95.55
Other	43	4.07	99.62
Unclear	4	0.38	100
Total	**1056**	**100**	**100**

Source: Survey data obtained from Jean Drèze and Reetika Khera, surveys supported by the Center for Development Economics, Delhi School of Economics.

primary occupation of more than half the sample is casual labor, followed by self-employment in agriculture. More than 90% of the population is covered by these two occupations. Table 4.2, drawn from the same data, shows the secondary occupation for those who stated self-employment in agriculture as their primary occupation (not everyone has a secondary occupation). As can be seen from the table, more than 92% reported casual labor as their secondary occupation. Longer-term contractual agreements between farmers and laborers, like tied labor-credit contracts, seem to be uncommon, at least in this sample.

Finally, the last feature of the NREGA that is of interest to us is the relationship between agricultural wages and NREGA wages at the time of the introduction of the program. Clearly a wage guarantee program will only have labor market effects if the wage it provides is greater than the equilibrium wage that prevailed before it was introduced. This does indeed seem to be the case in India in 2006, as borne out both by our own field observations as well as survey data. Table 4.3 (reproduced from (Khera, 2011)) reports the NREGA wage along with the average wages for other agricultural or casual work for both men and

Table 4.2: Secondary Occupation of Self-employed Agricultural Workers (survey data)

Occupation	Freq.	Percent	Cum.
Casual Labor	203	92.27	92.27
Self-employment (non-agriculture)	13	5.91	98.18
Regular Employment	1	0.45	98.63
Other (specify)	3	1.37	100
Total	**220**	**100**	**100**

Source: Survey data obtained from Jean Drèze and Reetika Khera, surveys supported by the Center for Development Economics, Delhi School of Economics.

women. It is clear from this table that the gap between the NREGA wage and the going agricultural wage was large. While this is data drawn from a small sample of workers in only six states, it is suggestive of the wedge between the going wage in the absence of the rural employment guarantee, and the NREGA wage.[4] In our models of Chapters 5 and 6, we model the NREGA as offering a wage that is deliberately pegged above the slack season agricultural wage that prevails in the absence of this program. We do not place any restrictions on the relationship between the NREGA wage and the peak season agricultural wage.

In the outline above we have discussed the main features of the Indian rural economy and of the NREGA. Of particular importance are seasonality in production and employment, the relationship between the NREGA wage and the agricultural wage, and the nature of the contractual arrangements between farmers and laborers. As will be seen in Chapters 5 and 6, these features will

[4]It is unclear from the context whether the agricultural wages are market-wide averages or those reported by the NREGA participants. If it is the latter, one might worry that those who were earning lower wages before the introduction of the NREGA were the ones who selected into the program - the so-called 'Ashenfelter's dip' phenomenon. It is also unclear whether these are season-specific wages or annual averages. In light of this the figures reported here are treated as being simply illustrative.

Table 4.3: Wages of Sample Workers

Average Wage (Rs/day)	Women	Men
Statutory Minimum wage	88*	88*
Agricultural Work	47	53
Other casual labor	58	71
NREGA work	85	85

Source: Reproduced from table 4.1 of (Khera, 2011).

* Unweighted average of the state minimum wages in the six survey states

inform our modeling decisions.

4.1.2 Rural-urban migration in India

There is a noticeable lack of information on the structure of the Indian urban labor market, so many of our assumptions regarding jobs and search strategies in Chapter 6 will be informed instead by our observations of how these markets work. However there are some papers that look at the linkages between rural and urban markets that arise out of migration, and these are also helpful in designing our model.

A large proportion of the migration in India is of a temporary nature, with workers moving to urban areas in search of work in the agricultural slack season and returning to the rural areas at the time of the harvest. This phenomenon of 'circular migration' - i.e. of workers who migrate in search of work but return to the rural areas when they can or when they have better job prospects - has been discussed in (Haan, 1997). The author shows that it is a broad middle strata

that emigrates, so not just the very poor or the very rich. Using NSS data he also shows that there is no evidence that migrants earn significantly different wages from other urban workers, or that their situation improves significantly over time relative to when they first arrived in the urban area. The migrants do retain strong ties with their villages. If the migrant owns a plot of land, even if small, this contributes to his return to his place of origin.

Just how important is short-term migration in India? NSS figures suggest that in 2007 8.5 million rural adults undertook short-term migration trips for work in urban areas ((Imbert and Papp, 2014)). In comparison, the net rural-urban migration during this period was only 2 million people, indicating that the volume of temporary migration is considerably larger than that of permanent migration. The influx of migrants is also large when seen relative to the number of urban residents engaged in short-term wage work, which the authors estimate at about 15 million. Temporary rural-rural migration in search of work used to be much more prevalent about twenty years ago, but with the increasing mechanization of agriculture and the decline in agricultural work, rural-urban migration has increased in importance over the last twenty years ((Deshingkar, 2005)).

(Banerjee, 1991) collects survey data on migrant workers in Delhi, and documents that workers migrate both for pre-arranged jobs as well as to search for jobs on arrival. Most migrants who arrive in Delhi without pre-arranged jobs quickly become employed, either in wage employment or in self-employment. For instance, about two thirds of the sample is employed within a week. There is not a great deal of on-the-job search. Those who want a non-manual job, and those workers who are more educated and older, have a greater probability of

migrating with a pre-arranged urban job. He also finds some evidence suggesting that those migrants who specified wanting a non-manual job resisted moving down the job ladder to a manual job and experienced a longer spell of unemployment as a result, as did workers with greater education levels.

In our model of the rural-urban labor market linkages in Chapter 6, we incorporate some of these observations. Our model has an urban and a rural sector. In the rural sector, there are two seasons, a slack and a peak season. Rural-born workers can either search for work in the rural area, or migrate to the urban area in search of work. The slack season precedes the peak season. Workers migrate to the urban area in the slack season, but those of them that do not find permanent high-paying jobs return to the rural area in the peak season - the 'circular migration' discussed above.

We also allow for a free-entry sector where anyone in the urban area who wants work can find it without delay. This assumption is supported by the speed with which (Banerjee, 1991) reports workers finding low-wage work in the urban areas upon arrival. Finally, we introduce a status cost for the free-entry sector which makes it preferable for an educated worker to be unemployed than to get a job in this sector. This also builds on the observation by (Banerjee, 1991) that workers who wanted a non-manual job and more educated workers prefer to be unemployed longer rather than accept a job that perhaps they deem to be 'beneath them'.

4.2 Further studies of the Indian labor market

Four strands of the literature are relevant to the development of our models, and we review these here. The first is the large number of empirical studies of the NREGA. The second strand is more limited, and consists of papers that have attempted to model features of the Indian rural labor market theoretically. The third strand includes empirical studies of the reasons for rural-urban migration and how it is affected by the introduction of an employment guarantee in India. The last strand consists of theoretical models of rural-urban migration, including some market-level models similar to the models we will develop in Chapter 6 of this dissertation. The empirical studies will help us determine the relevant features of our stylized model, while the theoretical models will highlight what has been discussed in the literature previously, and what our contribution is.

4.2.1 Empirical Evidence on the Impact of the NREGA

Most of the empirical work in the past few years has focused on the impact of the NREGA on agricultural wages and employment. In the previous section we presented some evidence of the wedge between the minimum wages workers received under the Act and the going agricultural wages at the time of the introduction of the program. Since the wage in the NREGA was greater than the going agricultural wage, economic theory predicts that the NREGA would attract workers away from agriculture, resulting in a fall in private agricultural employment and a rise in private agricultural wages.

These first-order effects have been the subject of several studies, using var-

ious different data sources. However these studies have delivered mixed and sometimes even contradictory evidence. (Imbert and Papp, 2011) use National Sample Survey (NSS) data and a difference-in-differences methodology. They find that following the introduction of the program, working-age laborers spend an additional .3 days per month in public employment, but 1.6% fewer days in private employment. Casual agricultural wages also increase by 4.5% in the districts where the NREGA was introduced early. The results are stronger in states with better implementation of the NREGA, and the effects are mostly confined to the slack season with low rainfall. These findings are in line with our intuition.

Using the same dataset and methodology but a different sampling frame, (Azam, 2012) finds that the NREGA has a positive impact on labor force participation, but that the effect is much stronger for women than for men. However (Zimmermann, 2013) also uses the NSS and a regression discontinuity design to conclude that the impact of the NREGA on employment and casual wages is small in magnitude and often statistically insignificant. She does show that take-up is higher following a bad rainfall shock, indicating that NREGA might act as a safety net for households. The fact that NREGA income acts as insurance against weather shocks was also pointed out by (Johnson, 2009).[5]

Another question that has been addressed is the question of who is working in the NREGA work-sites, and whether the program is indeed used more by poorer and more disadvantaged households. (Liu and Barrett, 2013) use NSS data to analyze which households actively seek NREGA work, which households are denied employment, and as a result of these two forces, what is the re-

[5]There is more evidence that the NREGA acts as implicit insurance, especially in the context of its impact on migration, which is often a risk-diversifying activity. These papers will be reviewed in the section on migration.

sulting participation profile across the household per capita expenditure distribution. The authors show that there is greater self-selection among those whom the Act was aimed at, i.e. scheduled castes and tribes and poorer households, but because of the variation across states in the extent of rationing of work, the overall effectiveness of the pro-poor targeting differs greatly across states.

In their study of households in Andhra Pradesh, (Ravi and Engler, 2009) find that while the program initially attracted non-agricultural labor, over time there was a shift towards households that would have participated in agricultural labor had the guarantee not existed. In their words "this suggests broader labor market distortions where NREGS is not just viewed as an employment assurance during the slack agricultural season but as an alternative to agricultural labor work". The setup of our rural-model also delivers the result that the NREGA attracts labor away from agriculture in the slack season.

There are also papers that have looked at broader effects of the NREGA which go beyond the first-order impacts on rural labor market outcomes. (Bhargava, 2014) shows that farmers in Indian agriculture increase their use of capital-intensive technologies after the NREGA is introduced, presumably because labor has become more expensive with the rise in agricultural wages. (Gehrke, 2014) shows that households in Andhra Pradesh increase the allocation of resources to riskier crops following the introduction of the program, suggesting that the NREGA is either providing implicit insurance or easing credit constraints. (Klonner and Oldiges, 2014) study the effect on consumption expenditure and poverty rates. They show that the program has had large and significant effects on smoothing consumption among households that are particularly reliant on agricultural wage labor, eradicating the large seasonal fluctuations

that were a feature of the pre-NREGA consumption patterns. They also find a large reduction in poverty among these households as a result of the NREGA.

4.2.2 Theoretical Models of the Indian rural labor market

(Mukherjee and Ray, 1991) develop a theoretical model of the Indian labor market that incorporates seasonality, and delivers a mechanism by which slack season involuntary unemployment can be sustained without a fall in the agricultural wages. In their model, every worker has a 'notional fair wage', and views a farmer as unfair if they do not pay their workers at least that wage in the slack season. Depending on the tightness of the peak season labor market and the proportion of 'fair' farmers in the total, workers can choose to punish unfair farmers by refusing to work for them in the relatively tight peak season. If a worker refuses to work for a particular farmer he waits for another one to make him an offer to work. If no farmer makes him an alternate offer or if he refuses all offers then he remains unemployed in the peak season. Since making an offer of employment to a worker takes time and losing time in the peak season is costly to the farmer, this threat of refusal to work has bite. With this implicit threat, it is possible to sustain equilibria where despite the existence of involuntary unemployment in the slack season, the wages in this season are greater than the workers' reservation wages.

(Eswaran and Kotwal, 1985) provide a reason for the existence of tied-labor contracts alongside casual labor contracts in the Indian framework. A tied contract in this context is an offer of employment for two seasons, at a constant wage in each season. Clearly for a risk-averse worker, a permanent contract of-

fers greater utility than a casual contract with a low wage in the slack season and a higher wage in the peak season. In their theoretical model crops take one year to be produced, and a year has two seasons of equal length. Different tasks in production require different levels of care and effort from workers. Tasks like weeding, threshing and so on are more routine, while tasks like the application of fertilizers and the maintenance of machines require more care. In this framework, permanent contracts are awarded to workers in order to elicit loyalty and trustworthiness, so that they can be entrusted with the tasks where worker effort is more crucial.

Despite reasons why tied labor contracts that offer consumption-smoothing would be preferable to risk-averse workers, these contracts have been declining over time, and are much less widespread than casual labor contracts in India. (Mukherjee and Ray, 1995) develop a theoretical explanation for this phenomenon. Their model incorporates seasonality and imperfect credit markets, both of which should make tied contracts more desirable from the point of view of risk-averse workers. However when an active casual labor market operates alongside the tied contracts, these features also encourage workers to deviate from their agreed upon implicit contracts, for example, when the spot wage in the casual labor market in the peak season exceeds the tied wage. Standard economic theory suggests that labor tying should be widespread so long as the markets display some seasonality. The authors show that with the possibility of default on the part of the worker, equilibrium levels of labor tying are actually much lower than standard theory would predict.

(Basu, 2013) builds on the same basic theoretical framework as in (Mukherjee and Ray, 1995). He models the impact of an EGA in a market with inter-seasonal

tied labor contracts as well as casual labor contracts. The farmer chooses the number of permanent contracts to offer at the start of the slack season at a fixed per-season wage, and supplements these permanent workers with casual workers in the peak season, at the going casual sector wage. The EGA is assumed to have a positive impact on agricultural productivity in the peak season. An increase in the wage paid under the EGA raises the wages paid to permanent workers and reduces the number of permanent workers hired. An increase in the productivity of labor in the EGA also raises the permanent contract wage through the productivity effect of the government program. Under an elasticity condition that ensures that agricultural production does not fall, raising the productivity of workers under the EGA also raises the number of permanent contracts offered. This in turn tightens the casual labor market and exerts an upward pressure on the casual wage. Thus an increase in the wage paid to workers working in the EGA, or in their productivity can have positive impacts on the wages received in both casual and permanent contracts, even while not reducing agricultural output.

(Basu et al., 2009) model the Employment Guarantee Scheme (EGS) as a 'wage-access pairing" - i.e. a combination of the wage offered, and the ease with which a worker can avail of the guarantee, for example, the distance to the workplace. They show theoretically how the wage and access parameters for the EGS should be designed in order to attain certain objectives, for example, maximizing private employment, or achieving a certain level of aggregate (private plus EGS) employment. The main role of the EGS in their framework is to increase contestability in labor markets, not to hire workers in public works, and its efficacy in achieving this goal depends on the credibility of the planner's commitment to the wage and access parameters he announces.

Finally, one of the main criticisms of the NREGA is that as a result of increasing agricultural wages and declining agricultural employment, the program has resulted in reduced farm production, and some authors have even blamed the NREGA for the recent increase in inflation in food prices. (Datta, 2014) develops a model of the Indian rural labor market to demonstrate that the critique that the NREGA has reduced agricultural production need not be true. In his model workers have a subsistence level of consumption, and receive an infinitely high dis-utility if they do not attain this level. In a market with this kind of discontinuity, he develops conditions on agricultural productivity and labor endowments under which the NREGA can improve worker welfare (measured by their utility) without decreasing agricultural production. Given differences in labor endowments and agricultural productivities across the country, this model suggests that the number of NREGA days each household is entitled to should not be fixed at a national level, but instead should be locally adjusted in order to ensure that agricultural production is not adversely affected.

4.2.3 Empirical studies of rural-urban migration in India

The very low permanent migration rates in India have been attributed to a number of factors - strong ties to the village, the absence of a land market which makes it difficult to sever ties and move permanently, the lack of steady employment opportunities in urban areas and so on. (Munshi and Rosenzweig, 2009) suggest that another possible reason for low migration is the existence of sub-caste networks within the village that provide mutual insurance to their members. Members of higher-income caste networks are less likely to both out-marry and out-migrate than members of lower-income caste networks, because

leaving the village could mean losing access to this insurance.

Other evidence on whether migration is used to mitigate risk comes from (Badiani and Safir, 2008). This paper studies whether households in rural India use migration as a way to cope with other aggregate shocks. Using ICRISAT data they find that an increase in aggregate rainfall, a positive shock, decreases temporary migration. However positive unanticipated idiosyncratic agro-climatic shocks increase temporary migration.

(Morten, 2013) models temporary migration and household risk-sharing decisions as being made jointly by rural households. She estimates her structural model using ICRISAT panel data, and finds that migration reduces risk sharing by 23%, and that risk-sharing in turn reduces migration by 60%. She then introduces an employment guarantee act along the lines of the NREGA into her model and demonstrates how such a policy substitutes for informal insurance and reduces risk-sharing, as well as increasing income within the village, thereby substituting for migration.

There are several papers that attempt to assess the causal impact of the NREGA on rural-urban migration, and through this channel, on urban labor market outcomes. (Imbert and Papp, 2014) use NSS data as well as survey information to study the effect of the NREGA on short-term migration from rural to urban areas in India. They find that the program significantly reduces rural-urban migration. Rural migrant workers compete with urban unskilled workers for unskilled short-term jobs, and so the decline in migration constitutes a leftward shift in the supply of labor curve for these jobs. With no effect on demand for labor, this shift in supply causes the wages for these short-term unskilled jobs to rise. They do not find any effect of the reduced migration on

the wages for urban salaried workers, as these workers do not participate in the same market as rural workers. Districts that had a 10% higher pre-NREGA in-migration rate from districts that received the NREGA early experienced about a 5% increase in short-term wages, and about a 0.5 percentage point increase in the time spent doing casual labor.

(Ravi et al., 2012) also study the impact of the NREGA on rural-urban migration and urban unemployment, using data from two rounds of the NSS. They compare districts that received the program with nearby districts in a difference-in-difference framework, thus exploiting the timing of the roll-out of the program. They find that the NREGA reduced rural-to-urban migration by close to 30%, and has also reduced urban unemployment by almost 40%. The effects they report are limited to rural uneducated households, and the migration decisions of rural educated households seem to largely unaffected by the program.

4.2.4 Theoretical models of the urban sectors and rural-urban migration in developing countries

Standard models discuss the rate of migration as a negative function of rural income, indicating that those workers who earn more in the rural areas are less likely to leave. (Banerjee and Kanbur, 1981) instead propose a model of rural-urban migration where the benefit to migrating to the urban area is an increasing function of the rural income of the worker. Their justification for this is that in the absence of perfect capital markets in rural areas, many rural workers do not have enough money to finance their migration move. As a result, current income becomes a constraint on migration, and only richer workers are able to

move. When this feature is introduced into the model, the returns to migration become a non-monotonic function of rural income. Their main theoretical predictions are that migration rates are higher in areas where there is greater land inequality, and rural poverty acts as a deterrent to migration, not an incentive.

There have also been a few market-level models along the lines of the now well-known Harris-Todaro model. (Gupta, 1993) models a labor market with formal and informal sub-sectors within the urban sector, and agricultural work in the rural sector. The level of output of the informal sector is determined by the demand from the urban formal sector, and the urban informal sector relies on rural migration for its labor supply. The key difference between this model and the Harris-Todaro model is that the size of the urban labor force is limited by the amount of food the rural sector produces. A policy that subsidizes rural employment increases the amount of food produced, which increases the amount of food available to the urban sector and allows the number of workers in this sector to expand. This has the effect of increasing urban unemployment, a result that is in direct opposition to the Harris-Todaro model. On the other hand, a wage subsidy to the urban formal sector does not change the availability of food to the urban sector and so the size of the urban workforce remains unchanged. Such a subsidy does, however, reduce the cost of employment and thus causes the urban unemployment rate to fall.

(Gupta, 1993) allowed for the production in the informal sector to depend on the demand from the urban formal sector, but did not allow aggregate demand to determine the output of the urban formal or rural sectors. (Chaudhuri, 2000) builds a model that allows aggregate demand to play a role in determining the level of production and employment in all three sectors. Now a price

or wage subsidy to the rural sector raises the aggregate income of the workers in the economy relative to no subsidy. This raises the aggregate demand for the output of all three sectors, and increases the level of employment in all three sectors. Thus rural development has the familiar effect of reducing urban unemployment, as in the classic Harris-Todaro model.

(Basu et al., 2013) extend the classic Harris-Todaro framework within an urban economy by adding overlapping generations of young and old workers, and a high-paying self-employment sector that only workers with some experience in the formal sector can enter. With these extensions they are able to capture the phenomenon of workers acquiring a skill in a formal sector job, and leaving to start their own businesses. They allow for heterogeneities in ability among workers as well. In this richer and more complex model they are able to answer many questions that cannot be addressed in the standard Harris-Todaro model, such as "what is the impact of increasing high-wage employment on employment among younger workers?", and "what proportion of the self-employed are able to become high-earning entrepreneurs?" Some of their conclusions also run counter to the classic Harris-Todaro model, such as their prediction that under certain conditions high-wage employment and urban unemployment can actually move in opposite directions.

Finally (Lal, 1988) develops a model of "bumping" - i.e of higher education workers replacing those workers with lower skills. He finds empirical evidence of this from the NSS. Such a use of education as a screening device is one possible explanation for open unemployment in urban India.

4.3 Features of the rural-only and rural-urban models of Chapters 5 and 6

In the previous two sections we discussed at length the features of the Indian rural and urban labor markets, and the nature of the linkages between the two. We then presented some of the empirical and theoretical papers from the literature on India in order to illustrate their key findings. The theoretical papers we discussed are also important because we will be borrowing some modelling features from them. Now that the groundwork has been laid, we are in a position to present the basic features we will be including in our models in Chapters 5 and 6.

4.3.1 The Basic Features of the Rural-only Model

Seasonality

The first feature of the rural-only model concerns the seasonality of agricultural work. In our model agriculture is characterized by two seasons - a slack season, and a peak season. The peak season is the busiest time of the year. Activities undertaken in the peak season involve plowing the land, sowing the seed for the crop to be grown, harvesting the grown crop, transporting the grain to the nearby market, and finally selling it at the end of this process. Empirical studies in India have shown that the peak season labor market is at or close to full employment, and that is how it will be modeled.

The slack season, on the other hand, is the part of the year when there is not

much work to be done on the farms. It consists largely of the time between the sowing of the crop and its maturation. During this period some labor is required for irrigation, weeding and the application of fertilizer to the crops - in general, tending to the well-being of the plants in their growing phase - however there are not a large number of tasks to be conducted. Thus the slack season will be characterized by less than full agricultural employment in the rural market.

In reality these two seasons do not follow one another in a neat chronological order, however for ease of exposition we will assume here that there is a clear time-line. The way it is modeled here the slack season comes first and is followed by the peak season, and both seasons are of equal length. We also assume that at the end of the peak season the cycle starts afresh, so it will suffice to study only one single year of two seasons. Importantly for our rural-only model, the amount of slack season labor hired positively affects the marginal productivity of the peak season laborers. As an example, we can think of this as better irrigation and better tended soil improving the quality and concentration of fruit on the trees, so that peak season workers do not need to inspect and discard as many fruit when they are collecting the harvest, thus making them more efficient.

Casual Labor Contracts

The second feature is that the main source of labor for the farmers is casual labor, where workers are generally hired on a day-to-day basis. Earlier in this chapter we presented evidence that tied labor contracts are less common than casual labor arrangements, and that tied contracts have been declining over time. As a result, in this and in the rural-urban model, there will not be any tied labor

contracts between farmers and workers. Farmers hire only as many workers as they need in each season separately.

Nature of the search for work

Laborers can search for agricultural work in nearby villages as well as in their own. In the slack season when employment is low there is not much benefit to travelling out of one's village in search of work. However in the peak season it is possible that laborers move to nearby rural areas to find employment. Since workers can move, we allow for a large number of laborers and a large number of farmers in our models. We model all workers and farmers as being identical.

Market Structure

Studies of the Indian labor market suggest that collusion on the part of the farmers is uncommon, but that there is some evidence of collusion among workers, particularly in resistance to wage cuts. For simplicity, however, we assume a competitive market, which means that no individual farmer (buyer of labor) or agricultural worker (seller of labor) can influence the wage (the price of labor), and that the wage adjusts to clear the market. We assume also that the daily wage is uniform across workers within a particular season.

The NREGA wage

The last stylized fact that we introduce is the relation of the NREGA wage to the prevailing slack season wage before its introduction. Earlier in this chapter we

discussed how the NREGA wage was higher than the agricultural slack season wage that prevailed before its introduction. Indeed the labor market effects that some authors have documented would not be present if the NREGA wage were lower than the slack season agricultural wage, as it would not attract workers away from agriculture. Thus we model the NREGA wage as being deliberately pegged at a level higher than the slack season agricultural wage.

4.3.2 Basic features of the rural-urban model

Locations and workers

There are two locations in this model, the urban sector and the rural sector. Workers are born into each of these sectors, and their location at the time of entering the labor force determines what types of jobs they can receive, and their wages in those jobs.

Education

We assume that workers can either be educated or uneducated, and are endowed with education at the time of entering the labor force. The level of education a worker possesses is not a choice variable in our model. All rural workers are uneducated, while urban workers can be either educated or uneducated. The access to certain jobs is determined by the level of education a worker possesses. Workers are distinguished then along two dimensions - where they are born (*rural* or *urban*) and the level of education they are endowed with (*educated* or *uneducated*).

Employment in the rural sector

There are two types of employment available in the rural sector - agricultural work, or work in the rural employment guarantee when it is provided. We assume that agricultural work can only be performed by those workers who were born into the rural areas, so urban workers do not move to the rural areas in search of employment in agriculture. This is meant to reflect the negligible urban-to-rural migration in India. However workers from the rural sector can migrate to the urban sector in search of work.

The rural sector in the model of Chapter 6 does not assume any inter-temporal spillovers in productivity between the slack and the peak seasons, but is otherwise similar in structure to the model in Chapter 5, with the features described above. We omit inter-temporal spillovers in order to maintain tractability, but can think of this rural-urban model as being a special instance of an all-India model that also incorporates these spillovers in productivity from one season to the next.

Jobs in the urban sector

There are three types of jobs in the urban area, which for ease of exposition we stylize as managers' jobs, office-workers' jobs and the free-entry sector jobs. Managers' jobs are only open to those workers with education, while office-work and free-entry sector work is open to both educated and uneducated workers. However educated workers are hired preferentially in the office-workers sector, so uneducated workers only get jobs if there is still work available after all the educated workers who want jobs in that sector have received

work. We assume that the number of educated workers is 'small' - such that there are still some office-workers' jobs available for uneducated workers. The condition that guarantees this will be discussed in Chapter 6.

Workers from both the urban and the rural sectors can take up work in the urban free-entry sector. Free-entry sector work consists of activities like manual labor in construction, selling tomatoes outside a housing colony, cutting hair on the side of the road, shining shoes and so on. Thus it is natural to think of this sector as having a fixed total product, but a declining average product - the more people there are selling tomatoes the fewer tomatoes each of them will sell.

We make the additional assumption that the total product that urban free-entry sector workers can earn is larger than the total product that rural free-entry sector workers can earn. This can be thought of as urban sector workers knowing where the best locations are for selling tomatoes, knowing the busiest intersections where the demand for magazines at traffic lights would be greatest and so on. There is only one period in this model, and so the rural workers who migrate to the urban sector do not have a chance to acquire the same knowledge as their urban counterparts.

The rural employment guarantee

The rural employment guarantees of Chapters 5 and 6 have the same features. The guarantee provides unskilled manual work for rural households at a fixed minimum wage, and is available only in the slack season. The wage it provides is higher than the slack season wage in the absence of the guarantee. Since this

is a *rural* employment guarantee, urban workers cannot participate in it.

This concludes our discussion of the features of the rural-only and the rural-urban models to follow. The search strategies of the workers, the equilibrium allocations of workers to the different jobs and to unemployment, and the impact of an employment guarantee will be discussed and developed in detail in Chapters 5 and 6.

4.4 Conclusion

In this chapter we have attempted to lay the ground work for the theoretical models that follow in Chapters 5 and 6. We began by outlining certain features of Indian rural labor markets and of the NREGA that will inform our models. Specifically, we will model a two-season rural labor market, with the NREGA wage being higher than the equilibrium slack season wage that prevailed before the introduction of the program. In each season farmers hire casual laborers based on their season-specific needs.

We also presented some facts regarding the nature of rural-urban migration in India that suggests that temporary or 'circular' migration is an important and widely prevalent feature. In the model where we link the rural and urban labor markets we allow for return migration in the peak season, and for a free-entry sector in the urban area that provides employment to those rural and urban workers who do not find permanent high-paying employment.

We then discussed the most relevant papers in various different (but related) strands of the literature - empirical models that deal with the impact of the

NREGA on rural labor market and other outcomes, theoretical models of the Indian rural labor market and employment guarantees, empirical studies of the movement of workers from rural to urban areas, and finally, theoretical market level models of rural-urban migration.

Finally we presented the basic features of the rural-only and the rural-urban models of Chapter 5 and 6. The next two chapters will take these features as given, and develop the analysis of the welfare impacts of the introduction of an employment guarantee act.

CHAPTER 5

FOR BETTER OR FOR WORSE? THE EFFECTS OF AN EMPLOYMENT GUARANTEE IN A SEASONAL AGRICULTURAL MARKET

This paper develops a theoretical framework in order to study the effects of the introduction of a rural employment guarantee on labor market outcomes in rural India. The rural employment guarantee has features of the National Rural Employment Guarantee Act (NREGA) that was introduced in India in 2006. The purpose of the Act was to improve the livelihoods of rural workers, to create useful assets like roads and canals, and to increase laborer incomes in the slack agricultural season when work is hard to come by. While the model allows for the possibility of workers being made better off with the introduction of the NREGA, it also delivers the interesting and somewhat counter-intuitive result that the introduction of such an Act could potentially have *negative* consequences on the welfare of workers.

We model a two-period seasonal agricultural market with a slack season followed by a peak season, and inter-temporal spillovers between the two periods. There is no urban sector in this model. The spillovers are of the productivity kind, as the amount of labor hired in the slack season to till the land or weed the fields affects the size and the quality of the peak season crop, and hence the marginal productivity of the peak season labor employed. A rural employment guarantee act that provides a high wage in the slack season attracts workers away from agricultural work. Since a smaller number of slack season workers means that the productivity of the peak season workers is reduced, the rural employment guarantee affects peak season labor demand, and peak season agricultural wage. We show that it is possible for the introduction of an em-

ployment guarantee to result in a fall in the peak season wage that reduces the overall annual welfare of the rural workers.

The basic features of this model were described in detail in subsection 4.3.1 of Chapter 4. For the rest of this chapter we will take those features as given. We should mention here that we develop and evaluate the effect of this employment guarantee in a partial equilibrium setup. We abstract away from the production side of the market, and assume that the relative prices of the agricultural and non-agricultural goods being produced in the economy do not change. One way to think of this is to model India as a small open economy, that takes the world prices for these goods as given. Secondly, we do not discuss the financing of the employment guarantee act. Currently the NREGA costs about 0.5% of India's GDP, which is only about 10% of India's government budget deficit, which was estimated at 4.5% of GDP in 2014. We can make the assumption that the employment guarantee is funded entirely by deficit financing.

The remaining sections of this paper are organized as follows: Section 5.1 analyses the pre-EGA model. Section 5.2 introduces a guarantee similar to the NREGA, and Section 5.3 provides a numerical example of the welfare effects of this program. Section 5.4 discusses some simple comparative static exercises, and then Section 5.5 concludes.

5.1 The Model

There are two main sets of actors in this agricultural market, the farmers and the laborers. In the following subsections we outline how we model them, and what their respective objective functions are.

5.1.1 Farmers

The first set of actors are the farmers, who own a certain amount of land that they grow their crops on. Farmers may or may not work on their own land, but regardless of this, they need to supplement family labor with hired labor, particularly in the peak season when tasks are many and timely completion of them is of the utmost importance. There are a large number of farmers, denoted by N.

The farmers hire labor on a season by season basis, not for the whole year at a time. The amount of labor the farmers choose to demand at each wage rate in each season is determined by maximizing their profit taking the wage rates as given, as detailed below. We assume here that all farmers are identical in terms of their land-holdings, availability of family labor, and their production functions. We also assume that farmers form rational expectations of next period's wages.

Since the amount of land held is assumed to be a fixed input it will not be explicitly mentioned henceforth, but one should keep in mind that all production functions mentioned below have both labor and land as inputs.[1]

5.1.2 Casual Laborers

The second set of actors are the casual laborers, who do not own any land of their own. Throughout this paper we will use the words *casual laborers* and *workers* interchangeably. There are a large number of workers relative to the

[1]The buying and selling of land is neither easy nor common in rural parts of India, so most landholdings are inherited.

number of farmers. Let us denote the number of workers by $L >> N$. We assume a certain degree of myopia on the part of the workers, in that their expectations of the next period's wage is equal to the wage that prevailed in the last period of the same type - i.e., they expect the peak season wage in the next period this year to be equal to the peak season wage last year.

The amount of labor each worker has available to supply in each season of the year is normalized to one. Workers can be hired by the farmers to perform agricultural tasks, but if they do not find work within the village they remain unemployed and engage in the next best use of their time. Workers are utility-maximizers, so faced with several economic opportunities they will supply their labor to that opportunity that provides them with the maximum amount of utility. We assume here that all the workers have identical reservation utilities of $\underline{u}$, representing either their valuation of leisure or of some other form of work that provides income (fishing, hunting, weaving baskets and so on).

Let $u(y)$ be a worker's per period utility as a function of his full income y in that period. The quantity y then includes not only wage income but also the valuation of each worker's leisure at his reservation wage.

Assumption 1. *The workers' utility functions satisfy the following conditions:*

1. *Utility is increasing in income:* $u'(y) > 0 \; \forall y > 0$.
2. *Utility is concave:* $u''() \leq 0 \; \forall y \geq 0$.

Let $\underline{w}$ be the wage rate at which if a worker supplies his entire one unit of labor to the farmer, he will receive utility of $\underline{u}$. In other words $u(\underline{w}) = \underline{u}$. Then $\underline{w}$ is each worker's reservation wage, i.e. the lowest wage the farmer can pay

a worker and still induce the worker to supply his labor to private agricultural work.

5.1.3 The Agents' Problems

In this section we will describe the maximization problems faced by the farmer and the workers in this market. Throughout we will use h to denote individual worker labor supply choices in each season, S to denote market labor supply, and L to denote the farmer's labor demand in each season. Subscripts s and p will denote the slack and the peak seasons respectively. For example, S_s and S_p stand for the market labor supply in the slack and the peak period. Actual equilibrium employment will be determined by the intersection of labor demand and supply, or by the short side of the market in case there is no intersection. We denote actual employment by E, with the same subscripts for the slack and peak seasons.

Farmers

At the start of the slack season, a farmer chooses the amount of slack and peak season labor to demand, denoted L_s and L_p respectively, taking as given the slack season wage w_s and his expectation of the wage that will prevail in the peak season, w_p^e. Since there is no uncertainty in this model, under the assumption of rational expectations the farmer's expectation of the peak season wage will be equal to the prevailing peak season wage, so we can replace w_p^e with w_p henceforth. The price of agricultural output produced in the peak season is given by p. Without loss of generality, we can normalize the output price p to

be 1, so that all wages are expressed in real terms. The farmer's annual profit maximization problem can then be written as:

$$\max_{L_s, L_p \geq 0} f(L_s, L_p) - w_s L_s - w_p L_p,$$

where f is the farmer's production function.

Assumption 2. *The production function $f(L_s, L_p)$ satisfies the following conditions:*

1. *Without any peak season labor input the agricultural output is zero, but even without slack season labor input, output can be produced so long as there are a strictly positive number of peak season workers. In mathematical terms, $f(L_s, 0) = 0 \; \forall L_s$, and $f(0, L_p) > 0$ for $L_p > 0$.*
2. *Output is strictly increasing in both arguments so long as peak season employment is not zero. Mathematically, $f_1(L_s, L_p) > 0, f_2(L_s, L_p) > 0 \; \forall L_p > 0$.*
3. *Slack and peak season labor inputs are complements to each other, which means that an increase in the amount of one input used increases the marginal productivity of the other input. Mathematically, $f_{12}(L_s, L_p) > 0 \; \forall L_p > 0$.*
4. *Lastly, the production function is strictly concave. This means that $f_{11}(L_s, L_p), f_{22}(L_s, L_p) < 0 \; \forall L_p > 0$, and $f_{11}(L_s, L_p) f_{22}(L_s, L_p) - f_{12}(L_s, L_p)^2 > 0$.*

The above problem can be thought of in the following intuitive manner. Labor hired in the slack season is used in tasks like weeding, irrigating, applying pesticides and so on. The greater the amount of labor that is spent in these tasks, the healthier the crop is and hence the larger the harvest. A plentiful harvest in turn requires more peak season laborers, as there are more fruits to be picked or bushels of corn to be gathered, more people needed to stack and store the harvest and to transport it to the nearby markets, and so on. Without workers

engaged in conducting the peak season activities, the fruits or grains grown rot in the fields, yielding no revenue to the farmer.

On the other hand, there will still be some output even if no workers are employed in the slack season. For example, rain could substitute for regular irrigation, enabling crops to grow despite being neglected during the slack season. However the *size* of the harvest will be smaller - perhaps because some part of it was consumed by pests, or perhaps because a poor monsoon meant that some of the more water-dependent crops did not survive to maturity. A smaller harvest in turn requires fewer peak season laborers.

Since the production function is concave, the farmer's problem has a unique solution for every combination of wages in the slack and the peak seasons. In addition we will be looking at a competitive equilibrium in this market. Thus it will suffice to look only at the problem of one single 'representative' producer.[2]

The first order condition for this problem gives us

$$\nabla f(L_s, L_p) = w, \tag{5.1}$$

where $\nabla f(L_s, L_p)$ is the gradient vector of the function f and w is the vector of input prices. These two first order conditions define the labor demanded in the slack and the peak season as functions of the two input prices, w_s and w_p. Under the assumptions stated above, we can derive that

[2]Under the conditions imposed above on the production function and the competitive nature of the market we can treat this problem as being one of a single representative agent. See (Mas-Colell et al., 1995), Chapter 5 for theory of aggregation of the firm.

$$\frac{\partial L_k}{\partial w_k} < 0 \text{ for } k = s, p,$$

and

$$\frac{\partial L_k}{\partial w_j} < 0 \text{ for } k, j = s, p; j \neq k.$$

In other words, an increase in the price of an input reduces the demand for that input. Since a reduction in the amount of that input also reduces the marginal product of the second input through our assumptions on the production function, the complementarity works to also reduce the demand for the second input at the same time.

Laborers

The laborers maximize their utility by choosing how much labor to supply to the farmer. This clearly depends on the wage being offered in that particular season. A given laborer chooses the amount of labor to devote to agricultural work and the amount to devote to other pursuits (in this case, enjoying leisure) in each season to solve the following problem:

$$\max_{h_k \geq 0} \sum_{k=s,p} [u(w_k h_k + \underline{w}(1 - h_k))],$$

where h_k is the amount of labor a worker supplies to the farmer in period $k = s, p$. We are assuming additive separability in each worker's utility function. Let us assume that in the absence of an employment guarantee each worker supplies his one unit of labor to the farmer if the wage in the agricultural market is equal to the worker's reservation wage. This is just a tie-breaking assumption.

The way we have written it each worker's problem can be analyzed on an annual basis. Under this assumption, the individual worker's labor supply to the representative farmer can be depicted as in Figure 5.1.

5.1.4 Equilibrium

We assume a competitive agricultural labor market in that no individual farmer or laborer can influence the price of labor in the market. Each farmer chooses the amount of labor to hire in each period to maximize his profits, taking wages as given. The workers choose how much labor to supply to the farmer in each season, also taking wage rates as given. The wage rates themselves adjust to equate labor demand and supply. The amount of labor employed in each season is denoted by E_k for $k = s, p$. Full employment in the peak season combined with less than full employment in the slack season implies that the equilibrium vector of wages and employment $((w_s^*, w_p^*), (E_s^*, E_p^*))$ would be such that

1. $L_s^*(\underline{w}, w_p^*) < L$, i.e even at the lowest possible wage the farmers can pay the workers and still induce them to work, they do not wish to hire all the laborers who are willing to supply their labor. Employment in the slack season is then determined by the short side of the market, the demand side, $E_s^* = L_s^* < L$.

2. $L_p^*(\underline{w}, w_p^*) = L$, i.e. in the peak season the farmer wants to hire all the labor that is available to him. Employment in the peak season is then given by $E_p^* = L$. The wage in the peak season adjusts to equate demand and supply, $w_p^* = f_2(L_s^*, L)$.

These two assumptions are depicted in Figures 5.2 and 5.3. The way we have depicted it, the peak season wage rate is greater than the workers' reservation wage of $\underline{w}$.

5.2 Introducing the Rural Employment Guarantee Act

This model now encompasses all the stylized features of a typical Indian rural labor market without an employment guarantee in place. In the slack season the wage rate is low, and there is less than full employment in agriculture as there simply is not enough demand for agricultural workers from the farmers. However in the peak period the demand is higher, there is full employment, and peak season wages are higher as well.

We are now ready to see what happens when an employment guarantee scheme is introduced into such a market. Workers' welfare is affected positively by the availability of NREGA work and the income derived therefrom, but affected negatively by the labor market effects of the reduced peak season labor demand brought about by reduced slack season agricultural employment. We shall show that under certain conditions an employment guarantee might actually lower workers' two-period welfare through the mechanism of intertemporal spillovers.

For simplicity, assume that the workers can only avail of this guarantee during the slack agricultural season, as the government wishes to interfere as little as possible with the agricultural market. Each worker can avail of $n < 1$ units of time on the rural employment guarantee works, at an exogenously given government wage of w_n (here n stands for NREGA). In line with the stylized facts

described earlier, we assume the following:

Assumption 3. *The NREGA wage is higher than the market clearing wage in the slack season in the absence of the scheme, i.e.* $w_n > w_s^* = \underline{w}$.

If instead we had that $w_n \leq \underline{w}$, then it should be clear that the rural employment guarantee scheme would have no effect on the labor market whatsoever. Workers would continue to supply as much labor as the farmers demanded at their reservation wage, farmers would still be able to hire all the labor they needed to complete the slack season tasks, and peak period labor demand would remain unaffected. In light of empirical observations that the introduction of the scheme attracted workers *away* from agricultural work, the scenario just described seems unlikely to have occurred. Note that we are not making any assumption about the relationship of the rural employment guarantee wage to the peak season wage; it will not matter in this model since we have assumed that the guarantee is available only in the slack season.

Now we need to say something about the timing of events in the slack season. At the start of the slack season, the government-stipulated employment scheme wage is announced, and workers first decide how many of their labor hours to supply to the scheme. Depending on how much they decide to supply, the farmer knows how much labor (or more correctly, how many units of labor time) are available for cultivation on the farm. The number of hours the guarantee provides then determines whether the agricultural labor markets are affected by the introduction of this scheme or not.

5.2.1 The Farmer's Problem

The farmer's problem does not change with the introduction of the NREGA. He still chooses the amount of labor to demand in each season in order to maximize his profit, taking the wage rate prevailing in that season as given.

5.2.2 The Worker's Problem

The worker's problem *does* change with the introduction of the guarantee, and so the equilibrium wage rate and the amount of labor employed in agriculture could also change, as described below. Without the rural employment guarantee, the typical worker had only one choice to make - how much labor to supply to agriculture in each season. The rest of his unit of time he spent in leisure, earning his reservation wage. Now he is faced with two choices in the slack season. He has to decide how much labor to supply to the rural employment guarantee, and how much to work on the farmer's land. He allocates his labor to that economic opportunity that provides him with the maximum income, and therefore utility. Let us assume here that faced with an option of working for the NREGA or working for the farmer at the same wage, he always chooses to work for the farmer. Similarly working for either the farmer or the NREGA is chosen over leisure if the wage he is paid for either type of work is equal to his reservation wage.[3]

The main assumption we are making in this setup is that the worker's labor supply choice in the slack season does not depend on his expectation of the wages that will prevail in the peak period to follow, because he assumes that the

[3]These are tie-breaking assumptions made only to simplify the analysis.

wage will be the same as in the peak season of the previous year. In other words, though we are modeling the farmers as completely rational, we are assuming a degree of myopia on the part of the workers. In justification it seems reasonable to think of these casual laborers as being concerned simply with their employment for this period and not thinking of how much the wages will change in the next season as a result of their collective choices in the current season.

To begin with we study the slack season. It should be clear that the farmer will never raise the agricultural wage so high that it exceeds the wage on the rural employment guarantee, as once the two wages are equal he can employ as many laborers as he chooses. Each worker chooses labor hours to allocate to the NREGA (denoted by h_n) and the amount to devote to agriculture (still denoted as h_s) in order to maximize his utility. This slack season problem can be written as:

$$\max_{0 \leq h_n, h_s \leq 1} [u(w_n h_n + w_s h_s + \underline{w}(1 - h_n - h_s))].$$

By assumption the wage the workers are being offered on the rural employment guarantee is higher than their reservation wage (the prevailing equilibrium slack season wage) and so the workers choose to allocate n units of their time to the guarantee. This leaves a maximum of $L(1 - n)$ units of labor for the farmer to use, unless he bids the wages up to w_n.

5.2.3 Equilibrium

There are now two cases to consider, depending on the number of hours the guarantee provides, n.

n is 'small'

The first case we consider is that where n is 'small', where 'small' means that even after the workers allocate n units of their time to the scheme, the number of labor hours available to the farmer is greater than the equilibrium number of labor units he employed in the market *without* the rural employment guarantee. In other words, $L(1-n) > L_s^*$. In this case, the rural employment guarantee clearly has an overall positive effect on worker welfare, and there is no change in the equilibrium levels of employment and wages. Since the farmer can employ as many workers as he needs at wage $\underline{w}$, equilibrium labor employment does not change ($E_s = L_s^*$), and neither does the slack season agricultural wage. This means that there is also no change in the peak season problem.

The slack season scenario is depicted in Figure 5.4. All workers (whether employed in agriculture in the slack season or not) receive wages w_n for n units of their time and $\underline{w}$ for the remaining $(1-n)$ units of time, and the same peak season wage as before, and are thus unambiguously better off. The change in worker welfare in the slack season can be written as:

$$[u(w_n n + \underline{w}(1-n)) - u(\underline{w})].$$

Since we have assumed that $w_n > \underline{w}$, this change in welfare is always positive.

However, empirical evidence suggests that in actuality this scenario is unlikely to have occurred. There is considerable evidence that the introduction of the rural employment guarantee led to an increase in the agricultural wage and a decline in private agricultural employment in the slack season. This is therefore the case we consider next.

n is 'large'

In the second case, the guaranteed amount of work on the employment scheme is so generous that after all the workers allocate n units of their time to the scheme the amount of labor hours available to the farmer is less than the equilibrium number of units he would have employed in the absence of the guarantee. This second case is depicted in Figure 5.5 or Figure 5.6. Here based on the original slack season labor demand curve of the farmer, the amount of labor he would have demanded at wage $\underline{w}$ is no longer available after the workers supply labor to the rural employment guarantee. If there were no inter-temporal linkages, then the wage would simply have risen to $\hat{w}$ to equate the amounts of labor demanded and supplied at $L(1-n)$. However with inter-temporal linkages the story is slightly more complicated.

Let us work through the logic in order to see what happens. Recall that the farmers are assumed to have rational expectations about the wage prevailing in the next period. Suppose the wage rose to $\hat{w}_s$ and the slack season labor employment fell to $L(1-n)$. This would have an effect on the amount of labor hired in the peak period as well, as is obvious from Equation 5.1. Due to the complementarity between slack and peak season labor in the farmer's production function, the peak season labor demand will shift to the left as a result of

the rural employment guarantee, leading to a fall in the peak season wage and possibly a fall in the peak season labor employed. Depending on how much the peak season labor demand curve shifts, therefore, there are two different possibilities or 'subcases' within this case of n being large. These are described below.

Subcase (i): Let us first consider the case where the amount of peak season labor demanded does indeed fall below the entire labor supply as a result of the change in the slack season employment. The farmer anticipates this in the slack season, and knows that the fall in the amount of peak season employment will decrease the marginal value of an extra unit of slack season labor. In other words, the slack season labor demand shifts to the left, causing the slack season wage to fall. Equilibrium is re-attained at the wage and employment combination where the slack season labor demand is optimal given the peak season wage and employment, which in turn is optimal given the slack season wage and employment. The final equilibrium in this case following the introduction of the NREGA is depicted in Figures 5.7 and 5.8.

Comparing the equilibrium with NREGA with the equilibrium without, the amount of labor employed in agriculture in both the slack and the peak seasons has fallen, and the peak season wage has also decreased to the workers' reservation wage. While the workers get the higher employment guarantee wage for a fraction of the slack season units of time, the fall in the peak season wage raises questions about the overall welfare effect of this policy. It is conceivable that in some scenarios the workers are actually made worse off by the introduction of this scheme, when the fall in peak season utility due to a lower peak season wage outweighs the rise in slack season utility due to increased earnings in that

period.

Subcase (ii): Now we consider the case where the equilibrium amount of peak season labor employed by the farmer does *not* change from full employment, L. In other words even after the fall in the amount of slack season labor employment, the farmer demands more labor than is available at the workers' reservation wage. The peak season wage falls to $w_p^{**} = f_2(L(1-n), L) < f_2(L_s^*, L) = w_p^*$. The slack season labor demand curve is defined by $f_1(L_s, L_p) = w_s$. Since the quantity of labor demanded in the peak season remains the same, the only movement is *along* this slack season demand curve to the point where employment is equal to $L(1-n)$ and the equilibrium slack season wage is given by w^{**}. In the new equilibrium, the slack season agricultural wage is higher than before the introduction of the NREGA, slack season agricultural employment is lower than before the introduction of the NREGA, peak season employment is unchanged, but the peak season wage has fallen relative to the pre-EGA model. This subcase is depicted in figures 5.9 and 5.10.

In the case of a large guarantee (n large), are the workers better or worse off following the introduction of the NREGA? The answer to this question depends on the relative magnitudes of the fall in peak season wage and the rise in the slack season wage. In the slack season, workers receive the NREGA wage for n units of time and the higher slack season agricultural wage for the remaining $(1-n)$ units of time. This sum is clearly higher than what they were receiving without the rural employment guarantee. However the wage they receive in the peak season has fallen; in extreme scenarios it is now equal to their reservation wage (Subcase (i)). Thus there are two effects going in opposite directions.

The change in annual worker welfare can be written as the sum of the slack

season utility change and the peak season utility change:

$$[u_s(w_n n + w_s^{**}(1-n)) - u_s(\underline{w})] + [u_p(w_p^{**}) - u_p(w_p^*)].$$

The first of these terms is always positive, while the second is always negative. Which one outweighs the other? There is no clear answer to this question - it depends on the particular slack and peak season labor demand curves. Interestingly, it is possible to construct scenarios where worker welfare actually *declines*, i.e. the fall in each worker's utility because of the decrease in the peak season wage is greater than the increase in his utility from the increased slack season wage:

$$|u_p(w_p^{**}) - u_p(w_p^*)| > [u_s(w_n n + w_s^{**}(1-n)) - u_s(\underline{w})].$$

This possibility has not been considered at all in the recent policy debates about the effect of the NREGA on worker welfare. We now turn to a simple numerical illustration of this welfare effect in the case of large n.

5.3 Numerical Example

We have claimed above that the result on worker welfare in the case where n is large is 'ambiguous' in the sense that it is possible to conceive of situations where worker welfare increases and other situations where it actually declines with the introduction of the rural employment guarantee. In order to illustrate this possibility, let us consider a numerical example that will show that in the

case of n large the effects on worker welfare can go both ways.

Let the production function be given by

$$f(L_s, L_p) = C(A + BL_s)^{\alpha} L_p^{\beta},\ \alpha + \beta < 1,\ \alpha > 0, \beta > 0,\ A, B, C > 0.$$

We can think of A as representing factors that allow for some output even when slack season labor employed is zero, for example, a good rainfall which substitutes for irrigation by slack season workers. B is a measure of the effectiveness of a unit of slack season labor. Parameter C can be thought of as land productivity.

This production function satisfies the conditions in Assumption 2. To see this, observe the following:

1. It is possible to produce output without using any slack season labor. However if no peak season labor is employed output is zero: $f(L_s, 0) = 0\ \forall L_s$, and $f(0, L_p) = CA^{\alpha} L_p^{\beta} > 0\ \forall L_p > 0$.
2. Output is increasing in both slack and peak season labor employed: $f_1(L_s, L_p) = BC\alpha(A + BL_s)^{\alpha-1} L_p^{\beta} > 0$ and $f_2(L_s, L_p) = C\beta(A + BL_s)^{\alpha} L_p^{\beta-1} > 0$ for $L_p > 0$.
3. Slack and peak season labor are complements to one another: $f_{12}(L_s, L_p) = BC\alpha\beta(A + BL_s)^{\alpha-1} L_p^{\beta-1} > 0$ for $L_p > 0$.
4. The production function is strictly concave: $f_{11}(L_s, L_p) = B^2 C\alpha(\alpha - 1)(A + BL_s)^{\alpha-2} L_p^{\beta} < 0$, and $f_{22}(L_s, L_p) = C\beta(\beta - 1)(A + BL_s)^{\alpha} L_p^{\beta-2} < 0$ for $L_p > 0$, and $f_{11}(L_s, L_p) f_{22}(L_s, L_p) - f_{12}(L_s, L_p)^2 = (1 - \alpha - \beta) B^2 C^2 \alpha\beta (A + BL_s)^{2(\alpha-1)} L_p^{2(\beta-1)} > 0$.

The farmer takes the slack and peak season wages as given and demands labor by equating the marginal product of labor to its price. The first order conditions for the farmer's maximization problem without the rural employment guarantee are given by

$$w_s = BC\alpha(A + BL_s)^{\alpha-1}L_p^{\beta} = f_1(L_s, L_p), \tag{5.2}$$

and

$$w_p = C\beta(A + BL_s)^{\alpha}L_p^{\beta-1} = f_2(L_s, L_p). \tag{5.3}$$

With this production function we can write the labor demands as functions of the input prices in the following way:

$$L_p = C^{\frac{1}{1-\alpha-\beta}}\left(\frac{\alpha B}{w_s}\right)^{\frac{\alpha}{1-\alpha-\beta}}\left(\frac{\beta}{w_p}\right)^{\frac{1-\alpha}{1-\alpha-\beta}},$$

and

$$L_s = \frac{1}{B}\left[C^{\frac{1}{1-\alpha-\beta}}\left(\frac{\alpha B}{w_s}\right)^{\frac{1-\beta}{1-\alpha-\beta}}\left(\frac{\beta}{w_p}\right)^{\frac{\beta}{1-\alpha-\beta}} - A\right].$$

It follows from the above expressions that

$$\frac{\partial L_k}{\partial w_j} < 0 \text{ for } k, j = s, p; j \neq k,$$

i.e. that slack and peak period labor are indeed complements.

Let us define a worker's full income as the sum of income he receives from working (either in agriculture or in the rural employment guarantee) and the valuation of his leisure at his reservation wage. We will call this simply income, and denote it by y. Let the workers' utility in each period as a function of that

period's income y be given by

$$u(y) = y^{\gamma}, y \geq 0, \gamma < 1.$$

This utility function is increasing and concave for positive income levels. Without an employment guarantee, each worker's income in any period $k = s, p$ is simply $y_k = w_k s_k + \underline{w}(1 - s_k)$, i.e. the total amount he receives either from his labor income or from leisure. With an employment guarantee in place, his slack season income is given by $y_s = w_n h_n + w_s h_s + \underline{w}(1 - h_n - h_s)$, i.e. the sum of incomes he receives from both the guarantee and agricultural work. The peak period income definition remains the same.

5.3.1 An example in which workers are worse off

To begin with, let us choose the following parameter values:

$$\alpha = 0.4, \beta = 0.46, n = .6, A = 50, B = 0.9, C = 300 \text{ and } \gamma = .9.$$

Let the number of workers be $L = 75$ and each workers' reservation wage be $\underline{w} = 47$. Workers supply labor in order to maximize their annual utility $u(y_s) + u(y_p)$ where y_s is the slack season income, and y_p is the peak period income. The workers will supply all their labor to the labor market for any wage greater than the reservation wage of 47.

As before, let h_s and h_p denote the individual worker's labor supply to the farmer, and S_s and S_p denote the market labor supplies. Let L_s and L_p denote the quantities of labor demanded, and E_s and E_p denote the actual levels of employment at the equilibrium. Throughout, let 0 denote the period without the rural employment guarantee, and 1 denote when the guarantee is present.

We assume full employment in the peak season, so $E_p^0 = L_p^0 = 75$. In the slack season, the quantity of labor demanded, L_s^0, at the reservation wage of 47 and full employment in the peak period can be determined by solving Equation 5.2, the solution of which is

$$\begin{aligned} L_s^0 &= \frac{1}{B}\left[\left(\frac{BC\alpha L_p^{0\beta}}{\underline{w}}\right)^{\frac{1}{1-\alpha}} - A\right] \\ &= \frac{1}{0.9}\left[\left(\frac{0.9 * 300 * 0.475^{0.46}}{47}\right)^{\frac{1}{0.6}} - 50\right] \\ &= 66.20 < 75. \end{aligned}$$

However the market quantity of labor supplied is $S_s^0 = 75$, as all laborers supply their one unit of labor so long as the wage is greater than $\underline{w}$. The short side of the market dictates equilibrium employment, so we have $E_s^0 = 66.20$. The wage in the slack season is driven down to the reservation wage, $w_s^0 = \underline{w} = 47$. Given these values of the parameters and the slack season levels of employment, we can use Equation 5.3 to solve for the peak season wage

$$\begin{aligned} w_p^0 &= C\beta(A + BL_s)^{\alpha} L_p^{\beta-1} \\ &= 300 * 0.46 * (50 + 0.9 * 66.2)^{0.4} * 75^{-.54} \\ &= 87.75. \end{aligned}$$

Each worker's utility is then equal to

$$\begin{aligned} u_0 &= \underline{w}^{\gamma} + (w_p^0)^{\gamma} \\ &= 47^{.9} + 87.75^{.9} \\ &= 88.07. \end{aligned} \tag{5.4}$$

Since we are defining income as also the valuation of leisure, this is the utility *all* workers earn, regardless of whether or not they are employed in agriculture.

Now the rural employment guarantee is introduced at the exogenous wage $w_n = 55 > w_s^0$. Since the size of the rural employment guarantee in hours is $n = .6$, the amount of labor available to the farmer after the introduction of the guarantee is $L(1-n) = 75(1-.6) = 30$, unless the agricultural wage rises till w_n. Accordingly the market labor supply curve shifts as in Figure 5.9. The farmer's labor demands at the new slack and peak season wages are still given by Equations 5.2 and 5.3.

Assuming that the peak season labor equilibrium employment remains at full employment, $E_p^1 = L_p^1 = 75$, the slack season wage that would prevail if the farmers simply hired the remaining laborers, $L(1-n)$, would be

$$\begin{aligned} w_s &= BC\alpha(A + BL(1-n))^{\alpha-1}L_p^{\beta} \\ &= 0.9 * 300 * 0.4 * (50 + 0.9 * 30)^{-0.6}75^{0.46} \\ &= 58.08 > w_n. \end{aligned}$$

This is therefore a situation where the slack season demand curve intersects the horizontal line through $L(1-n)$ at a wage *greater* than the rural employment guarantee wage (see Figure 5.5 for a graphical illustration). However the farmers can hire as many workers as they like at the wage paid by the rural employment guarantee, so they will not pay the workers a wage higher than this. At the rural employment guarantee wage of 55 the farmers will want instead to hire

$$\begin{aligned} L_s^1 &= \frac{1}{B}\left[\left(\frac{BC\alpha L_p^{1\beta}}{w_n}\right)^{\frac{1}{1-\alpha}} - A\right] \\ &= \frac{1}{0.9}\left[\left(\frac{0.9 * 300 * 0.4 * 75^{0.46}}{55}\right)^{\frac{1}{0.6}} - 50\right] \\ &= 38.14 \end{aligned}$$

number of workers. The prevailing slack season wage will simply be equal to

the rural employment guarantee wage at $w_s^1 = 55$. This constitutes a movement *along the same slack season demand curve* to a higher equilibrium wage and lower equilibrium employment than in the baseline case with no employment guarantee.

As a result of the fall in the slack season amount of labor employed, the new peak season wage will fall to

$$\begin{aligned} w_p^1 &= C\beta(A + BL_s)^{\alpha} L_p^{\beta-1} \\ &= 300 * 0.46 * (50 + 0.9 * 38.14)^{0.4} * 75^{-0.54} \\ &= 79.02. \end{aligned}$$

Each worker's slack season income is given by the sum of his earnings from the rural employment guarantee and the agricultural work, $y_s^1 = w_s^1(1-n) + w_n n$.

Each worker's utility is then equal to

$$\begin{aligned} u^1 &= (w_s^1(1-n) + w_n n)^{\gamma} + (w_p^1)^{\gamma} \\ &= 55^{.9} + 79.02^{.9} \\ &= 87.88. \end{aligned} \tag{5.5}$$

Comparing Equation 5.5 to the original utility without the rural employment guarantee (88.07 in Equation 5.4), we can see that in this example that each worker's utility actually *declines* as a result of the introduction of the rural employment guarantee. As can be seen from the example, this is because the fall in the peak season wage is so large that it outweighs the rise in the slack season wage.

5.3.2 An example in which workers are better off

In the second case let us keep all the parameter values the same, except for the value of B, which we change to $B = 0.8$. For completeness, the list of the parameter values we use is

$$\alpha = 0.4, \beta = 0.46, n = .6, A = 50, B = 0.8, C = 300 \text{ and } \gamma = .9.$$

Again, we assume full employment in the peak season, $E_p^0 = L_p^0 = 75$ and the same reservation wage of $\underline{w} = 47$. In the slack season, farmers demand only

$$\begin{aligned} L_s^0 &= \frac{1}{B}\left[\left(\frac{BC\alpha L_p^{0\beta}}{\underline{w}}\right)^{\frac{1}{1-\alpha}} - A\right] \\ &= \frac{1}{0.8}\left[\left(\frac{0.8 * 300 * 0.4 * 75^{0.46}}{47}\right)^{\frac{1}{0.6}} - 50\right] \\ &= 50.06 \end{aligned}$$

number of workers at the reservation wage. The market labor supply is, however, 75. Thus in equilibrium the agricultural employment is determined by the short side of the market, in this case, the demand: $E_s^0 = 50.06$. The slack season wage is driven down to the reservation wage $w_s^0 = \underline{w} = 47$, but the peak season wage is higher at

$$\begin{aligned} w_p^0 &= C\beta(A + BL_s)^{\alpha} L_p^{\beta-1} \\ &= 300 * 0.46 * (50 + 0.8 * 50.06)^{0.4} * 75^{-.54} \\ &= 81.12. \end{aligned}$$

Each worker's utility is given by

$$\begin{aligned} u_0 &= \underline{w}^{\gamma} + (w_p^0)^{\gamma} \\ &= 47^{.9} + 81.12^{.9} \\ &= 84.24. \end{aligned} \tag{5.6}$$

The rural employment guarantee is introduced into this economy, and a worker can avail of $n = .6$ units of time at the same exogenous wage of $w_n = 55$. Now the maximum amount of labor available for the farmer is again 30, unless he pays workers the same wage as the guarantee. The new slack season wage is given by the intersection of the slack season labor demand curve and the new labor supply curve, which has shifted to the left for wages between the reservation wage and the rural employment guarantee wage (see Figure 5.9 for a diagrammatic representation).

The slack season wage if the farmer were to hire only $L(1-n)$ workers would be given by

$$\begin{aligned} w_s &= BC\alpha(A + BL(1-n))^{\alpha-1} L_p^{\beta} \\ &= 0.8 * 300 * 0.4 * (50 + 0.8 * 30)^{-.6} * 75^{0.46} \\ &= 52.87. \end{aligned}$$

With the guarantee the slack season wage has risen from its pre-guarantee level, though not as high as the rural employment guarantee wage, so the slack season employment is simply 30 workers. The new peak season wage can be calculated

from Equation 5.3:

$$\begin{aligned} w_p^1 &= C\beta(A + BL_s)^{\alpha} L_p^{\beta-1} \\ &= 300 * 0.46 * (50 + 0.8 * 30)^{0.4} * 75^{-0.54} \\ &= 74.99. \end{aligned}$$

As before, the peak season wage falls following the introduction of the NREGA.

Each worker's utility after the introduction of the guarantee is given by

$$\begin{aligned} u^1 &= (w_s^1(1-n) + w_n n)^{\gamma} + (w_p^1)^{\gamma} \\ &= (52.8 * 0.4 + 55 * 0.6)^{0.9} + 74.99^{0.9} \\ &= 85.02. \end{aligned} \tag{5.7}$$

Comparing the utility with the guarantee (85.02 in Equation 5.7) to the utility without the guarantee (84.24 in Equation 5.6) we can see that in this example each worker is made *better off* with the introduction of the rural employment guarantee. The only change between the two examples is that the efficiency of slack season labor (as captured by parameter B) has fallen, which dampens the inter-temporal spillover effects on the peak season wage.

5.3.3 Summary

With the help of a numerical example we have shown that it is possible for workers to be made worse off with the introduction of such a scheme, depending on the nature of the production function chosen and the values of parameters like the size of the rural employment guarantee, n, and the efficiency of slack season labor, B.

5.4 Discussion

The counter-intuitive result of this paper - that worker welfare can actually decline with the introduction of a program like the NREGA - has been demonstrated using a realistic production function and specific numerical examples. However it has only been demonstrated 'theoretically', as we have not necessarily chosen plausible values for the parameters of interest, nor have we estimated a production function for agriculture. The results here are therefore intended simply to present one possible scenario where the effects of the rural employment guarantee act on private agricultural wages may move in an unintended direction, thus having adverse welfare effects.

Having made this caveat, it still might be interesting to see how the change in utility responds to changes in the parameter values, and this will be dealt with briefly in this section. In tables 5.1 and 5.2 we vary the values of the efficiency of slack season labor (B) and the size of the guarantee (n) and look at how the change in utility (utility with the NREGA minus utility in the baseline case with no NREGA) is affected by this variation. Other parameter values are retained at

$$\alpha = 0.4, \beta = 0.46, A = 50, C = 300 \text{ and } \gamma = .9.$$

Negative values of the change in utility mean that each worker is worse off under the introduction of the NREGA than without.

We can see from the tables that as the value of B and n fall the changes in utility become more positive. In the case of a reduction in n, this makes intuitive sense. The smaller the value of n, the less the effects on the private agricultural market, as workers can spend less time on the rural employment guarantee and so farmers can still hire as many workers as they would want at the reservation

wage.

The reduction in B can also be explained in the following way. B captures the effective units of slack season labor - one physical worker hired is really only worth B, where B is less than 1 in all the examples here. (This is not a restriction, though it might be more realistic if slack season labor performs less important tasks than peak season labor does.) A reduction in B means that slack season labor is worth less and less to the farmer, or equivalently that the link between the amount of peak season labor he demands and the amount of slack season labor he hired is less strong. As a result of the weakening of these inter-temporal spillovers, the effects on peak season wages are muted, and overall welfare does indeed improve with the introduction of the NREGA.

So we can see that the changing of the parameters of the model has results along the lines of what we would expect, given the story we have told about inter-temporal spillovers between the slack and peak seasons.

Table 5.1: Changing the efficiency of slack season labor (B), keeping the size of the rural employment guarantee (n) at 0.6.

B	Δ Utility
1	-0.348
0.9	-0.188
0.8	0.78
0.7	2.766
0.6	2.92

5.4.1 The persistence of negative welfare impacts

We made the assumption earlier while describing the workers' problem that the workers are completely myopic, but the farmers are far-sighted. The farmers

Table 5.2: Changing the size of the rural employment guarantee (n), keeping the efficiency of slack season labor (B) at 0.9.

n	Δ Utility
0.6	-0.188
0.5	-0.188
0.4	0.29
0.3	0.59
0.2	0.66
0.1	0.53

can foresee the effect of the workers' choices on the peak season wage in the next period, and adjust their slack season labor demand accordingly, but the workers are 'taken by surprise' and can be rendered worse off as a result. However this does not mean that our result on worker welfare is necessarily limited to only one period, after which expectations correct themselves.

There could be several reasons why this equilibrium could persist. The first is that workers have adaptive expectations, and so it takes several periods for them to adjust their expectations of the peak season wage to the actual post-EGA level. The second is that the workers do not take the actions of the other workers into account when they make their decisions, but act in their own individual best interest. Coordination failures amongst workers can lead to the persistence of the equilibrium where workers are worse off. Each worker chooses to work on the EGA work-site in the slack season in order to earn the higher EGA wage, without realizing that all other workers are making the same choice and that these collective decisions will ultimately cause the peak season agricultural wage to fall. Finally, we haven't taken into account any credit market failures. Workers are taken by surprise by the fall in the peak season wage in the first period, and have to borrow in that period in order to finance their consumption. In the slack season of the following year they work longer on the higher-paying

EGA work-sites in order to pay back this debt, which reduces the peak season agricultural wage in the following season, and the cycle continues.

These are just a few of the ways in which this simple static model can be made dynamic, and the negative impact of the EGA on worker utility can be sustained over several periods.

5.5 Conclusion

This paper contributes to the debate about the desirability of an employment guarantee such as the NREGA introduced recently in India by developing a theoretical framework within which is it possible to generate either positive or negative welfare effects for casual laborers. While policy makers might well have limited patience with the complaints of big farmers on rising wages and the reduction in the availability of labor at the previous going agricultural wage, there is also the possibility of potentially harmful effects even on the casual laborers the Act was instituted to help. We do not know whether such an adverse effect is a likely possibility or not. However it is an interesting theoretical result that has not been mentioned in the literature as far as we are aware.

5.6 Figures

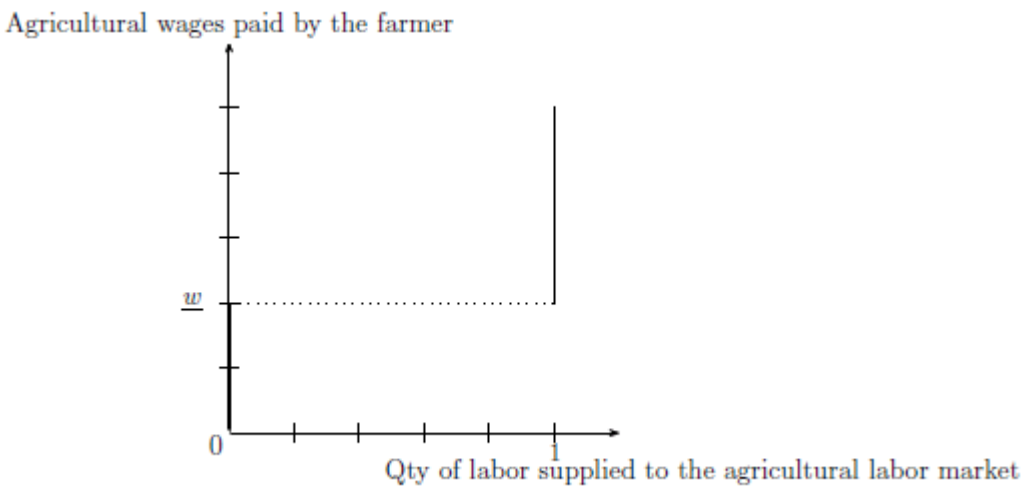

Figure 5.1: An individual worker's labor supply in period k

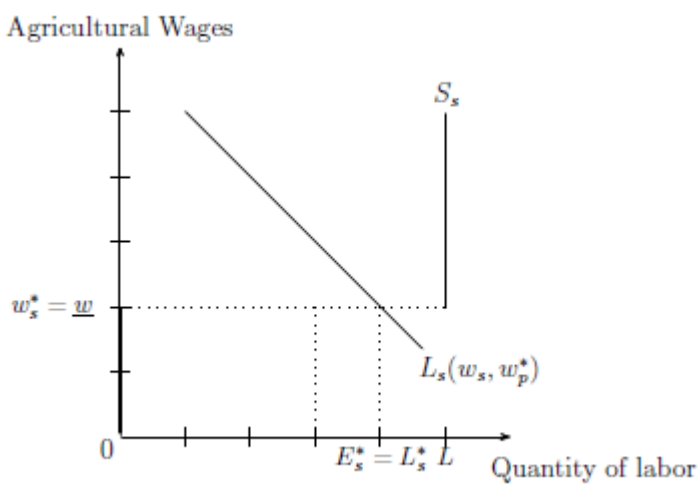

Figure 5.2: Agricultural Market Supply and Demand - slack season

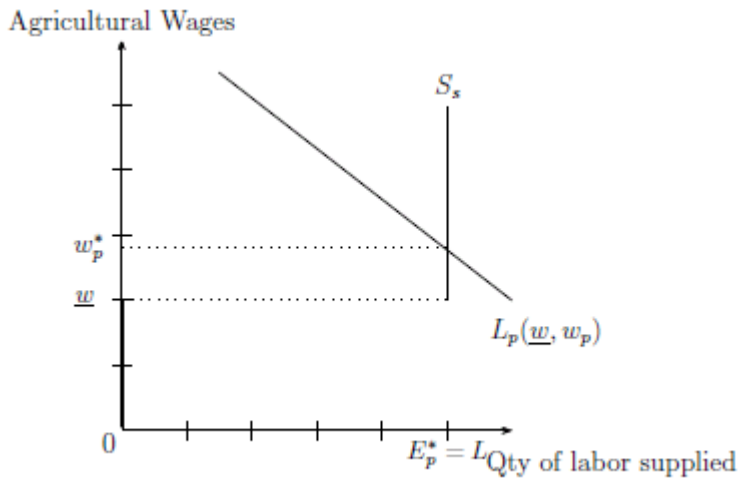

Figure 5.3: Agricultural Market Supply and Demand - Peak Period

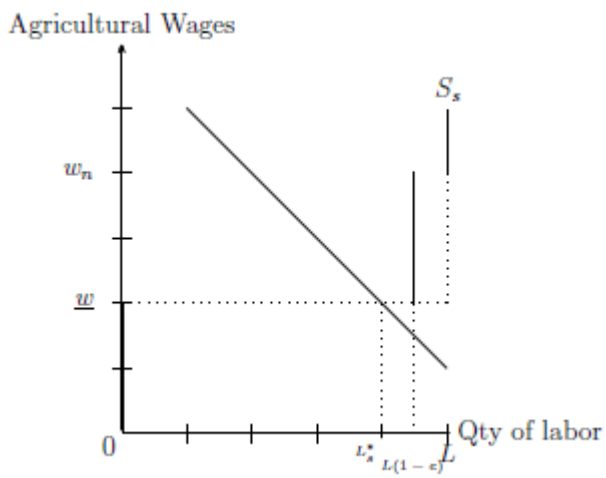

Figure 5.4: Slack Season in the Agricultural Market with Employment Guarantee - *n* small

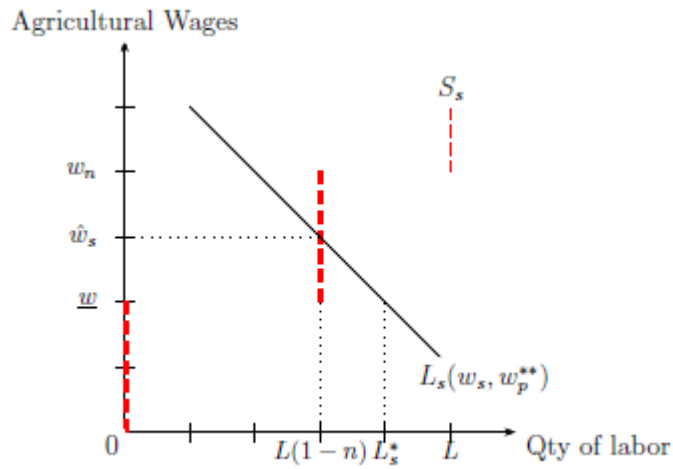

Figure 5.5: Slack Season in the Agricultural Market with Employment Guarantee - *n* large

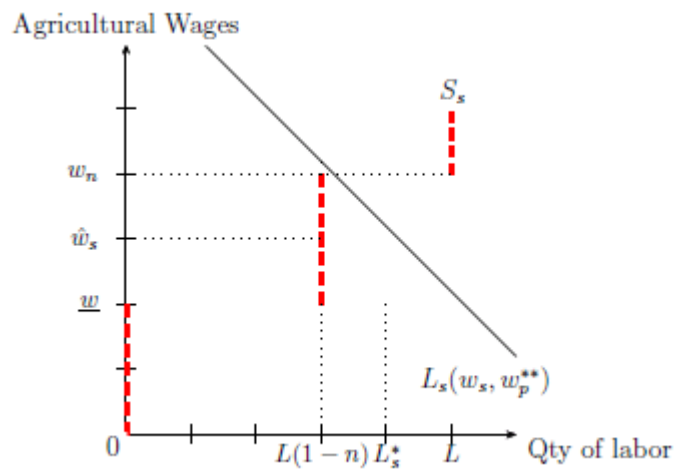

Figure 5.6: Slack Season in the Agricultural Market with Employment Guarantee - *n* large

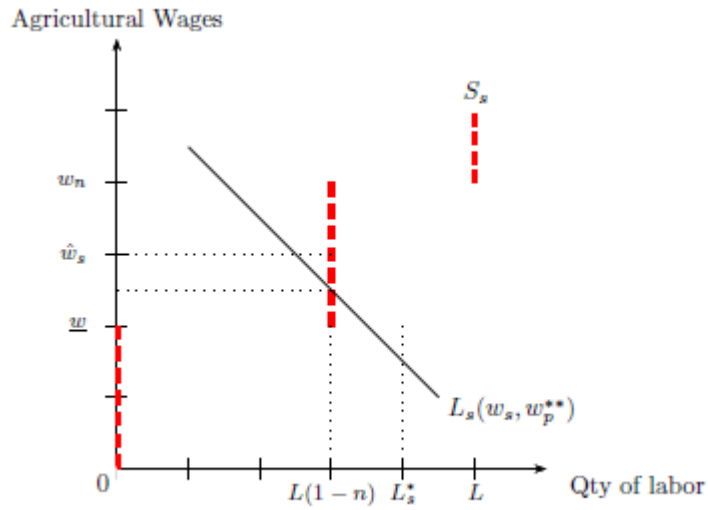

Figure 5.7: *n* large, subcase (i) - Slack Season Final Equilibrium

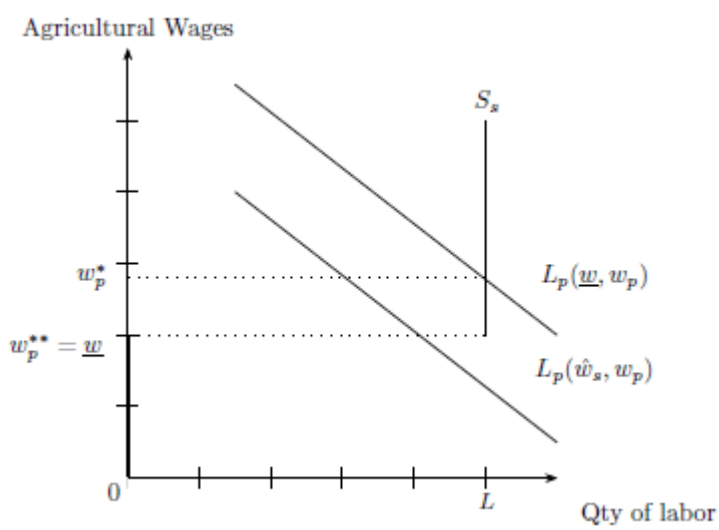

Figure 5.8: *n* large, subcase (i) - Peak Season Final Equilibrium

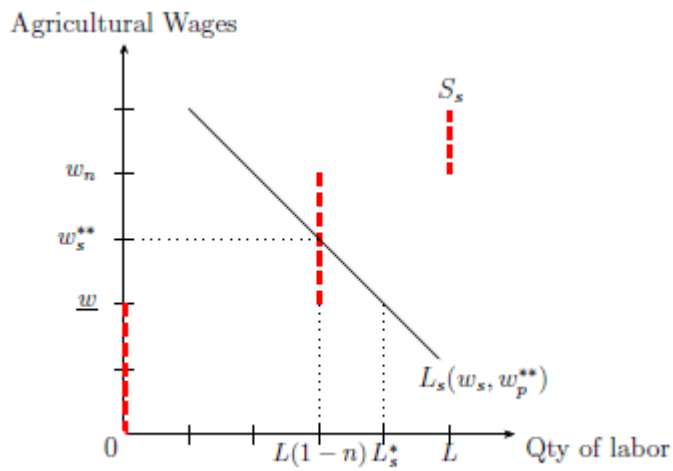

Figure 5.9: *n* large, subcase (ii): Slack Season Final Equilibrium

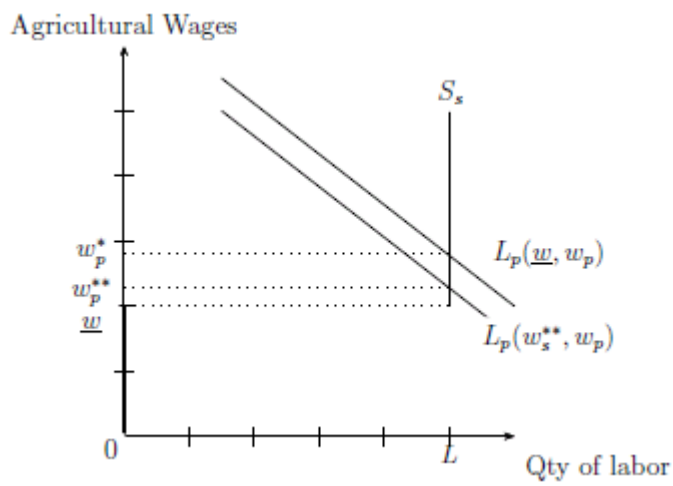

Figure 5.10: *n* large, subcase (ii): Peak Season Final Equilibrium

CHAPTER 6

A MULTI-SECTOR MODEL OF THE INDIAN LABOR MARKET

In Chapter 5 we developed a model of Indian rural labor market which incorporated seasonality in agricultural production and inter-temporal spillovers to deliver some interesting and counter-intuitive insights into the effect of an employment guarantee act on worker welfare. In this chapter we extend that model to include an urban sector and linkages between the rural and urban labor markets. Rural-born workers can now migrate to the urban sector in search of work. We allow for several different types of employment in the urban sector, and for different education levels among the workers. After discussing the search strategies of the workers from different sectors, we outline the ex-ante choices of workers among search strategies and the ex-post allocation of workers among different types of employment and unemployment in both seasons. Then we introduce a rural employment guarantee into this model, and discuss the welfare implications of such a program.

As in the rural-only model of Chapter 5 there are two seasons in this model as well, which are governed by the seasonality of agricultural work in the rural areas. The slack season is the time of year when agricultural work is hard to come by, and many workers migrate to the urban areas in search of work. The peak season is the time of harvest, when there is a higher need for agricultural labor in the farms, and some of those workers who had moved to the urban areas in the slack season migrate back to the rural areas. Rural-born workers who migrate to urban areas and get 'good' jobs in high-paying sectors do not return to the rural area in the peak season.

We find that the introduction of the rural employment guarantee does not unambiguously increase the welfare of workers. The main effect of the introduction of the rural employment guarantee is that it provides an alternative to migration in the slack season. This reduces the number of people who migrate to the urban areas in the slack season, but also reduces the number of rural-born workers who get the high-paying jobs. As a result the number of workers returning to the rural sector to look for jobs in the agricultural sector in the peak season rises, which depresses the peak season agricultural wage.[1] At the same time the guarantee does increase the slack season wage. These two effects work in opposite directions, and this results in an ambiguous effect on worker welfare.

The features of the rural-urban model have been described in detail in Section 4.3.2 in Chapter 4. As a reminder, workers are differentiated along two dimensions, where they are born and what their education levels are. Let the urban-born workers be denoted by capital letters, and the rural-born workers by lower-case letters. The total number of workers in each category in the economy will be denoted by script letters. The total urban population is given then by L, the rural population by l, and the overall number of workers in the economy by $\mathcal{L}$. Workers without education (or with a low level of education) are denoted by a superscript u, and those with education by a superscript e. All the rural-born workers are assumed to be uneducated, while the urban-born workers can be either educated or not educated. Then $L^u + L^e = L$, $l^u = l$, $\mathcal{L}^e = L^e$ and $\mathcal{L}^u = l^u + L^u$.

The next section will discuss in greater detail the various sectors and the wages workers can earn in these sectors in both the rural and urban areas.

[1]This result of a fall in the peak season agricultural wage is similar to the result in the model in Chapter 5, though it operates through migration, and not inter-temporal productivity spillovers.

6.1 Description of the types of employment and search strategies

In this section we will describe the different types of jobs available to urban-born and rural-born workers in each of the two sectors - rural or urban. We will then describe the various search strategies, and how the number of workers adopting each strategy is determined.

6.1.1 The Urban Sector

Free entry sector jobs

Let us start with the simplest sector, that of the urban free-entry sector F. The jobs workers perform when employed in this sector include construction, vegetable vending, clothes washing and ironing services and so on. We assume that the demand for this sector comes from the urban educated workers, and since their size remains constant, so does the total product in this free-entry sector. The greater the number of people who enter this sector, the less work there is for each person, and so it is natural to think of wages in the free-entry sector as being declining average products.

Urban-born workers know the best intersections to sell their vegetables, or the best locations for a clothes ironing service, and hence can earn a higher total product than rural-born workers who move to the urban areas. As a result, the wage paid to a worker in this sector is different depending on whether the worker comes from an urban area or from a rural area.

In keeping with our general naming conventions, let the total product that is to be divided among the urban-born workers be F and among the rural-born workers be f. This total product is then divided up equally among those workers who choose to take up work in this sector. Workers in this sector do not require any educational qualifications.

Office-worker jobs

The second occupation is that of the office workers. These are the lower level clerks, secretaries or messengers in government offices, schools and universities. The wage in this sector is given by w_o, which is exogenously given and fixed (this could be because of minimum wage legislation, unions or some other institutional considerations), and the number of jobs available in this sector is also fixed at E_o. While there is no educational requirement for this sector, those workers with higher education are hired preferentially, and only if there are jobs available once all the educated job-seekers have found employment will the uneducated job-seekers be taken on. The potential supply of work to this sector is thus the entire labor force in the economy. The sector where office-workers work is called the O-sector for simplicity.

Managers' jobs

The third and final occupation is that of the managers. In this sector only those workers who have education will be considered for a job. The wages in this sector are given by $w_m > w_o$ for those workers with high education, and the number of jobs in this sector is fixed at E_m. The potential supply of workers

to this sector is L^e, the total number of educated workers. The sector where managers work is called the M-sector for simplicity.

6.1.2 The Rural Sector

In the pre-EGA model there is only one type of work in the rural sector, which is agricultural work. However agricultural work in India is characteristically seasonal, with certain times of year being more busy than others. Wages and employment levels differ considerably over the course of the year as a result of this seasonality. We model this as a slack season followed by a peak season. The slack season is where wages are lower and so is employment - we can think of this as being the season in which there is only some weeding or fertilizing work to be done. The peak season consists of harvest work as well as other labor-intensive work like planting and transplanting. As a result of the time-sensitive nature of these activities, labor demand and the wages paid to employed workers are both higher in this season than in the slack season.

Let the wage in agriculture in the slack season be given by w_s and in the peak season by w_p. If the number of people employed in season $k = s, p$ is E_k then the agricultural wage in that season is given by the marginal product of labor, i.e.

$$w_k = \frac{\partial F_k(E_k, a_k)}{\partial E_k},$$

where F_k is the agricultural production function for season k, and a_k is a production shifter that also changes with the season under consideration (for example, rainfall, where the effects of rainfall may be larger in the peak season

than in the slack season and so on).

6.1.3 Search strategies

Workers born in different areas and with different levels of education can search for different types of jobs. There is one sector of overlap between workers with high education and those with low education, which is the O-sector, and here the chances of a low education worker getting a job depend on the number of high *and* low education workers seeking a job in this sector. We will study the search strategies of the uneducated rural, uneducated urban and the educated workers separately. In all scenarios the workers are assumed to be risk-neutral income maximizers who simply choose to adopt that strategy that delivers them the highest expected income.

Search strategies for educated urban-born workers

For educated workers there are three search strategies: searching for a job in the M-sector, searching for a job in the O-sector, and the safe strategy of accepting a job in the free-entry sector. However we assume that there is a status cost of an educated worker accepting a job in the free-entry sector. This status cost, s, is high enough to deter educated urban-born workers from accepting employment in this sector even if they are otherwise unemployed. This assumption is reasonable given that educated workers are not found selling tomatoes by the side of the road, or working in construction. It is also one of the reasons why educated unemployment has been on the rise in India. A sufficient condition on s for educated workers to never want to work in the free-entry sector is $s > F$,

though of course this is not a necessary condition.

Suppose educated workers compete for jobs in the O-sector. If there are more jobs available than the supply of educated workers then all workers are hired, if there are fewer jobs than the number of job-seekers then each worker has an equal chance of getting hired to a job. If instead they search for jobs in the M-sector then they earn wage w_m if successful and are unemployed if not. Let us call these two strategies M- and O-search for simplicity.

The number of educated workers seeking jobs in the free entry sector is given by L_f^e, in the O-sector is given by L_o^e, and the number adopting the M-search strategy and looking for the higher paying jobs is given by L_m^e. Then $L_m^e + L_o^e + L_f^e = L^e$.

The timing in this model is as follows: educated workers choose their search strategies at the start of the period. Once the number of educated workers in the O-sector is determined, the number of jobs available for the uneducated workers is known. Uneducated workers then choose their search strategies accordingly.

Expected income to the educated worker from

1. Adopting the O-search strategy: $V_o^e = \eta_e w_o + (1 - \eta_e)0$, where η_e is the probability of an educated worker getting a job in the O-sector, and $1 - \eta_e$ is the probability of an educated worker being unemployed and earning nothing.
2. Adopting the M-search strategy: $V_m^e = \pi w_m + (1 - \pi)0$, where π is the probability of getting a job in the M-sector.
3. Safe search: $V_f^e = w_f = \frac{F}{L_f^e + L_f^u} - s$.

Our condition on s will ensure that no educated urban-born workers ever chooses the safe strategy of looking for a job in the free-entry sector, even if it means being unemployed instead.

The probability of an educated worker getting a job in the O-sector does not depend on the choices of the uneducated workers, but only on the number of jobs available and the total number of educated job seekers. So we have

$$\eta_e = min\left\{\frac{E_o}{L_o^e}, 1\right\}.$$

In other words, if the number of jobs is greater than the number of educated job seekers, then due to preferential hiring all educated workers find a job and so the O-search strategy is a safe strategy. If however the number of jobs available is less than the number of educated job-seekers, then every educated worker has an equal chance of receiving a job, so the probability of getting a job is just the number of jobs divided by the total number of job-seekers.

The probability of getting the M-sector job, π, is also just the total number of jobs available divided by the number of job seekers:

$$\pi = min\left\{\frac{E_m}{L_m^e}, 1\right\}.$$

The expected return equalization between the O and M-sectors for the urban educated workers is shown in Figure 6.1. Along the x-axis we measure the amount of labor employed in each sector. Starting from the origin labeled O_m and moving to the right we measure the number of workers searching in the M-sector. Starting from the origin labeled O_o and moving left we measure the number of workers searching in the O-sector. Along the vertical axis on the left we measure the expected returns to a worker from searching for a job in the M-sector, and similarly the right vertical axis measures the returns to searching in

the O-sector. Since every worker in the O-sector just receives the same wage of w_o that is depicted as a horizontal line. However the returns to a worker searching in the M-sector decline with the number of workers searching for those jobs, as shown by the downward sloping line. Where the two lines intersect is the point of expected income equalization, and determines the number of workers searching for jobs in the two sectors, L_m^e and L_o^e. The number of workers searching for jobs in the M-sector is larger than the number of jobs available, E_m, which is also depicted on the x-axis.

Search strategies for uneducated urban-born workers

Each uneducated urban worker has access to two search strategies - risky and safe. The safe strategy is when these workers look for jobs in the free-entry sector, and are guaranteed employment at the going wage in that sector. If instead they opt for a risky strategy then they compete for jobs in the O-sector, and earn w_o if they are employed, and 0 if not. Let the number of uneducated urban-born workers who search for work in the O-sector be given by L_o^u and the number who search for work in the free entry sector be given by L_f^u. Then $L^u = L_f^u + L_o^u$.

Expected income to the uneducated worker from

1. Adopting the O-search strategy: $V_o^u = \eta_u w_o + (1 - \eta_u)0$, where η_u is the probability of an uneducated worker getting a job in the O-sector.
2. Safe search: $V_f^u = \frac{F}{L_f^e + L_f^u}$.

The probability of an uneducated worker getting an O-sector job is given by the ratio of the number of jobs available to the number of people searching for

those jobs, i.e. by

$$\eta_u = \frac{E_o - L_o^e}{l_o^u + L_o^u}.$$

Given the earlier condition on the number of educated workers, we know that this probability is definitely positive. Equalization of expected income would mean

$$\begin{aligned} \frac{(E_o - L_o^e)w_o}{L_o^u + l_o^u} &= \frac{F}{L_f^e + L_f^u} \\ &= \frac{F}{L^e - L_m^e + L^u - L_o^u} \\ &= \frac{F}{L - L_m^e - L_o^u}. \end{aligned} \tag{6.1}$$

The expected return equalization between the O and F sectors for the urban uneducated workers is shown in Figure 6.2. This figure follows the same logic as the similar graph for urban educated workers, except the two sectors are now the free-entry and the office-workers' sectors. In both sectors the expected return to a worker searching for a job declines with the total number of workers searching in that sector. Again the point of intersection between the lines determines the actual numbers of workers searching in both sectors, L_o^u and L_f^u. Let us denote $E_o - L_o^e$ as J_o^u, the number of jobs in the office-workers sector that are available to the uneducated workers in the economy. The number of jobs available to the urban uneducated workers is then $J_o^u - l_o^u$, also depicted on the x-axis in the figure.

Search strategies for uneducated rural-born workers

Each uneducated rural worker can search for three types of jobs. He can stay in the agricultural sector and earn the going agricultural wage, he can migrate

to the city and search for an O-sector job, or he can migrate to the urban sector and simply accept a free-entry sector job at the average wage. Again let the subscripts denote the type of search strategy that each worker adopts - so the number of workers who choose to work in agriculture is given by l_s^u for the slack season and l_p^u for the peak season, the number of these workers searching for jobs in the O-sector is given by l_o^u and the number of workers searching for jobs in the free-entry sector is l_f^u. Again, $l^u = l_f^u + l_k^u + l_o^u$ for each season $k = s, p$. We assume that migration is costless.

Expected income to the rural uneducated worker from

1. Taking a job in agriculture: $v_k^u = w_k = \frac{\partial F_k(E_k, a_k)}{\partial E_k}\Big|_{(E_k = l_k^u)}$ for $k = s, p$.
2. Taking a job in the free-entry sector: $v_f^u = \frac{f}{l_f^u}$.
3. Searching for a job in the O-sector: $v_o^u = \eta_u w_o + (1 - \eta_u)0$.

Expected income equalization means

$$\frac{\partial F_k(E_k, a_k)}{\partial E_k} = \frac{f}{l_f^u} = \frac{(E_o - L_o^e)w_o}{L_o^u + l_o^u}. \tag{6.2}$$

6.2 The Pre-EGA Model without an Employment Guarantee

Now that we have outlined the search strategies in the pre-EGA model, we can describe the slack and peak season equilibrium outcomes. There is a distinction to be made between ex-ante and ex-post outcomes. Since some of the search strategies are risky, the number of people searching for a particular type of job does not necessarily have to be the number of people who receive that job. We

discuss both the ex-ante number of workers searching for each type of job as well as the ex-post number of workers in each position.

6.2.1 Slack Season Equilibrium

We start by making some assumptions about the relative sizes of the labor pools in order to capture certain stylized facts. There are two features we want to capture. The first is that the jobs in the office-workers' sector are in fact competed over by both educated and uneducated workers, so that this sector is indeed one of overlap. The second feature is that the office-workers' jobs are indeed available only in limited supply, so that not everyone who wants one of these jobs can get one. We formulate both these stylized fact now.

Stylized Fact 1. *The number of educated workers in the economy is 'small', or*

$$L^e < \frac{E_m w_m + E_o w_o}{w_o}.$$

This stylized fact guarantees that the number of educated workers is not so large as for them to occupy all the jobs in the O-sector - as a result there are some of jobs left in that sector for the uneducated workers to compete over as well. This also ensures that educated workers are not engaged in free-entry sector jobs. Given the stylized fact above regarding the size of the educated population, it follows that $\eta_e = 1$. The expected income equalization condition then dictates that

$$\frac{E_m w_m}{L_m^e} = w_o,$$

or

$$L_m^e = \frac{E_m w_m}{w_o} \text{ and } L_o^e = L^e - \frac{E_m w_m}{w_o} < E_o. \tag{6.3}$$

Stylized Fact 2. *The number of jobs in the office-workers' sector is smaller than the number of people who are searching for these jobs, or*

$$E_o < l_o^u + L_o^u + L_f^e,$$

This stylization means that not every uneducated worker who wants a job in the O-sector can find one, so that for these uneducated workers searching in this sector is indeed a risky strategy.

Ex-ante outcomes: How many people search for a job in each occupation?

We start with characterizing the slack season equilibrium. Since there is no seasonality in the wages and number of jobs in the urban sector, nothing changes from season to season for the urban educated workers, and the solution to their problem in both seasons is given by equation 6.3. For completeness, the equations are reproduced here:

$$\begin{aligned} L_m^e &= \frac{E_m w_m}{w_o} \\ L_o^e &= L^e - \frac{E_m w_m}{w_o} < E_o. \end{aligned} \tag{6.4}$$

For the uneducated workers, however, the equilibrium changes depending on the season we consider. In order to be able to solve for the equilibrium, let us give the agricultural production function the following form: $F_k(E_k, a_k) = a_k log E_k$for $k = s, p$. Then for the slack season

$$\frac{\partial F_s(E_s, a_s)}{\partial E_s} = \frac{a_s}{E_s}.$$

From equation 6.2 we then get the following equation that characterizes the interior equilibrium for the rural uneducated workers

$$\frac{a_s}{l_s^u} = \frac{f}{l_f^u} = \frac{(E_o - L_o^e)w_o}{L_o^u + l_o^u}.$$

Since the problem for the educated workers is self-contained, the quantity J_o^u is known. The above equation can be solved to give

$$l_f^u = \frac{f(L_o^u + l^u)}{J_o^u w_o + a_s + f}, \tag{6.5}$$

$$l_s^u = \frac{a_s}{f} l_f^u = \frac{a(L_o^u + l^u)}{J_o^u w_o + a_s + f}, \text{ and} \tag{6.6}$$

$$l_o^u = l^u - l_s^u - l_f^u = \frac{J_o^u w_o l^u - (a_s + f)L_o^u}{J_o^u w_o + a_s + f}. \tag{6.7}$$

In each of the above series of equations, there are two unknowns (L_o^u and the left-hand-side variable) and only one equation to solve for them. In order to obtain the closed form solutions we now consider equation 6.1 for the urban uneducated workers. Rearranging this equation, we get the following solution for L_o^u,

$$L_o^u = \frac{J_o^u w_o L^u - F l_o^u}{J_o^u w_o + F}. \tag{6.8}$$

Combining equations 6.7 and 6.8, we can now solve for the two unknowns L_o^u and l_o^u, which tells us the number of rural and urban uneducated workers who search for the office-worker jobs. The expressions can be derived as

$$L_o^u = \frac{(J_o^u w_o + a_s + f)L^u - F l^u}{J_o^u w_o + a_s + f + F}, \text{ and}$$

$$l_o^u = \frac{(J_o^u w_o + F)l^u - (a_s + f)L^u}{J_o^u w_o + a_s + f + F}.$$

The total number of uneducated workers from either the rural or the urban

sectors who are searching for the O-sector jobs is then given by

$$l_o^u + L_o^u = \frac{J_o^u w_o (L^u + l^u)}{J_o^u w_o + a_s + f + F}$$

Using equation 6.8 we can solve for the number of urban uneducated workers who search for jobs in the free-entry sector

$$L_f^u = L^u - L_o^u = \frac{F(L^u + l^u)}{J_o^u w_o + a_s + f + F}.$$

Now that we have a closed form expression for l_o^u, we can also solve for l_f^u and l_s^u using equations 6.5 and 6.6 respectively. Doing this gives us the following expressions for the number of uneducated rural-born workers searching for agricultural jobs and free-entry sector jobs in the slack season:

$$l_f^u = \frac{f(l^u + L^u)}{J_o^u w_o + a_s + f + F} \text{ and}$$
$$l_s^u = \frac{a_s(l^u + L^u)}{J_o^u w_o + a_s + f + F}.$$

With the last two expressions we have now solved for all the ex-ante variables - L_m^e, L_o^e, L_o^u, L_f^u, l_o^u, l_f^u and l_s^u - in other words, we know exactly how many workers of each education level are adopting each search strategy. However since not every worker who applies for a job in the managers' and office-workers' occupations is employed, the ex-post outcomes of how many workers actually end up employed in these occupations (and hence the incomes of the workers) still needs to be analyzed. A similar distinction between ex-ante choices and ex-post outcomes can be found in (Harris and Todaro, 1970).

Ex-post outcomes: How many people end up in each employment state?

The number of jobs in the managers' and office-workers' sectors are fixed at E_m and E_o respectively. However more than this number of workers search for these jobs, precisely because workers are risk-neutral income maximizers, and the wages in these jobs are 'high'. As a result after making the decision of what jobs to search for, workers may or may not be employed. We make the assumption that if a worker does not manage to get a job in the sector he applies for, he remains unemployed for the rest of that period (but can search for a job again in the next period if he so chooses to).

The total number of people searching for jobs in the O-sector is given by $l_o^u + L_o^u + L_o^e$. We have already made the assumption that the number of educated workers searching for these jobs is less than the number of such jobs available, and so all L_o^e workers are employed in the O-sector. Once these workers have been preferentially hired, the number of jobs left for the other uneducated workers from the rural and urban areas is $E_o - L_o^e$. Let us assume that workers from the rural and the urban areas receive these jobs in proportion to their representation in the searching population. Then we have that the probability of any given rural or urban worker receiving an O job is given by the following expressions:

$$\text{Rural uneducated worker: } \frac{l_o^u}{l_o^u + L_o^u}$$
$$\text{Urban uneducated worker: } \frac{L_o^u}{l_o^u + L_o^u}.$$

Then the fraction of the available jobs in the O-sector that go to the rural-born

or urban-born workers is given by

$$O\text{-sector jobs to rural-born workers: } \frac{l_o^u}{l_o^u + L_o^u}(E_o - L_o^e)$$
$$O\text{-sector jobs to urban-born workers: } \frac{L_o^u}{l_o^u + L_o^u}(E_o - L_o^e).$$

Using the above expressions, the number of unemployed rural-born workers, u^u, is given by

$$\begin{aligned} u^u &= \max(0, l_o^u - \frac{l_o^u}{l_o^u + L_o^u}(E_o - L_o^e)) = \max\left(0, \left[\frac{l_o^u + L_o^u - J_o^u}{l_o^u + L_o^u}\right] l_o^u\right) \\ &= \max\left(0, [(J_o^u w_o + F)l^u - (a_s + f)L^u]\left[\frac{1}{J_o^u w_o + a_s + f + F} - \frac{1}{w_o(L^u + l^u)}\right]\right), \end{aligned}$$

and the number of unemployed uneducated urban-born workers, U^u, is given by

$$\begin{aligned} U^u &= \max(0, L_o^u - \frac{L_o^u}{l_o^u + L_o^u}(E_o - L_o^e)) = \max\left(0, \frac{l_o^u + L_o^u - J_o^u}{l_o^u + L_o^u} L_o^u\right) \\ &= \left(0, [(J_o^u w_o + a_s + f)L^u - Fl^u]\left[\frac{1}{J_o^u w_o + a_s + f + F} - \frac{1}{w_o(L^u + l^u)}\right]\right). \end{aligned}$$

Finally we can use equation 6.4 to show that the number of unemployed educated workers is simply given by

$$U^e = L_m^e - E_m = E_m \frac{w_m - w_o}{w_o}.$$

Now we have completed the discussion of the slack season equilibrium, both in terms of the number of workers that adopt each strategy and the number of workers who actually end up employed in the various jobs, or who end up unemployed. The summary of this is given in Table 6.1.

6.2.2 Peak Season Equilibrium

The analysis of the peak season proceeds just as in the case of the slack season. The one change that does have effects is the change in the agricultural production function. Now the production function is given by

$$\frac{\partial F_p(E_p, a_p)}{\partial E_p} = \frac{a_p}{E_p}.$$

Stylized Fact 3. *We assume that the marginal product of a unit of labor in the peak season is greater than it is in the slack season, i.e. that $a_p > a_s$.*

This makes intuitive sense as the peak season is the period when time-sensitive tasks like harvesting, planting and transplanting take place. Without these activities occurring in a timely and organized fashion the quality of the harvest and hence the payoff to the farmers is greatly reduced. Slack season activities like weeding or watering the crops regularly also improve the quality of the harvest, but with lower returns as they can be partially compensated for by natural elements like rainfall.

What the assumption that the marginal product of a unit of labor in the peak season is greater than it is in the slack season guarantees is that

$$\left.\frac{\partial F_p(E_p, a_p)}{\partial E_p}\right|_{E_p = l_s^u} > \left.\frac{\partial F_s(E_s, a_s)}{\partial E_s}\right|_{E_s = l_s^u} = \frac{f}{l_f^u}$$

In other words, if the distribution of workers were to remain the same in the peak season as in the slack season, then the rural-born workers in agriculture would be earning more in the peak season than the rural-born workers in the free-entry sector. Since the number of jobs in the O-sector does not change, this means there is some reshuffling of workers between the free entry sector and

the agricultural sector in order to bring the expected wages back in line with one another. It also tells us that the *direction* of the adjustment is such that that there will be some *reverse-migration* of workers from the urban areas to the rural areas in the peak season.

We assume that those uneducated urban and rural-born workers who were unemployed can re-enter the market and search for new jobs in the peak season. Those rural-born workers who receive O-sector jobs in the slack season do not give up these jobs and return to the rural sector, so there is no additional migration from the rural sector to the urban sector in search of positions in the office-workers' sector. Instead the rural-born workers who were unemployed in the slack season (because their search for a job in the O-sector was unsuccessful) search for positions in peak season agriculture and the free-entry sector. The urban unemployed workers can only get jobs in the free-entry sector.

Urban educated workers who were unemployed in the slack season could potentially find jobs in the urban free-entry sector, but the status cost s is so high that they do not choose to do so, even if it means remaining unemployed for an additional season.

Ex-ante outcomes

The overall 'active' rural population in the peak season is u^u. These are the rural-born workers who were unemployed in the slack season, and are actively searching for jobs in the peak season. Since office-worker and manager jobs have already been filled, these rural-born unemployed workers search for work either in agriculture or in the urban free-entry sector. Let the total number of

people working in agriculture in the peak season be denoted l_p^u. Then the 'additional' people employed in agriculture in the peak season over the slack season is given by $l_p^u - l_s^u$. This is the portion of the active population that enters agriculture. The remaining members of the active population enter the free-entry sector. Let the new free-entry sector rural population be denoted by $l_{f'}^u$. It follows then that:

$$l_f^u + u^u + l_s^u = l_{f'}^u + l_p^u.$$

Now the number of workers who are engaged in agriculture in the peak season is determined by the equation of the expected wages in the free-entry sector and the agricultural sector and can be solved from the following equation:

$$\begin{aligned}\frac{a_p}{l_p^u} &= \frac{f}{l_{f'}^u} \\ &= \frac{f}{l_f^u + u^u + l_s^u - l_p^u}.\end{aligned}$$

Since all the other quantities other than l_p^u are known, this equation can be solved for the unknown, and the solution is

$$l_p^u = \frac{a_p}{a_p + f}(l_f^u + u^u + l_s^u) = \frac{a_p}{a_p + f}(l_f^u + u^u) + \frac{a_p}{a_p + f}(l_s^u). \tag{6.9}$$

We can easily show that indeed the number of workers in peak season agriculture is larger than the number of workers in slack season agriculture, i.e. that $l_p^u > l_s^u$. In order for this result to hold, we would need:

$$\begin{aligned}l_p^u &= \frac{a_p}{a_p + f}(l_f^u + u^u) + \frac{a_p}{a_p + f}(l_s^u) > l_s^u \\ &\implies \frac{a_p}{a_p + f}(l_f^u + u^u) > \frac{f}{a_p + f}(l_s^u) \\ &\implies l_f^u + u^u > \frac{f}{a_p} l_s^u.\end{aligned}$$

Now from wage equalization between agriculture and the free-entry sector in the slack season and the assumption that $a_p > a_s$ we know that

$$\frac{f}{l_f^u} = \frac{a_s}{l_s^u} < \frac{a_p}{l_s^u}.$$

Rearranging this equation we get the result

$$l_f > \frac{f}{a_p} l_s^u,$$

so it follows that agricultural employment in the peak season is indeed greater than in the slack season.

Using equation 6.9, and substituting the expressions we already have for l_s^u, l_f^u and u^u, we can solve for peak season agricultural employment as

$$l_p^u = \frac{a_p}{a_p + f}\left[l^u + \left(\frac{(a_s + f)L^u - (J_o^u w_o + F)l^u}{w_o(l^u + L^u)}\right)\right],$$

and so the peak season agricultural wage is given by

$$w_p = \frac{a_p}{l_p^u} = \frac{w_o(a_p + f)(l^u + L^u)}{l^u\left[w_o(l^u + L^u) - (J_o^u w_o + F)\right] + (a_s + f)L^u}.$$

The number of rural uneducated workers in the O-sector remains the same as in the slack season, at

$$\frac{(J_o^u w_o + a_s + f)L^u - Fl^u}{w_o(L^u + l^u)}$$

(see Table 6.1).

Since we know that

$$\frac{a_p}{l_p^u} = \frac{f}{l_{f'}^u},$$

the peak season number of rural-born workers in the free-entry sector can be backed out as being

$$l^u_{f'} = \frac{f}{a_p + f}\left[l^u + \left(\frac{(a_s + f)L^u - (J^u_o w_o + F)l^u}{w_o(l^u + L^u)}\right)\right].$$

We have solved for the number of rural-born workers in agriculture, the free-entry sector and the office-workers' sector in the peak season.

For the urban uneducated workers, the only opportunity is to move into the free-entry sector, so the new free-entry sector urban population is given by

$$l^u_{F'} = U^u + L^u_f = \frac{w_o(L^u + l^u) + Fl^u - (J^u_o w_o + a_s + f)L^u}{w_o(L^u + l^u)}.$$

The wage they earn in the peak season is no longer the same as the wage earned by the rural uneducated workers in the free-entry sector, and simply falls as more and more people are absorbed from unemployment into free-entry sector employment. This new free-entry sector peak season wage is given by

$$w_{F'} = \frac{F}{L^u_f + U^u} = \frac{Fw_o(L^u + l^u)}{w_o(L^u + l^u) + Fl^u - (J^u_o w_o + a_s + f)L^u}.$$

The final outcomes in the peak season are summarized in Table 6.2.

The Annual Income Distribution without the Rural Employment Guarantee

Now we have the income earned by the workers in both periods, we can work out the total annual income earned by them, which is essential in order to make any pre- and post-policy intervention welfare comparisons of the kind we would like to. In order to do this, we need an exhaustive list of the types of job combinations a worker can have. Since there is some overlap in the types

of jobs workers from the rural and the urban areas can have, let's denote all rural job types by lower-case letters and urban job types by upper-case letters. So the combination of unemployment in the first period and free-entry sector work in the second would be (u, f) for rural-born workers, and (U, F) for urban-born workers.

The first and most obvious of these are those educated workers who keep their managers' jobs in both periods, we call this combination (M, M). Then those who are employed in the office-worker sector for both periods are in combination (O, O). Since the wages for these occupations are fixed, the annual income calculation is fairly straightforward.

The next set of combinations to consider all earn the same annual income, despite having completely different 'career paths'. These are those rural-born workers who stay in agriculture for both periods and those who stay in the free-entry sector in both periods. Since worker movement equalizes the expected wages in the free-entry and the agricultural sectors, these workers all earn the same amount within a period. These combinations are called (s, p) and (f, f) respectively - with s and p standing for slack and peak agriculture, and f standing for rural free-entry sector employment.

Then there are the urban uneducated workers who stay in the free-entry sector for both periods - these workers earn the annual income $w_s + w_{F'}$, and number L^u_f, and this combination is called (F, F). Then there are the workers who were unemployed in the slack season, and who earned nothing in that period - their annual income is either $w_{F'}$ or w_p depending on whether they were from the urban or the rural sector, and their numbers are U^u and u^u respectively. These combinations are named (U, F) and (u, f). And then finally there are those

workers who were unemployed for both periods - the (U, U) job combination - the educated urban population who did not want a free-entry sector job in the following period either. These number $\frac{E-m(w_m-w_o)}{w_o}$ in total. These numbers and the corresponding annual incomes are summarized in Table 6.3.

6.3 The Model with a Rural Employment Guarantee Act

Now that we have the pre-EGA model in place, we consider the welfare implications of introducing an rural employment guarantee act (EGA) into this economy. The EGA is only available in the slack season, and to rural laborers. It provides employment in the rural areas at the exogenously fixed wage of $\tilde{w}_n$. The introduction of this act in India in 2006 involved the government pegging the wage it paid to laborers to be greater than the going agricultural wage, but the wage was not at the same level as the office jobs in the cities. As a result we make the following assumptions about the guaranteed wage:

Stylized Fact 4. *The wage in the rural employment guarantee is larger than the slack season agricultural wage but lower than the wage workers could earn in the office-workers' sector, i.e.*

1. $\tilde{w}_n > w_s$, *and*
2. $\tilde{w}_n < w_o$.

We denote all quantities specific to the post-EGA equilibrium with tildes, for example the number of rural-born workers seeking work in the office-workers' sector in the slack season when an EGA is in place is given by $\tilde{l}_o^u$.

6.3.1 Slack Season Equilibrium

Search Strategies

There is no change in the problem faced by the urban-born workers, as they continue to have the same work options as before. However the rural-born workers now have an additional choice in the slack season. They can choose to stay in the rural areas and work for either the EGA or for the agricultural work, or they can migrate to the cities in search of work. The payoffs to these search strategies are as follows:

1. Taking a job in agriculture: $v_s^u = \tilde{w}_s = \left.\frac{\partial F_s(E_s, a_s)}{\partial E_s}\right|_{(E_s = \tilde{l}_s^u)}$
2. Taking a job in the free-entry sector: $v_f^u = \frac{f}{\tilde{l}_f^u}$.
3. Searching for a job in the O-sector: $v_o^u = \eta_u w_o + (1 - \eta_u)0$.
4. Taking a job in the EGA, sector N: $v_n^u = \tilde{w}_n$.

The number of workers choosing each of these strategies is given by $\tilde{l}_s^u$, $\tilde{l}_f^u$, $\tilde{l}_o^u$ and $\tilde{l}_n^u$ respectively, and again we have $\tilde{l}_n^u + \tilde{l}_o^u + \tilde{l}_f^u + \tilde{l}_s^u = l^u$.

Ex-ante outcomes

We analyze the case of an interior solution, with the same formulation for the agricultural production function as in the previous discussion. The problem for the urban educated workers remains the same, and so does the solution. As a

result we have

$$\tilde{L}_m^e = \frac{E_m w_m}{w_o} (= L_m^e) \text{ and}$$
$$\tilde{L}_o^e = L^e - \frac{E_m w_m}{w_o} (= L_o^e) < E_o.$$

Since the number of urban educated workers choosing the various strategies hasn't changed, the number of jobs in the office-workers' sector available to the rural and urban uneducated workers, J_o^u, also does not change. To avoid confusion we do not denote this quantity with a tilde.

For rural uneducated workers, the interior solution means that the expected wages from the various different strategies are equalized. This reduces to the following set of equations:

$$\frac{a_s}{\tilde{l}_s^u} = \frac{f}{\tilde{l}_f^u} = \tilde{w}_n = \frac{J_o^u}{\tilde{L}_o^u + \tilde{l}_o^u} w_o. \tag{6.10}$$

As a result of the introduction of the EGA, the agricultural wages in the slack season have to be at least as high as $\tilde{w}_n$. Using 6.10 we can solve for the following:

$$\tilde{l}_f^u = \frac{f}{\tilde{w}_n}$$
$$\tilde{l}_s^u = \frac{a_s}{\tilde{w}_n}$$
$$\tilde{L}_o^u + \tilde{l}_o^u = \frac{J_o^u w_o}{\tilde{w}_n}.$$

For the urban unemployed workers, the interior solution means the equation of the wages from the free-entry sector and the office-workers' jobs, and can be written as

$$\frac{J_o^u w_o}{\tilde{L}_o^u + \tilde{l}_o^u} = \frac{F}{\tilde{L}_f^u} = \frac{F}{L^u - \tilde{L}_o^u},$$

which can be rearranged in order to express $\tilde{L}_o^u$ in terms of $\tilde{l}_o^u$ as

$$\tilde{L}_o^u = \frac{J_o^u w_o L^u - F\tilde{l}_o^u}{F + J_o^u w_o}.$$

Substituting this into the expression for $\tilde{L}_o^u + \tilde{l}_o^u$ we can solve for the following:

$$\begin{aligned}
\tilde{l}_o^u &= \frac{F + J_o^u w_o - \tilde{w}_n L^u}{\tilde{w}_n}, \\
\tilde{L}_f^u &= \frac{F}{\tilde{w}_n}, \\
\tilde{L}_o^u &= L^u - \frac{F}{\tilde{w}_n}, \text{and} \\
\tilde{l}_n^u &= l^u + L^u - \frac{J_o^u w_o + f + a_s + F}{\tilde{w}_n}.
\end{aligned}$$

We have solved for the number of workers adopting each search strategy in the slack season with the introduction of the EGA.

Ex-post outcomes

The number of workers that actually end up in each state of employment is determined after the choice of search strategy has been made, just as in the previous case. Nothing changes from the earlier discussion in the case of urban educated workers - the number of managers' jobs remains at E_m and so the number of unemployed urban educated workers is given by $U^e = L_m^e - E_m = \frac{E_m(w_m - w_o)}{w_o}$. However both the urban uneducated and the rural uneducated workers are affected by the introduction of this policy.

The number of O-sector jobs available to the uneducated workers is given by J_o^u, and the number of workers searching for these jobs is given by $L_o^u + l_o^u = \frac{J_o^u w_o}{\tilde{w}_n}$. We assume once more that the jobs are occupied by the rural and urban-born workers in proportion to their representation in the searching population, which

means that the number of jobs each group gets is given by

$$\text{Rural: } \frac{\tilde{l}_o^u}{\tilde{l}_o^u + \tilde{L}_o^u} J_o^u = \frac{F + J_o^u w_o - \tilde{w}_n L^u}{w_o}, \text{ and}$$
$$\text{Urban: } \frac{\tilde{L}_o^u}{\tilde{l}_o^u + \tilde{L}_o^u} J_o^u = \frac{\tilde{w}_n L^u - F}{w_o}.$$

This then means that the number of unemployed uneducated workers is given by

$$\text{Rural:} \qquad \tilde{u}^u = \tilde{l}_o^u - \frac{\tilde{l}_o^u}{\tilde{l}_o^u + \tilde{L}_o^u} J_o^u = \frac{w_o - \tilde{w}_n}{\tilde{w}_n w_o}(F + J_o^u w_o - \tilde{w}_n L^u)$$
$$\text{Urban:} \qquad \tilde{U}^u = \tilde{L}_o^u - \frac{\tilde{L}_o^u}{\tilde{l}_o^u + \tilde{L}_o^u} J_o^u = \frac{w_o - \tilde{w}_n}{\tilde{w}_n w_o}(L^u \tilde{w}_n - F).$$

We can summarize the above discussion in tabular form as in Table 6.5. The main points of interest are the following:

Firstly, the introduction of the rural employment guarantee scheme does not change the outcomes for the educated urban-born workers at all, and so their incomes remain the same. The reason is that there is no sector of the economy where the educated urban-born workers compete with the rural-born workers for jobs. Since the EGA only affects the search strategies of the rural-born workers, it has no effect on the educated urban-born workers whatsoever.

Secondly, it reduces the number of rural-born workers who work in agriculture in the slack season (from $\frac{a_s}{w_s}$ to $\frac{a_s}{\tilde{w}_n}$) and the number of rural-born workers who enter the free entry sector as well (from $\frac{f}{w_s}$ to $\frac{f}{\tilde{w}_n}$).

It reduces the number of urban uneducated workers who search for free-entry sector jobs (from $\frac{F}{w_s}$ to $\frac{F}{\tilde{w}_n}$), and thus increases the number who search for office-worker jobs ($\tilde{L}_o^u > L_o^u$).

The total number of uneducated workers, rural and urban, that search for office-worker jobs falls with the introduction of the EGA. In the absence of the EGA, this number is given by

$$l_o^u + L_o^u = \frac{J_o^u w_o}{w_s},$$

and with the EGA it is given by

$$\tilde{l}_o^u + \tilde{L}_o^u = \frac{J_o^u w_o}{\tilde{w}_n}.$$

Since $\tilde{w}_n > w_s$, it follows that the right hand side of the second expression is smaller than that of the first expression, and hence that $\tilde{l}_o^u + \tilde{L}_o^u < l_o^u + L_o^u$.

Since $\tilde{l}_o^u + \tilde{L}_o^u$ has fallen but $\tilde{L}_o^u$ has increased from when the EGA was not present, it follows that the number of rural uneducated workers searching for a job in the office-workers' sector, $\tilde{l}_o^u$ must have fallen. The rural-born workers who would have chosen to migrate to the urban areas in search of office-worker jobs now choose to work in the EGA instead.

The number of jobs in the office-workers' sector that are available to the uneducated workers has not changed. For urban uneducated workers, the number of unemployed workers is given by

$$\tilde{U}^u = \tilde{L}_o^u \left(1 - \frac{J_o^u}{\tilde{l}_o^u + \tilde{L}_o^u}\right).$$

It is clear from the above equation that unemployment in this population depends positively on the number of people searching for these jobs ($\tilde{L}_o^u$), and negatively on the probability that each of them will get a job ($\frac{J_o^u}{\tilde{l}_o^u + \tilde{L}_o^u}$). The number of people searching has gone up, but since the total number of uneducated people searching for these jobs has fallen, the probability of getting a job has also

gone up. As a result the effect of the rural employment guarantee on unemployment among these workers is ambiguous.

How many urban uneducated workers get jobs in the office-workers' sector? This number is given by

$$\frac{L_o^u J_o^u}{L_o^u + l_o^u} = L_o^u \frac{w_s}{w_o}$$

before the introduction of the EGA, and by

$$\frac{\tilde{L}_o^u J_o^u}{\tilde{L}_o^u + \tilde{l}_o^u} = \tilde{L}_o^u \frac{\tilde{w}_n}{w_o}$$

after the introduction of the EGA. Since $\tilde{w}_n > w_s$, it follows that the number of urban uneducated workers who actually get jobs in the office-workers' sector goes up with the introduction of the EGA.

We know that the number of jobs in the O-sector remains the same. These jobs are divided between the urban uneducated workers and the rural uneducated workers who search in the O-sector. Since the number of jobs received by urban uneducated workers increases following the introduction of the EGA, the number of jobs received by rural uneducated workers must fall.

For rural uneducated workers, the number of unemployed workers is given by

$$\tilde{u}^u = \tilde{l}_o^u \left(1 - \frac{J_o^u}{\tilde{l}_o^u + \tilde{L}_o^u}\right).$$

and since both portions of the right-hand side of this equation have fallen, unemployment among rural uneducated workers definitely falls with the introduction of the EGA.

6.3.2 Peak Season Equilibrium

The analysis of the peak season equilibrium in the economy which had an EGA in the slack season follows the same general idea as in the previous baseline case. We start with looking at the rural uneducated population. The active workers in this population that are looking for jobs are those who were unemployed and those who were employed by the EGA in the slack season. Thus the active population is given by

$$\tilde{l}_n^u + \tilde{u}^u = l^u - \tilde{l}_s^u - \tilde{l}_f^u - \tilde{l}_o^u \frac{\tilde{w}_n}{w_o}$$

Let the number of people in peak season agriculture in the model with the rural employment guarantee be given by $\tilde{l}_p^u$. The active rural population searches for jobs in agriculture or in the free-entry sector. So at the end of the peak season, everyone in the rural population who did not get a job in the O-sector in the slack season is employed either in agriculture or in the free entry sector. This means that

$$\tilde{l}_p^u + \tilde{l}_{f'}^u = \tilde{u}^u + \tilde{l}_s^u + \tilde{l}_n^u + \tilde{l}_f^u = l^u - \left(\frac{J_o^u w_o + F - \tilde{w}_n L^u}{w_o} \right). \tag{6.11}$$

In the case of an interior equilibrium, the wages in the agricultural sector and in the free-entry jobs are equalized:

$$\frac{a_p}{\tilde{l}_p^u} = \frac{f}{\tilde{l}_{f'}^u}. \tag{6.12}$$

Now using equations 6.11 and 6.12 we can solve for $\tilde{l}_p^u$ and $\tilde{l}_{f'}^u$ as

$$\tilde{l}_p^u = \frac{a_p}{a_p + f} \left(l^u + \frac{\tilde{w}_n L^u}{w_o} - \frac{J_o^u w_o + F}{w_o} \right),$$

and

$$\tilde{l}_{f'}^u = \frac{f}{a_p + f} \left(l^u + \frac{\tilde{w}_n L^u}{w_o} - \frac{J_o^u w_o + F}{w_o} \right).$$

The peak season agricultural wage (equal to the free-entry sector wage for rural-born workers, $\tilde{w_{f'}}$) is given by

$$\tilde{w_p} = \frac{a_p}{\tilde{l_p^u}} = \frac{w_o(a_p + f)}{w_o l^u + L^u \tilde{w_n} - (J_o^u w_o + F)}.$$

Urban uneducated workers who were unemployed in the slack season now take up jobs in the free-entry sector, which means the new free-entry urban uneducated population is given by

$$\tilde{L_{f'}^u} = \tilde{U^u} + \tilde{L_f^u} = \frac{L^u(\tilde{w_n} - w_o) + F}{w_o},$$

and the new free-entry sector wage they earn is simply

$$\tilde{w_{F'}} = \frac{F}{L_{f'}^u} = \frac{w_o F}{L^u(\tilde{w_n} - w_o) + F}.$$

The peak season outcomes are summarized in Table 6.6. One important effect of the rural employment guarantee is that it lowers the peak season wage relative to the pre-EGA model. It is easy to show this. In the peak season of the pre-EGA model, the number of rural uneducated workers in the free-entry sector and in agriculture is given by

$$l_p^u + l_{f'}^u = l^u - (\text{rural-born workers who got jobs in the } O\text{-sector}) = l^u - l_o^u \frac{J_o^u}{l_o^u + L_o^u}.$$

In the peak season of the model with the rural employment guarantee, the number of rural-born workers in the free-entry sector and in agriculture is similarly given by

$$\tilde{l_p^u} + \tilde{l_{f'}^u} = l^u - (\text{rural-born workers who got jobs in the } O\text{-sector}) = l^u - \tilde{l_o^u} \frac{J_o^u}{\tilde{l_o^u} + \tilde{L_o^u}}.$$

From Table 6.5 we have the result that the number of rural-born workers who got jobs in the O-sector falls with the introduction of the EGA. This means

that

$$\tilde{l}^u_p + \tilde{l}^u_{f'} > l^u_p + l^u_f. \tag{6.13}$$

From the equalization of wages of rural-born workers between the free-entry sector and agriculture in the peak season, we know that

$$\begin{aligned} l^u_{f'} &= \frac{f}{a_p} l^u_p, \text{and} \\ \tilde{l}^u_{f'} &= \frac{f}{a_p} \tilde{l}^u_p. \end{aligned} \tag{6.14}$$

Combining equations 6.13 and 6.14, we see that

$$\begin{aligned} \tilde{l}^u_p + \tilde{l}^u_{f'} &> l^u_p + l^u_f \\ \implies \frac{a_p}{a_p + f} \tilde{l}^u_p &> \frac{a_p}{a_p + f} l^u_p \\ \implies \tilde{l}^u_p &> l^u_p. \end{aligned}$$

Since the agricultural wage in the peak season is inversely related to the size of the population in the agricultural sector, it follows that the peak season wage in the model with the rural employment guarantee, $\tilde{w}_p$, is lower than in the pre-EGA model, w_p. By providing a viable alternative to the urban free-entry sector in the slack season, the rural employment guarantee discourages rural-born workers from searching for office-worker jobs, and this results in a decline in the number of them who are successful in finding these jobs. This in turn increases the number of rural-born workers who look for jobs in agriculture in the next period, which depresses the agricultural wage.

The final income distribution for the model with the rural employment guarantee is given in Table 6.7.

What we know about the effect of the rural employment guarantee on

wages, unemployment and the number of workers in each search strategy is summarized in Table 6.8.

6.4 The Welfare Impact of the Rural Employment Guarantee

We have the final annual income distributions for the models with and without the rural employment guarantee. The next question is then - what can we say about how these distributions compare? The answer is - "it depends". We do not know enough about the way the two distributions compare to be able to make a general statement about whether or not the introduction of the EGA improves welfare unambiguously. However we can present some evidence once we make assumptions about the relationships of wages.

What is it we know with certainty in this model? We know that the peak season agricultural wage and the peak season rural free-entry sector wage fall with the introduction of the EGA, because the presence of the EGA reduces rural-urban migration and the number of rural-born workers who get the office-worker jobs. We also know that the unemployment among rural-born workers falls, and that wage equalization ensures that the slack season rural free entry sector wage, agricultural wage and the urban free-entry sector wage all rise to equate the EGA wage of $\tilde{w}_n$. However we do not know what happens to the quantities $\tilde{w}_f + \tilde{w}_{f'}$ and $\tilde{w}_F + \tilde{w}_{F'}$ - the first of these terms rises and the second falls, so the overall sum cannot be signed. Nor do we know how $\tilde{w}_p$ compares to $\tilde{w}_{F'}$.

To begin with, let us simply plot the cumulative distribution function of income for both before and after the introduction of the EGA. In order to do so, we

make the assumption that $\tilde{w_p} > \tilde{w_{F'}}$. Figure 6.3 depicts the cumulative income distribution for the pre-EGA model. On the x-axis we measure annual income. On the y-axis we measure the number of workers who earn at most a given income level, and this axis clearly ranges from 0 to the total number of workers in the economy, $\mathcal{L}$. Because the income distribution is discrete in this case, the cumulative distribution function 'jumps' at each new income level precisely by the number of workers who earn exactly that income.

From Table 6.7 we know that there are seven levels of income, so there are six 'jumps'. At the lowest level, 0, there are only those urban educated workers who searched unsuccessfully for managers' jobs in the slack season. There are $E_m \frac{w_m - w_o}{w_o}$ of them, which is why the distribution function jumps up by that amount at the lowest point. The next highest income is $w_{F'}$, earned by U^u workers. Then w_p, earned by u^u workers. The rest of the distribution function can be constructed in the same way. Just as an illustration, at wage $w_s + w_p$ there are a total of $l_s^u + l_f^u + l_F^u + U^u + u^u + E_m \frac{w_m - w_o}{w_o}$ workers who earn *at most* that income. All workers in the economy earn at or less than the highest level of income, $2w_m$, so once the income reaches that level the number of workers on the y-axis is simply $\mathcal{L}$.

Similarly, we can construct the cumulative distribution function for the model with the rural employment guarantee. This is depicted in Figure 6.4. The construction follows the same logic as above. The important thing to note between the figures 6.3 and 6.4 is that they coincide at three places - the lowest level of income, and at the highest two levels of income. This is because the introduction of the EGA does not affect either the number of workers in these three points of the income distribution, nor the wages they receive.

With the assumption we made that $\tilde{w_p} > \tilde{w_{F'}}$, there are still two possibilities of how the cumulative distribution functions before and after EGA introduction will compare. The differences between the two result from the fact that the impact of the EGA on unemployment among the urban uneducated population is ambiguous. Here we outline the two possibilities, or 'cases', depending on the relationships between unemployment before and after the introduction of the rural employment guarantee.

6.4.1 $\tilde{U}^u + \tilde{u}^u > U^u$.

The comparison of the two income distributions under this case is shown in Figure 6.5[2]. For the remainder of these welfare comparisons the thick dashed line will depict the distribution with the rural employment guarantee, and the thin solid line the income distribution without the rural employment guarantee. A welfare improvement under the EGA would be depicted by the distribution function with the EGA lying completely below the distribution without the EGA - meaning that at every income level, there are *fewer* workers who receive that income *or less* when the rural employment guarantee is present than when it is not. A crossing of the two distribution functions means that we cannot say anything unambiguously about the welfare comparison.

Here we can see that the two distributions do indeed cross, so the welfare implications of the EGA are ambiguous. Let's take a closer look to see why this is happening. Firstly, note that the two distributions only differ in the center, because both at the highest and lowest income levels the number of workers

[2] A sufficient condition for the above case classification is of course that urban uneducated unemployment increases with the introduction of the EGA.

and the wages they earn have not changed. We know that the wage earned by those rural-born workers who were unemployed in the slack season but work in the free-entry sector or in agriculture in the peak season falls, i.e. $\tilde{w_p} < w_p$. This is shown by the thick dashed line falling short of w_p. Because we are in the case where $\tilde{U^u} + \tilde{u^u} > U^u$, the sum of the second and third 'jumps' in the thick dashed distribution are greater than the second 'jump' in the thin solid distribution, which leads to a crossing of the two distributions. The welfare change due to the EGA is therefore, ambiguous.

6.4.2 $\tilde{U^u} + \tilde{u^u} < U^u$.

In the second case, however, we can actually show that welfare does increase with the introduction of the EGA. This is graphically depicted in Figure 6.6. In simple terms, in this case, the sum of the second and third 'jumps' in the thick dashed distribution are *less than* the second 'jump' in the thin solid distribution. As a result the distribution with the EGA lies completely below the distribution without the EGA. We know that rural uneducated unemployment falls with the introduction of the EGA, but a necessary condition for this case to hold is that urban uneducated unemployment also falls, and by a large enough amount.

We prefaced this whole discussion of these two distinct cases with the assumption that $\tilde{w_p} > \tilde{w_{F'}}$. What happens when we reverse that assumption? Well in that case the effect of the rural employment guarantee is *always* ambiguous. Why is that? We know that the introduction of the EGA depresses the peak season agricultural wage, and also causes a fall in the number of rural uneducated unemployed. Figure 6.7 depicts this case. Now $\tilde{w_p} < \tilde{w_{F'}}$, so the first portion

of the distribution function with the EGA must lie to the left of the distribution without the EGA. However the number of rural uneducated workers who receive this annual income also falls, $\tilde{u}^u < u^u$, which is shown by the fact that the first 'jump' in the thick dashed distribution is smaller than the first jump in the thin solid distribution. As a result the distribution of income in the model with the EGA *must* cross the distribution of income from the model without the EGA, rendering the welfare comparison ambiguous. It is important to note that this conclusion holds regardless of what happens to unemployment among the urban uneducated population.

6.5 Conclusion

This paper develops a simple yet powerful model of the Indian labor market, modeling the linkages between the rural and urban sectors, while also incorporating the seasonality that is such an important characteristic of Indian agriculture. It delivers the result of slack season migration to the urban areas followed by peak season reverse migration from the urban to the rural areas, another feature which has been widely acknowledged in the literature on the Indian labor market.

Into this stylized model we introduce a rural employment guarantee along the lines of the National Rural Employment Guarantee Act (NREGA) that was introduced in India in 2006. This EGA is only available in the slack season and only to rural-born workers, but provides a viable alternative to rural-urban migration since the wage in this guarantee was deliberately pegged by the government at a level higher than the going slack season agricultural wage prior to its

introduction. We study the changes in the search strategies, ex-ante choices and ex-post outcomes in this model with the rural employment guarantee. We show that one of the results of this program is a *depression* of the peak season agricultural wage. The reason this occurs is that the EGA encourages more rural-born workers to remain in the rural areas in the slack season, and results in a fall in the number of rural-born workers who get jobs in the office-workers' sector. This means the rural population that gets jobs in agriculture in the peak season is greater in the model with the rural employment guarantee than in the model without it, thereby depressing the peak season wage.

It turns out that comparing the income distributions for the cases with and without the rural employment guarantee is not straightforward. Under certain assumptions on the effect of the EGA on urban unemployment and on the relationships between certain levels of wages it is possible to use income distributions to show that the EGA unambiguously increases welfare. However there are several cases when the impact of the guarantee is ambiguous. The main reason this happens is that there is one section of the population - those rural uneducated workers who are unemployed in the slack season and employed in agriculture or in the free-entry sector in the peak season - who are made worse off by the guarantee because the introduction of the guarantee causes the rural wage in the peak season to fall.

The model we have proposed is tractable and easy to understand, but still powerful enough to deliver predictions about the effect of the rural employment guarantee on various different sections of the population, as well as on the wages in various sectors.

6.6 Figures and Tables

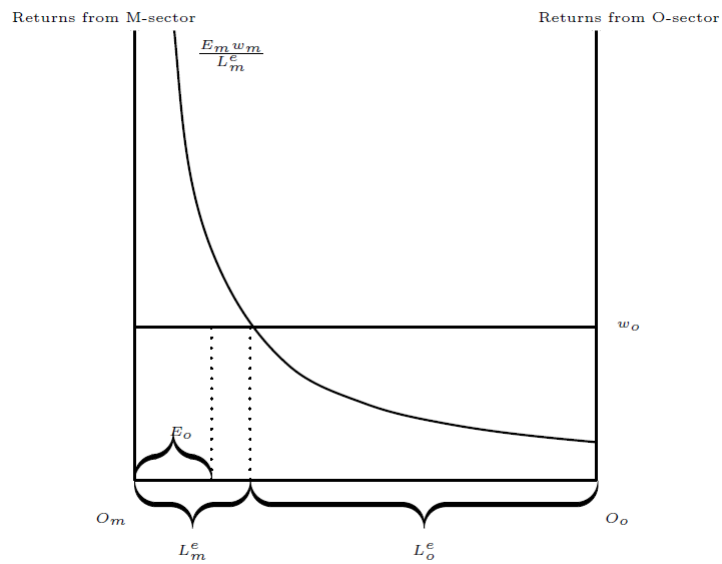

Figure 6.1: Search strategies for urban educated workers

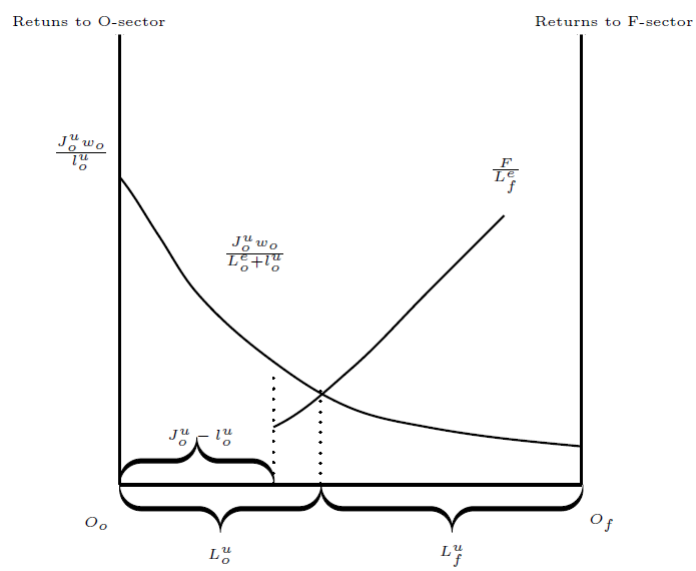

Figure 6.2: Search strategies for urban uneducated workers

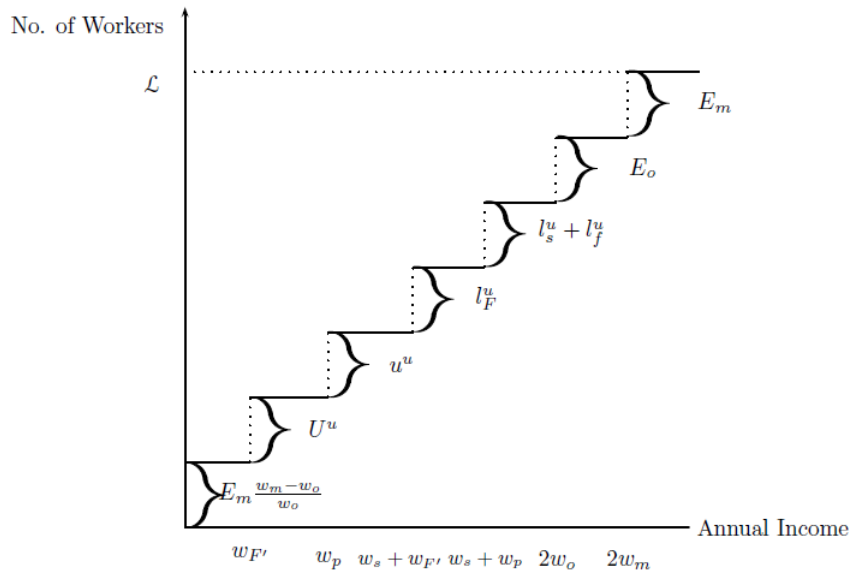

Figure 6.3: Cumulative annual income distribution - pre-EGA model ($w_p > w_{F'}$)

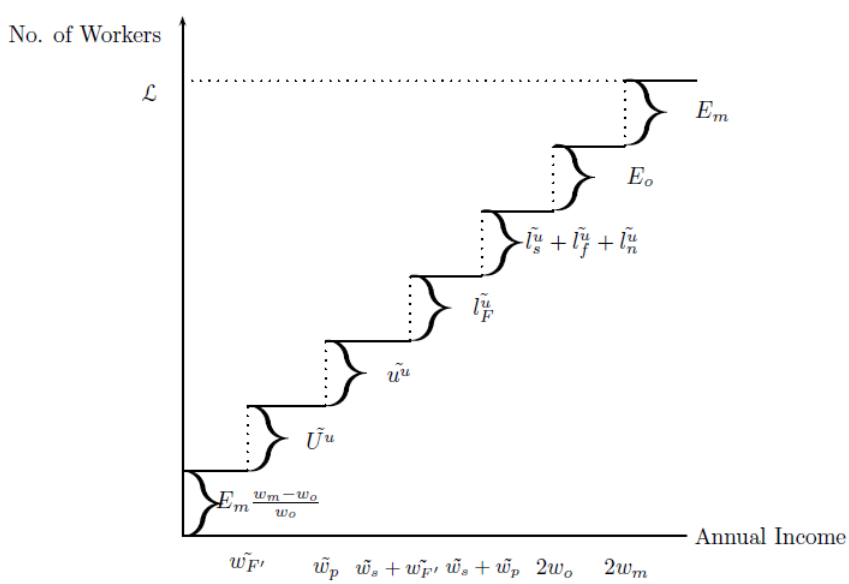

Figure 6.4: Cumulative annual income distribution - model with EGA ($w_p > w_{F'}$)

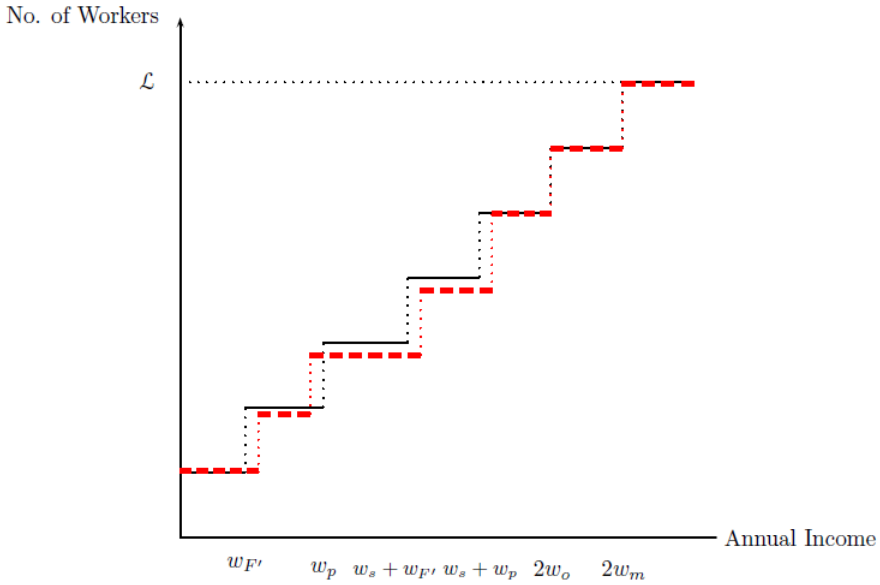

Figure 6.5: Welfare comparisons yield ambiguous results in the case where $\tilde{U}^u + \tilde{u}^u > U^u$ and $\tilde{w_p} > \tilde{w_{F'}}$.

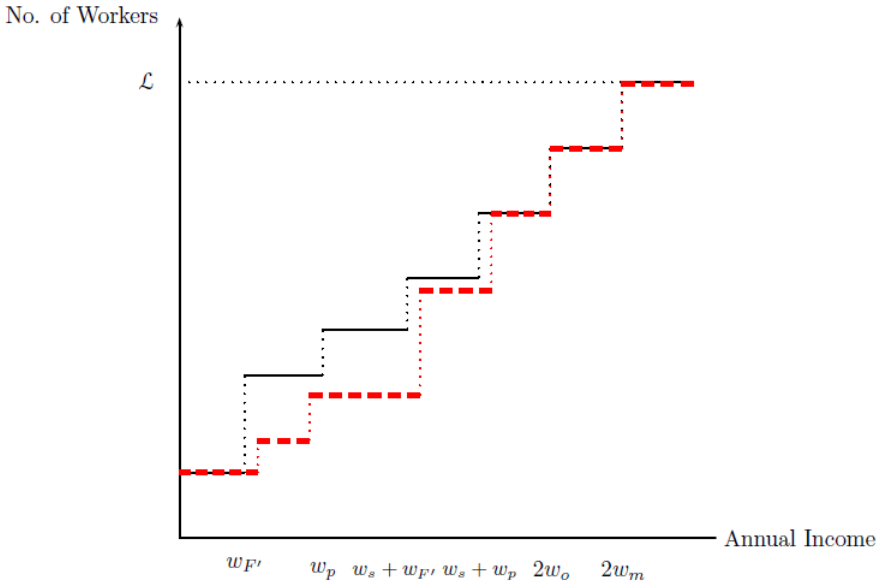

Figure 6.6: Welfare comparisons yield the result that the EGA is unambiguously better in the case where $\tilde{U}^u + \tilde{u}^u < U^u$ and $\tilde{w_p} > \tilde{w_{F'}}$.

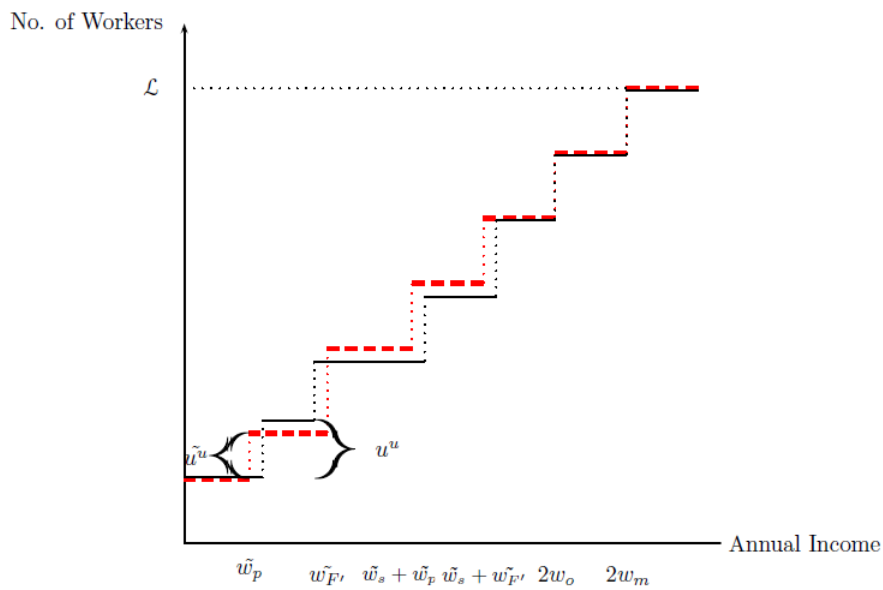

Figure 6.7: Welfare comparisons yield the result that the effect of the EGA is ambiguous in the case where $\tilde{w}_p < \tilde{w}_{F'}$.

Table 6.1: Slack Season Final Outcomes

Job	**Wage**	**Ed. urban-born workers**	**Uned. urban-born workers**	**Uned. rural-born workers**
Managers	w_m	E_m	0	0
Office Wkrs	w_o	$L^e - \frac{E_m w_m}{w_o}$	$\frac{(J_o^u w_o + a_s + f)L^u - F l^u}{w_o(L^u + l^u)}$	$\frac{(J_o^u w_o + F)L^u - (a_s + f) l^u}{w_o(L^u + l^u)}$
Agriculture	$\frac{J_o^u w_o + a_s + f + F}{L^u + l^u}$	0	0	$\frac{a_s(L^u + l^u)}{J_o^u w_o + a_s + f + F}$
Free-entry	$\frac{J_o^u w_o + a_s + f + F}{L^u + l^u}$	0	$\frac{F(L^u + l^u)}{J_o^u w_o + a_s + f + F}$	$\frac{f(L^u + l^u)}{J_o^u w_o + a_s + f + F}$
Unemployed	0	$\frac{E_m(w_m - w_o)}{w_o}$	U^u	u^u

Table 6.2: Peak Season Employment Patterns

Job	**Wage**	**Ed. urban-born workers**	**Uned. urban-born workers**	**Uned. rural-born workers**
Managers	w_m	E_m	0	0
Office-wkr	w_o	$L^e - \frac{E_m w_m}{w_o}$	$\frac{(J_o^u w_o + a + f)L^u - F l^u}{w_o(L^u + l^u)}$	$\frac{(J_o^u w_o + F)L^u - (a + f) l^u}{w_o(L^u + l^u)}$
Free-entry urban sector	$w_{F'}$	0	$L^u - \left(\frac{(J_o^u w_o + a + f)L^u - F l^u}{w_o(L^u + l^u)}\right)$	0
Agriculture	w_p	0	0	$\frac{a_p}{a_p + f}\left[l^u + \left(\frac{(a_s + f)L^u - (J_o^u w_o + F) l^u}{w_o(l^u + L^u)}\right)\right]$
Free-entry rural sector	w_p	0	0	$\frac{f}{a_p + f}\left[l^u + \left(\frac{(a_s + f)L^u - (J_o^u w_o + F) l^u}{w_o(l^u + L^u)}\right)\right]$
Unemployed	0	$\frac{E_m(w_m - w_o)}{w_o}$	0	0

Table 6.3: Annual Income Distribution without the EGA

Job combination	Annual Income	Number of workers
(M,M)	$2w_m$	E_m
(O,O)	$2w_o$	E_o
(F, F)	$w_s + w_{F'}$	$\frac{F(L^u+l^u)}{J_o^u w_o+a_s+f+F}$
(f,f), (f, p), (s, p)	$w_s + w_p$	$\frac{(a_s+f)(L^u+l^u)}{J_o^u w_o+a_s+f}$
(U, F)	$w_{F'}$	U^u
(u, f), (u, p)	w_p	u^u
(U, U)	0	$\frac{E_m(w_m-w_o)}{w_o}$

Table 6.4: Wages in various occupations and seasons

Wage	Description	Expression
w_s	Slack season agricultural wage	$\frac{J_o^u w_o+a_s+f+F}{l^u+L^u}$
w_f	Slack season rural free entry sector wage	
w_F	Slack season urban free entry sector wage	
w_p	Peak season agricultural wage	$\frac{w_o(a_p+f)(l^u+L^u)}{l^u\left[w_o(l^u+L^u)-(J_o^u w_o+F)\right]+(a_s+f)L^u}$
$w_{f'}$	Peak season rural free entry sector wage	
$w_{F'}$	Peak season urban free entry sector wage	$\frac{Fw_o(L^u+l^u)}{w_o(L^u+l^u)+Fl^u-(J_o^u w_o+a_s+f)L^u}$

Table 6.5: Slack Season Final Outcomes - With EGA

Type of job	**Wage**	**Ed. urban-born workers**	**Uned. urban-born workers**	**Uned. rural-born workers**
Managers	w_m	E_m	0	0
Office Wkrs	w_o	$L^e - \frac{E_m w_m}{w_o}$	$\frac{\tilde{w}_n L^u - F}{w_o}$	$\frac{F+J_o^u w_o-\tilde{w}_n L^u}{\tilde{w}_n}$
Agriculture	$\tilde{w}_n$	0	0	$\frac{a_s}{\tilde{w}_n}$
EGA	$\tilde{w}_n$	0	0	$l^u + L^u - \frac{J_o^u w_o+a_s+f+F}{\tilde{w}_n}$
Free-entry	$\tilde{w}_n$	0	$\frac{F}{\tilde{w}_n}$	$\frac{f}{\tilde{w}_n}$
Unemployed	0	$\frac{E_m(w_m-w_o)}{w_o}$	$\frac{w_o-\tilde{w}_n}{\tilde{w}_n w_o}(L^u\tilde{w}_n - F)$	$\frac{w_o-\tilde{w}_n}{\tilde{w}_n w_o}(F + J_o^u w_o - \tilde{w}_n L^u)$

Table 6.6: Peak Season Employment Patterns - With EGA

Job	**Wage**	**Ed. urban-born workers**	**Uned. urban-born workers**	**Uned. rural-born workers**
Managers	w_m	E_m	0	0
Office Wkrs	w_o	$L^e - \frac{E_m w_m}{w_o}$	$\frac{\tilde{w_n} L^u - F}{w_o}$	$\frac{J_o^u w_o + F - \tilde{w_n} L^u}{\tilde{w_n}}$
Free-entry urban sector	$\tilde{w_{F'}}$	0	$\frac{L^u(\tilde{w_n} - w_o) + F}{w_o}$	0
Agriculture	$\tilde{w_p} = \tilde{w_{f'}}$	0	0	$\frac{a_p}{a_p + f}\left(l^u + \frac{\tilde{w_n} L^u}{w_o} - \frac{J_o^u w_o + F}{w_o}\right)$
Free-entry rural sector	$\tilde{w_p} = \tilde{w_{f'}}$	0	0	$\frac{f}{a_p + f}\left(l^u + \frac{\tilde{w_n} L^u}{w_o} - \frac{J_o^u w_o + F}{w_o}\right)$
Unemployed	0	$\frac{E_m(w_m - w_o)}{w_o}$	0	0

Table 6.7: Annual Income Distribution with the EGA

Job combination	Annual Income	Number of workers
(M,M)	$2w_m$	E_m
(O,O)	$2w_o$	E_o
(F, F)	$\tilde{w_n} + \frac{w_o F}{L^u(\tilde{w_n} - w_o) + F}$	$\frac{F}{\tilde{w_n}}$
(f,f), (n,f), (s, p), (n,p)	$\tilde{w_n} + \frac{w_o(a_p + f)}{w_o l^u + L^u \tilde{w_n} - (J_o^u w_o + F)}$	$(l^u + L^u) - \left(\frac{J_o^u w_o + F}{\tilde{w_n}}\right)$
(U, F)	$\frac{w_o F}{L^u(\tilde{w_n} - w_o) + F}$	$\frac{w_o - \tilde{w_n}}{\tilde{w_n} w_o}(L^u \tilde{w_n} - F)$
(u, f), (u, p)	$\frac{w_o(a_p + f)}{w_o l^u + L^u \tilde{w_n} - (J_o^u w_o + F)}$	$\frac{w_o - \tilde{w_n}}{\tilde{w_n} w_o}(F + J_o^u w_o - \tilde{w_n} L^u)$
(U, U)	0	$\frac{E_m(w_m - w_o)}{w_o}$

Table 6.8: The effect of the EGA on various outcomes

Quantity	**Description**	**Effect of the EGA**
l_s^u	Rural uned. workers in agriculture	$\downarrow$
l_f^u	Rural-born uned. workers in the free-entry sector	$\downarrow$
l_o^u	Rural uned. workers searching in the O-sector	$\downarrow$
$l_o^u \frac{J_o^u}{l_o^u + L_o^u}$	Rural uned. workers who get jobs in the O-sector	$\downarrow$
u^u	Rural uned. unemployment	$\downarrow$
w_s	Slack season agricultural wage	$\uparrow$
w_f	Slack season free-entry sector wage for rural-born workers	$\uparrow$
w_p	Peak season agricultural wage	$\downarrow$
$w_{f'}$	Peak season free-entry sector wage for rural-born workers	$\downarrow$
$l_o^u + L_o^u$	Total uned. workers searching for jobs in the O-sector	$\downarrow$
L_o^u	Urban uned. workers searching for jobs in the O-sector	$\uparrow$
L_f^u	Urban-born uned. workers in the free-entry sector	$\downarrow$
$L_o^u \frac{J_o^u}{l_o^u + L_o^u}$	Urban uned. workers who get jobs in the O-sector	$\uparrow$
U^u	Urban uned. unemployment	ambiguous
w_F	Slack season free-entry sector wage for urban-born workers	$\uparrow$
$w_{F'}$	Slack season free-entry sector wage for urban-born workers	ambiguous

CHAPTER 7

CONCLUSION

Chapters 2 through 6 presented essays in labor and development economics, in the context of the Indian economy. The chapters can be divided into those essays that use household- or district-level data and empirical analysis, and those essays that develop theoretical models of the Indian labor market.

In Chapter 2, we showed that the introduction of an employment guarantee can have an impact on the production decisions of credit- and insurance-constrained agents. Specifically, the access to sure income increases the riskiness of the portfolio of crops grown at a district level, measured by various different quantities and with different sample restrictions. We provided a decomposition of changes in land allocation into risk-reducing and risk-increasing and study the impact of the program on these two measures. In Chapter 3 we showed that an increase in the opportunity cost of a mother's time due to positive rainfall shocks has a negative impact on health investments in children, and on their chances of survival. Districts with high female labor force participation that experience positive rainfall shocks see a significant decline in the probability that the child of the household will receive certain important vaccinations, or will be breastfed. Given the impact early-life illness can have on later-life outcomes, this is an important effect of rainfall shocks that has not been adequately explored before.

Chapter 4 discussed the stylized features of Indian labor markets in preparation for the models of the next two chapters. It also reviewed the relevant empirical and theoretical literature in the Indian context. Then in Chapter 5, we developed a rural-only model of India. This model incorporated several

of the widely-accepted stylized features of the Indian rural labor market, and then studied the introduction of an employment guarantee act along the lines of the NREGA. This model includes inter-temporal spillovers between the slack and peak seasons, and shows that by attracting labor away from agriculture in the slack season, the EGA has a negative effect on the peak season agricultural wage. Chapter 6 extended this model to include an urban sector, and modeled the seasonal patterns in migration between the rural and urban areas. This paper also introduced an EGA, and studied the effects of this act on various parameters in the model, and then finally also on overall worker welfare. In the paper with migration, the EGS provides a viable alternative to rural-urban migration, and keeps people in the rural area in the slack season. This increases the number of rural-born workers searching for jobs in agriculture in the peak season, and thus depresses the peak season wage. We showed in both models that the effect of the employment guarantee on welfare was ambiguous. In both cases this is because the peak season agricultural wage falls with the introduction of the EGA, though the cause of this decline is not the same in each paper.

This chapter concludes the dissertation.

BIBLIOGRAPHY

A. Adhvaryu, A. V. Chari, and S. Sharma. Firing costs and flexibility: Evidence from firms' employment responses to shocks in India. *Review of Economics and Statistics*, 95(3):725–740, 2013.

H. Alderman, J. Hoddinott, and B. Kinsey. Long term consequences of early childhood malnutrition. *Oxford Economic Papers*, 58:450–474, 2006.

A. Alesina, P. Giuliano, and N. Nunn. On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 2013. doi: 10.1093/qje/qjt005. URL `http://qje.oxfordjournals.org/content/early/2013/02/19/qje.qjt005.abstract`.

D. Almond. Is the 1918 influenza pandemic over? Long-term effects of in-utero influenza exposure in the post-1940 U.S population. *Journal of Political Economy*, 114(4):672–712, August 2006.

D. Almond and J. Currie. Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, 25(3):153–72, 2011.

J. W. Anderson, B. M. Johnstone, and D. T. Remley. Breast-feeding and cognitive development: A meta analysis. *American Journal of Clinical Nutrition*, 70(4): 525–535, October 1999.

F. Arnold, Minja Kim Choe, and T.K.Roy. Son Preference, the family-building process and child mortality in India. *Population Studies*, 52:301–315, 1998.

E. Artadi. Going into labor: Earnings vs. infant survival in rural Africa. Technical report, Harvard University, 2005. URL `http://www.econ.upf.edu/docs/seminars/vilartadi.pdf`.

Association for Indian Development. Key observations on NREGA work in Andhra Pradesh. Technical report, 2009. URL http://aidindia.org/main/content/view/788/279/.

M. Azam. The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment. Technical report, IZA Discussion Paper Series 6548, May 2012. URL http://www.iza.org/en/webcontent/publications/papers/viewAbstract?dp_id=6548.

R. Badiani and A. Safir. Coping with aggregate shocks: Temporary migration and other labor responses to climatic shocks in rural India. Working paper, 2008. URL http://www.webmeets.com/files/papers/ESWC/2010/2531/Temporary%20Migration%20India%20Badiani%20Safir.pdf.

B. Banerjee. The determinants of migrating with a pre-arranged job and of the initial duation of urban unemployment: An analysis based on Indian data on rural-to-urban migrants. *Journal of Development Economics*, 36(2):337–351, 1991.

B. Banerjee and S. Kanbur. On the specification and estimation of macro rural-urban migration functions: with an application to Indian data. *Oxford Bulletin of Economics and Statistics*, 43(1):7–29, 1981.

P. K. Bardhan. On life and death questions. *Economic and Political Weekly*, 9 (32/34), August 1974.

P. K. Bardhan. Wages and unemployment in a poor agrarian economy: A theoretical and empirical analysis. *Journal of Political Economy*, 87:479–500, 1979.

R. Basant. Agricultural technology and employment in India: A survey of re-

cent research. *Economic and Political Weekly*, 22(31):pp. 1297–1308, 1987. ISSN 00129976. URL http://www.jstor.org/stable/4377314.

A. K. Basu. Impact of rural employment guarantee schemes on seasonal labor markets: Optimum compensation and workers' welfare. *Journal of Economic Inequality*, 1:1–34, 2013.

A. K. Basu, N. H. Chau, and R. Kanbur. A theory of employment guarantees: Contestability, credibility and distributional concerns. *Journal of Public Economics*, 93:482–497, 2009.

A. K. Basu, N. H. Chau, G. S. Fields, and R. Kanbur. A multi-sector labor market model for developing countries. Working paper, 2013.

J. R. Behrman and A. B. Deolalikar. The intrahousehold demand for nutrients in rural south India: Individual estimates, fixed effects, and permanent income. *The Journal of Human Resources*, 25(4):665–696, 1990.

J. R. Behrman and M. R. Rosenzweig. Returns to birthweight. *The Review of Economics and Statistics*, 86(2):586–601, May 2004.

C. R. Belfield and I. R. Kelly. The benefits of breastfeeding across the early years of childhood. Working Paper 16496, NBER, October 2010.

E. Berg, S. Bhattacharyya, R. Durgam, and M. Ramachandra. Can rural public works affect agricultural wages? Evidence from India. CSAE Working Paper, May 2012. URL http://www.csae.ox.ac.uk/workingpapers/pdfs/csae-wps-2012-05.pdf.

S. Bhalotra. Fatal fluctuations? Cyclicality in infant mortality in India. *Journal of Development Economics*, 93:7–19, 2010.

A. K. Bhargava. The impact of India's rural employment guarantee on demand for agricultural technology. Technical report, IFPRI Discussion Paper, July 2014. URL http://akbhargava.weebly.com/uploads/2/1/2/5/21256332/akb_2014-07-15.pdf.

S. E. Black, P. J. Devereaux, and K. G. Salvanes. From the cradle to the labour market: The effect of birth weight on adult outcomes. *Quarterly Journal of Economics*, 122(1):409–439, 2007.

M. L. Burton and K. Reitz. The plow, female contribution to agricultural subsistence and polygyny: A log linear analysis. *Cross-Cultural Research*, 16:275–305, 1981.

E. Carranza. Soil endowments, production technologies and missing women in India. World Bank Policy Research Working Paper 5974, World Bank, February 2012.

A. Case, A. Garrib, A. Menendez, and A. Olgiati. Paying the piper: The high cost of funerals in South Africa. *Economic Development and Cultural Change*, 62 (1), 2013.

P. Chakraborty. Implementation of the National Rural Employment Guarantee Act in India: Spatial dimensions and fiscal implications. Technical report, National Institute of Public Finance and Policy and The Levy Economics Institute of Bard College, 2007.

P. Chakraborty. Implementation of the National Rural Employment Guarantee Act in India: Spatial dimensions and fiscal implications. Technical report, National Institute of Public Finance and Policy and The Levy Economics Institute of Bard College, 2007.

P. Chatterji and K. Frick. Does returning to work after childbirth affect breastfeeding practices? *Review of Economics of the Household*, 3:315–335, 2005.

S. Chaudhuri. Rural-urban migration, the informal sector, urban unemployment and development policies: A theoretical analysis. *Review of Development Economics*, 4(3):353–364, 2000.

A. Datta. A simple model of NREGA, subsistence and agricultural labour markets. Working paper, Indira Gandhi Institute of Development Research, 2014. URL `http://dev.onesquarepixel.com/igidr/public/downloadPaper/24`.

R. Dehejia and A. Lleras Muney. Booms, busts, and babies' health. *The Quarterly Journal of Economics*, 119(3):1091–1130, August 2004.

S. Dercon and L. Christiaensen. Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics*, 96: 159–173, 2011.

P. Deshingkar. Maximizing the benefits of internal migration for development. Presented at the regional conference on migration and development in Asia, 2005. URL `http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/2358.pdf`.

S. M. Dev. Small farmers in India: Challenges and opportunities. Technical report, Indira Gandhi Institute of Development Research, 2012. URL `http://www.igidr.ac.in/pdf/publication/WP-2012-014.pdf`.

N. Dey and J. Drèze. Employment Guarantee Act: A Primer, 2007. URL `http://www.righttofoodindia.org/data/ega\_english.pdf`.

S. Dey. Evaluating the impact of the National Rural Employment Guarantee Scheme: The case of Birbhum district, West Bengal, India. Technical report, International Institute of Social Studies, The Hague, January 2010. URL http://www.iza.org/en/webcontent/publications/papers/viewAbstract?dp_id=6548. ISS Working Paper No. 490.

C. M. Dieterich, J. P. Felice, E. O'Sullivan, and K. M. Rasmussen. Breastfeeding and health outcomes for the mother-infant dyad. *Pediatric Clinics of North America*, 60:31–48, 2013.

J. Drèze and A. Mukherjee. *The Balance between Industry and Agriculture in Economic Development vol.3: Manpower and Transfers*, chapter Labour Contracts in rural India: Theories and evidence. Macmillan Press, London, 1989.

P. Dutta, R. Murgai, M. Ravallion, and D. van de Walle. Does India's employment guarantee scheme guarantee employment? Technical report, 2012. World Bank Policy Research Working Paper 6003.

M. Eswaran and A. Kotwal. A theory of two-tier labor markets in agrarian economies. *The American Economic Review*, 75(1):pp. 162–177, 1985. ISSN 00028282. URL http://www.jstor.org/stable/1812711.

M. Fafchamps and J. Pender. Precautionary saving, credit constraints, and irreversible investment: Theory and evidence from semiarid India. *Journal of Business & Economic Statistics*, 15(2):180–194, 1997.

A. D. Foster and M. R. Rosenzweig. Comparative advantage, information and the allocation of workers to tasks: Evidence from an agricultural labor market. *The Review of Economic Studies*, 63(3):347–374, 1996. doi: 10.2307/

2297887. URL http://restud.oxfordjournals.org/content/63/3/347.abstract.

A. D. Foster and M. R. Rosenzweig. Microeconomics of technology adoption. *Annual Review of Economics*, 2(1):395–424, 2010.

E. Gehrke. An employment guarantee as risk insurance? Assessing the effects of the NREGS on agricultural production decisions. Technical report, BGPE Discussion Paper No. 152, May 2014. URL http://www.die-gdi.de/uploads/media/152_Gehrke.pdf.

A. K. Ghose. Wages and employment in Indian agriculture. *World Development*, 8:413–428, 1980.

X. Gine and S. Klonner. Credit constraints as a barrier to technology adoption by the poor. Research Paper 2006/104, UNU-WIDER, September 2006.

P. Glewwe and E. M. King. The impact of early-childhood nutritional status on cognitive development: Does the timing of malnutrition matter? *The World Bank Economic Review*, 15(1):81–113, 2001.

M. Grimm, R. Hartwig, and J. Lay. Investment decisions of small entrepreneurs in a context of strong sharing norms. Workng paper, World Bank Social Protection and Laor, October 2011.

M. R. Gupta. Rural-urban migration, informal sector and development policies: A theoretical analysis. *Journal of Development Economics*, 41:137–151, 1993.

A. D. Haan. Rural-urban migration and poverty: The case of India. *IDS Bulletin*, 28(2):35–47, 1997. URL http://search.proquest.com/docview/57580465?accountid=10267.

J. R. Harris and M. P. Todaro. Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60:126–142, 1970.

J. Hoddinott and B. Kinsey. Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, 63(4):409–436, 2001.

C. Imbert and J. Papp. Government Hiring and Labor Market Equilibrium: Evidence from India's Employment Guarantee. Presented at the ISI conference, New Delhi, December 2011. URL `http://www.isid.ac.in/\~pu/conference/dec\_11\_conf/Papers/ClementImbert.pdf`.

C. Imbert and J. Papp. Labor market effects of social programs: Evidence from India's employment guarantee. Technical report, Centre for the Study of African Economies, University of Oxford, 2013. CSAE Working Paper Series 2013-03.

C. Imbert and J. Papp. Short-term migration, rural workfare programs and urban labor markets: Evidence from India. Job market paper, October 2014. URL `http://sites.bu.edu/neudc/files/2014/11/paper_70.pdf`.

S. Jayachandran. Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575, June 2006.

R. Jha, R. Gaiha, and M. K. Pandey. Net transfer benefits under India's Rural Employment Guarantee Scheme. *Journal of Policy Modelling, Elsevier*, 34(2): 296–311, 2012.

D. Johnson. Can Workfare serve as a Substitute for Weather Insurance? The Case of NREGA in Andhra Pradesh. Technical report, Institute for Financial Management and Research, Centre for microfi-

nance, 2009. URL http://ifmr.ac.in/cmf/publications/wp/2009/32_Johnson_NREGA_Bad\%20Weather.pdf.

R. C. Johnson and R. F. Schoeni. The influence of early-life events on human capital, health status, and labor market outcomes over the life course. PSC Research Report 07-616, January 2007.

D. Karlan, R. Osei, I. Osei Akoto, and C. Udry. Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129(2): 597–652, 2012.

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

S. Klonner and C. Oldiges. Safey net for India's poor or waste of public funds? Poverty and welfare in the wake of the world's largest job guarantee program. Discussion paper 564, University of Heidelburg, May 2014.

M. S. Kramer, F. Aboud, E. Mironova, I. Vanilovich, R. W. Platt, L. Matush, S. Igumnov, E. Fombonne, N. Bogdanovich, and T. Ducruet et al. Breastfeeding and child cognitive development: New evidence from a large randomized trial. *Archives of general psychiatry*, 65(5):578–584, 2008.

D. Lal. The determinants of urban unemployment in India. *Indian Economic Review*, 23(1):61–81, 1988.

E. Ligon, J. P. Thomas, and T. Worrall. Informal insurance arrangements with limited commitment: Theory and evidence from village economies. *The Review of Economic Studies*, 69(1):209–244, 2002.

Y. Liu and C. Barrett. Heterogeneous pro-poor targeting in india: Mahatma

Gandhi National Rural Employment Guarantee Scheme. *Economic and Political Weekly*, XLVIII(10), 2013.

S. Maccini and D. Yang. Under the weather: Health, schooling and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–1026, 2009.

O. Mahul, N. Verma, and D. J. Clarke. Improving farmers' access to agricultural insurance in India. Working paper no. 5987, The World Bank, March 2012. URL `http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-5987`.

A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic Theory*. Oxford University Press, 1995.

I. M. Mbiti. Moving women: Household composition, labor demand and crop choice. Technical report, Brown University, November 2006. URL `http://paa2007.princeton.edu/papers/72117`.

K. F. Michaelsen, L. Lauritzen, M. H. Jrgensen, and E. L. Mortensen. Breastfeeding and brain development. *Scandinavian Journal of Nutrition*, 47(3):147–151, 2003.

B. Millett and M. Shah. Could droughts improve human capital? Evidence from India. Technical report, September 2012. URL `http://www.frbsf.org/economic-research/files/Shah\_Steinberg.pdf`.

A. M. Mobarak and M. Rosenzweig. Selling formal insurance to the informally insured. Working paper, Yale University, 2014.

J. Morduch. Risk, production, and saving: Theory and evidence from Indian households. Unpublished, Harvard University, 1990.

M. Morten. Temporary migration and endogenous risk sharing in village India. Working paper, September 2013. URL http://web.stanford.edu/~memorten/Melanie_Morten_files/India_Sep2013.pdf.

C. M. Moser and C. B. Barrett. The complex dynamics of smallholder technology adoption: The case of SRI in Madagascar. *Agricultural Economics*, 35:373–388, 2006.

D. L. Mowbray. From field to lab and back: Women in rice farming systems. Technical report, Consultative Group on International Agricultural Research, 1995. URL http://www.worldbank.org/html/cgiar/publications/gender/field.pdf.

A. Mukherjee and D. Ray. Wages and involuntary unemployment in the slack season of a village economy. *Journal of Development Economics*, 37:227–264, 1991.

A. Mukherjee and D. Ray. Labor tying. *Journal of Development Economics*, 47: 207–239, 1995.

L. Müller and U. Schindler. Soil moisture and workability of heavy arable soils. *Archives of Agronomy and Soil Science*, 44(2):161–174, 1999.

K. Munshi and M. Rosenzweig. Why is mobility in India so low? Social insurance, inequality, and growth. Working Paper 14850, National Bureau of Economic Research, April 2009. URL http://www.nber.org/papers/w14850.

K. Munshi and M. Rosenzweig. Networks and misallocation: Insurance, migration and the rural-urban wage gap. Working paper, October 2013.

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

P. H. Nyangal, F. H. Johnsen, and T. H. Kalinda. Gendered impacts of conservation agriculture and paradox of herbicide use among smallholder farmers. *International Journal of Technology and Development Studies*, 3(1):1–24, 2012.

W. Olsen and S. Mehta. A pluralist account of labour participation in India. Technical Report GPRG-WPS-042, Global Poverty Research Group, May 2006. URL `http://www.gprg.org/pubs/workingpapers/pdfs/gprg-wps-042.pdfs`.

P. Oreopoulos, M. Stabile, R. Walld, and L. Roos. Short, medium, and long term consequences of poor infant health: An analysis using siblings and twins. NBER Working Paper 11998, January 2006.

Planning Commission. Report of the Task Force: Identification of districts for wage and self-employment programmes. Technical report, 2003a.

Planning Commission. Backward Districts Initiative - Rashtriya Sam Vikas Yojana - the Scheme and Guidelines for preparation of District Plans. Technical report, 2003b. MLP division.

N. Qian. Missing women and the price of tea in China: The effect of relative female income on sex imbalance. CEPR discussion paper no. 5986, 2005. URL `http://sticerd.lse.ac.uk/seminarpapers/special24022005.pdf`.

M. Ravallion. *The Political Economy of Hunger*, volume 2, chapter 5, Market Responses to Anti-Hunger Policies: Effects on Wages, Prices and Employment. Clarendon, Oxford, 1990.

S. Ravi and M. Engler. Workfare in low income countries: An effective way to fight poverty? The case of the NREGS in India. Technical report, February 2009. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1336837.

S. Ravi, M. Kapoor, and R. Ahluwalia. The impact of NREGS on urbanization in India. SSRN working paper, August 2012. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2134778.

D. I. Rees and J. J. Sabia. The effect of breast feeding on educational attainment: Evidence from sibling data. *Journal of Human Capital*, 3(1):pp. 43–72, 2009. ISSN 19328575. URL http://www.jstor.org/stable/10.1086/599861.

T. Romero-Paris. Women's roles and needs in changing rural Asia with emphasis on rice-based agriculture. Technical report, International Rice Research Institute (IRRI), 2009. URL http://www.fftc.agnet.org/htmlarea_file/library/20110725165454/eb593.pdf.

M. F. Rosales. Impact of early life shocks on human capital formation: Evidence from El Niño floods in Ecuador. Job market paper, University of Chicago, April 2014. URL http://home.uchicago.edu/~mrosales/Research_files/Mrosales_JMP_Apr302014.pdf.

E. Rose. Consumption smoothing and excess female mortality in rural India. *Review of Economics and Statistics*, 81(1):41–49, 1999.

M. Rosenzweig. Risk, implicit contracts and the family in rural low-income countries. Working paper, Economic Development Center, University of Minnesota, November 1986.

M. Rosenzweig and H. Binswanger. Wealth, Weather risk and the composition and profitability of agricultural investments. *The Economic Journal*, 103(416): 56–78, 1993.

M. Rosenzweig and O. Stark. Consumption smoothing, migration, and marriage: Evidence from rural India. *Journal of Political Economy*, 97(4):905–926, 1989.

H. Royer. Separated at birth: US twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1):49–85, 2009.

C. Ruhm and W. E. Black. Does drinking really decrease in bad times? *Journal of Health Economics*, 21:659–678, 2002.

C. J. Ruhm. Are recessions good for your health? *The Quarterly Journal of Economics*, 115(2):617–650, May 2000.

C. J. Ruhm. Healthy living in hard times. Working Paper 711, IZA, February 2003.

K. J. S. Satyasai. Access to rural credit and input use: An empirical study. *Agricultural Economics Research Review*, Conference Number:461–471, 2012.

E. Skoufias, K. Vinha, and H. V. Conroy. The impacts of climate variability on welfare in rural Mexico. Technical report, The World Bank, February 2011. URL `http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2011/02/01/000158349_20110201090740/Rendered/PDF/WPS5555.pdf`.

A. V. Swamy. A simple test of the nutrition-based efficiency wage model. *Journal of Development Economics*, 53:85–98, 1997.

A. Szott. Rainfall shocks and human capital. Technical report, University of Maryland, College Park, 2012. URL http://econweb.umd.edu/$\sim$szott/SzottJMP.pdf. Job Market Paper.

T. Q. Thai and E. M. Falaris. Child schooling, child health and rainfall shocks: Evidence from rural Vietnam. Technical Report WP 2011-011, Max-Planck Institute for Demographic Research, 2011. URL www.demogr.mpg.de/papers/working.wp-2011-011.pdf.

T. Q. Thai and M. Myrskylä. Rainfall shocks, parental behavior and breastfeeding: Evidence from rural Vietnam. Technical Report WP 2012-009, Max-Planck Institute for Demographic Research, 2012. URL www.demogr.mpg.de/papers/working.wp-2012-009.pdf.

D. Thomas, K. Beegle, E. Frankenberg, B. Sikoki, J. Strauss, and G. Teruel. Education in a crisis. *Journal of Development Economics*, 74(1):53–85, 2004.

L. M. Thompson and F. R. Troeh. *Soils and Soil Fertility*. McGraw-Hill, 1978. ISBN 0-07-064411-X.

R. M. Townsend. Risk and insurance in village India. *Econometrica*, 62(3):539–91, 1994.

L. J. Unnevehr and M. L. Stanford. Technology and the demand for women's labor in Asian rice farming. In *Proceedings of a conference on Women in Rice Farming Systems, The International Rice Research Institute, 26-30 September 1983*, Manila, Philippines, 1983. Gower Publishing Company. ISBN 0 566 05105 2.

S. N. Wadood and R. L. Lamb. Choice of crops and employment uncertainty in the off-farm labor market. 2006. URL http://mpra.ub.uni-muenchen.de/10779/.

L. Zimmermann. Why guarantee employment? Evidence from a large Indian public-works program. Technical report, University of Michigan, October 2013.