Does the Indian Employment Guarantee reduce households' risk exposure? Assessing the effects of the NREGA on crop choice

Esther Gehrke*

June 17, 2013

Preliminary draft. Please do not cite. Comments welcome.

Abstract

This paper assesses the role of risk constraints in households' production decisions. It shows that the introduction of the Indian Employment Guarantee affects input allocation in agricultural production, particularly crop choice, by reducing households' uncertainty about future income streams. Taking advantage of the fact that the Indian National Rural Employment Guarantee Scheme was rolled out in three phases, I show that the key innovation of the Indian public works programme, namely giving households the right to work, encourages agricultural households to increase the share of risky but profitable crops in their portfolio. By triggering considerable increases in the fraction of inputs allocated to more profitable crops, especially to cotton, chillies and turmeric, the NREGA has the potential of raising the productivity of small and marginal farmers. The results of this paper are robust to a range of alternative specifications and to the inclusion of weather data as well as to changes in household income. Linking the Employment Guarantee to risk considerations is the key innovation of this paper. Therewith, it makes two contributions. First, it contributes to the empirical evidence on the validity of the theory of decision making under uncertainty. Second, it provides additional theoretical and empirical evidence to the ongoing debate on the effects of the NREGA on agricultural productivity.

Keywords: Uncertainty, Employment Guarantee, Crop choice.
JEL Codes: I38, O12, Q16.

*Esther Gehrke is researcher at German Development Institute / Deutsches Institut fuer Entwicklungspolitik (DIE) and PhD candidate at University of Passau, e-mail: *esther.gehrke@die-gdi.de*.

1 Introduction

Which constraints households in developing countries face in their decision making processes, has long been debated in the academic literature. Getting these constraints right, might help policy makers, development practitioners and governments to better design policies aimed at improving households' wellbeing.

One particularly relevant sector for studying such constraints is the agricultural sector. In most developing countries, poverty reduction will to a large extent depend on the ability of governments in providing viable income strategies to households living in rural areas and depending on agricultural production as major source of income.

Existing literature shows that farmers in developing countries are constrained in their production and investment decisions (Foster and Rosenzweig 2010; Duflo, Kremer and Robinson 2008; Suri 2011). This, mainly economic, literature has traditionally focused on learning processes (Besley and Case 1993; Munshi 2004; Conley and Udry 2010), lack of human capital (Foster and Rosenzweig 1996), risk (Rosenzweig and Binswanger 1993; Dercon and Christiaensen 2011) and credit constraints (Rosenzweig and Wolpin 1993; Fafchamps and Pender 1997; Gine and Klonner 2006) as potential explanations for delayed technology adoption, lack of investment in productive capital and for conservative crop choices.

While policy recommendations that can be derived from evidence on learning processes and educational constraints are straightforward, disentangling the effects of credit constraints and risk is much more difficult and at the same time highly relevant for policy recommendations. One of the major challenges for disentangling both constraints is to get exogenous variation in one or both constraints. Observational studies mainly have to rely on proxies for both constraints, but face the challenge that indicators that represent a household's ability to cope with risk are mainly the same indicators that predict access to credit and or own financing possibilities.

In this paper, I explore exogenous variation in the exposure to risk, in order to test the relevance of risk constraints in household production decisions. More specifically, I test whether the introduction of the Indian National Rural Employment Guarantee Act (NREGA) has effects on households' input allocation, because it reduces uncertainty about future income streams. The unique feature of the programme is the legal entitlement of households to a hundred days of work, which is crucial in helping households not only to cope with shocks ex-post, but also affects households' expectations about future income streams and about possibilities to smooth consumption over time.

The paper's main focus lies on crop choice for two reasons. On one side, fertilizer adoption and use of high yielding variety seeds are quite common in rural India since the Green Revolution, which suggests that most households are not heavily constrained in applying these technologies. On the other side, the production of highly profitable crops is still mainly restricted to large-scale farmers. Uncertainty regarding prices and yield developments, obliges many small-scale farmers to restrict themselves to the cultivation of low-risk low-profitability crops.

In order to test the outlined research question, I build a household model of crop choice under uncertainty. The effects of the NREGA are modeled as increase in harvest season labour market wages. The model clearly predicts an increase in the share of risky crops in the overall portfolio, if off-farm employment opportunities are increased.

The model predictions are tested using the Young Lives Survey panel dataset conducted in 2002, 2007 and 2009 in Andhra Pradesh (India). For the current analysis, I use rounds 2 (2007) and 3 (2009-10) of the survey and restrict the sample to households with non-zero agricultural production in both rounds. Empirical estimation is based on matched panel data regression methods, taking advantage of the sequenced roll-out of the National Rural Employment Guarantee Scheme.

I show that the key innovation of the Indian public works programme, namely giving households the right to work, encourages agricultural households to increase the share of risky but profitable crops in their portfolio, in particular cotton, chillies and turmeric. The results are robust to a range of alternative specifications and the inclusion of weather data and changes in household income.

The results of this paper suggest that reductions in households risk exposure could trigger important gains in agricultural productivity in the medium term. These gains go much beyond the direct effect that the provision of employment in agricultural lean seasons has on available income of rural households. That increases in productivity and therewith in households' incomes can be triggered via reductions in risk exposure alone, is a very important lesson to be learned for other countries with planned or ongoing public works programmes.

Linking the employment guarantee to risk considerations is the key innovation of this paper. By doing so, this paper makes two contributions. First, it contributes to the empirical evidence on the validity of the theory of decision making under uncertainty. Second, it provides additional theoretical and empirical evidence to the ongoing debate on the effects of the NREGA on agricultural productivity.

The remainder of this paper proceeds as follows. Sections 2 and 3 give an overview about existing empirical evidence on the role of uncertainty in production decision and about the implementation of the NREGA respectively. Section 4 presents a theoretical framework for analysing the effects of uncertainty on crop choice. The estimation strategy is outlined in Section 5. Sections 6 and 7 proceed with data description and empirical results, while Section 8 concludes.

2 Empirical evidence on the effects of uncertainty on crop choice

Limited ability of households to cope with shocks, may have negative effects on crop choice, if certain crops are more sensitive to rainfall fluctuations but at the same time more profitable on average. Similarly some crops may require higher investment prior to the full realization of uncertainty, therefore increasing overall risk. Although theoretically sound, there is still significant scope for empirical research on the extent to which risk affects the choice of crops.

The lack of empirical evidence can mainly be attributed to difficulties in measuring risk exposure and to methodological challenges arising from the necessity to rely on proxies such as wealth levels and buffer stocks. First, such proxies are very imprecise measures of risk exposure. Second, unobserved heterogeneity, such as land quality, farmer's ability and risk aversion, is likely to affect both, wealth levels and crop choice. Third, wealth levels are likely to be correlated with other potential explanations for suboptimal crop choice such as educational levels, access to credit and information or own financing possibilities.

Nonetheless, a range of innovative approaches exist, that deserve to be mentioned. Although they do not all address crop choice per se, the results of these studies are based on similar theoretical considerations and confirm the existence of risk-related barriers in production and investment decisions.

The first and path breaking exercise was done by Rosenzweig and Binswanger (1993). They show that the composition of asset portfolios is determined by a farmer's exposure to rainfall variability and its own risk coping ability. They find that poorer households hold asset portfolios that are less influenced by rainfall and in consequence have lower profits. Other authors have looked at the effects on risk and credit constraints on fixed capital investment. Particularly relevant are studies by Rosenzweig and Wolpin (1993); Fafchamps and Pender (1997) and Grimm, Hartwig and Lay (2011).

A few studies then specifically address technology adoption. Dercon and Christiaensen (2011) explore the empirical importance of risk avoidance in fertilizer adoption in Ethiopia. They build a model of risky input choice that explicitly accounts for both, the possible impact of seasonal credit constraints and of risk constraints on input adoption. Risk constraints are mainly related to the necessity to smooth consumption over time. In their empirical analysis, they construct first an indicator, that proxies each household's risk exposure. This indicator calculates for each household by how much consumption levels would fall under adverse weather shocks, taking into account existing buffer stocks as well as the probability distribution of rainfall. Then, the authors test whether this indicator is related to a household's probability of adopting fertilizer. They find a statistically significant link between both variables, suggesting that the expectation of low consumption outcomes in case of rainfall shortages reduces a household's likelihood to adopt fertilizer.

The most recent evidence on technology adoption has been provided by Karlan et al. (2012). In their experimental study, cash grants, subsidized insurance or the combination of both are randomly assigned to households in Ghana. They find that households who received insurance grants increased chemicals' application and total production levels. In contrast, households who received cash grants only, did not change their investment strategies. The results are highly relevant for understanding household rationing under uncertainty: they suggest that risk constraints are much more important than credit constraints in delaying technology adoption. But the analysis suffers from the fact that changes in investment behaviour could only be observed for households with subsidized access to insurance and not for households who purchased insurance coverage at actuarially fair premiums. Additionally, insurance take-up was much lower when insurance policies were sold at actuarially fair prices (41%) or even market prices (18%).[1] This finding is in line with previous literature on insurance take-up (Cole et al. 2012) and severely limits the policy implications that can be derived from this experiment.

Crop choice, finally, has to the best of my knowledge only been addressed in two papers. The first author to address the effects of risk on crop choice was Dercon (1996). Using data on rural households in Tanzania, he shows that wealthier households are more likely to grow more profitable but also riskier crops and argues that the absence of risk management strategies could be the reason for observed poverty traps in rural areas. Although he provides very convincing empirical evidence for the existence of such poverty traps, attributing these to risk arguments is relatively difficult, as many reasons exist, why wealth could affect crop choice. For instance, household wealth is often correlated with higher educational levels and with better access to credits and information. These channels might be equally relevant.

The second approach to address the effects of risk on crop choices has been proposed by Wadood and Lamb (2006). The authors use the ICRISAT data to test the effects of dysfunctional off-farm labour markets on crop choice. They argue, that off-farm employment opportunities are a crucial tool for consumption smoothing and should therefore reduce risk exposure. However, the empirical approach suffers from data limitations and endogeneity problems. Changes in agricultural conditions over time, particularly in input and output prices, strongly affect the profitability of certain crops and are likely to affect off-farm employment availability in rural areas as well, especially in contexts in which the main off-farm employment opportunities

[1] At subsidized rates, from 1/3 of the market price to full subsidisation, insurance take-up lay between 65% and 100%.

are on the fields of large-scale farmers.

3 Context

The NREGA is the largest public works programme in the world. In the financial year 2010-11, it provided work to close to 55 Million households in rural districts in India (MoRD-GoI 2012). A total of 2.5 Billion person-days of employment were generated in the same year.

The Indian National Rural Employment Guarantee Act was notified in September 2005 and started to be implemented in 2006. Implementation was sequenced, giving priority to the 200 poorest districts of India and subsequently extending it to the remaining districts.[2] The last districts introduced NREGA in 2008. The Act entitles every household living in rural areas to up to a hundred days of work per year, at state minimum wages and to be provided by the Block Officer within 14 days after the application for work has been made.[3]

Implementing a programme of the size of NREGA did not go without problems and a huge strand of literature shows that many challenges remain to be addressed in the implementation of NREGA. Among these challenges the most pressing ones seem to be delays in wage payments, corruption, the lack of awareness of many participants about their rights as well as an employment generation that continues to be supply-driven, instead of reacting to the demand for work (Aiyar and Samji 2009; Raabe et al. 2010).

In order to test the questions outlined above, I use household level data from Andhra Pradesh. I argue that this state is particularly suited to study the relation of interest because it is one of the best performing states in terms of number of workdays generated per household and in meeting the demand for work (Dutta et al. 2012). In terms of outreach, only Chhattisgarh, West Bengal, Madya Pradesh and Rajastan were able to reach a higher share of rural households in the financial year 2009-10 (MoRD-GoI 2012).[4]

Nonetheless, the programme continues to be implemented in a top-down manner in Andhra Pradesh. Usually, work is not being generated upon demand, rather applications for work are only being accepted if there is work available. This could potentially limit the analysis proposed in this paper. If the programme does not sufficiently respond to households' demand, then

[2] India has a total of 655 districts, of which 625 have introduced NREGA to date. The 30 remaining district are urban districts.

[3] The Block Officer is the NREGA official at the block (in Andhra Pradesh: mandal) level. The block/ mandal is the administrative unit below the district.

[4] At the same time, Andhra Pradesh has been forerunner in terms of innovative approaches to the implementation of NREGA. First, it has a long experience with performing social audits in order to increase accountability within the scheme. Second, it was one of the first states that cooperated with IT enterprises in order to strengthen the efficiency of administrative processes. For increased transparency, entries on muster rolls, number of workdays generated per jobcard holder etc. are publicly accessible.

these cannot rely on the availability of additional income generating opportunities in the case of a shock. As I will show later, this reliability is crucial for reducing households' risk exposure and thereby triggering changes in production decisions. But because Andhra Pradesh is still the state with one of the highest numbers of workdays generated per household, very few households report having received less work than they were entitled to. This is why I feel confident that testing the outlined questions in this particular state is feasible.

4 Theoretical framework

Providing additional employment opportunities to a total of 55 Million households has brought about considerable changes in the social and economic realities of rural areas in India.

The NREGA affects households in rural areas through a range of effects. The most obvious and so far most intensely researched effect, is the increase in available income once households participate in the programme. This effect is most pronounced for households with surplus labour, thus for households with little or no own land and without significant entrepreneurial activities. The increase in income has been shown to raise consumption levels (Jha, Gaiha and Pandey 2012), expenditure for education (Afridi, Mukhopadhyay and Sahoo 2012) and to drive women empowerment (Pankaj and Tankha 2010). Note here, that increases in available income might also positively influence investment behaviour and risk taking capacity, although this effect has not been analysed so far.

The second effect, which is by far less well understood, is the insurance effect: by giving households the right to work, the NREGA greatly influences their ability to smooth income in the case of a shock. In case this insurance effect holds, households could change their production decisions, take more risk and reach higher expected incomes. If then shocks realise, households can smooth part of the loss by working under the scheme, whereas they would not do so or maybe to a lesser extent, if the shock did not realise.

Finally, all changes will probably affect general equilibria in the village economy. Increases in consumption and labour market opportunities were shown to trigger wage increases (Azam 2012; Imbert and Papp 2012; Berg et al. 2012; Basu 2013). These changes again affect production levels. Analysing the overall effect on the local economy is beyond the scope of this paper.

This paper concentrates on the insurance effect as described above. In order to do so, this section develops a theoretical model of household decision making under uncertainty that shows how the introduction of NREGA could affect crop choice via the insurance effect vs. the income effect.

The model mainly builds on Dercon and Christiansen (2011). Taking into account the ideas outlined by Fafchamps (1993) and Van Den Berg (2002), I explore in particular how the sequencing of input allocation, shock realisation and harvesting influences production decisions. The possibility to smooth consumption over time is therein constrained by two main factors: the lack of adequate risk management strategies and the limited access to credit.

Crop choice is first modeled in a world without risk but with constrained credit markets, and then extended to a world with uncertainty. This allows to isolate the effects of uncertainty and risk aversion on production decisions. Finally, I will show how the introduction of the NREGA could affect input allocation decisions in both scenarios.

4.1 General setup

I assume that the household, who engages in agricultural production, has the choice between two agricultural products Q^d and Q^s. Both products are well known to the farmer and have been produced in the region for some time, so that I can abstract from learning and other sunk costs. These products are produced with two different production functions: one is deterministic, the other stochastic.[5] Both products can be sold on local markets at the same price p.

$$
\begin{aligned}
&Q^d = f^d(a^d, l_1^d, i^d) \\
&Q^s = f^s(a^s, l_1^s, i^s, \epsilon) \qquad E[\epsilon] = 1 \\
&\alpha(Q^d + Q^s) = l_2
\end{aligned}
$$

Agricultural production takes place over two periods, the planting and the harvesting season. Input allocation at planting stage defines total yield, which has to be harvested in the second stage. Total yield of both products therefore depends on land a, labour l_1 and input i allocation in period 1.[6] Inputs i are defined as a bundle of variable inputs such as seeds, fertilizer and pesticides. I assume that the first period production function is a Cobb-Douglas production function. Total yield of the risky product additionally depends on the realisation of a (multiplicative) random, serially uncorrelated shock ϵ at the end of the first period. Expected value of this shock is 1, thus in expectation, the production function of the risky crop is just $f(a^s, l_1^s, i^s)$. It is also assumed that the risky crop is more productive on average. Labour

[5] The assumption, that one product is deterministic and the other stochastic is rather extreme. Rather, one would expect both products to depend on the realisation of random shocks, although to different extent. I simplify the model because the results would be the same while the analysis would get more complicated.

[6] So far, I abstract from fixed capital because the marginal effect of productive capital was found to be relatively low.

required for harvesting in the second period l_2 is just linear function of realised yield, where α is a parameter indicating how much labour is needed for harvesting relative to realised yield. Because labour allocation is linear in realised yield, it will be profitable to harvest either the entire crop or nothing at all (depending on wage levels and output prices), thus allowing only for corner solution outcomes. This assumption is in line with earlier work on the sequencing of agricultural production by Fafchamps (1993) and Dillon (2010).

Now, assume that the household maximises utility from consumption C in both periods, the planting and the harvesting period. The utility function is additive over both periods and future utility is discounted by the factor δ. The utility function satisfies the usual properties: it is twice differentiable and increases in C but at decreasing rates, $\partial U/\partial C > 0$ and $\partial^2 U/\partial C^2 < 0$. This also implies that the household is risk averse.[7] I abstract from leisure in this model because it would not change the results of choice under uncertainty.[8] The household generates income from wage employment on local labour markets and from agricultural production. Building on the full-income approach, the household maximisation problem can be summarised as follows:

$$
\begin{aligned}
max \qquad V =& U_1(C_1) + \delta E U_2(C_2) \\
& s.t. \\
& \quad C_1 \leq w_1(T_1 - l_1^d - l_1^s) - g(i^d + i^s) + B \\
& \quad C_2 \leq (p - \alpha w_2)(Q^d + Q^s) + w_2 T_2 - (1+r)B \\
& \quad B \leq B^m \\
& \quad a^d + a^s \leq 1
\end{aligned}
$$

Total time endowment is represented by T_1 and T_2. In both periods, total time can be allocated between working on the labour market or working on own fields. In the first period, the household obtains income from labour market work at wage levels w_1 and from borrowings B. Inputs for agricultural production can be purchased at prices g. In the second period, the household obtains income from its own agricultural production $Q = Q^d + Q^s$ and from labour market work at wage w_2. Note here that the household will have to allocate labour to harvesting in order to generate

[7] Concave utility functions always imply risk aversion, however the exact coefficients of risk aversion may differ. The most common functions such as the CES also imply constant relative risk aversion

[8] By dropping leisure, I ignore possible income effects of increases in wage levels on a household's time allocation between labour and leisure. But since my main interest lies in crop choice rather than in production levels, I feel that ignoring leisure is not of major concern. Similar approaches can be found in Rosenzweig and Binswanger (1993); Fafchamps and Pender (1997) and Dercon and Christiansen (2011).

income from agricultural production, it is therefore useful to replace l_2 in the budget constraint by $\alpha(Q^d + Q^s)$.

Incurred debts will have to be repayed in the second period at interest rate r. Input credits are relatively common in rural Andhra Pradesh, although it seems that the amount of credit that is conceded is limited by households wealth. In the sample, around 18% of the households, who applied for credit, report not having received the total amount of credit they needed. Therefore, B^m describes the maximum amount a household can borrow for productive purposes. In contrast to input credit, consumption credit is much more difficult to obtain and highly expensive, as households have to rely on local moneylenders as source of consumption credit. Because households would opt for that source of credit only under extreme circumstances, this model does not allow for any borrowing beyond the harvesting period.

Local labour markets are assumed to be functioning with the option to hire labour in as well as out. In fact, in the sample most households report a range of income sources and casual labour features prominently among them. However, one should keep in mind that agricultural wages are very low in rural India, and that they vary strongly with covariate shocks such as rainfall shortages (Jayachandran 2006). For most small and marginal farmers this means, that they can only form expectations about harvest stage wages and that they face double risk from rainfall fluctuations: first, their own harvest is going to fail in case of rain shortages and second, they will not be able to find work at adequate wage levels on local labour markets.

Finally, $a^d + a^s = 1$ describes the restrictions on allocable land. I assume that land markets are not functioning and that owned land is used for own agricultural production or left fallow. This is obviously a simplifying assumption that will not hold everywhere in India. Nonetheless, observed levels of land renting are relatively low in rural Andhra Pradesh and land sales are almoust absent.[9]

The model described so far deviates from standard neoclassical models in that credit and land markets are assumed to be dysfunctional. Given these constraints the separability of households' production and consumption decisions will not hold even in the absence of risk.

4.2 Deterministic case

First, consider a scenario without uncertainty. In such a world, each household maximises utility by maximising profits from agricultural production

[9]Part of this is due to very restrictive legal environment, that discourages land owners from renting out their land, even if it is otherwise left fallow. Also, land prices are very high, which combined with low levels of credit availability makes land acquisition impossible for the majority of households. Those who could afford this, rather seek to diversify out of agriculture and move to urban areas.

plus income from wage employment. Identical results would be obtained if the household was risk neutral. Because both production functions are deterministic in this scenario, optimal land, input and labour allocation is achieved when their marginal product equals respective prices. In the deterministic case, the Lagrange can be summarised as follows:

$$\begin{aligned}\mathcal{L} =& U_1(C_1) + \delta U_2(C_2) \\ &+ \lambda(w_1(T_1 - l_1^d - l_1^s) - g(i^d + i^s) + B - C_1) \\ &+ \mu[(p - \alpha w_2)(Q^d + Q^s) + w_2 T_2 - (1 + r)B - C_2] \\ &+ \varphi(B^m - B) \\ &+ \rho(1 - a^d - a^s)\end{aligned}$$

Differentiating the Lagrange with respect to the choice variables, leads to the following decision rules for the allocation of variable inputs to each of the crops. The main focus of this paper lies on input allocation, but similar results can be obtained for the allocation of labour as much as for the allocation of land to each of the crops. A detailed derivation of all decision rules can be found in the Mathematical Appendix.

$$\frac{\partial Q^d}{\partial i^d} = \frac{g}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{1}$$

$$\frac{\partial Q^s}{\partial i^s} = \frac{g}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{2}$$

Equations (1) and (2) show the optimal allocation of inputs to each of the crops in the first stage. Since decision rules are equal for both crops, optimal allocation will imply that the marginal product of inputs in d is equal to the marginal product of inputs in s. Because realised yield has to be harvested in the second period, input allocation does not only depend on current prices, but also on future wage levels and on the relation between marginal utility of consumption today vs. tomorrow.

$$\frac{\partial U_1}{\partial C_1} = \delta(1 + r)\frac{\partial U_2}{\partial C_2} + \varphi \tag{3}$$

Finally, eq. (3) describes the optimal consumption rule over both periods given credit constraints: if the credit constraint is binding, φ is greater than zero and the marginal utility from consumption in the planting period will be greater than the marginal utility from consumption in the harvesting period (after accounting for the time discount factor δ and the interest rate r). This means that consumption in the planting stage will be lower than optimal.

Including eq. (3) into eq. (1) also reveals the effect of the credit constraint on input allocation:

$$\frac{\partial Q^d}{\partial i^d} = \frac{g(1+r)}{(p-\alpha w_2)} + \frac{g\varphi}{(p-\alpha w_2)\delta\frac{\partial U_2}{\partial C_2}} \tag{4}$$

If the credit constraint was not binding, $\varphi = 0$, the marginal product of input allocation would be lower and input allocation higher. The same effect holds for input allocation to Q^s, as well as for labour allocation to each of the crops.

4.3 Introducing uncertainty

When introducing uncertainty, the Lagrange becomes the following:

$$\begin{aligned}\mathcal{L} =& U_1(C_1) + \lambda(w_1(T_1 - l_1^d - l_1^s) - g(i^d + i^s) + B - C_1)) \\ &+ E[\delta U_2(C_2) + \mu[(p-\alpha w_2)(Q^d + Q^s) + w_2 T_2 - (1+r)B - C_2]] \\ &+ \varphi(B^m - B) \\ &+ \rho(1 - a^d - a^s)\end{aligned}$$

Note here that the household forms expectations not only about the utility he derives from consumption in period 2, but also about the level of consumption that can be achieved. When differentiating the Lagrange with respect to the choice variables, the optimal consumption rule becomes:

$$\frac{\partial U_1}{\partial C_1} = (1+r)\delta\frac{\partial EU_2}{\partial C_2} + \varphi \tag{5}$$

And the following decision rules for input allocation can be derived:

$$\frac{\partial Q^d}{\partial i^d} = \frac{g}{(p-\alpha w_2)}\frac{\frac{\partial U_1}{\partial C_1}}{\delta\frac{\partial EU_2}{\partial C_2}} \tag{6}$$

$$E[\frac{\partial Q^s}{\partial i^s}] = \frac{g}{(p-\alpha w_2)}\frac{\frac{\partial U_1}{\partial C_1}}{\delta\frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \frac{\partial Q^s}{\partial i^s})}{\frac{\partial EU_2}{\partial C_2}} \tag{7}$$

Eq. (6) shows the allocation rule for inputs to the safe crop. It looks similar to eq. (1), only now the household maximises expected utility of consumption in the harvesting period. Since for any expected consumption level C_2, expected utility $EU_2(C_2)$ is lower than utility of the expected value $U_2(E(C_2))$, marginal utility will be higher under uncertainty. This means that the right hand side term will be lower than in the deterministic case, implying that the household applies more inputs to the safe crop than he would in the absence of risk. The difference in input allocation between a scenario with no uncertainty and a scenario with uncertainty will be higher the lower consumption levels and the higher initial input allocation i^d (in terms of intensity per unit of land).

Eq. (7) shows the effect of risk on input allocation to the risky crop. Here the effect is less clear: again the expected utility is smaller than utility of the expected value, implying higher input allocation. However, the covariance between marginal utility of consumption and marginal product of variable inputs allocation is strictly negative and this term increases the value of the right hand side of the equation, reducing input allocation to the risky crop in the stochastic case.[10] Which of the two effects is stronger, depends on the degree of risk aversion of the household, expected consumption levels C_2 as well as on the coefficient of covariance between marginal utility and marginal product of inputs. Since the covariance will be larger, the lower wages in period 2 and the higher interest rate r, I am relatively confident that the net effect of uncertainty on input allocation will be negative in this context.

It can also be clearly seen that input allocation to the safe crop i^d will always be larger than input allocation to the risky crop i^s relative to their respective productivity. The first part of the right hand side of eq. (7) is equal to $\partial Q^d/\partial i^d$ and since the covariance term is strictly negative, $\partial Q^s/\partial i^s$ will always be greater than $\partial Q^d/\partial i^d$.

Again, equations (6) and (7) can be reformulated to include the credit constraint. Here for input allocation to the risky crop:

$$E[\frac{\partial Q^s}{\partial i^s}] = \frac{g(1+r)}{(p-\alpha w_2)} + \frac{g\varphi}{(p-\alpha w_2)\delta\frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \frac{\partial Q^s}{\partial i^s})}{\frac{\partial EU_2}{\partial C_2}} \tag{8}$$

We can see from eq. (8) that both constraints, the risk and credit constraint, go in the same direction: both reduce the allocation of inputs to the risky crop. More importantly, it also shows, that uncertainty reduces input allocation to the risky crop, even if credit constraints are not binding.

4.4 Potential impact of the NREGA

The effects of the National Rural Employment Guarantee Act on households' rationing can best be represented by an increase in agricultural wages in the second period. This holds for two reasons. First, a range of studies have shown that agricultural wages, particularly harvest wages, have increased due to NREGA (Azam 2012; Imbert and Papp 2012; Berg et al. 2012). Second, for many households with labour-surplus, the possibility to find employment in harvest periods is on other farmers fields, where wages tend to be low and in case of major weather shocks they have to expect

[10] In a bad state of the world, consumption would be lower the higher the input allocation to Q^s, thus the covariance between both marginal effects is negative.

that no employment can be found at all. Because the NREGA provides reliable income opportunities throughout the year, households can expect to be able to generate income in the harvest period at higher wages, in good years, because they have an alternative to working on other farmers fields at minimum wages, and in bad years, because they can find work at all (Jayachandran 2006). The comparative statics in this section show that the introduction of NREGA affects optimal input allocation differently under certainty than under uncertainty.

In the deterministic case, the optimal allocation of input to both crops is given by:

$$\frac{\partial Q^s}{\partial i^s} = \frac{g}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{9}$$

An increase of harvest period wages w_2, would affect optimal allocation through two channels and differently depending on whether the household hires labour in or out (net). First, for all households an increase in w_2 raises production costs, and thus should lower optimal output levels, as well as input allocation. Second, the increase in wages has an effect on consumption levels that can be realised in the second period by rotating the budget constraint. For households who hire labour in (net), the effect of increased wages on consumption in the harvest period is negative, however for those households who hire labour out, thus whose land is too small to produce at higher levels, the effect on consumption is positive. One would thus see a decrease in input allocation for net lenders of labour, because C_2 increases, therefore reducing $\partial U_2/\partial C_2$ and increasing the second part of the right hand side in our equation. Intuitively, the effect of increased wages levels through consumption levels can be understood as a substitution effect: because working in the harvest period becomes more profitable for households with little cultivated land, the allocation of inputs to those lands should decrease, from very high levels, to more efficient ones.

An entirely different effect can be observed if uncertainty reduces input allocation to risky crops as given by equation (10):

$$E[\frac{\partial Q^s}{\partial i^s}] = \frac{g}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \frac{\partial Q^s}{\partial i^s})}{\frac{\partial EU_2}{\partial C_2}} \tag{10}$$

If wages increase, we observe the same effects on cost of production as well as on marginal utility of consumption as in the deterministic case. However, under uncertainty the negative covariance term reduced input allocation to the risky crop and this effect is now partially offset by increasing wages. If income generating possibilities on the labour market increase, the effects of allocating a high share of inputs to the risky crop translate less strongly to reduced consumption in the harvesting period in bad states of the world.[11] Because the household knows that in case of negative production shocks, he can improve income by spending the time he otherwise would have spend with harvesting by working under NREGA, the possibilities to smooth income are strongly increased. The more the covariance term on the right had side of our equation approaches 0, the more the ratio of inputs allocated to the risky crop (vs. the safe crop) approaches the scenario under certainty. This means that even if total input (or similarly labour) allocation is reduced after wages increase, the share of total inputs allocated to each of the crops approaches the ratio in the deterministic scenario. Interestingly, this effect holds independently of whether credit constraints reduce total input allocation or not.

5 Estimation strategy

The key prediction of the model is that the introduction of the NREGA increases expected wages in the harvesting period, which, *ceteris paribus*, has a positive effect on the inputs allocated to riskier crops if households were previously constrained in crop choice by high levels of uncertainty regarding output levels and dysfunctional insurance markets. The model to be estimated would thus be:

$$i_{it}^s/(i_{it}^d + i_{it}^s) = \beta_0 + \beta_1 w_t^h + \beta_2 X_{it} + u_i + v_{it} \tag{11}$$

The dependent variable is the ratio of inputs allocated to risky crops. The key explanatory variable is harvest stage wages. In the empirical specification, w_2 is substituted by w^h because time subscripts are needed to distinguish between both rounds of interviews. Let X_{it} be a set of exogenous household characteristics affecting preferences and income, and u_i be time-constant unobserved household level characteristics, such as risk aversion, farming ability and land quality. Also, as described earlier, village level characteristics and changes will probably influence both: labour market wages and farmers' crop choice. Because these unobservable characteristics are assumed to be correlated with the dependent variable, as well as with the wage levels, households can achieve on local labour markets, estimating eq. (11) in OLS will yield biased results.

Instead, I argue, that the introduction of NREGA leads to exogenous increases in expected wages in the harvesting period and make use of the sequenced introduction of the NREGA at district level as identification strategy. I define those households as "treated" that live in districts in which NREGA was introduced in 2006, thus phase I districts. This allows to

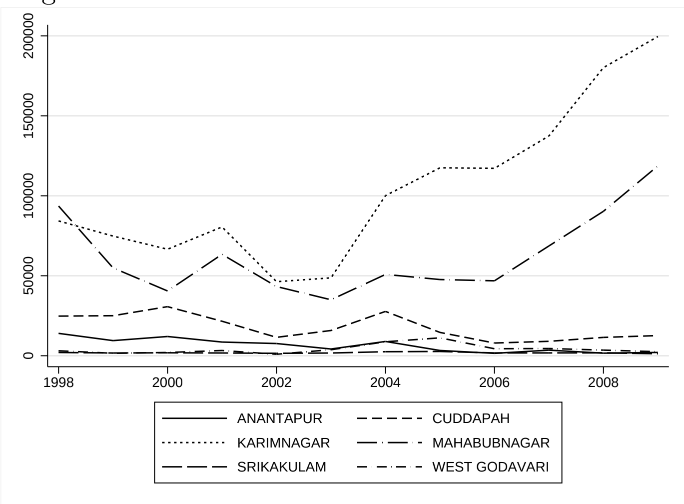
[11] Of course the covariance will never become zero, if the household assumed that he would be able to generate as much income on the local labour markets as from its own agricultural production, he would not engage in agricultural production at all.

ignore self selection and to account for the fact, that many households in rural Andhra Pradesh form expectations about income opportunities through NREGA, not only those who are already registered with the scheme. The equation to be estimated then becomes:

$$i^s_{it}/(i^d_{it} + i^s_{it}) = \beta_0 + \beta_1 D_{it} + \beta_2 X_{it} + u_i + \upsilon_{it} \tag{12}$$

Here, D_{it} represents the introduction of NREGA in a household's district.[12] In this model, β_1 is identified, if two assumptions are fulfilled: the parallel trend assumption and the assumption that treatment is not correlated with potential outcomes.

Figure 1: District-wise land use statistics for cotton



Source: Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, GoI

In this setting both assumptions seem to hold. The parallel trends assumption could be tested if the first round of the YLS data included information on crop choice, which unfortunately is not the case. Instead, I have to rely on the Land Use Statistics provided by the Ministry of Agriculture. Time series by district for cotton are displayed in Figure 1. Showing that treatment is not correlated with potential outcomes is also straightforward. Although, NREGA should have been introduced in the most backward districts first, general economic characteristics are not significantly different

[12]In the fixed effects model, the coefficient of D_{it} is only identified if D_{it} equals 0 for all households in the first period and 1 for some households in the second period. This might seem arbitrary because NREGA was introduced in the first districts from April 2006 onwards, thus before round 2 interviews were conducted. However, the questions in the survey relevant to our analysis all refer to decisions taken between July 2005 and May 2006, which suggests that at this point in time households could not yet have felt the effects of NREGA.

between treatment and control districts in this sample. Table 1 shows that GDP per capita in 2007 was almost equal between treatment and control districts.

In order to estimate equation (12), I combine matching methods with household fixed effects regression models. The main goal of combining both methods is to increase the comparability between households in the two groups, especially in order to account for geographical differences, while controlling for unobserved time-constant household and village level characteristics that might influence the outcome variable.

The first step matching method builds on entropy balancing as developed by Hainmueller (2012). Hainmueller's matching algorithm was shown to outperform most existing matching algorithms in terms of balance reached on the entire set of relevant covariates. Matching occurs on the mean of key variables, that could influence post-treatment outcomes, such as cost incurred in agricultural production, total cultivated area, percentage of area irrigated, fertilizer application, wealth levels and off-farm income, as well as on dummies indicating whether the household cultivated any cotton or chillies and turmeric in the past agricultural year. The resulting covariate balance is displayed in Table 2. Since I estimate the model on a balanced sample, the same weights can be applied to the 2009-10 round of interviews.

In the Difference-in-Difference (DID) scenario, combining matching with regression was suggested by Heckman, Ichimura and Todd (1998) and Abadie (2005).[13] In a cross-section, identification then relies on the conditional independence assumption. By combining matching and DID strategies, it is possible to additionally account for unobservable characteristics that might influence programme participation and potential outcomes. Causal effects can in such cases be isolated as long as the additive separability between observable and unobservable characteristics holds.

The validity of matching on the dependent variable is still under debate in the current literature.[14] By matching on lagged outcomes, one gives more weight to those households that already planted the crop it the first round. This is potentially problematic if those households are also more likely than other households to react to the introduction of NREGA by further increasing the share or inputs allocated to risky crops.

It seems important at this point in time to be specific about the causal effects that could result in such bias. It is plausible that the probability of planting a certain crop is influenced by past experience with the crop. Also,

[13] The household fixed effects estimator does not differ from the the DID estimator as long as we consider only two time periods.

[14] While Imai (2008) as well as Imbens and Wooldridge (2009) suggest that this should increase comparability of treatment and control groups, Lechner (2010) argues that matching on lagged outcomes is similar to including lagged dependent variables in the regression equation, which would violate the strict exogeneity assumptions made in fixed as well as in random effects models.

the share of this crop in the overall portfolio can be correlated with past experience, if the farmer subsequently allocates more land and inputs to the crop as he gets more familiar with it. The importance of such learning effects on input allocation, especially with regard to fertilizer adoption, has been highlighted *inter alia* by Besley and Case (1993) and Munshi (2004). Nonetheless, cotton production is not a novelty in rural Andhra Pradesh. In contrast, cotton has been produced for centuries and is common in most regions. Learning effects thus seem not to play a role in the production of cotton anymore. Since learning effects seem not to play a role, I conclude that the probability of planting cotton and the share of input allocated to this crop is influenced either by unobservables that are constant over time such as ability, land quality, risk aversion etc, or by past experience with crop output. Since weather-related outcomes could have either been very good or very bad, serial correlation due to past experience can be positive and negative and should not systematically bias the results. This is why I think that matching on crops planted in the pre-treatment period should only increase the comparability between treatment and non-treatment households in terms of product portfolio.

The second step of the estimation consist of applying a weighted fixed effects regression. In the estimated model, a time dummy is included to control for state wide changes in input and output prices, weather trends that are not captured by rainfall data and other changes at the state level that could influence farmers' crop choice.[15] I additionally control for self-reported shocks as well as mandal (block) level rainfall data.

The choice between fixed effects and random effects models relies mainly on assumptions made regarding the unobserved variables. The key assumption to be made is as to whether or not one assumes the unobserved heterogeneity to be correlated with the observed explanatory variables. In this case, it seems straightforward to assume that the key explanatory variable, D_{it}, is not correlated with individual heterogeneity. Thus both random and fixed effects models could be applied. However, random effects does not support probability weights in Stata, which is why I rely mainly on fixed effects estimation. The results in fixed and random effects models are very similar (not reported here).

Additionally to the fixed effect model, I also apply the fractional logit model suggested by Papke and Wooldridge (2008). The assumptions in their model are essentially the same as in fixed effects, however their estimator is able to take into account the boundedness of the dependent variable between 0 and 1.

[15] Note here, that it is not possible to control for time varying community or district level effects, because the treatment variable is the same for all households living in the same district. The inclusion of district-year dummies would thus unintentionally capture the effect of the introduction of NREGA.

6 Data

The model specified above is tested using the Young Lives Survey (YLS) data for Andhra Pradesh (India). The dataset covers 3019 households living in six different districts. Three rounds of interviews were conducted so far, in 2002, 2007 and 2009-10. Panel attrition is relatively low: in 2009-10, 2910 households could be revisited, which gives an attrition rate of 3.6% (Galab et al. 2011).

Selection of districts under the YLS ensured that all three geographical regions were represented in the survey as well as poor and non-poor districts of each region. Classification of districts was done along economic, human development and infrastructure indicators (Galab et al. 2011). This sample design ensures that the YLS is broadly representative for the population of Andhra Pradesh, although maybe not suited for monitoring of outcome indicators at the state level.

The current analysis is restricted to those households with non-zero agricultural production in 2007 and 2009-10, which reduces the dataset to 1118 households (2236 observations). Furthermore, only the second (2007) and third (2009-10) round are considered for reasons of comparability. Summary statistics of general household characteristics are reported in Table 3.

The vast majority of sampled households are headed by males. Table 3 also reveals, that the average household consists of six members, whose head is around 41 years old. Schooling levels are somewhat higher for men than for women. Typically, the best educated male household member achieved 6.3 years of schooling, while the best educated woman had on average 1.9 years less schooling. Most households generate income from both: own farming and off-farm activities, with income from off-farm activities being slightly higher than income from own agricultural production on average.

Table 4 reports summary statistics of farming characteristics. Paddy rice is by far the most popular crop: 57% of the households planted at least some paddy in 2007 and 62% in 2009. It is followed by Grams and Pulses (27% in 2007) and Groundnuts (26% in 2007). Households irrigated around 17% of their land in the dry season. Fertilizer adoption is relatively high (94% in both rounds), as is the use of high yielding variety (HYV) seeds (73% in 2007). Unfortunately, the data do not provide any information about the quantity of fertilizer or seeds applied. The only information available is the total expenditure on variable inputs per crop. This variable includes expenditure for seeds, fertilizer and pesticides. The total amount in logs as well as the share in total expenditure on variable inputs for each crop are reported in Table 4. The highest share of inputs is allocated to paddy rice, which is not surprising, given the predominance of this crop in the sample. A large share of inputs is also allocated to groundnuts (19% in 2007). Interestingly, the share of inputs allocated to cotton has increased considerably from 2007 (9%) to 2009 (14%).

Table 5 reports the occurence of different shocks in both groups and both periods. Although lagged rainfall is lower in 2007 in the treatment group than in the control group and more households report having experienced drought, this is reverted between 2007 and 2009, with similar trends for treatment and control groups.

In order to identify viable strategies for households to improve their income from agricultural production, I estimate a standard Cobb-Douglas production function, linking the total value of agricultural production to input allocation, plot size and the relevance of different types of crops in a household's product portfolio. The equation is estimated as random effects and fixed effects models in order to account for unobserved heterogeneity that could bias the estimates. In this case, both models provide similar results.

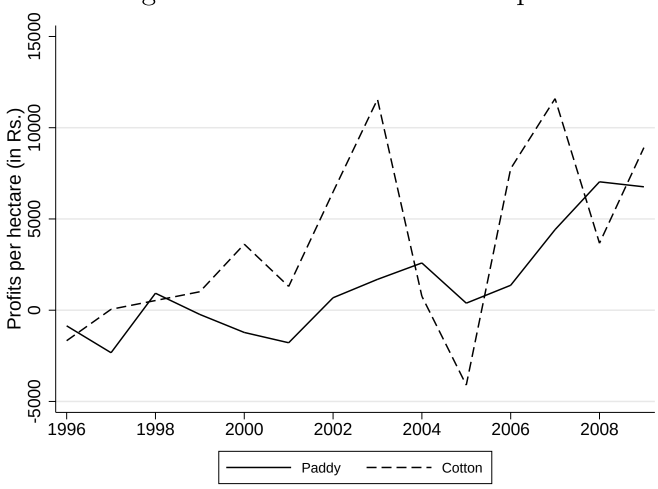
Table 6 shows, that the most important determinant of agricultural output is the level of inputs applied. Additionally to inputs allocated, total area cultivated and the share of area that is under irrigation are crucial determinants of output levels. The dummies indicating whether or not the household applied fertilizer or high yielding variety (HYV) seeds are both not statistically significant. This might seem somewhat surprising, but since the expenditure for fertilizer and seeds is included in variable inputs, one should not attribute to much weight to this finding. Finally, Table 6 reveals that households could significantly improve the value of their agricultural production if they increased the share of inputs allocated to cotton, chillies, turmeric or to other commercial crops relative to food crops.[16] Producing fruits could also lead to considerably higher income from agricultural production. In contrast, producing a higher share of oilseeds or groundnuts would reduce the total value of agricultural production.

Because looking only at two time periods might lead to misleading conclusions, I additionally consider the Cost of Cultivation Statistics for major crops in India. Unfortunately, these statistics are only available at state level and only for very few crops. Figure 2 reports average profits per hectare in Andhra Pradesh between 1996 and 2009 for cotton and paddy rice. Cotton clearly outperforms paddy for most years, but in some years profits from cotton cultivation are much smaller than profits from rice cultivation. Thus, households planting cotton generate higher profits on average, although it seems that returns are also more variable.

Understanding the risk involved in producing higher profitability crops is critical for the analysis. I therefore compute the variance of yield per acre of different crops. The data for this come from the District-wise Crop Production Statistics of the Ministry of Agriculture. Figure 3 shows that

[16] Foodgrains were used as reference category in the estimation. Commercial crops include coffee, tobacco, sugarcane, flowers, eucalyptus, ginger, garlic, black pepper and other spices.

Figure 2: Profits of selected crops



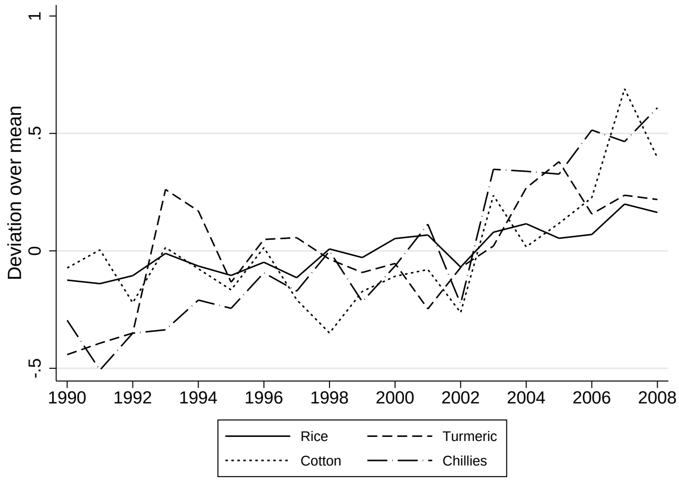
Source: Cost of Cultivation Statistics, Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, GoI

yields of commercial crops such as cotton, chillies and turmeric deviate much more from their mean than paddy rice. I only consider paddy as foodcrop here, but yield fluctuations of other foodcrops are very similar to paddy. It thus seems that risk considerations might be a major factor in explaining the relatively limited production of cotton and other commercial crops in the sample.

Another explanation why households do not cultivate more profitable crops, could be differences in costs involved in producing these crops. If paid-out costs are higher for commercial crops, many poor farmers might be constrained by the limited availability and relatively high cost of credit. Again, the cost of cultivation statistics of paddy rice and cotton can be helpful to understand the importance of paid-out costs. Table 7 reports the cost of cultivating 1 hectare of cotton vs. paddy in 2006-07 and 2009-10, respectively. It also reports the average share of different items in total expenditures. Table 7 clearly reveals that the cultivation of cotton does not imply higher costs per hectare than the cultivation of paddy rice.

Finally, differential developments in prices between districts could drive observed effects on the total production of certain crops. Figure 4 displays the district wise development of nominal prices of paddy vs. commercial crops such as cotton, chillies and turmeric. Cotton prices changed considerably during the period of reference. However, while farm harvest prices were traditionally somewhat higher in the treatment districts than in the control districts, prices increased more pronouncedly in the control districts during the period of interest. In general, it seems as if price trends of most crops are very similar across districts, which suggests that other factors must be

Figure 3: Yield of selected crops in Andhra Pradesh



Source: District wise crop production statistics, Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, GoI

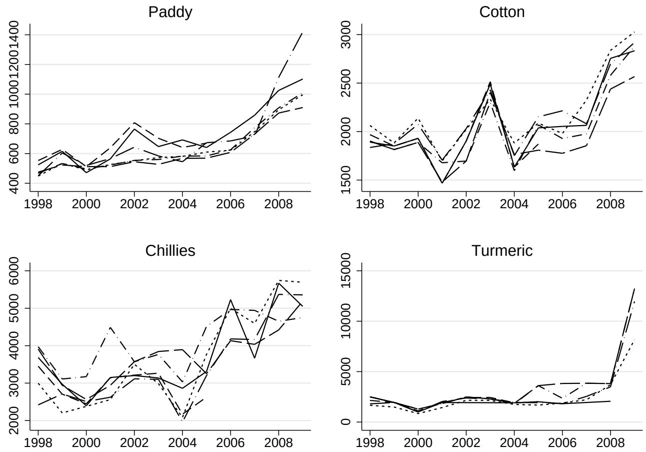
driving observed changes in production trends.

7 Results

If risk is a relevant constraint for households' production decisions, the provision of an employment guarantee should enable households to grow a higher share of risky crops. This is the main outcome of the theoretical model presented in this paper. Central to this model is the expectation each household forms about future income generating opportunities. These expectations will to a large extend depend on each household's experience with the National Rural Employment Guarantee Scheme. So, it can be expected that the NREGA can have an insurance effect only if provision of work is sufficiently able to cover households demand and more importantly to react to increases in demand for work in case of shocks. Thus, it depends on the reliability of the programme in the treatment districts, whether effects of the employment guarantee on production decision and particularly on the choice of crops can be observed.

Therefore, I first test if the employment guarantee holds. Because one can expect households' demand for work to sharply increase in case of shocks, I test whether the number of days a household works with NREGA responds to the realisation of such shocks. In particular, I estimate if the deviation of cumulative rainfall in the agricultural year from its mean, as well as self reported shocks on the household level, drive changes in the number of days worked under NREGA. The model is estimated as fixed effects model in order to account for fixed household and village level characteristics. At

Figure 4: District-wise farm harvest prices of selected crops



Source: Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, GoI

household level the most important fixed characteristics to influence work take-up are proximity to the work sites, wealth levels, education as well as unobservables such as ability, personal networks and access to other risk management tools. At the village level, the fixed effects model captures time-constant effects such as differential commitment of local officials, that could drive differences in the number of workdays generated between districts and communities. Because rainfall is a covariate shock, that affects households living in the same village to similar extend, standard errors are clustered at community level. The estimation is also restricted to phase one districts, thus only households who experienced NREGA in both survey rounds are considered. Results are reported in Table 8.

The results show that the number of days worked under NREGA changes considerably with variation in rainfall levels. It does so most strongly for lagged rainfall levels, thus cumulative rainfall in the agricultural season preceding the period of reference. The coefficient of the lagged rainfall variable is large and negative, which implies that households worked more under NREGA if lagged rainfall levels were below the mean and that they worked less if rainfall was above the mean. This supports the assumption of NREGA acting as insurance, because households use the programme to smooth income ex-post, after harvest is made and after agricultural products are sold. Similar evidence has been provided by Johnson (2010), who finds that the number of days worked under NREGA increases if rainfall levels are below normal.

Table 8 also reveals how important maturation of the programme is. A large share of changes in days worked under NREGA can be explained by time alone. In contrast, wealth levels seem not to be valid predictors of the number of days households work under NREGA. The size of cultivated area is also not statistically significant in the fixed effects model, but this can probably be attributed to limited variation of this variable over time. If estimated in random effects, this variable is positive and statistically significant at the 5% level.[17]

Interestingly, self-reported shocks seem not to influence the dependent variable either. The variables capturing different self-reported shocks are mostly not statistically significant, not even when combined in one single index. This might be so for three reasons. First, self-reported shocks refer to any shock in the 4 years previous to the date of interview, which might be simply to imprecise to capture the effect of shocks on individual labour allocation.[18] The second reason could be potentially more problematic. Because the number of days a household works with NREGA does not only depend on each household's demand for work, but also on the provision of work, it may be that the provision of work reacts to major covariate shocks such as droughts whereas it does not respond to individual changes in demand for work after idiosyncratic shocks occurred. It is thus possible that households would have liked to smooth income after idiosyncratic shocks occurred by working under the scheme, but that provision of labour did not react to this sufficiently. Thirdly, one could also imagine, that households rely more strongly on informal risk coping strategies for idiosyncratic shocks. All explanations presented above are possible and more data would certainly be necessary to generate more reliable insights to this question.

Nonetheless, it seems that employment generation under NREGA in the sample districts does sufficiently respond to labour demand as to allow households to smooth consumption in case of weather related shocks. Since a large share of agricultural production is rain fed in rural India and in Andhra Pradesh in particular, rainfall fluctuations are among the most important sources of risk for rural households. This is why I feel confident to conclude that the NREGA can indeed have an insurance function for rural households in Andhra Pradesh.

Given that this precondition holds, the introduction of NREGA should affect farmers' crop choice if risk constraints are relevant in their production decisions. I test equation (12) in a linear fixed effects model. Table 9 reports the effects of the introduction of the NREGA on households' share of inputs allocated to cotton. Because changes observed in one household could be correlated with observed changes in other households living in the same

[17]A positive coefficient could be a sign for programme capture of wealthier households. But further investigating this is beyond the scope of the paper

[18]Additional difficulties to estimate coefficients correctly could be the relatively low sample size combined with the total amount of different indicator variables included: Households are asked to report on a total of 30 different shocks, of which 11 are included in the estimation.

community, I cluster standard errors again at the community level, which gives a total of 83 clusters.

The results clearly show a positive and statistically significant effect of the introduction of the NREGA on households' crop choice. In nominal terms, the results suggest that the share of inputs allocated to cotton is 9% higher for households living in disctricts that introduced NREGA in 2006. Considering that the average share of inputs allocated to cotton did not exceed 9% in 2006, the magnitude of this effect is striking. The effect is also robust to the inclusion of a range of control variables, such as total inputs applied, area cultivated, wealth levels, household off-farm income, actual and lagged rainfall levels and area under irrigation. The magnitude of the effect is even larger when weighting the sample with entropy balancing weights as described in Section 5.

Because the distribution of the dependent variable is far from a normal distribution, I additionally fit a fractional logit model as proposed by Papke and Wooldridge (2008). Because coefficients of logit models are difficult to interpret, I also compute average partial effects (APE), which for the explanatory variable of interest would just be the average treatment effect (ATE) given that it is a binary variable. Table 10 shows that the estimated effects are even larger when the empirical model is able to account for the boundedness of the dependent variable.

Of course one might be concerned that the observed effects are driven by other factors than the reduction in risk exposure attributed to NREGA. I cannot fully exclude the possibility of unobserved time-varying factors driving these results. Nonetheless, I can try to account for as much potential confounders as possible. In Section 6, the possibility of heterogeneous district level changes in farm harvest prices driving these results could already be excluded. Additionally, I conduct cluster level regressions in order to assess the importance of outliers. If I exclude the two most dominant cluster from the analysis, the observed effects is still positive and statistically significant at the 10% level (not reported here).

In order control whether the results might be driven by district level changes in the profitability of cotton cultivation, I additionally test if the NREGA affects allocation of inputs to other risky crops besides cotton. Table 11 reports the determinants of the share of inputs allocated to cotton, chillies and turmeric as well as to total commercial crops. Again the estimated effects are positive and statistically significant. Effects are larger for the subset of cotton, chillies and turmeric only, which might seem arbitrary. However, while we know that cotton, chilies and turmeric are more volatile in yields per hectare and therefore riskier than food crops (as was shown in Section 6), very little information exists with regards to yield fluctuations, cost of cultivation etc. of other commercial crops. Therefore, it is difficult to assess whether risk considerations are relevant constraints to the cultivation of these crops.

Finally, I test whether households who report to be working with NREGA change their input allocation decision. I find that households who registered with NREGA in 2006-07 already are more likely to grow risky crops such as cotton. In contrast, households who registered only later seem not to have changed their input allocation yet (Table 12). I perform the same analysis for phase I districts and early registrants only, as reported in Table 13. In that specification, I cannot cluster standard errors at community level, because the number of clusters is too low. I still find positive and statistically significant effects.

Given the size of the estimates and number of robustness checks, I feel confident to conclude that NREGA seems indeed to reduce households' risk exposure and thereby enables them to grow more profitable but at the same time riskier crops. Recall, that the introduction of the NREGA can influence production decision not only by reducing risk exposure but also by increasing available income that can be spend on inputs. But because changes in crop choice are robust to levels of input allocation and to changes in income from off-farm activities, I feel confident to attribute these changes to reductions in risk exposure rather than to changes in income.

8 Conclusions

This paper assesses the role of risk constraints in households' production decisions. It provides theoretical and empirical evidence that the right to work, as households have been entitled to under the NREGA, reduces households' exposure to a range of risks by guaranteeing income opportunities in areas and time periods where they previously did not exist. This paper also provides evidence that risk constrained households choose suboptimal production strategies. The introduction of the NREGA enables them to generate higher incomes from agricultural production by switching their production towards higher profitability products. With this finding, the paper provides empirical evidence for the validity of the theory of choice under uncertainty as much as it contributes important evidence to the ongoing debate on the effects of the NREGA on agricultural productivity.

The effects of the NREGA on households' production and investment decisions are similar to those that were attributed to the provision of insurance by Karlan et al. (2012). But in contrast to purchasing insurance, the registration with NREGA provides little ex-ante cost. Since trust-related considerations continue to limit the take-up of insurance products in many countries, providing public works schemes could be better suited to protect households against agricultural production risks. One should not forget however, that the take-up of work is always associated with opportunity costs. In countries or regions with well functioning off-farm labour markets, providing public works might not be necessary. A food-for-work programme or cash-for-work programme can always only be effective in areas and time-periods where labour is in surplus. Also, public works programmes such as the NREGA are in no way able to protect households against many other risks than agricultural production risks, particularly against risks that reduce a household's member's ability to work such as sickness and work accidents.

The results of this paper reveal that public works programmes can have welfare effects that go much beyond the immediate effects of employment generation. This finding entails important lessons learned for the ongoing debates on the effectiveness of the NREGA as well as for other countries considering the implementation of such schemes.

First, for the insurance effect to unfold, the design of the public works programme is crucial. An employment guarantee that is entitled by law combined with adequate grievance redressal mechanisms provides households with the necessary protection against agricultural production risks that enables them to take more risk in their production and investment decisions. Additionally, it is crucial not to limit the number of workdays too drastically, otherwise the potential of such a scheme as risk coping strategy cannot be realised.

Second, implementation matters. The data analysed in this paper cover only the state of Andhra Pradesh. This is *inter alia* because the performance of the NREGA in terms of number of workdays generated per eligible household varies immensely across states and even across districts in India. Andhra Pradesh is one of the best performing states and it goes without saying that many of the effects captured in this paper might not unfold in all Indian states.

The paper contributes to ongoing debates in and outside India on the effects of the NREGA on agricultural productivity in the country. Current discussions are mainly driven by considerations regarding the trade-off between providing minimum income to poor households on one side and ensuring that production costs in the agricultural sector do not rise to drastically because of increased agricultural wages on the other side. As this paper shows, these discussions are missing one important aspect. Because the number of workdays each household is entitled to additionally affects its risk managing capacity, the amount of risk each household is willing to take will depend on the number of workdays he can expect to be able to work in case of production shocks. Considerations about decreasing agricultural productivity levels in the country so far have mainly ignored this fact. The assumption that only large scale farmers can drive increases in agricultural productivity is still mainstream. Including the effects of NREGA on households' risk exposure and the resulting changes in production decisions in the discussion might change the overall picture.

Obviously a range of open questions remain and much more research is needed to get conclusive answers to these questions. First, internal and external validity of the results could be improved with more data, especially if the analysis could be extended to the whole country. Second, effects of the programme on total levels of inputs applied as well as on investments in fixed capital could be very interesting topics to study. Similarly, effects on households' willingness to take-up entrepreneurial activity need to be assessed. Third, effects of the programme are likely to be heterogeneous. With more data it would be possible to assess to which extent the programme effects vary between different subgroups of the population, particularly with size of land holdings.

Acknowledgments

The author is grateful to Michael Grimm for continuous guidance and advice, as well as to Markus Loewe and Tilman Altenburg for many helpful comments and suggestions. The author takes full responsibility for all errors. Funding of the German Federal Ministry for Economic Cooperation and Development is gratefully acknowledged. The data used in this publication come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh), Peru and Vietnam (www.younglives.org.uk). Young Lives is core-funded by UK aid from the Department for International Development (DFID) and co-funded from 2010 to 2014 by the Netherlands Ministry of Foreign Affairs. The views expressed here are those of the author. They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.

References

Abadie, Alberto. 2005. "Semiparametric difference-in-differences estimators." *The Review of Economic Studies*, 72: 1–19.

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

Aiyar, Yamini, and Salimah Samji. 2009. *Transparency and Accountability in NREGA: A Case Study of Andhra Pradesh.* Vol. 1 of *AI Working Paper*, New Delhi.

Azam, Mehtabul. 2012. *The impact of Indian job guarantee scheme on labor market outcomes: Evidence from a natural experiment.* Vol. 6548 of *Discussion Paper*, Bonn.

Basu, Arnab K. 2013. "Impact of Rural Employment Guarantee Schemes on Seasonal Labor Markets: Optimum Compensation and Workers' Welfare." *Journal of Economic Inequality*, 11(1): 1 – 34.

Berg, Erlend, Sambit Bhattacharyya, Rajasekhar Durgam, and Manjula Ramachandra. 2012. *Can rural public works affect agricultural wages? Evidence from India.* Vol. 2012-05 of *CSAE Working Papers*, Oxford.

Besley, Timothy, and Anne Case. 1993. "Modeling technology adoption in developing countries." *The American Economic Review*, 83(2): 396–402.

Cole, Shawn, Xavier Gine, Jeremy Tobacman, Robert Townsend, Petia Topalova, and James Vickery. 2012. "Barriers to household risk management: evidence from India." 5(1): 104–135.

Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a new technology: Pineapple in Ghana." *The American Economic Review*, 100(1): 35–69.

Dercon, Stefan. 1996. "Risk, crop choice, and savings: Evidence from Tanzania." *Economic Development and Cultural Change*, 44(3): 485–513.

Dercon, Stefan, and Luc Christiaensen. 2011. "Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia." *Journal of Development Economics*, 96(2): 159–173.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2008. "How high are rates of return to fertilizer? Evidence from field experiments in Kenya." *The American economic review*, 98(2): 482–488.

Dutta, Puja, Rinku Murgai, Martin Ravallion, and Dominique van den Walle. 2012. *Does India's Employment Guarantee Scheme Guarantee Employment?* Vol. 6003 of *Policy Research Working Paper*.

Fafchamps, Marcel. 1993. "Sequential labor decisions under uncertainty: An estimable household model of West-African farmers." *Econometrica*, 61(5): 1173–1197.

Fafchamps, Marcel, and John Pender. 1997. "Precautionary Saving, Credit Constraints, and Irreversible Investment: Theory and Evidence From Semiarid India." *Journal of Business & Economic Statistics*, 15(2): 180–194.

Foster, Andrew D., and Mark R. Rosenzweig. 1996. "Technical change and human-capital returns and investments: evidence from the green revolution." *The American economic review*, 86(4): 931–953.

Foster, Andrew D., and Mark R. Rosenzweig. 2010. "Microeconomics of technology adoption." *Annual Review of Economics*, 2(1): 395–424.

Galab, S., S. Vijay Kumar, P. Prudvikhar Reddy, Renu Singh, and Uma Vennam. 2011. *The Impact of Growth on Childhood Poverty in Andhra Pradesh: Initial Findings from India.*

Gine, Xavier, and Stefan Klonner. 2006. *Credit Constraints as a Barrier to Technology Adoption by the Poor.* Vol. 2006/104 of *Research Paper.*

Grimm, Michael, Renate Hartwig, and Jann Lay. 2011. *Investment Decisions of Small Entrepreneurs in a Context of Strong Sharing Norms. Working Paper.*

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1998. "Matching as an econometric evaluation estimator." *The Review of Economic Studies*, 65(2): 261–294.

Imai, Kosuke. 2008. *Causal Inference with Repeated Measures in Observational Studies. Causal Inference Lecture Notes.*

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent developments in the econometrics of program evaluation." *Journal of Economic Literature*, 47(1): 5–86.

Imbert, Clement, and John Papp. 2012. *Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee. Working Paper.*

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

Jha, Raghbendra, Raghav Gaiha, and Manoj K. Pandey. 2012. "Net transfer benefits under India's rural employment guarantee scheme." *Journal of Policy Modeling*, 34(2): 296–311.

Johnson, Doug. 2010. *Can Workfare Serve as a Substitute for Weather Insurance? The Case of NREGA in Andhra Pradesh.* Vol. 32 of *Working Paper*, Chennai.

Karlan, Dean, Robert Osei, Isaak Osei-Akoto, and Christopher Udry. 2012. *Agricultural Decisions after Relaxing Credit and Risk Constraints.* Vol. 23 of *ILO Research Paper.*

Lechner, Michael. 2010. "The Estimation of Causal Effects by Difference-in-Difference Methods. Estimation of Spatial Panels." *Foundations and Trends in Econometrics*, 4(3): 165–224.

MoRD-GoI, (Ministry of Rural Development-Government of India). 2012. *MGNREGA Sameeksha: An Anthology of Research Studies on the Mahatma Gandhi National Rural Employment Guarantee Act, 2005. 2006-2012.* New Delhi:Orient BlackSwan.

Munshi, Kaivan. 2004. "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution." *Journal of Development Economics*, 73(1): 185–213.

Pankaj, Ashok K., and Rukmini Tankha. 2010. "Empowerment Effects of the NREGS on Women Workers: A Study in Four States." *Economic and Political Weekly*, XLV(30): 45–55.

Raabe, Katharina, Regina Birner, Madhushree Sekher, K.G. Gayathridevi, Amrita Shilpi, and Eva Schiffer. 2010. *How to overcome the governance challenges of implementing NREGA.* Vol. 00963 of *IFPRI Discussion Paper.*

Rosenzweig, Mark R., and Hans P. Binswanger. 1993. "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." *Economic Journal*, 103(416): 56–78.

Rosenzweig, Mark R., and Kenneth I. Wolpin. 1993. "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India." *Journal of Political Economy*, 101(2): 223–244.

Suri, Tavneet. 2011. "Selection and Comparative Advantage in Technology Adoption." *Econometrica*, 79(1): 159–209.

Van Den Berg, Marrit. 2002. "Do public works decrease farmers' soil degradation? Labour income and the use of fertilisers in India's semi-arid tropics." *Environment and Development Economics*, 7(3): 487–506.

Wadood, Syed N., and Russell L. Lamb. 2006. *Choice of Crops and Employment Uncertainty in the Off-farm Labor Market.* Vol. 10779 of *MPRA Paper.*

A Mathematical Appendix

A.1 Deterministic Case

In the deterministic case, the Lagrange can be summarised as follows:

$$
\begin{aligned}
\mathcal{L} =& U_1(C_1) + \delta U_2(C_2) \\
&+ \lambda(w_1(T_1 - l_1^d - l_1^s) - g(i^d + i^s) + B - C_1) \\
&+ \mu[(p - \alpha w_2)(Q^d + Q^s) + w_2 T_2 - (1 + r)B - C_2] \\
&+ \varphi(B^m - B) \\
&+ \rho(1 - a^d - a^s)
\end{aligned}
$$

Differentiating the Lagrange with respect to the choice variables, leads to the following first order conditions:

$$\frac{\partial \mathcal{L}}{\partial C_1} = \frac{\partial U_1}{\partial C_1} - \lambda = 0 \tag{A.1}$$

$$\frac{\partial \mathcal{L}}{\partial C_2} = \delta \frac{\partial U_2}{\partial C_2} - \mu = 0 \tag{A.2}$$

$$\frac{\partial \mathcal{L}}{\partial l_1^d} = -\lambda w_1 + \mu(p - \alpha w_2)\frac{\partial Q^d}{\partial l_1^d} = 0 \tag{A.3}$$

$$\frac{\partial \mathcal{L}}{\partial l_1^s} = -\lambda w_1 + \mu(p - \alpha w_2)\frac{\partial Q^s}{\partial l_1^s} = 0 \tag{A.4}$$

$$\frac{\partial \mathcal{L}}{\partial i^d} = -\lambda g + \mu(p - \alpha w_2)\frac{\partial Q^d}{\partial i^d} = 0 \tag{A.5}$$

$$\frac{\partial \mathcal{L}}{\partial i^s} = -\lambda g + \mu(p - \alpha w_2)\frac{\partial Q^s}{\partial i^s} = 0 \tag{A.6}$$

$$\frac{\partial \mathcal{L}}{\partial a^d} = \mu(p - \alpha w_2)\frac{\partial Q^d}{\partial a^d} - \gamma = 0 \tag{A.7}$$

$$\frac{\partial \mathcal{L}}{\partial a^s} = \mu(p - \alpha w_2)\frac{\partial Q^s}{\partial a^s} - \gamma = 0 \tag{A.8}$$

$$\frac{\partial \mathcal{L}}{\partial B} = \lambda - \mu(1 + r) - \varphi = 0 \tag{A.9}$$

Rearranging the first order conditions (A.1) and (A.2) gives:

$$\lambda = \frac{\partial U_1}{\partial C_1} \tag{A.10}$$

$$\mu = \delta \frac{\partial U_2}{\partial C_2} \tag{A.11}$$

And including (A.10) and (A.11) into (A.3)-(A.9) gives our decision rules:

$$w_1 \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial Q^d}{\partial l_1^d} = 0 \Leftrightarrow \frac{\partial Q^d}{\partial l_1^d} = \frac{w_1}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.12}$$

$$w_1 \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial Q^s}{\partial l_1^s} = 0 \Leftrightarrow \frac{\partial Q^s}{\partial l_1^s} = \frac{w_1}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.13}$$

$$g \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial Q^d}{\partial i^d} = 0 \Leftrightarrow \frac{\partial Q^d}{\partial i^d} = \frac{g}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.14}$$

$$g \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial Q^s}{\partial i^s} = 0 \Leftrightarrow \frac{\partial Q^s}{\partial i^s} = \frac{g}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.15}$$

$$\frac{\partial Q^d}{\partial a^d} = \frac{\partial Q^s}{\partial a^s} \tag{A.16}$$

$$\varphi = \frac{\partial U_1}{\partial C_1} - \delta(1 + r)\frac{\partial U_2}{\partial C_2} \tag{A.17}$$

A.2 Stochastic Case

When introducing uncertainty, the Lagrange becomes the following:

$$\begin{aligned}
\mathcal{L} =& U_1(C_1) + \lambda(w_1(T_1 - l_1^d - l_1^s) - g(i^d + i^s) + B - C_1)) \\
&+ E[\delta U_2(C_2) + \mu[(p - \alpha w_2)(Q^d + Q^s) + w_2 T_2 - (1 + r)B - C_2]] \\
&+ \varphi(B^m - B) \\
&+ \rho(1 - a^d - a^s)
\end{aligned}$$

Note here that the household forms expectations not only about the utility he derives from consumption in period 2, but also about the level of consumption that can be achieved. Differentiating the Lagrange with respect to the choice variables, leads to the following first order conditions:

$$\frac{\partial \mathcal{L}}{\partial C_1} = \frac{\partial U_1}{\partial C_1} - \lambda = 0 \tag{A.18}$$

$$\frac{\partial \mathcal{L}}{\partial C_2} = E[\delta \frac{\partial U_2}{\partial C_2} - \mu] = 0 \tag{A.19}$$

$$\frac{\partial \mathcal{L}}{\partial l_1^d} = -\lambda w_1 + E[\mu](p - \alpha w_2)\frac{\partial Q^0}{\partial l_1^0} = 0 \tag{A.20}$$

$$\frac{\partial \mathcal{L}}{\partial l_1^s} = -\lambda w_1 + E[\mu(p - \alpha w_2)\frac{\partial Q^1}{\partial l_1^1}] = 0 \tag{A.21}$$

$$\frac{\partial \mathcal{L}}{\partial i^d} = -\lambda g + E[\mu](p - \alpha w_2)\frac{\partial Q^d}{\partial i^s} = 0 \tag{A.22}$$

$$\frac{\partial \mathcal{L}}{\partial i^s} = -\lambda g + E[\mu(p - \alpha w_2)\frac{\partial Q^s}{\partial i^d}] = 0 \tag{A.23}$$

$$\frac{\partial \mathcal{L}}{\partial a^0} = E[\mu](p - \alpha w_2)\frac{\partial Q^0}{\partial a^0} - \gamma = 0 \tag{A.24}$$

$$\frac{\partial \mathcal{L}}{\partial a^1} = E[\mu(p - \alpha w_2)\frac{\partial Q^1}{\partial a^1}] - \gamma = 0 \tag{A.25}$$

$$\frac{\partial \mathcal{L}}{\partial B} = \lambda - E[\mu](1 + r) - \varphi = 0 \tag{A.26}$$

Rearranging (A.18) and (A.19) gives:

$$\lambda = \frac{\partial U_1}{\partial C_1} \tag{A.27}$$

$$E[\mu] = \delta \frac{\partial EU_2}{\partial C_2} \tag{A.28}$$

And the optimal consumption rule becomes:

$$\frac{\partial U_1}{\partial C_1} = (1 + r)\delta \frac{\partial EU_2}{\partial C_2} + \varphi \tag{A.29}$$

Including (A.27) and (A.28) into (A.20)-(A.25) gives our decision rules for l_1^d,

$$w_1 \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2)\delta \frac{\partial EU_2}{\partial C_2}\frac{\partial Q^d}{\partial l_1^d} = 0$$

$$\Leftrightarrow \frac{\partial Q^d}{\partial l_1^d} = \frac{w_1}{(p - \alpha w_2)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial EU_2}{\partial C_2}} \tag{A.30}$$

for l_1^s,

$$\begin{aligned}
&w_1\frac{\partial U_1}{\partial C_1}-(p-\alpha w_2)\delta E[\frac{\partial U_2}{\partial C_2}\frac{\partial Q^s}{\partial l_1^s}]=0\\
&\Leftrightarrow (p-\alpha w_2)\delta[\frac{\partial EU_2}{\partial C_2}E[\frac{\partial Q^s}{\partial l_1^s}]+cov(\frac{\partial U_2}{\partial C_2},\frac{\partial Q^s}{\partial l_1^s})]=w_1\frac{\partial U_1}{\partial C_1}\\
&\Leftrightarrow E[\frac{\partial Q^s}{\partial l_1^s}]=\frac{w_1}{(p-\alpha w_2)}\frac{\frac{\partial U_1}{\partial C_1}}{\delta\frac{\partial EU_2}{\partial C_2}}-\frac{cov(\frac{\partial U_2}{\partial C_2},\frac{\partial Q^s}{\partial l_1^s})}{\frac{\partial EU_2}{\partial C_2}} \qquad \text{(A.31)}
\end{aligned}$$

for i^d,

$$\begin{aligned}
&g\frac{\partial U_1}{\partial C_1}-(p-\alpha w_2)\delta\frac{\partial EU_2}{\partial C_2}\frac{\partial Q^d}{\partial i^d}=0\\
&\Leftrightarrow \frac{\partial Q^d}{\partial i^d}=\frac{g}{(p-\alpha w_2)}\frac{\frac{\partial U_1}{\partial C_1}}{\delta\frac{\partial EU_2}{\partial C_2}} \qquad \text{(A.32)}
\end{aligned}$$

for i^s,

$$\begin{aligned}
&g\frac{\partial U_1}{\partial C_1}-(p-\alpha w_2)\delta E[\frac{\partial U_2}{\partial C_2}\frac{\partial Q^s}{\partial i^s}]=0\\
&\Leftrightarrow (p-\alpha w_2)\delta[\frac{\partial EU_2}{\partial C_2}E[\frac{\partial Q^s}{\partial i^s}]+cov(\frac{\partial U_2}{\partial C_2},\frac{\partial Q^s}{\partial i^s})]=g\frac{\partial U_1}{\partial C_1}\\
&\Leftrightarrow E[\frac{\partial Q^s}{\partial i^s}]=\frac{g}{(p-\alpha w_2)}\frac{\frac{\partial U_1}{\partial C_1}}{\delta\frac{\partial EU_2}{\partial C_2}}-\frac{cov(\frac{\partial U_2}{\partial C_2},\frac{\partial Q^s}{\partial i^s})}{\frac{\partial EU_2}{\partial C_2}} \qquad \text{(A.33)}
\end{aligned}$$

for a^d,

$$\delta\frac{\partial EU_2}{\partial C_2}(p-\alpha w_2)\frac{\partial Q^d}{\partial a^d}=\gamma$$

and a^s,

$$\begin{aligned}
&\delta E[\frac{\partial U_2}{\partial C_2}\frac{\partial Q^s}{\partial a^s}](p-\alpha w_2)=\gamma\\
&\Leftrightarrow \delta(p-\alpha w_2)\frac{\partial EU_2}{\partial C_2}E[\frac{\partial Q^s}{\partial a^s}]+cov(\frac{\partial U_2}{\partial C_2},\frac{\partial Q^s}{\partial a^s})=\gamma
\end{aligned}$$

resulting in:

$$\frac{\partial EU_2}{\partial C_2}E[\frac{\partial Q^s}{\partial a^s}]+cov(\frac{\partial U_2}{\partial C_2},\frac{\partial Q^s}{\partial a^s})=\frac{\partial EU_2}{\partial C_2}\frac{\partial Q^d}{\partial a^d} \qquad \text{(A.34)}$$

B Tables

Table 1: District-level statistics

		Treatment	Control
GDP per capita (2006-07)		783487	776179.5
Rural population (2001 census)		80.54	84.64
SC/ST population (2001 census)		20.50	18.36
Literacy rate (2001 census)		54.6	64.4
Cropping Intensity (2007-08)		1.238	1.505
Avr. wage rate agric. labourers (2007)	Men	70.26	82.92
	Women	54.91	57.23

Source: Districts at a glance, Directorate of Economics & Statistics, Govt. of AP

Table 2: Weighted summary statistics

	Treatment	Control	
		Not matched	Matched
Value of agr. production	30165.9	25123.2	28399.4
Variable inputs	15310.9	14988.1	15310.9
Area cultivated (acres)	4.24	2.71	4.24
Irrigated area (% of total)	0.19	0.15	0.19
Fertilizer (dummy)	0.98	0.87	0.98
Produced any: Cotton	0.16	0.055	0.16
Produced any: Chillies & Turmeric	0.040	0.089	0.040
Annual income, off-farm activities	26647.1	21055.3	26647.1
Wealth index	0.35	0.39	0.35
Observations	770	348	348

Table 3: General household characteristics

	2007		2009	
	Mean	Std. Dev.	Mean	Std. Dev.
Household head is male	0.97	0.18	0.96	0.20
Age of household head	41.77	12.15	41.34	10.26
Household size	6.00	2.48	5.96	2.64
Highest grade: males	6.32	4.32	7.13	4.07
Highest grade: females	4.47	3.73	5.58	3.77
Wealth index	0.36	0.16	0.46	0.15
Housing quality index	0.49	0.28	0.54	0.27
Consumer durables index	0.25	0.16	0.29	0.17
Housing services index	0.36	0.19	0.54	0.15
Income, off-farm activities (log)	9.61	1.37	10.20	1.34
Value of agr. production (log)	9.47	1.31	10.02	1.21

Table 4: Farming characteristics

	2007		2009	
	Mean	Std. Dev.	Mean	Std. Dev.
Value of agr. production (log)	9.47	1.31	10.02	1.21
Variable inputs (log)	8.87	1.21	9.40	1.14
Area cultivated (acres, log)	0.90	0.98	0.97	0.91
Irrigated area (% of total)	0.18	0.32	0.16	0.29
Fertilizer (dummy)	0.95	0.23	0.94	0.23
HYV seeds (dummy)	0.74	0.44	0.58	0.49
Share inputs: Paddy rice	0.40	0.42	0.39	0.40
Share inputs: Grams and Pulses	0.03	0.11	0.03	0.11
Share inputs: Cotton	0.09	0.25	0.14	0.30
Share inputs: Groundnuts	0.19	0.36	0.21	0.36
Share inputs: Maize	0.02	0.12	0.05	0.17
Share inputs: Jowar	0.03	0.12	0.02	0.08
Share inputs: Chillies & Turmeric	0.03	0.12	0.02	0.10
Share inputs: Foodgrains	0.03	0.12	0.01	0.08
Share inputs: Oilseeds	0.08	0.21	0.05	0.16
Share inputs: Commercial crops	0.05	0.18	0.05	0.16
Share inputs: Fruits	0.02	0.10	0.01	0.07
Share inputs: Vegetables	0.02	0.12	0.02	0.11
Share inputs: Other crops	0.01	0.09	0.01	0.09
Produced any: Paddy rice	0.57	0.49	0.61	0.49
Produced any: Grams and Pulses	0.28	0.45	0.18	0.39
Produced any: Cotton	0.14	0.34	0.21	0.41
Produced any: Groundnuts	0.27	0.44	0.30	0.46
Produced any: Maize	0.04	0.20	0.09	0.28
Produced any: Jowar	0.14	0.34	0.08	0.27
Produced any: Chillies & Turmeric	0.06	0.24	0.06	0.24
Produced any: Foodgrains	0.11	0.32	0.05	0.22
Produced any: Oilseeds	0.18	0.39	0.12	0.33
Produced any: Commercial crops	0.12	0.32	0.12	0.33
Produced any: Fruits	0.05	0.21	0.03	0.18
Produced any: Vegetables	0.07	0.26	0.07	0.26
Produced any: Other crops	0.02	0.14	0.04	0.20

Table 5: Shocks (sample mean)

	Treatment		Control	
	2007	2009	2007	2009
Rainfall (deviation over mean)	0.32	0.03	-0.05	0.03
Rainfall (deviation over mean, lag)	-0.39	0.28	-0.12	0.23
Reported: Increases in input prices	0.12	0.25	0.08	0.13
Reported: Decreases in output prices	0.10	0.25	0.03	0.03
Reported: Death of livestock	0.12	0.19	0.09	0.06
Reported: Drought	0.59	0.16	0.16	0.04
Reported: Flooding	0.14	0.04	0.09	0.05
Reported: Erosion	0.00	0.01	0.00	0.00
Reported: Hailstorms	0.01	0.01	0.00	0.01
Reported: Pest or Diseases	0.21	0.20	0.09	0.10
Reported: Crop failures	0.32	0.54	0.33	0.18
Reported: Storage losses	0.01	0.01	0.01	0.02
Reported: Theft	0.08	0.08	0.03	0.01

Table 6: Agricultural Production Function

	RE	FE	RE	FE
Variable inputs (log)	0.570***	0.505***	0.570***	0.506***
	(0.042)	(0.039)	(0.042)	(0.039)
Area cultivated (acres, log)	0.327***	0.248***	0.327***	0.249***
	(0.042)	(0.054)	(0.042)	(0.054)
Irrigated area (% of total)	0.207**	0.095	0.201**	0.098
	(0.072)	(0.102)	(0.071)	(0.101)
Fertilizer (dummy)	-0.098	-0.110	-0.084	-0.109
	(0.084)	(0.145)	(0.085)	(0.145)
HYV seeds (dummy)	0.048	0.084+	0.050	0.082+
	(0.038)	(0.047)	(0.037)	(0.047)
Share inputs: Cotton, chillies & turmeric	0.251*	0.260+		
	(0.105)	(0.152)		
Share inputs: Cotton			0.219**	0.293*
			(0.083)	(0.137)
Share inputs: Chillies & Turmeric			0.379	0.046
			(0.337)	(0.514)
Share inputs: Groundnuts	-0.402**	-0.288*	-0.403**	-0.286*
	(0.135)	(0.127)	(0.135)	(0.126)
Share inputs: Oilseeds	-0.521***	-0.620**	-0.528***	-0.606**
	(0.108)	(0.186)	(0.107)	(0.184)
Share inputs: Commercial crops	0.512***	0.536***	0.513***	0.511**
	(0.096)	(0.156)	(0.097)	(0.171)
Share inputs: Fruits	0.439+	0.622*	0.439+	0.602*
	(0.246)	(0.251)	(0.247)	(0.260)
Share inputs: Vegetables	0.221	0.238	0.212	0.252
	(0.216)	(0.317)	(0.211)	(0.302)
Year 2009 (dummy)	0.235**	0.208***	0.237**	0.205***
	(0.074)	(0.048)	(0.074)	(0.047)
Observations	2068	2068	2068	2068

Clustered standard errors in parentheses

Shocks and cluster dummies included, but not reported. Foodgrains is reference category.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Cost of Cultivation

	Paddy			Cotton		
	2005-06	2008-09	Share	2005-06	2008-09	Share
Operational Cost	**18196**	**28285**	**61.4%**	**18879**	**27903**	**64.6%**
Human Labour	9291	15800	33.1%	8490	14672	32.0%
Animal Labour	662	564	1.6%	2013	2147	5.7%
Machine Labour	2306	4515	9.0%	1338	2197	4.9%
Seed	891	1129	2.7%	1679	2879	6.3%
Fertilizer & Manure	2963	3411	8.4%	2733	3786	9.0%
Insecticides	900	1546	3.2%	2030	1403	4.7%
Irrigation Charges	665	546	1.6%	132	120	0.3%
Miscellaneous	41	47	0.1%	6	1	0.0%
Interest: Working Capital	477	726	1.6%	457	697	1.6%
Fixed Costs	**11061**	**18165**	**38.6%**	**8746**	**16853**	**35.4%**
Rental Value: Own Land	9338	15747	33.1%	7296	13614	28.9%
Rent: Leased-in-Land	369	1186	2.1%	0	768	1.1%
Land Rev., Taxes, Cesses	2	1	0.0%	3	2	0.0%
Depreciation	220	193	0.5%	192	468	0.9%
Interest: Fixed Capital	1132	1039	2.9%	1254	2002	4.5%
Total Cost	**29257**	**46450**	**100.0%**	**27625**	**44757**	**100.0%**

Source: Directorate of Economics and Statistics, Dep. of Agriculture and Cooperation, Ministry of Agriculture, GoI

Notes: Share in total cost, average 2005-06 & 2008-09, Prices in nominal Rs. per Hectare

Table 8: Number of days worked with NREGA (Fixed Effects)

	NREGA days	NREGA days (log)
Rainfall (deviation, lag)	-53.027*	-1.753***
	(22.993)	(0.317)
Rainfall (deviation)	-25.222**	-0.665***
	(9.173)	(0.179)
Area cultivated (acres, log)	3.963	0.050
	(2.849)	(0.059)
Wealth index	-8.982	-0.074
	(25.561)	(0.545)
Year 2009 (dummy)	54.501***	1.643***
	(12.044)	(0.183)
Constant	26.047*	2.575***
	(11.984)	(0.256)
Observations	941	941

Clustered standard errors in parentheses

Self reported shocks included, but not reported

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Inputs allocated to Cotton (Fixed Effects)

	Not matched	Matched
NREGA	0.092**	0.114**
	(0.032)	(0.037)
Variable inputs (log)	0.068***	0.051**
	(0.019)	(0.019)
Area cultivated (acres, log)	-0.003	-0.005
	(0.014)	(0.014)
Irrigated area (% of total)	-0.041	-0.073
	(0.022)	(0.047)
Annual income, off-farm activities (log)	-0.005	0.000
	(0.004)	(0.004)
Wealth index	-0.076	-0.063
	(0.064)	(0.077)
Rainfall (deviation, lag)	-0.104	0.025
	(0.056)	(0.072)
Year 2009 (dummy)	0.017	-0.089
	(0.030)	(0.051)
Observations	2236	2236

Clustered standard errors in parentheses

Self reported shocks included, but not reported

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Inputs allocated to Cotton (Fractional logit, matched)

	Coefficients		APE	
NREGA	1.024***	(0.256)	0.153*	(0.069)
Variable inputs (log)	0.390***	(0.116)	0.058**	(0.020)
Area cultivated (acres, log)	-0.067	(0.090)	-0.010	(0.014)
Irrigated area (% of total)	-0.631	(0.352)	-0.094	(0.057)
Annual income, off-farm activities (log)	-0.017	(0.029)	-0.002	(0.004)
Wealth index	-0.222	(0.589)	-0.033	(0.088)
Year 2009 (dummy)	-0.999***	(0.297)		
Observations	2236		2236	

Clustered standard errors in parentheses

Self reported shocks included, but not reported

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Inputs allocated to Commercial Crops (Fixed Effects)

	(1) Cotton, chillies & turmeric	(2) Total comm. crops
NREGA	0.110**	0.066+
	(0.034)	(0.038)
Variable inputs (log)	0.079***	0.101***
	(0.019)	(0.018)
Area cultivated (acres, log)	0.001	-0.007
	(0.014)	(0.013)
Irrigated area (% of total)	-0.032	-0.037
	(0.021)	(0.026)
Annual income, off-farm activities (log)	-0.010*	-0.009+
	(0.005)	(0.005)
Wealth index	-0.058	-0.011
	(0.071)	(0.083)
Rainfall (deviation, lag)	-0.127*	-0.047
	(0.057)	(0.061)
Year 2009 (dummy)	0.009	-0.028
	(0.030)	(0.035)
Observations	2236	2236

Clustered standard errors in parentheses

Self reported shocks included, but not reported

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Inputs allocated to Cotton (Fixed Effects)

	(1)	(2)
NREGA registered in 2006	0.063*	
	(0.030)	
NREGA registered in 2009		0.001
		(0.016)
Variable inputs (log)	0.070***	0.074***
	(0.019)	(0.020)
Area cultivated (acres, log)	-0.004	-0.004
	(0.013)	(0.013)
Irrigated area (% of total)	-0.043*	-0.040
	(0.021)	(0.021)
Annual income, off-farm activities (log)	-0.005	-0.005
	(0.004)	(0.004)
Wealth index	-0.056	-0.033
	(0.065)	(0.065)
Rainfall (deviation, lag)	-0.077	-0.044
	(0.053)	(0.048)
Year 2009 (dummy)	0.028	0.032
	(0.031)	(0.029)
Constant	-0.484**	-0.510**
	(0.166)	(0.175)
Observations	2236	2236

Clustered standard errors in parentheses

Self reported shocks included, but not reported

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Inputs allocated to Commercial Crops (Fixed Effects)

	(1)	(2)
	Cotton	Cotton, chillies & turmeric
NREGA registered in 2006	0.033*	0.030*
	(0.015)	(0.015)
Variable inputs (log)	0.102***	0.119***
	(0.011)	(0.011)
Area cultivated (acres, log)	-0.030**	-0.023*
	(0.011)	(0.011)
Irrigated area (% of total)	-0.034+	-0.006
	(0.018)	(0.019)
Annual income, off-farm activities (log)	-0.019**	-0.020***
	(0.006)	(0.005)
Rainfall (deviation, lag)	-0.151***	-0.129***
	(0.035)	(0.035)
Year 2009 (dummy)	0.105***	0.080**
	(0.027)	(0.028)
Observations	1540	1540

Standard errors in parentheses

Self reported shocks included, but not reported. Phase I districts only, matched sample.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$