An employment guarantee as risk insurance? Assessing the effects of the NREGS on agricultural production decisions

Esther Gehrke*

November 22, 2015

Abstract

This paper assesses inhowfar employment guarantees can support households in managing agricultural production risks. Using representative panel data for Andhra Pradesh, India, it analyzes the effects of the National Rural Employment Guarantee Scheme (NREGS) on households' crop choices. This paper shows that the introduction of the NREGS reduces households' uncertainty about future income streams because it provides reliable employment opportunities in rural areas independently of weather shocks and crop failure. Households with access to the NREGS can therefore shift their production towards riskier but also more profitable crops. These shifts in agricultural production can considerably raise the incomes of smallholder farmers. Linking the employment guarantee to risk considerations is the key innovation of this paper. Therewith, it provides empirical evidence that employment guarantees can, similarly to crop insurance, help households in managing agricultural productions risks and contributes to the ongoing debate on the effects of the NREGS on agricultural productivity.

Keywords: Uncertainty; Employment Guarantee; Crop choice
JEL: I38; O12; Q16

*German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE) and University of Passau, e-mail: *esther.gehrke@die-gdi.de*. The author is grateful to Michael Grimm for continuous guidance and advice and to Denis Cogneau; Andrew Foster; Clement Imbert; Markus Loewe; Abhiroop Mukhopadhyay; Dominique Van de Walle and the participants of the German Economic Association AEL Conference (Munich), the AEL PhD seminar (Zürich), the ECINEQ Conference (Bari), the IZA/World Bank Conference on Employment and Development (Bonn), the PEGNet Conference (Copenhagen) and the Brown University Applied Microeconomics lunch for their many helpful comments and suggestions. This work was supported by the German Federal Ministry for Economic Cooperation and Development [grant number 2011 95 809]. The data used in this publication comes 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). 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.

1 Introduction

Previous research suggests that farmers in developing countries are constrained in their production and investment decisions. Evidence of delayed technology adoption, low investment in fixed capital, a preference for conservative crop choices and, more generally, a lack of innovative capacity is by now well established (Foster and Rosenzweig, 2010b; Duflo, Kremer, and Robinson, 2008; Suri, 2011). This has potentially severe and long-lasting effects on income and well-being in developing countries as a large share of their populations still rely on agricultural production as a major source of income.

Empirical evidence suggests that uninsured risk prevents farmers from adopting new technologies. A number of studies have used randomized variation in the availability of index-based agricultural insurance to estimate the importance of uninsured risk in production decisions. These studies show that crop insurance is critical in stimulating fertilizer application (Karlan et al., 2013), risky crop choice (Cole, Gine, and Vickery, 2013) and risk taking in agriculture more generally (Mobarak and Rosenzweig, 2013).

However, trust- related considerations and basis risk continue to limit the uptake of agricultural micro-insurance in many developing countries (Cole et al., 2013; Carter et al., 2014). Given these limitations, it seems worthwhile to explore other policy options that could help farmers to cope with shocks and manage risks.

This paper aims at contributing to the empirical evidence on the importance of risk management in farmers' production decisions. But instead of exploring variance in the availability of insurance, as do the studies cited above, it examines variation in the access to an alternative mechanism that could improve a household's risk management: an employment guarantee. The main argument is that public works programs or employment guarantees could help households to cope with income shocks by providing additional employment opportunities. This idea is not new; the potential of public works schemes in helping households to smooth income in the case of shocks has been highlighted *inter alia* by Barrett, Holden, and Clay (2005) and Binswanger-Mkhize (2012). However, to the best of my knowledge, no empirical evidence on the insurance effect of an employment guarantee on households' production decisions has been provided so far.

In this paper I present evidence that the introduction of the National Rural Employment Guarantee Scheme (NREGS) reduces households' uncertainty about future income streams and enables them to produce a higher share of high-risk, high-profit crops. The National Rural Employment Guarantee Act (NREGA) was passed in India in September 2005; the implementation thereof began in 2006. The NREGA entitles every rural household to up to a 100 days of work per year at the state minimum wage, which is to be provided within 14 days of the application for work being made.The NREGS is the largest public works program in the world. In the financial year 2010/11 it provided work to close to 55 million rural households (MoRD, 2012). A total of 2.5 billion person-days of employment were generated in the same year.

The hypothesis described above is tested using a household-level panel data set that is representative of the state of Andhra Pradesh in southern India. The quality of implementation of the NREGS has been shown to vary immensely across India (Dutta et al., 2012). In most states the provision of work under NREGS is far too unpredictable to completely offset the effects of a shock. Under such circumstances, the NREGS would not affect households' risk expectations. Andhra Pradesh, however, is one of the states with the highest number of days of employment generated per rural household. I find that the provision of work in Andhra Pradesh does effectively respond to changes in household demand and thus supports households in managing agricultural production risks.

The estimation strategy employed here builds on the sequenced introduction of the NREGS. Using the introduction of the NREGS at district level, it explores the fact that the scheme was introduced in four out of the six survey districts in 2006 and in the remaining two districts in 2008 and 2009. Because this approach relies heavily on the parallel trends assumption, I perform a number of robustness checks. The use of alternative treatment variables (e.g. block-level spending and employment days generated under the NREGS, as well as households' registration with NREGS) does not change the results. The results of several robustness checks support the hypothesis that the observed effect can indeed be attributed to an insurance function of the NREGS.

I find that the key innovation of the Indian public works program (i.e. giving households the right to work) encourages agricultural households to increase the share of risky but profitable crops in their portfolios. The results of this paper suggest that employment guarantees can trigger important gains in agricultural productivity in the medium term. These gains go far beyond the direct income effect that the provision of employment in agricultural lean seasons has on the wellbeing of rural households. That increases in productivity and, in turn, in households' incomes can be triggered solely through the insurance effect of an employment guarantee is a very important lesson for other countries with planned or ongoing public works programs.

The remainder of this paper proceeds as follows: Section 2 introduces a theoretical framework for analyzing the effects of an employment guarantee on crop choice. Section 3 presents the data and summary statistics. Section 4 outlines the estimation strategy. Section 5 presents the empirical results, and Section 6 concludes.

2 Risk management and households' crop choices: A theoretical framework

Providing additional employment opportunities to a total of 55 million households has brought about considerable changes in the social and economic realities in India. The NREGS affects households in rural areas through various channels. The most obvious and so far most intensely researched effect is the increase in available income and wealth of those households participating in the program. This wealth effect is most pronounced for households with surplus labor - namely households whose labor supply exceeds labor demand - and in regions where regular labor markets fail to absorb this excess. The increase in income resulting from NREGS participation has been shown to increase consumption levels (Jha, Gaiha, and Pandey, 2012) and to reduce poverty (Klonner and Oldiges, 2014). Increases in disposable income and wealth might also positively influence the capacity to take risks and investment behavior.[1]

Another effect, which is much less well understood, is the insurance effect. It is

[1]This effect is different from the insurance effect, which is the main focus of this paper. I address the robustness of my findings to this alternative mechanism in Section 5.3.

particularly relevant for households that are highly exposed to covariate shocks such as droughts, floods or large-scale crop diseases. In rural areas of India wages were shown to fall with covariate shocks (Jayachandran, 2006). Such wage fluctuations severely limit households' possibilities to cope with shocks through the labor market. By giving households the right to work and making employment opportunities available independently of shocks, the NREGS greatly influences households' ability to smooth income in the case of a shock. If the insurance effect holds, households could change their production decisions, take more risks and reach higher expected incomes. If a shock then occurs, households can cope with the shock by working for the scheme. Without the shock, it is unlikely that all of these households would participate in the NREGS, because their shadow wages exceed the wage rate paid in the scheme.

Finally, the NREGS is expected to affect wage levels through general equilibrium effects in the village economy. The NREGS was shown to raise wage levels in the private sector because wages under the NREGS are in many cases higher than the wages paid for casual work and households consequently shift their labor supply from the private sector towards the public works program (Imbert and Papp, 2015; Berg et al., 2012). Increases in wages could also affect production levels or crop choice in agriculture because they raise production costs, particularly for large-scale farmers.

In this paper, I focus specifically on the insurance effect. The idea is that households with access to the NREGS are able to shift their agricultural production towards high-risk, high-profit crops, because their capacity to cope with production shocks is improved due the NREGS. I focus on input allocation as indicator for the importance of different crops in a household's portfolio.[2] To show the effect of the NREGS on crop choice more systematically, I develop a theoretical model of household decision-making under uncertainty that shows how the introduction of NREGS can affect crop choice via the insurance effect. The model primarily builds on Dercon and Christiaensen (2011). Taking into account the ideas outlined by Fafchamps (1993) and Van Den Berg (2002), I particularly

[2] This is because of data constraints (land allocation was not consistently collected) and because input allocation seems to be least influenced by the wage setting effect of the NREGS. In Section 5.5, I explore the robustness of my findings to alternative mechanisms, such as the wealth and wage effects.

explore how the sequencing of input allocation, shock realization 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 limited access to credit. Crop choice is first modeled in a world without risk but with imperfect credit markets and then extended to a world with uncertainty. This allows for the isolation of the effects of uncertainty and risk aversion on production decisions. Finally, I will show how the introduction of the NREGS can affect input allocation decisions in both scenarios.

2.1 General setup

Assume that a household engaging in agricultural production has the choice between two agricultural products Q^d and Q^s. Given that both products are well known to the farmer and have been produced in the region for some time, we can abstract from learning and other sunk costs. These products are produced with two different types of production functions: one is deterministic and the other stochastic.[3] It is also assumed that the risky crop is more productive on average. Both products can be sold at local markets at the same price p.

Agricultural production takes place over two periods, the planting and the harvesting seasons. Input allocation at the planting stage defines total yield Q, which has to be harvested in the second stage (such as in Fafchamps, 1993):

$$Q^d = f^d(k^d, l_1^d, a^d) \tag{1}$$

$$Q^s = \epsilon f^s(k^s, l_1^s, a^s) \qquad E[\epsilon] = 1 \tag{2}$$

The total yield of both products depends on land a, labor l_1 and input k allocation in period one.[4] Inputs k are defined as a bundle of variable inputs such as seeds, fertilizer

[3] The assumption, that one production function is deterministic and the other stochastic is rather extreme. Instead, one would expect both production functions to depend on the realization of random shocks, although to a different extent. However, this simplification is without major impact on the results obtained here.

[4] I have abstracted from fixed capital because the marginal effect of productive capital was found to be close to zero.

and pesticides. I assume that the first period production function is a Cobb-Douglas type of production function. The total yield of the risky product additionally depends on the realization of a multiplicative, random, serially uncorrelated shock ϵ at the end of the first period. The expected value of this shock is 1; thus in expectation, the production function of the risky crop is just $f^s(a^s, l_1^s, k^s)$. The labor required for harvesting in the second period l_2 is a linear function of realized yields, e.g. $l_2 = \alpha(Q^d + Q^s)$, where α is a parameter indicating how much labor is needed for harvesting given any realized yield.[5]

I assume that the household maximizes utility from consumption C in both the planting and the harvesting periods. 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. I abstract from leisure in this model because it will not change the choice under uncertainty.[6] The household generates income from wage employment on local labor markets and from agricultural production. Building on the full-income approach, the household maximization problem can be described as follows:

$$
\begin{aligned}
max \qquad V =& U_1(C_1) + \delta U_2(C_2) \\
& s.t. \\
& \quad C_1 \leq w_1(T_1 - l_1^d - l_1^s) - g(k^d + k^s) + B \\
& \quad C_2 \leq p(Q^d + Q^s) + w_2(T_2 - l_2) - (1 + r)B \\
& \quad B \leq B^m \\
& \quad a^d + a^s \leq 1
\end{aligned} \tag{3}
$$

Total time endowment is represented by T_1 and T_2. In both periods total time can

[5]Because labor allocation is linear in realized yields, it will be profitable to harvest either the entire crop or nothing at all (depending on wage levels and output prices), thus only allowing for corner solution outcomes.

[6]By dropping leisure, I ignore possible income effects of increases in wage levels on a household's time allocation between labor and leisure. But since my main interest lies in crop choice rather than in production levels, 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).

be allocated between working in the labor market and working in own fields. In the first period, the household obtains income from wage work at level w_1 and from borrowing B. Inputs for agricultural production can be purchased at price g. In the second period, the household obtains income from the sale its own agricultural production $p(Q^d + Q^s)$ and wage work at level w_2. Note here that the household will have to allocate labor to harvesting in order to generate income from its agricultural production. Because it seems plausible that the household will always prioritize its own harvest over wage employment, I assume that the household deems the cost of harvesting to equate to reservation wages rather than market wages. It is therefore useful to replace the wage cost of harvesting $w_2 l_2$ in the budget constraint with $\alpha w_2^r(Q^d + Q^s)$, where w_2^r is the reservation wage and $\alpha(Q^d + Q^s)$ is the effort necessary for harvesting expressed in units of realized yield.

Incurred debts will have to be repaid in the second period at an interest rate of r. Input credits are relatively common in rural Andhra Pradesh, although it seems that the amount of credit conceded is limited by a household's wealth. In the sample around 18% of the households that applied for credit reported not receiving the total amount of credit they applied for. 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. Because households are expected to opt for that source of credit only under extreme circumstances, this model does not allow for any borrowing beyond the harvesting period.

In this setting local labor markets are assumed to function with the option to hire labor in as well as out. In fact, most households in the sample report a range of income sources - of which casual labor features prominently. However, harvest stage wages are assumed to be stochastic and to covary with covariant shocks, such as rainfall shortages. This was shown in the case of rural India by Jayachandran (2006). For most farmers, this means that they can only form expectations about harvest stage wages and face a double risk from rainfall fluctuations: First, their own harvest is likely to fail if there is a rain shortage. Second, they will not be able to find work at adequate wage levels in local labor markets.

Finally, $a^d + a^s = 1$ describes the restrictions on allocable land. I assume that there are no functioning land markets 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 virtually absent.[7]

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.

2.2 Deterministic case

First, consider a scenario without uncertainty. In such a world each household maximizes utility by maximizing profits from agricultural production plus income from wage employment. Identical results would be obtained if the household were risk neutral. Because both production functions are deterministic in this scenario, optimal land, input and labor allocation are achieved when their marginal products equal respective prices. In the deterministic case, the Lagrange can be written as follows:

$$\begin{aligned}\mathcal{L} =& U_1(C_1) + \delta U_2(C_2) \\ &+ \lambda(w_1(T_1 - l_1^d - l_1^s) - g(k^d + k^s) + B - C_1) \\ &+ \mu[(p - \alpha w_2^r)(Q^d + Q^s) + w_2 T_2 - (1+r)B - C_2] \\ &+ \varphi(B^m - B) \\ &+ \rho(1 - a^d - a^s)\end{aligned} \tag{4}$$

Solving the household maximization problem leads to the following decision rules for

[7] Part of this is due to a 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.

the allocation of variable inputs to each of the crops:[8]

$$\frac{\partial f^d}{\partial k^d} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{5}$$

$$\frac{\partial f^s}{\partial k^s} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{6}$$

Equations (5) and (6) show that decision rules are equal for both crops, and optimal allocation will imply that the marginal product of inputs in d is equal to the marginal product of inputs in s. Because realized yield is harvested in the second period, input allocation does not only depend on input and output prices but also on future reservation wages and on the intertemporal marginal rate of substitution in consumption.

Finally, equation (7) describes the optimal consumption rule over both periods given credit constraints:

$$\frac{\partial U_1}{\partial C_1} = \delta(1+r)\frac{\partial U_2}{\partial C_2} + \varphi \tag{7}$$

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 discounted marginal utility from consumption in the harvesting period. This means that consumption in the planting stage will be lower than what could be achieved if the credit constraints were not binding. Including equation (7) into equation (5) also reveals the effect of the credit constraint on input allocation:

$$\frac{\partial f^d}{\partial k^d} = \frac{g(1+r)}{(p - \alpha w_2^r)} + \frac{g\varphi}{(p - \alpha w_2^r)\delta \frac{\partial U_2}{\partial C_2}} \tag{8}$$

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

[8] As mentioned earlier, the main focus of this paper is on input allocation, but similar results can be obtained for the allocation of labor and land to each of the crops. A detailed derivation of all decision rules can be found in the Supplementary Appendix, Section A.

2.3 Introducing uncertainty

When introducing uncertainty, the Lagrange is written as follows:

$$
\begin{aligned}
\mathcal{L} = & U_1(C_1) + \lambda(w_1(T_1 - l_1^d - l_1^s) - g(k^d + k^s) + B - C_1)) \\
& + E[\delta U_2(C_2) + \mu[(p - \alpha w_2^r)(Q^d + Q^s) + w_2 T_2 - (1+r)B - C_2]] \\
& + \varphi(B^m - B) \\
& + \rho(1 - a^d - a^s) \qquad (9)
\end{aligned}
$$

The household faces uncertainty with respect to the realized yield of the risky crop Q^s and the wage levels in the harvest period w_2. This affects the expectations a household forms about the level of consumption that can be achieved in the second period. When differentiating the Lagrange with respect to the choice variables, the optimal consumption rule is:

$$
\frac{\partial U_1}{\partial C_1} = (1+r)\delta\frac{\partial EU_2}{\partial C_2} + \varphi \qquad (10)
$$

The consumption rule - equation (10) - changes slightly when introducing uncertainty because for any expected consumption level C_2, expected utility $EU_2(C_2)$ will be lower than the utility of the expected value $U_2(E(C_2))$, and marginal expected utility will be higher than the marginal utility of the expected value. Since all other variables remain constant, C_2 has to be higher relative to C_1 under uncertainty for the identity to hold. This is equivalent with the well-known argument that risk decreases current consumption levels and enhances savings.

The decision rules for input allocation under uncertainty are the following:

$$
\frac{\partial f^d}{\partial k^d} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta\frac{\partial EU_2}{\partial C_2}} \qquad (11)
$$

$$
\frac{\partial f^s}{\partial k^s} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta\frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \epsilon)}{(p - \alpha w_2^r)\delta\frac{\partial EU_2}{\partial C_2}} \qquad (12)
$$

Equation (11) shows the allocation rule for inputs to the safe crop. It looks similar to equation (5), except that now the household maximizes expected utility of consumption in the harvest period. Again, marginal expected utility is higher than marginal utility of the expected value. Thus, under uncertainty, the right-hand side term will be lower than in the deterministic case, implying that the household allocates more inputs to the safe crop than it would in the absence of risk. This reflects the greater weight households put on future consumption than on current consumption as described above.

Equation (12) shows the effect of uncertainty on input allocation to the risky crop. Here the decision rule changes considerably and the overall effect is less clear. Again, marginal expected utility is higher than marginal utility, thus implying higher input allocation to the risky crop also. However, the covariance between marginal utility of consumption and the random shock ϵ is strictly negative.[9] This term increases the value of the right-hand side of equation (12), which means that input allocation to the risky crop is lower under uncertainty. Which of the two effects is stronger depends on the degree of risk aversion of the household, expected consumption levels C_2 and the amount of covariance between marginal utility and the random shock. Since the covariance will be greater with lower wages in period two and with a higher interest rate r, the net effect of uncertainty on input allocation can be expected to be negative in this context.

Irrespective of total levels of input allocation, it can be clearly seen that under uncertainty, input allocation will shift towards the safe crop k^d relative to the risky crop k^s. Thus under uncertainty, the share of risky crops in a household's portfolio will always be lower than in the deterministic scenario.

Again, equations (11) and (12) can be reformulated to include the credit constraint. Then, input allocation to the risky crop is determined as follows:

$$\frac{\partial f^s}{\partial k^s} = \frac{g(1+r)}{(p-\alpha w_2^r)} + \frac{g\varphi}{(p-\alpha w_2^r)\delta\frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \epsilon)}{(p-\alpha w_2^r)\delta\frac{\partial EU_2}{\partial C_2}} \tag{13}$$

We can see from equation (13) that both risk and credit constraints go in the same direction and reduce the input allocation to the risky crop. More importantly, it also shows

[9] In a bad state of the world ($\epsilon = 0$) consumption in the second period will be lower and marginal utility higher than in a good state of the world. Conversely, a high ϵ will lead to higher consumption in period 2 and to lower marginal utility of consumption.

that uncertainty reduces input allocation to the risky crop relative to the deterministic crop even if credit constraints are not binding.

2.4 The insurance effect of an Employment Guarantee

The insurance effect of an employment guarantee, such as the National Rural Employment Guarantee Scheme (NREGS), on households' allocation rules are best represented by an increase in expected harvest stage wages.[10] For households with a labor surplus, other farms offer the best possibility of finding employment during harvest periods; in the case of major weather shocks, they have to expect to not find any employment at all (Jayachandran, 2006; Kaur, 2014). Because the NREGS provides reliable income opportunities throughout the year, households can expect to find employment in the harvest period even in bad years. In other words, the NREGS increases wage levels in bad years and therewith reduces the covariance between harvest stage wage levels and covariant shocks. The comparative statics in this section show that the introduction of NREGS affects optimal input allocation under certainty differently than under uncertainty.

Without uncertainty, the optimal allocation of input to both crops is determined as follows:

$$\frac{\partial f^{d,s}}{\partial k^{d,s}} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{14}$$

An increase in average harvest period wages w_2 affects optimal input allocation by increasing consumption levels that can be realized in the second period. Households that hire labor out (i.e. those whose land is too small to produce at higher levels) will increase consumption. One will thus see a decrease in input allocation for net lenders of labor because of increases in C_2, which will reduce $\partial U_2/\partial C_2$ and increase the second part of the right-hand side of equation (14). The effect of increased wages on agricultural production levels (through consumption) can be understood as a substitution effect. Because working outside the farm becomes more profitable for households with little cultivated land, the

[10] Of course, in a scenario without uncertainty, expected wage levels need to be replaced by average wage levels.

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 (15):

$$\frac{\partial f^s}{\partial k^s} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \epsilon)}{(p - \alpha w_2^r) \delta \frac{\partial EU_2}{\partial C_2}} \tag{15}$$

If harvest stage wages increase, we will observe the same effects on marginal utility of consumption as in the deterministic case. Under uncertainty, however, the negative covariance term reduces input allocation to the risky crop, and this effect is now partially offset by the introduction of an employment guarantee. If possibilities to generate market income improve, shocks will have less effect on consumption in the harvesting period. Because the household knows that it can earn additional income in instances of negative production shocks by spending more time working for the NREGS, the possibilities to smooth income increase significantly. The more the covariance term on the right-hand side of equation (15) approaches zero, the more the ratio of inputs allocated to the risky crop (versus the safe crop) will approach the deterministic scenario. This means that even if total input (or similarly labor) allocation is reduced due to the employment guarantee, the share of total inputs allocated to each of the crops will approach the ratio of the deterministic scenario. Interestingly, this effect holds independently of whether credit constraints reduce total input allocation or not.

3 Data

When estimating the insurance effect of the NREGS, one must take into account considerable variation in the quality of implementation of the program across states (Dutta et al., 2012). The section above highlighted the importance of households' expectations about future income streams. Therefore it seems plausible to observe insurance effects only in states in which the demand for employment has been sufficiently met, already in the early years of program implementation.

Given these considerations, the model specified above is tested using the Young Lives Survey (YLS) data for Andhra Pradesh. Andhra Pradesh is particularly suited to studying the question of interest because it is one of the best performing states in India in terms of the number of workdays generated per household and meeting the demand for work (Dutta et al., 2012). Regarding outreach, only Chhattisgarh, West Bengal, Madya Pradesh and Rajastan reached higher proportions of rural households in the financial year 2009/10 (MoRD, 2012).[11]

The YLS data set covers 3,019 households living in six different districts. Three rounds of interviews have been conducted so far (2002, 2007 and 2009/10). Panel attrition is relatively low: 2,910 households were revisited in 2009/10, giving an attrition rate of 3.6% (Galab et al. 2011). For reasons of comparability, only the second (2007) and third (2009/10) rounds are considered in the current analysis. Furthermore, the analysis is restricted to households with non-zero agricultural production in 2007 and 2009/10.

The selection process of districts for the YLS ensured that all three geographical regions were surveyed, as too were the poor and non-poor districts of each region, such that the YLS is broadly representative of the population of Andhra Pradesh (Galab et al., 2011).[12] Out of the six survey districts, four introduced the NREGS in phase one (2006) (the treatment group) and the other two districts in phase two (2007) and three (2008) (the control group). The data is clustered on 87 villages in 17 sub-districts (blocks). This data is complemented by secondary data for the calculation of the dependent variable as well as for a number of controls.

For the calculation of the dependent variable - a risk index of each households' crop portfolio - data on input allocation to each crop from the questionnaire is combined with District-level crop production statistics. The time series of crop production statistics are

[11] At the same time, Andhra Pradesh has been a forerunner in terms of innovative approaches to the implementation of the NREGS. First, it has a lot of experience with performing social audits to increase accountability within the scheme. Second, it was one of the first states to cooperate with IT enterprises to strengthen the efficiency of administrative processes. To increase transparency, entries on muster rolls and the number of workdays generated per job card holder, inter alia, are publicly accessible. Nonetheless, the program continues to be implemented in a top-down manner in Andhra Pradesh. Usually, work is not generated upon demand, rather work applications are only accepted if there is work available.

[12] This is in reference to the State of Andhra Pradesh in 2013, prior to its division into the states of Andhra Pradesh and Telangana.

used to calculate the variability of each crop's yield. With this information, a risk index R_i can be constructed for each household given its reported allocation of inputs to each of the crops.[13] The risk index for household i given input allocation k to crop m is defined as $R_i = \sum r_m k_m / \sum k_n$, where r_m is the coefficient of variation of the yield of crop m. The risk index of a household's crop portfolio is thus the weighted average of each crop's coefficient of variation in yield.[14] Note here, that r_m is only available for a subset of all crops n (26 out of 42), such that $m \subseteq n$. Still, $\sum k_m$ represents roughly 90% of the total allocation of inputs in the sample. To reduce potential bias, I drop all observations from the sample which have no crop in their portfolio for which risk information is available, e.g. $\sum k_m = 0$ or $R_i = 0$, in one or both of the survey rounds.[15]

Table 1 reports summary statistics for the main variables and controls used in the paper. I split the sample in treatment and control group and over the two survey rounds.[16] Agricultural production levels are higher among treatment group households than among control group households. The average amount spent on variable inputs (such as seeds, fertilizer and pesticides), cultivation areas and irrigation levels are all higher in the treatment group than in the control group. The risk index at baseline is also higher in the treatment group (0.36) than in the control group (0.26). Table 1 also summarizes the occurrence of different shocks in both groups and in both periods. Rainfall deviation and rainfall deviation (lag) describe the deviation of annual cumulative rainfall levels from their long-term average.

Table 1 also summarizes the four different treatment variables used in the analysis. First, as discussed above, I explore the universal nature of the NREGS by coding as

[13]Allocation of inputs refers to the share in total variable inputs such as seeds, fertilizer and pesticides that is allocated to each crop in a household's portfolio. This is the only information collected in the survey that gives information about the relative importance of each crop in a household's production. Obviously this information is not a perfect approximation of the amount of land allocated to each crop, and could be biased towards crops with higher input costs. However, the focus of the analysis is on changes in this index over time, such that this measure should nevertheless be able to capture changes in the relative importance of risky crops in a household's portfolio.

[14]The distribution of the risk index as well as of the change in this variable between survey rounds is plotted in Figures (C.1) and (C.2) respectively in the Supplementary Appendix.

[15]Section B of the Supplementary Appendix provides more information on how the variable is constructed. The robustness of my findings to alternative aggregation methods is discussed in Section 5.4.

[16]For a detailed discussion of data sources and the construction of variables refer to the Supplementary Appendix, Section B.

‘treated’ those households based in districts where the NREGS was introduced in 2006.[17] Second, I use lagged block-level disbursements under the program as an indicator of the intensity of treatment, arguing that households living in blocks with higher past disbursements have higher expectations about the availability of employment in situations of need. Third, following the same logic, I use the lagged annual total of employment person-days generated per job card at the block-level. Fourth, I explore the self-selection of households into the program in order to increase the robustness of my results.

4 Estimation strategy

The key prediction of the model described in Section 2 is that an increase in expected labor market wages in the harvesting period, *ceteris paribus*, increases the share of inputs allocated to risky crops if households were previously constrained in their crop choice by high levels of uncertainty regarding output levels and dysfunctional insurance markets. It is not possible to test this hypothesis directly for two reasons. First, households’ expectations with regard to wages depend on a range of individual factors (such as perceived access to the labor market) that would not be captured by observed village-level wages. Second, a range of unobserved village characteristics may change over time and those changes will probably influence both expected labor market wages and farmers’ crop choice.

To circumvent the problems mentioned above, I explore the availability of the NREGS as a source of exogenous variation in expected labor market wages during the harvest period. As argued in Section 2.4, the introduction of NREGS increases expected wages in the harvest period because employment opportunities through the NREGS do not depend on favorable weather outcomes and hence do not covary with village-level shocks.

It is important to notice here that the NREGS does not only affect households’ crop choices through the insurance effect - which is the main focus of this paper. Because

[17] Given the size of the program and the huge awareness campaigns undertaken at the beginning of implementation, it seems valid to assume that households in rural Andhra Pradesh form expectations about income opportunities through the NREGS based on the local availability of the program and not only based on being registered with the program.

increases in available income and wealth due to the NREGS might also influence a household's ability to cope with shocks, their access to credit and their willingness to take risks, it is essential to control for these changes in order to isolate the insurance effect.[18] The outcome equation can be written as follows:

$$R_{ijt} = \beta_1 D_{ijt} + \beta_2 X_{it} + \beta_3 Z_{jt} + u_i + \gamma_j + \delta_t + \upsilon_{ijt} \tag{16}$$

The dependent variable is the risk index of household i's crop portfolio at time t. D_{ijt} represents a household's access to the NREGS. Let X_{it} be a set of time-varying household characteristics that affect preferences and crop choice (such as education, wealth, income and past experience with shocks) and u_i be time-constant unobserved household characteristics (such as risk aversion, farming ability and land quality). Z_{it} is a set of time-varying village-level characteristics (e.g. weather trends, extension services, prices, etc.), γ_j are time-constant village characteristics (such as the land's suitability for certain crops), δ_t is a time fixed-effect and υ_{ijt} is the error term.

Taking the first difference removes unobserved household and village level characteristics that are constant over time:[19]

$$\triangle R_{ij} = R_{ij,t+1} - R_{ij,t} = \beta_1 \triangle D_{ij} + \beta_2 \triangle X_i + \beta_3 \triangle Z_j + \triangle\delta + \triangle\upsilon_{ij} \tag{17}$$

For β_1 to have a causal interpretation, it must be the case that the differences in the change of the risk index between the treatment and control groups are entirely due to the NREGS.

This assumption could be violated for a number of reasons. First, since the access to the NREGS is non-random, treatment could be correlated with potential outcomes of R_{ijt}. Second, households in the treatment and control group may not be following parallel trends in their crop choices. The remainder of this section discusses how I address these points.

[18] A number of robustness checks that address alternative explanations for observed outcomes are presented in Section 5.5.

[19] With two time periods, taking the first differences is essentially the same as estimating the model in fixed effects.

As mentioned earlier, different variables can be used as treatment indicators. At the district level, the NREGS should have been introduced in the poorest districts first. This could potentially bias the estimates downwards because poorer districts are less likely to have extension services and marketing structures in place that would enable households to seize the opportunity to plant more profitable cash crops. However, in most states - and in Andhra Pradesh in particular - the prioritization of the poorest districts was not systematically implemented. In this sample, the general economic characteristics of treatment and control districts do not differ greatly.[20] The treatment intensity at the block level should also be exogenous to potential outcomes. Estimates could be biased if funds allocated to blocks responded to rainfall shocks and if these rainfall shocks also affected a household's input allocation decision. However, the amount of funds to be sanctioned per block is defined between December and March for the following financial year (April to March). Since I am using lagged values of disbursed funds, these amounts are fixed 14 to 18 months before household's decide on their input allocation.[21] Lastly, I explore differences in crop choices across households who registered with the NREGS or not. Here, the possibility that unobserved shocks affect both the decision to register and a household's crop choice cannot be ruled out. I employ matching techniques to reduce selection bias, but of course I cannot account for unobserved shocks that affect the decision to register with the NREGS and households' crop choices.

The parallel trends assumption could be violated due to differences in crop productivity which cause the share of those crops in total input allocation to increase independently of the NREGS. Given the small number of districts in the sample, this could significantly bias the results. District-wise time trends in the risk index of crop production are displayed in Figure 1. One of the treatment districts (Mahaboobnagar) displays a decreasing trend in the risk index, while all other districts seem to be following the same trend.

Another - more subtle - violation of the parallel trends assumption could emerge from mean reversion in the dependent variable. Why might households with riskier crop port-

[20] See Section B.4 and Table D.1 in the Supplementary Appendix for more information.

[21] It is also fixed between 6 and 8 months before the start of the monsoon, which could affect next years input allocation through time-lags in the effect of shocks. For more information on the time line see Supplementary Appendix, Section B.

folios display a negative change in the risk index? The reason could be effects of lagged shocks on current input choices which are rooted in the non-separability of production and consumption decision of agricultural households (Sadoulet and De Janvry, 1995). In a world with imperfect credit markets and risk, past shocks will affect current wealth and therefore also current input allocation decisions. If household wealth is perfectly captured by the data, controlling for changes in wealth should eliminate any bias. If wealth is, however, also reflected in soil nutrition, which is affected by weather shocks and not captured in the data (Foster and Rosenzweig, 2010a), then controlling for wealth will not be sufficient.

Assume that the risk index of each household's crop portfolio follows a modified AR(1) process, where - in the absence of a shock - the risk index at time $t+1$, R_{t+1}, is equal to a linear transformation of the risk index of the previous period plus some random noise, e.g $\rho R_t + \epsilon_{t+1}$.[22] In contrast, if a shock occurs, households with higher risk in their crop portfolio will also face higher losses in agricultural production. This will lead them to choose a more conservative crop portfolio in the following period. Formally, this process can be described as follows:

$$R_{t+1} = \rho R_t + \delta u_t + g(R_t)u_t + \epsilon_{t+1} \tag{18}$$

The shock u_t has expected value zero and $g(R_t)$ is a flexible function of input allocation, which allows shocks to have a differential effect on next seasons crop choice, depending on the level of R_t. In the absence of any program effect, the observed change in crop choice would be the following:

$$\begin{aligned}\triangle R &= R_{t+1} - R_t \\ &= \rho(R_t - R_{t-1}) + \delta(u_t - u_{t-1}) + g(R_t)u_t - g(R_{t-1})u_{t-1} + \epsilon_{t+1} - \epsilon_t \\ &= (\rho - 1)R_t + \delta u_t + g(R_t)u_t + \epsilon_{t+1} \end{aligned} \tag{19}$$

In expectation this change would be $E(\triangle R) = (\rho - 1)R_t$. A placebo treatment effect

[22]For expositional purposes, I drop all subscripts except the time subscript.

would be zero in expectation only if the process approached a random walk (e.g. $\rho = 1$) or if the distribution of R_t were equal in treatment and control groups. The placebo treatment effect will be even higher if the occurrence of lagged shocks u_t is different in both groups. The low number of districts used in this analysis warrants special attention to this phenomenon. As discussed earlier, baseline levels of risk as well as the occurrence of shocks are substantially different between treatment and control groups. I estimate the importance of mean reversion in the control group only and find estimates of $\rho - 1$, δ and $g(R_t)u_t$ equal to -0.61, 0.03 and -0.24 respectively.[23]

Following Chay, McEwan, and Urquiola (2005), I account for shock induced mean reversion by adjusting equation (17) in a way that eliminates sources of correlation between $\triangle D_{ij}$ and $(\upsilon_{ij,t+1} - \upsilon_{ij,t})$. Using equation (19) to rewrite eq. (17) yields:

$$\begin{aligned} \triangle R_{ij} = {} & \beta_1 \triangle D_{ij} + \beta_2 \triangle X_i + \beta_3 \triangle Z_j + \triangle \delta \\ & + (\rho - 1) R_{ijt} + \delta u_{jt} + g(R_{ijt}) u_{jt} + \triangle \upsilon_{ij} \end{aligned} \tag{20}$$

I estimate a simplified version, such as:

$$\begin{aligned} \triangle R_{ij} = {} & \beta_1 \triangle D_{ij} + \beta_2 \triangle X_i + \beta_3 \triangle Z_j + \triangle \delta \\ & + \beta_4 R_{ijt} + \beta_5 u_{jt} + \beta_6 R_{ijt} u_{jt} + \triangle \upsilon_{ij} \end{aligned} \tag{21}$$

Before proceeding, one last empirical challenge needs to be addressed: within cluster correlation in $\triangle \upsilon_{ij}$. Studies that work with a small number of clusters always face the challenge of adequately adjusting standard errors for potential within cluster correlation of errors. Throughout the paper, I calculate Eicker-White standard errors clustered at the sub-district (block) level. However, since the number of blocks is fairly small, these standard errors are likely to be downward biased (Cameron, Gelbach, and Miller, 2008). In cases of very few clusters, Cameron, Gelbach, and Miller (2008) suggest using a wild cluster-bootstrap with Rademacher weights. This approach was applied, inter alia, by

[23] I use the level and the square of R_t as approximation for $g(R_t)$. Detailed results can be found in the Supplementary Appendix, Table D.2.

Adrianzen (2014) to data clustered in 26 villages and by Akosa Antwi, Moriya, and Simon (2013) to 28 quarter-year groups. In a more recent paper, Cameron and Miller (2015) suggest the use of Webb's (2014) weights if the number of clusters is smaller than ten, which seems reasonable when using a district level treatment variable. P-values of the respective treatment variable using both versions of the bootstrap with 4,999 replications are reported at the bottom of Table 4.[24]

5 Results

This section starts by presenting estimates for an agricultural production function. It proceeds by assessing the extent to which the NREGS can actually support households in this sample in coping with shocks, which is the precondition for expecting any insurance effect. This section then analyzes the effects of the NREGS on households' crop choices and presents a number of robustness checks.

5.1 Identifying profitable production strategies

To understand inhowfar households' crop choice can improve their income from agricultural production, I estimate an agricultural production function, linking the total value of agricultural output Q_{ijt} to input allocation K_{ijt}, labor L_{ijt}, plot size A_{ijt} and risky crop choice R_{ijt}. I estimate agricultural output assuming a Cobb-Douglas production function, in which the choice of crops affects output multiplicatively in the following manner:

$$Q_{ijt} = (K_{ijt}^{\beta_1} L_{ijt}^{\beta_2} A_{ijt}^{\beta_3}) e^{g(R_{ijt})} \tag{22}$$

I allow R_{ijt} to affect output non-linearly because it seems very likely that increasing the average risk in a crop portfolio is only beneficial to a certain extent, beyond which risk is simply too high to increase output. The production function described in equation (22)

[24]The wild cluster-bootstrap calculates t-statistics for each bootstrap sample and estimates rejection rates based on the resulting distribution of t-statistics. Therefore, I report clustered standard errors throughout the text. Implementation in Stata is done based on the do-file written by Douglas Miller, which can be accessed online: http://www.econ.ucdavis.edu/faculty/dlmiller/statafiles/.

can be estimated by log-transforming the data and controlling for shocks Z_{ijt}, unobserved characteristics γ_{ij} and time effects δ_t. Again, I use the level and the square of R_{ijt} as approximation for $g(R_{ijt})$:

$$\begin{aligned} \ln(Q_{ijt}) = \beta_0 + \beta_1 \ln(K_{ijt}) + \beta_2 \ln(L_{ijt}) + \beta_3 \ln(A_{ijt}) + \beta_4 R_{ijt} + \beta_5 R_{ijt}^2 \\ + \beta_6 Z_{ijt} + \gamma_{ij} + \delta_t + \upsilon_{ijt} \end{aligned} \tag{23}$$

The equation is estimated in OLS, random effects and fixed effects. As can be seen in Table 2, all models generate similar results. Columns (1) and (2) report OLS estimates for the survey round of 2007. These show that estimates are not affected by the exclusion of labor from the agricultural production function.[25] In columns (4) and (6), I additionally allow the effect of rainfall to vary with the amount of risk in a household's crop portfolio.

The estimates in Table 2 suggest that households could significantly raise the value of their agricultural production if they were to increase the share of inputs allocated to riskier crops. However, this is only true up to a certain level. The square of the risk index is statistically significant at the 5% level in all specifications that use both rounds of data. With estimates of column (6), Figure 2 plots predicted agricultural output as function of the risk index in the crop portfolio. A can be clearly seen, an increase in risk increases the value of agricultural production, reaching its maximum at a risk index of 0.42. Beyond this point, a further increase in risk would reduce total agricultural output. Average risk levels in households' crop portfolios are well below this value; in the survey round of 2007 the average risk index was 0.36 in the treatment group and 0.26 in the control group (c.f. Table 1).

Other variables, such as the level of inputs allocated, total cultivated area and labor have the expected sign and are all statistically significant.[26] Additionally, the share of area under irrigation seems to increase output levels. In contrast, the dummies indicating whether or not a household applied fertilizer or high yielding variety (HYV) seeds are

[25] I cannot control for labor in the panel data models, because time information was only collected in 2007 and not in 2009/10.

[26] The level of inputs allocated is the total amount spent on variable inputs, such as seeds, fertilizer, pesticides and so forth. Manual labor is accounted for if hired in.

not statistically significant. This might seem somewhat surprising, but since expenditure on fertilizer and seeds is included in variable inputs, one should not attribute too much weight to this estimate.

Interestingly the interaction term of rainfall and the risk index is positive and statistically significant at the 5% level. Computing marginal effects at different levels of risk show that rainfall affects agricultural output only at very high risk levels. At the optimal risk level of 0.42, the marginal effect of rainfall is as high as 0.24 with a standard error of 0.17.

If households are able to increase the value of their agricultural production by producing a greater share of risky crops, this raises the question why they do not do so. First of all, it may not be possible to generalize these results to extended periods of time if, for instance, the two survey years were exceptionally dry or exceptionally productive. Therefore, I additionally consider state-level statistics on the returns per hectare for major crops between 1996 and 2009.[27] Figure 3 plots the average returns of different crops against the standard deviation of these returns for the years 1996 to 2006 in Andhra Pradesh. The scatter plot shows a clear positive relationship between average returns and their volatility, indicating again that the riskiness of crops is strongly correlated with returns to producing these crops.

5.2 Does the NREGS support households in coping with shocks?

If risk is a relevant constraint for households' production decisions, the provision of an employment guarantee should enable households to grow a larger share of risky crops. This is the main prediction of the theoretical model presented in this paper. Crucial to this prediction are households' expectations about opportunities to smooth income in the event of a shock. Therefore, we need to understand the extent to which the NREGS helps households in coping with shocks.

I test whether deviations from mean rainfall levels, as well as households' self-reported shocks, drive changes in the number of days households work for the NREGS. I argue

[27]Unfortunately, these statistics are only available at state level and only for very few crops.

that the NREGS will have an insurance effect only if work provision sufficiently reacts to increasing demand in the case of a shock. This question is tested in a fixed effects model. The results are reported in Table 3. In the first two columns, the total number of days worked in the past 12 months is the dependent variable; in the last two columns it is the log of this variable. The estimation is also restricted to phase one districts; thus only households who had access to the NREGS in both survey rounds are considered.

The results suggest that the number of days worked for the NREGS changes considerably with variation in rainfall levels. The greatest change is observed for lagged rainfall levels - that is, cumulative rainfall in the agricultural year preceding the period of reference. The coefficient of the lagged rainfall variable is large and negative, which implies that households worked more for the NREGS if lagged rainfall levels were below average and worked less if lagged rainfall was above average. This supports the assumption that the NREGS helps households in coping with shocks, because households use the program to smooth income ex post - for instance, after harvest and after agricultural products have been sold. Similar evidence is provided by Johnson (2009), who finds that the number of days households work for NREGS increases if rainfall levels are lower than average.

Table 3 also shows how important maturation of the program is. A large share of the variance in the number of days worked for the NREGS can be explained by time alone. In contrast, wealth levels do not seem to influence the dependent variable, and the size of the cultivated area is only statistically significant in one specification. This is probably due to the limited variation of this variable over time.[28] Self-reported shocks also seem to increase the number of days worked for the NREGS.[29]

To quantify the contribution of the NREGS to households' risk coping, I compare agricultural losses due to rainfall shortages with income gains through the NREGS. The agricultural production function estimated in Section 5.1 (col. 6) suggests that a deviation from average annual rainfall by negative 25%, would reduce agricultural output by 5.9% at the optimal level of the risk index. For the average household, this implies a nominal

[28] A positive coefficient could indicate program capture by wealthier households. But a further investigation of this issue is beyond the scope of this paper.

[29] The variable is coded as one of a household reported any of 12 self-reported shocks related to agricultural production.

loss of about INR 1,740 (or US$ 37.5 in constant July 2006 values). The same deviation in lagged rainfall would lead households to work about 15.8 more days for the NREGS, which would generate an additional income of INR 1,020 (US$ 22) at mean wages observed in the sample. The NREGS thus allows households to compensate about 58% of agricultural production losses caused by rainfall shortages. Since rainfall fluctuations are among the most important sources of risk for rural households, these results suggest that the NREGS could indeed have an insurance effect in Andhra Pradesh.

5.3 The effects of the NREGS on households' crop choices

In this section I estimate the effect of the NREGS on households' input allocation decisions. Table 4 reports the effects of the NREGS on the risk index of a household's crop portfolio. As described in Section 4, I estimate all equations in first differences and control for initial condition in columns (2), (4) and (6). To isolate the insurance effect described in Section 2.4, I control for variables that might be affected by the NREGS and might influence a household's crop choice through effects other than the insurance effect. These variables include household off-farm income and wealth, as well as key farming characteristics, such as the size of cultivated land, irrigation and total value of variable inputs allocated.[30] In all specifications I also control for self reported shocks, access to other government programs and rainfall levels (current and lagged). Additionally, 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 a household's crop choice.

The results show a positive and statistically significant effect of the NREGS on the riskiness of households' production decisions. Consistent with the higher prevalence of shocks in the treatment districts and higher initial values of the risk index, controlling for mean reversion increases the estimated effect of the NREGS. All three treatment vari-

[30]Household off-farm income consists, inter alia, of income generated through the NREGS in the past 12 months. Optimally, this should be a lagged value because input allocation decisions are taken at the beginning of the season, while the income variable refers to the time period shortly after these allocative decisions were taken. Unfortunately, the survey does not include this information. Table D.3 in the Supplementary Appendix shows that the results are not influenced by changes in income or changes in total input allocation.

ables generate positive and statistically significant results. Results presented in column (2) suggest that the risk index in households' crop portfolios increased by 7.2 percentage points due to the introduction of the NREGS at the district level. Given that the risk index in the treatment group was 36% at baseline, the introduction of the NREGS raised the average risk index to 0.43 (absent any shock induced mean reversion), which is remarkably close the optimal risk index identified in Section 5.1. In columns (3) to (6), I also test for the effects of cumulative expenditure and total employment generated per job card under the NREGS at block level. These variables are lagged by one year, to avoid correlation of the treatment intensity with past shocks.

In terms of economic relevance, the results suggest that per additional day of employment generated in the block, each household would increase the risk index by 0.15 percentage points (col. 5). One standard deviation increase in the number of person-days generated per job card (6.9) would increase a household's risk index by 1.07 percentage points and raise net income from agricultural production, *ceteris paribus*, by about INR 480 (or US$ 10.4 in constant July 2006 values). This is particularly interesting from a cost-benefit perspective, since these net income gains per household are slightly higher than the wage cost (evaluated at the sample average of observed NREGS wages) of creating 6.9 days of employment under the NREGS, e.g. INR 467 (US$ 10). Of course, wage costs make up for only a part of overall program costs and not all of the NREGS participants own their own land, but nevertheless the magnitude of this effect is striking.

5.4 Robustness checks

Because the results presented so far rely heavily on the parallel trends assumption, I perform a number of robustness checks. The first set of robustness checks is intended to rule out the possibility that the observed effects is not due to the NREGS. The second set of robustness checks, which will be presented in the following section is intended to rule out potential alternative mechanisms through which the NREGS could affect crop choices.

As a first robustness check, I test whether households that registered with the NREGS change their input allocation more strongly than households who are not registered with the NREGS. To account for potential self-selection bias, I match households on their probability to register with the NREGS by using entropy balancing, a method developed by Hainmueller (2012). Entropy balancing seems to outperform most existing matching algorithms in terms of the balance reached on the entire set of relevant covariates (Hainmueller, 2012). The matching algorithm assigns weights to all observations in the control group such that the distribution of selected variables matches the observed distribution in the treatment group. These weights can then be used as sampling weights in the estimation.[31] I match households on the mean and the variance of variables that determine a household's registration with the NREGS and potentially influence post-treatment outcomes, such as cost incurred in agricultural production, total cultivated area, percentage of area irrigated, a dummy indicating whether a household participates in labor sharing in agriculture, wealth levels and off-farm income, and household characteristics, e.g. education, age and sex of the household head, indebtedness, and the ability to raise INR 1,000 (US$ 21.6) in one week. The resulting covariate balance is displayed in Table 5. As can be seen, this method reaches a perfect balance on all variables included.

Table 6 reports the effects of registering with the NREGS on the risk index of households' crop portfolios. I find that households that already registered with the NREGS in 2007 are more likely to grow a higher share of risky crops in the follow-up period. Five different specifications are presented: Columns (1) and (2) show estimates without matching, where column (2) additionally controls for initial conditions. Column (3) excludes all households that did not register with the NREGS by 2009.[32] Column (4) shows estimates for the matched sample. As we can see, the effects are only slightly smaller when matching households on their probability to register with the NREGS. Overall, the effects are of a similar size in most specifications though somewhat lower than the estimates presented in Table 4, column (2). Column (5) shows the estimation results for

[31]Since I estimate the model on a balanced sample, the same weights can be applied to the 2009/10 round of interviews.

[32]This is to exclude all households from the sample that - either because they consider it socially undesirable or because they have other means of risk coping - would probably never register with the NREGS.

the full sample without matching. Here, being registered by 2009 is the main explanatory variable. As we would expect, households that registered with the NREGS only shortly before or even after deciding on their crop portfolio, did not alter their input allocation in a meaningful way.

Finally, I test if the results presented so far are sensitive to the choice of the dependent variable. Table D.4 in the Supplementary Appendix shows that the results are robust to selecting alternative dependent variables, such as the weighted average of the standard deviation of crop returns, but also to different methods of aggregating the risk index.[33]

5.5 Alternative explanations

As mentioned before, the NREGS can affect household decisions via different mechanisms. This section seeks to understand if the observed effect of the NREGS is indeed an insurance effect and not due to alternative mechanisms such as the increase in income of participating households or the change in agricultural wages. If, for example, risky crops are also more capital intensive, then observed outcomes could also be driven by increases in income and wealth or better access to credit of participating households. Likewise, if risky crops are also less labor intensive, then observed outcomes could be driven by wage changes due to the NREGS instead of its insurance effect.

I start by testing if the NREGS has effects on the labor and cost intensity of households' crop portfolios. Labor intensity per crop is calculated as the share of expenditures on labor in total production costs. Cost intensity is defined as the total production cost that has to be incurred per Hectare for each crop. Both estimates are based on the crop-wise Cost of Cultivation Statistics, published by the Ministry of Agriculture (see Supplementary Appendix for more details). Table 7, col. (1) to (4) shows estimates of the effect of the NREGS on the labor intensity of households' crop portfolios. Columns (5) to (8) summarize estimates of the effect of the NREGS on cost intensity in households' crop portfolios.[34] The coefficient on labor intensity is positive, indicating that

[33]More information to the calculation of these indices can be found in the Supplementary Appendix, Section B.2.

[34]Columns (2), (4), (6) and (8) additionally control for initial conditions in rainfall, a household's risk index and the interaction term of rainfall and risk, similar to the estimates presented in Section 5.3.

the NREGS, if anything, increases the labor intensity of crop portfolios. The coefficient on cost intensity is also positive, suggesting that households are able to spend more on their agricultural production. However, only two out of 8 specifications are statistically significant, and as soon as I control for initial conditions the effect of the NREGS on labor and cost intensity disappears. This suggests, that the NREGS acts through the insurance effect more than through the wage or income mechanism.

The presence of alternative mechanisms through which the NREGS could affect production decisions also means that households might register with the NREGS for different reasons. For some households, consumption needs are a much more important reason for registering with the program than the insurance effect. We would thus expect households to react differently to the availability of the NREGS depending on whether the program can contribute to smoothing their incomes in the case of a shock. Households that need to work for the NREGS as much as possible to satisfy their consumption needs - even in good years - are unlikely to cultivate higher risk crops despite working for the NREGS. In contrast, households that rely on the NREGS mainly in the case of a shock are expected to react differently. One option to separate these two groups is to condition the treatment effect on the main reason for households in the second group to register with the program: the experience of a shock to agricultural production. Since rainfall fluctuations are among the most important risks to agricultural production, I interact the treatment variable with the lagged deviation in rainfall levels. Table 8, columns (1) and (2) report the results of the regression with interaction terms. For better visualization, the marginal effect of registering with the NREGS conditional on lagged rainfall is plotted in Figure 4. It shows that the treatment effect is large and statistically significant for households that experienced rainfall shocks before registering with the NREGS and diminishes for households who registered despite more favorable rainfall levels. This suggests that households that registered with the NREGS to cope with a shock are much more likely to adjust their input allocation towards more profitable crops, which is exactly what we would expect in case of an insurance effect. In contrast, households that had already registered with the NREGS in 2007 even though they experienced a good year in their agricultural production do not adjust their production decisions. These are probably households that rely on the NREGS for additional income rather than as a risk-coping instrument.

In addition to the robustness checks presented so far, I also test the extent to which treatment effects vary depending on the initial presence of other government programs, such as watershed development projects, crop insurance schemes or public works programs other than the NREGS. The results are reported in Table 8, columns (3) to (5). I find that treatment effects are smaller in villages with existing watershed development projects, crop insurance programs and public works schemes, although only the coefficient on watershed development projects is statistically significant at the 10% level. These results again support the hypothesis that the NREGS has an insurance function for households because observed effects on input allocation are smaller if households already have access to other insurance or risk mitigation mechanisms.

6 Conclusions

This paper presents theoretical and empirical evidence that an employment guarantee, such as the NREGS in India, improves households' ability to cope with shocks in agriculture by guaranteeing income opportunities in areas where and time periods when they previously did not exist. By improving the risk management of households, the NREGS enables households to switch their production towards riskier but also higher profitability products and to generate higher incomes from agricultural production.

The results of this paper show that public works programs can have welfare effects that go beyond immediate income effects. The insurance effect of the NREGS on agricultural productivity is similar to the effects of rainfall insurance analyzed by Karlan et al. (2013), Cole, Gine, and Vickery (2013) and Mobarak and Rosenzweig (2013). But in contrast to purchasing insurance, registration with the NREGS provides little ex ante cost. Since trust-related considerations continue to limit the uptake of insurance products in many countries, providing public works schemes - combined with an employment guarantee - could be an alternative option with which to protect households against agricultural production risks and to enable productivity gains in agriculture.

Current discussions regarding the effects of the NREGS on agricultural productivity focus mainly on the trade-off between providing minimum income to poor households, on one hand, and ensuring that production costs in the agricultural sector do not rise too drastically due to increased agricultural wages, on the other hand. As this paper shows, these discussions have failed to consider the following key aspect: because the number of workdays each household is entitled to additionally affects its risk management capacity, the amount of risk each household is willing to take in tis own agricultural production - and therewith potential productivity gains - will crucially depend on the number of days each household can expect to be able to work in the case of production shocks. Thus, increasing the number of days each household is entitled to work with the NREGS could increase agricultural productivity - an argument that has been largely ignored so far. The assumption that only large-scale farmers can raise agricultural productivity is still a mainstream one. Including in the discussion the effects of the NREGS on households' risk management and the resulting changes in production decisions might change the overall picture.

The findings here contain some lessons for the ongoing debates on the effectiveness of the NREGS and for other countries considering the implementation of such schemes. First, for the insurance effect to unfold, the design of a public works program is crucial. An employment guarantee that is entitled by law and entails adequate grievance redress mechanisms provides households with the necessary protection against agricultural production risks to enable them to take more risks in their production and investment decisions. Additionally, it is crucial not to severely limit the number of workdays, otherwise such a scheme's potential as a risk-coping instrument cannot be realized. Second, implementation matters. The data analyzed in this paper cover only the state of Andhra Pradesh. This is, *inter alia*, because the performance of the NREGS in terms of the 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 in the implementation of the NREGS, so it goes without saying that many of the effects captured in this paper might not be found in all Indian states. Third, working for a public works scheme is always associated with opportunity costs. In countries or regions with well functioning off-farm labor markets, providing public works schemes might not be necessary. A food-for-work program or cash-for-work program will always only be effective in areas and time periods where labor is in surplus.

Obviously, a number of open questions remain, and more research is required to provide conclusive answers to these questions. First, the internal and external validity of the results here could be improved with more data - especially if the analysis were extended to the whole country or to other countries with similar programs. Second, the effects of the program on total levels of inputs allocated and on investments in fixed capital could prove to be very interesting topics of study. Similarly, the effects on households' willingness to engage in entrepreneurial activity need to be assessed. Third, heterogeneity in treatment effects could be assessed in more detail with more data.

References

Adrianzen, M.A. 2014. "Social Capital and Improved Stoves Usage Decisions in the Northern Peruvian Andes." *World Development* 54:1 – 17.

Akosa Antwi, Y., A.S. Moriya, and K. Simon. 2013. "Effects of Federal Policy to Insure Young Adults: Evidence from the 2010 Affordable Care Act's Dependent-Coverage Mandate." *American Economic Journal: Economic Policy* 5:1–28.

Barrett, C., S. Holden, and D.C. Clay. 2005. "Can food-for-work programmes reduce vulnerability?" In S. Dercon, ed. *Insurance against poverty*. New York: Oxford University Press, pp. 361–386.

Berg, E., S. Bhattacharyya, R. Durgam, and M. Ramachandra. 2012. *Can rural public works affect agricultural wages? Evidence from India*. No. 2012-05 in CSAE Working Papers.

Binswanger-Mkhize, H.P. 2012. "Is there too much hype about index-based agricultural insurance?" *The Journal of Development Studies* 48:187–200.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90:414– 427.

Cameron, A.C., and D.L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50:317–372.

Carter, M.R., A. de Janvry, E. Sadoulet, and A. Sarris. 2014. *Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling*. No. P111 in FERDI Working Papers.

Chay, K.Y., P.J. McEwan, and M. Urquiola. 2005. "The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools." *American Economic Review* 95:1237–1258.

Cole, S., X. Gine, J. Tobacman, R. Townsend, P. Topalova, and J. Vickery. 2013. "Barriers to household risk management: evidence from India." *American Economic Journal: Applied Economics* 5:104–135.

Cole, S., X. Gine, and J. Vickery. 2013. *How does risk management influence production decisions? Evidence from a field experiment*. No. 6546 in World Bank Policy Research Working Paper.

Dercon, S., and L. Christiaensen. 2011. "Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia." *Journal of Development Economics* 96:159–173.

Duflo, E., M. Kremer, and J. Robinson. 2008. "How high are rates of return to fertilizer? Evidence from field experiments in Kenya." *The American Economic Review* 98:482–488.

Dutta, P., R. Murgai, M. Ravallion, and D. Van de Walle. 2012. *Does India's Employment*

Guarantee Scheme Guarantee Employment?. No. 6003 in World Bank Policy Research Working Paper.

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

Foster, A.D., and M.R. Rosenzweig. 2010a. *Is There Surplus Labor in Rural India?*. No. 85 in Yale University, Department of Economics Working Papers.

—. 2010b. "Microeconomics of technology adoption." *Annual Review of Economics* 2:395–424.

Galab, S., S. Vijay Kumar, P. Prudvikhar Reddy, R. Singh, and U. Vennam. 2011. *The Impact of Growth on Childhood Poverty in Andhra Pradesh: Initial Findings from India*. Young Lives Round 3 Survey Report: India.

Hainmueller, J. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis* 20:25–46.

Imbert, C., and J. Papp. 2015. "Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee." *American Economic Journal: Applied Economics* 7:233–63.

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

Jha, R., R. Gaiha, and M.K. Pandey. 2012. "Net transfer benefits under India's rural employment guarantee scheme." *Journal of Policy Modeling* 34:296–311.

Johnson, D. 2009. *Can Workfare Serve as a Substitute for Weather Insurance? The Case of NREGA in Andhra Pradesh*. No. 32 in Centre for Micro Finance Working Paper Series.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry. 2013. "Agricultural Decisions after Relaxing Credit and Risk Constraints." *Quarterly Journal of Economics* 129:597–652.

Kaur, S. 2014. *Nominal Wage Rigidity in Village Labor Markets*. Columbia University Working Paper.

Klonner, S., and C. Oldiges. 2014. *Safety Net for India's Poor or Waste of Public Funds? Poverty and Welfare in the Wake of the World's Largest Job Guarantee Program*. No. 0564 in University of Heidelberg, Department of Economics Working Papers.

Mobarak, A.M., and M.R. Rosenzweig. 2013. "Informal Risk Sharing, Index Insurance, and Risk Taking in Developing Countries." *American Economic Review* 103:375–80.

Ministry of Rural Development (MoRD), 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.

Sadoulet, E., and A. De Janvry. 1995. *Quantitative development policy analysis*. Johns Hopkins University Press Baltimore.

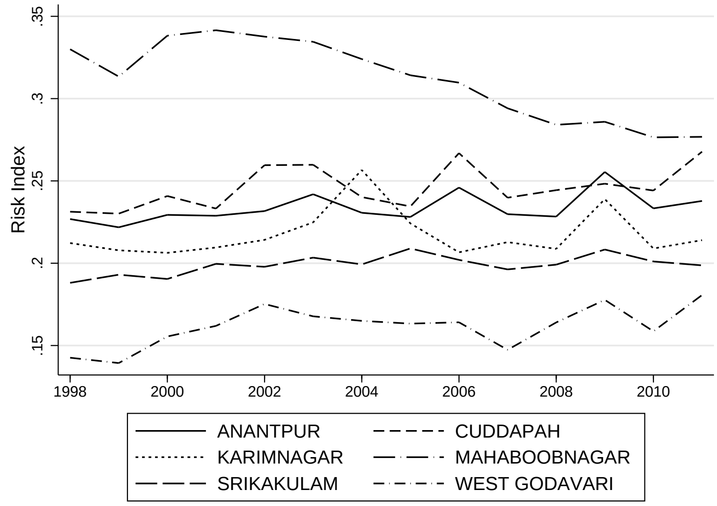
Suri, T. 2011. "Selection and Comparative Advantage in Technology Adoption." *Econometrica* 79:159–209.

Van Den Berg, M. 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* 7:487–506.

Webb, M.D. 2014. *Reworking Wild Bootstrap Based Inference for Clustered Errors*. No. 1315 in Queen's University Working Paper.

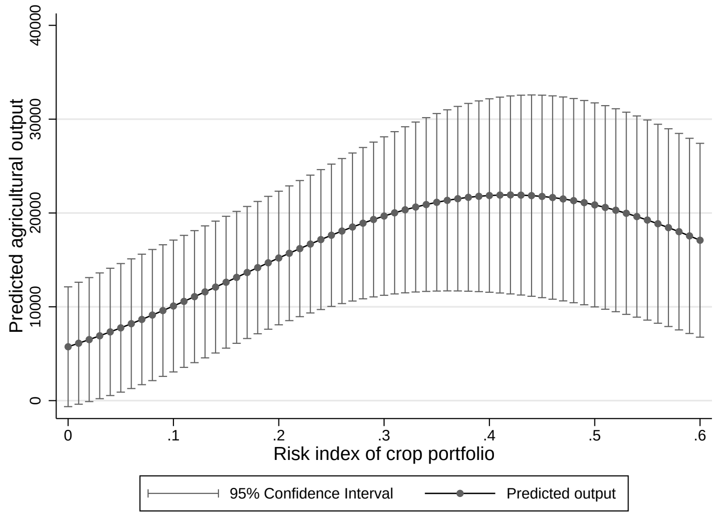
Figures

Figure 1: District-wise risk-index of land use



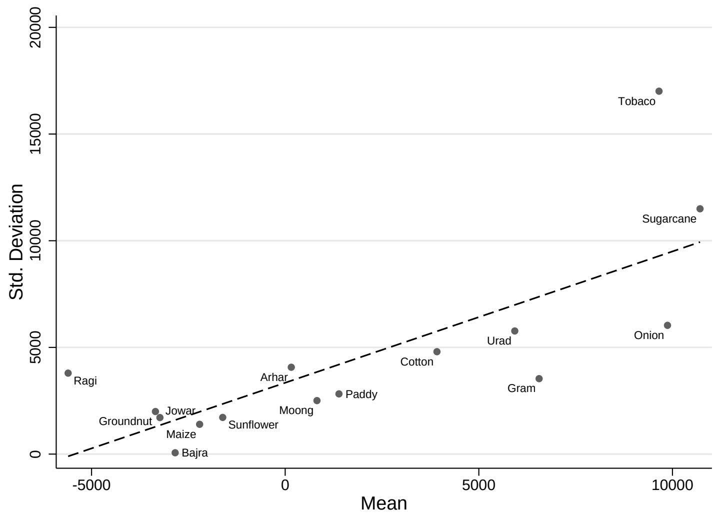
Source: Own estimation based on the Land use statistics and District-wise crop production statistics, Ministry of Agriculture, GoI.

Figure 2: Agricultural output as function of the riskiness of crops



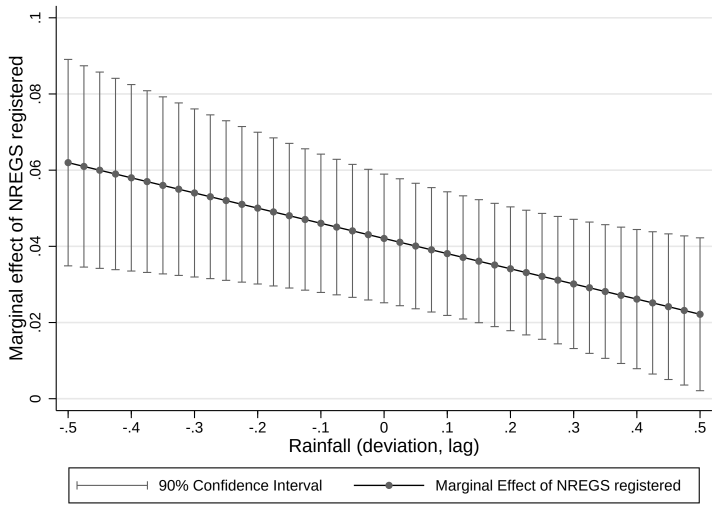
Source: Own estimation based on the Young Lives data.

Figure 3: Returns per hectare of major crops



Source: Own estimation based on the Cost of cultivation statistics for Andhra Pradesh, Ministry of Agriculture, GoI.

Figure 4: Effect of the NREGS on risk index conditional on lagged rainfall



Source: Own estimation based on the Young Lives data.

Tables

Table 1: General household characteristics

	Treatment				Control			
	2007		2009/10		2007		2009/10	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Male household head	0.96	(0.20)	0.95	(0.21)	0.97	(0.18)	0.96	(0.21)
Age of hh head	41.93	(12.13)	41.48	(10.44)	41.01	(11.83)	41.28	(9.67)
Household head is literate	0.32	(0.47)	0.32	(0.47)	0.25	(0.43)	0.25	(0.43)
Household size	6.10	(2.62)	6.01	(2.77)	5.61	(2.07)	5.51	(2.06)
Wealth index	0.39	(0.13)	0.46	(0.13)	0.38	(0.20)	0.45	(0.19)
Hh benefits from credit/training	0.62	(0.49)	0.79	(0.41)	0.58	(0.49)	0.76	(0.43)
Annual income, off-farm activities	24.70	(24.82)	32.41	(35.81)	19.81	(26.13)	24.34	(27.24)
Any serious debts	0.63	(0.48)	0.39	(0.49)	0.47	(0.50)	0.27	(0.45)
Able to raise INR 1000 in one week	0.61	(0.49)	0.51	(0.50)	0.33	(0.47)	0.58	(0.49)
Value of agr. production	28.49	(45.76)	34.36	(56.22)	24.38	(124.96)	25.60	(98.61)
Value of variable inputs	14.51	(21.34)	17.67	(20.48)	14.42	(69.52)	14.37	(92.25)
Area cultivated (acres)	4.14	(4.57)	4.33	(4.43)	2.73	(5.47)	2.58	(3.28)
Risk index of crop portfolio	0.36	(0.12)	0.36	(0.10)	0.26	(0.08)	0.25	(0.07)
Irrigated area (% of total)	0.18	(0.32)	0.18	(0.30)	0.14	(0.30)	0.10	(0.26)
Fertilizer (dummy)	0.98	(0.15)	0.99	(0.10)	0.87	(0.34)	0.83	(0.38)
HYV seeds (dummy)	0.77	(0.42)	0.61	(0.49)	0.63	(0.48)	0.52	(0.50)
Participated in labor sharing (dummy)	0.75	(0.43)	0.66	(0.47)	0.78	(0.41)	0.77	(0.42)
Total hours hh worked in agriculture	2085	(2280)	.	(.)	1365	(1310)	.	(.)
Rainfall (deviation)	0.33	(0.28)	0.02	(0.29)	-0.06	(0.16)	0.03	(0.29)
Rainfall (deviation, lag)	-0.39	(0.10)	0.28	(0.28)	-0.12	(0.10)	0.23	(0.23)
Self-reported shock	0.81	(0.39)	0.78	(0.41)	0.51	(0.50)	0.38	(0.49)
Hh registered with NREGS	0.66	(0.47)	0.77	(0.42)	0.00	(0.00)	0.78	(0.42)
Hh generated income from NREGS	0.54	(0.50)	0.71	(0.46)	0.00	(0.00)	0.75	(0.43)
Income, NREGS	1.24	(2.39)	2.75	(3.69)	0.00	(0.00)	2.97	(3.61)
Cumulative expend. NREGS (lag)	0.00	(0.00)	142.76	(96.42)	0.00	(0.00)	0.00	(0.00)
Cumulative expend. NREGS	0.00	(0.00)	506.40	(210.72)	0.00	(0.00)	194.03	(152.01)
Observations	750		750		338		338	

Notes: All values in constant INR 1000 (July 2006). One US$ is equivalent to 46.38 INR (July 2006). NREGS expenditure is in current INR 100,000. Variable definitions and sources are described in the Online Appendix, Section B.

Table 2: Agricultural Production Function

	2007 OLS		Random Effects		Fixed Effects	
	(1)	(2)	(3)	(4)	(5)	(6)
Risk index of crop portfolio	6.020*	5.752*	6.652**	6.603**	6.670*	6.281*
	(2.595)	(2.787)	(2.085)	(2.063)	(2.828)	(2.744)
Risk index of crop portfolio (squared)	-6.878^{+}	-7.091^{+}	-8.444**	-8.627**	-7.873*	-8.154*
	(3.585)	(3.929)	(2.833)	(2.785)	(3.364)	(3.273)
Variable inputs (log)	0.793***	0.864***	0.910***	0.907***	0.736***	0.743***
	(0.108)	(0.109)	(0.087)	(0.087)	(0.117)	(0.118)
Area cultivated (acres, log)	0.621***	0.783***	0.867***	0.876***	0.655***	0.670***
	(0.141)	(0.155)	(0.140)	(0.139)	(0.177)	(0.176)
Labor (hours, log)	0.224***					
	(0.054)					
Irrigated area (% of total)	0.423	0.497^{+}	0.336^{+}	0.348^{+}	0.113	0.118
	(0.286)	(0.286)	(0.195)	(0.196)	(0.288)	(0.291)
Fertilizer (dummy)	-0.295	-0.306	-0.398	-0.417	0.450	0.399
	(0.388)	(0.423)	(0.266)	(0.270)	(0.322)	(0.324)
HYV seeds (dummy)	0.052	0.026	-0.004	0.014	0.169	0.237
	(0.193)	(0.198)	(0.120)	(0.122)	(0.157)	(0.157)
Self-reported shock	-0.010	0.045	-0.284*	-0.278*	-0.322^{+}	-0.326^{+}
	(0.197)	(0.197)	(0.128)	(0.127)	(0.164)	(0.168)
Rainfall (deviation)	-0.439	-0.393	-0.194	-0.711	-0.018	-1.161^{+}
	(0.291)	(0.294)	(0.155)	(0.546)	(0.181)	(0.629)
Rainfall (deviation) × Risk index				1.515		3.327*
				(1.359)		(1.567)
Rainfall (deviation, lag)	0.556	0.532	0.614*	0.648*	0.262	0.421
	(0.638)	(0.631)	(0.249)	(0.262)	(0.298)	(0.328)
Year 2009 (dummy)			-0.443*	-0.448*	-0.136	-0.196
			(0.196)	(0.198)	(0.230)	(0.240)
Observations	1088	1088	2176	2176	2176	2176
R^2	0.315	0.292			0.129	0.132

Notes: Dep. var.: Log(total income from agricultural production). Controls in col. (1) & (2) include household characteristics: age, sex, education of household head and houshold size. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3: Number of days worked with NREGS (Fixed Effects)

	(1)	(2)	(3)	(4)
	NREGS days		NREGS days (log)	
Rainfall (deviation, lag)	-65.423**	-63.104**	-2.945**	-2.873**
	(18.461)	(19.151)	(0.887)	(0.845)
Rainfall (deviation)	-28.798**	-31.004**	-0.800	-0.876
	(8.700)	(8.935)	(0.554)	(0.497)
Self-reported shock	1.543	1.437	0.174*	0.179*
	(3.631)	(3.899)	(0.069)	(0.077)
Area cultivated (acres, log)		6.962*		0.220
		(2.935)		(0.158)
Wealth index		-11.202		0.022
		(19.906)		(0.576)
Hh benefits from credit/training		10.571*		0.367*
		(3.805)		(0.139)
Year 2009 (dummy)	54.750***	62.716**	2.455***	2.757**
	(7.749)	(14.905)	(0.331)	(0.656)
Observations	1490	1490	1490	1490

Notes: Clustered standard errors in parentheses. Region-time trends included, but not reported. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Effect of the NREGS on risky crop choices (First Differences)

	(1)	(2)	(3)	(4)	(5)	(6)
D.NREGS introduced in district	0.038**	0.072***				
	(0.010)	(0.016)				
D.Cumulative expend. NREGS (log, lag)			0.007**	0.014***		
			(0.002)	(0.003)		
D.Employment per JC generated, NREGS (log, lag)					0.002+	0.002
					(0.001)	(0.002)
D.Variable inputs (log)	-0.016	-0.005	-0.015	-0.004	-0.013	-0.002
	(0.008)	(0.006)	(0.007)	(0.005)	(0.007)	(0.005)
D.Area cultivated (acres, log)	0.017	0.011	0.017	0.011	0.017	0.013
	(0.009)	(0.007)	(0.009)	(0.006)	(0.009)	(0.006)
D.Irrigated area (% of total)	-0.033*	-0.024	-0.034*	-0.025	-0.034*	-0.025
	(0.014)	(0.013)	(0.014)	(0.013)	(0.014)	(0.012)
D.Fertilizer (dummy)	-0.042	-0.028*	-0.042	-0.028*	-0.041	-0.027
	(0.025)	(0.012)	(0.025)	(0.013)	(0.026)	(0.016)
D.HYV seeds (dummy)	0.013	0.025**	0.014	0.026***	0.015	0.026**
	(0.006)	(0.007)	(0.007)	(0.006)	(0.008)	(0.008)
D.Participated in labor sharing (dummy)	-0.003	-0.002	-0.003	-0.002	-0.004	-0.003
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
D.Annual income, off-farm activities (log)	-0.000	-0.003	-0.000	-0.003	-0.000	-0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
D.Hh benefits from credit/training	-0.011	-0.013**	-0.011	-0.013**	-0.011	-0.012*
	(0.006)	(0.004)	(0.006)	(0.004)	(0.006)	(0.004)
D.Self-reported shock	0.009	0.001	0.009	0.001	0.010	0.004
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)
D.Rainfall (deviation)	0.025**	0.050*	0.024*	0.035	0.015	0.026
	(0.008)	(0.020)	(0.009)	(0.020)	(0.009)	(0.024)
D.Rainfall (deviation, lag)	-0.044***	-0.042*	-0.045**	-0.036*	-0.036	-0.013
	(0.011)	(0.017)	(0.013)	(0.015)	(0.018)	(0.025)
Rainfall (deviation) at baseline		0.121		0.109		0.175*
		(0.069)		(0.072)		(0.080)
Risk index at baseline		-0.603***		-0.582***		-0.512***
		(0.050)		(0.058)		(0.068)
Risk index × rainfall (baseline)		-0.109		-0.152		-0.296
		(0.126)		(0.132)		(0.156)
Bootstrap p-value of main treatment variable						
Rademacher weights:	0.107	0.047	0.072	0.015	0.326	0.388
Webb weights:	0.099	0.045	0.062	0.013	0.315	0.391
Observations	1088	1088	1088	1088	1088	1088
R^2	0.067	0.443	0.066	0.435	0.058	0.400

Notes: Clustered standard errors in parentheses. P-values are obtained by performing a wild cluster-t bootstrap with 4999 replications and two alternative weights. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5: Weighted summary statistics (2007)

	Treatment		Control			
			(not matched)		(matched)	
	Mean	SD	Mean	SD	Mean	SD
Value of variable inputs	12.81	(16.41)	15.88	(55.71)	13.95	(27.45)
Total area owned	3.36	(3.53)	3.53	(5.49)	3.36	(3.54)
Irrigated area (% of total)	0.14	(0.29)	0.19	(0.33)	0.14	(0.28)
Participated in labor sharing (dummy)	0.79	(0.41)	0.73	(0.44)	0.79	(0.41)
Annual income, off-farm activities	23.68	(21.48)	22.76	(28.15)	24.82	(28.31)
Male household head	0.96	(0.19)	0.96	(0.20)	0.96	(0.19)
Age of hh head	41.20	(12.07)	42.01	(12.01)	41.23	(11.72)
Household head is literate	0.32	(0.47)	0.28	(0.45)	0.32	(0.47)
Wealth index	0.37	(0.11)	0.40	(0.19)	0.37	(0.13)
Household size	6.02	(2.56)	5.88	(2.40)	6.02	(2.46)
Able to raise INR 1000 in one week	0.56	(0.50)	0.49	(0.50)	0.57	(0.50)
Any serious debts	0.67	(0.47)	0.50	(0.50)	0.67	(0.47)
Observations	496		592		592	

Notes: All values in constant INR 1000 (July 2006). One US$ is equivalent to 46.38 INR (July 2006). Variable definitions and sources are described in the Online Appendix, Section B.

Table 6: Effect of the NREGS on risk index of crop portfolio by registration status (First Differences)

	(1)	(2)	(3)	(4)	(5)
D.NREGS registered (2007)	0.019^{+}	0.035**	0.034**	0.027*	
	(0.010)	(0.010)	(0.010)	(0.009)	
D.NREGS registered (2009)					0.007
					(0.006)
D.Variable inputs (log)	-0.014^{+}	-0.003	-0.001	-0.001	-0.002
	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)
D.Area cultivated (acres, log)	0.018^{+}	0.013*	0.018*	0.011	0.013^{+}
	(0.010)	(0.006)	(0.007)	(0.007)	(0.007)
D.Irrigated area (% of total)	-0.032*	-0.025*	-0.019	-0.036*	-0.025*
	(0.013)	(0.012)	(0.013)	(0.016)	(0.011)
D.Fertilizer (dummy)	-0.039	-0.027^{+}	-0.025	-0.007	-0.025
	(0.026)	(0.015)	(0.016)	(0.012)	(0.016)
D.HYV seeds (dummy)	0.013	0.024**	0.022**	0.024**	0.023*
	(0.007)	(0.007)	(0.007)	(0.006)	(0.008)
D.Participated in labor sharing (dummy)	-0.004	-0.003	0.001	-0.003	-0.003
	(0.004)	(0.004)	(0.004)	(0.007)	(0.003)
D.Annual income, off-farm activities (log)	0.000	-0.002	-0.001	-0.005	-0.002
	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)
D.Hh benefits from credit/training program	-0.010	-0.011*	-0.009^{+}	-0.012*	-0.011^{+}
	(0.006)	(0.004)	(0.004)	(0.005)	(0.005)
D.Self-reported shock	0.010	0.005	0.003	0.000	0.005
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
D.Rainfall (deviation)	0.016	0.038^{+}	0.045^{+}	0.033	0.040
	(0.011)	(0.020)	(0.023)	(0.022)	(0.024)
D.Rainfall (deviation, lag)	-0.029^{+}	-0.015	-0.014	-0.017	-0.005
	(0.016)	(0.022)	(0.024)	(0.028)	(0.029)
Rainfall (deviation) at baseline		0.206**	0.194**	0.212**	0.227***
		(0.055)	(0.055)	(0.063)	(0.056)
Risk index at baseline		-0.515***	-0.572***	-0.472***	-0.500***
		(0.062)	(0.059)	(0.078)	(0.061)
Risk index × rainfall (baseline)		-0.343*	-0.276*	-0.387*	-0.342*
		(0.129)	(0.128)	(0.155)	(0.136)
Observations	1088	1088	839	1087	1088
R^2	0.057	0.414	0.459	0.388	0.395

Notes: Clustered standard errors in parentheses. Col. (1), (2) & (5) present results for the full sample without matching. Col. (3) restricts the sample to households who have registered with the NREGS by 2009. Col. (4) matches households in the full sample based on baseline characteristics. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 7: Effect of NREGS on labor and cost intensity (First Differences)

	Labor intensity				Cost intensity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D.NREGS introduced in district	0.011*	0.007			1766.336+	482.360		
	(0.004)	(0.007)			(997.043)	(1057.434)		
D.NREGS registered (2007)			0.003	0.004			774.037	436.143
			(0.005)	(0.006)			(523.065)	(442.335)
D.Area cultivated (acres, log)	-0.021***	-0.020***	-0.020**	-0.019**	-1574.977**	-1448.857*	-1472.847*	-1422.896*
	(0.005)	(0.005)	(0.005)	(0.005)	(538.583)	(558.003)	(564.904)	(560.745)
D.Irrigated area (% of total)	-0.009	-0.008	-0.008	-0.008	46.274	-176.023	81.716	-184.972
	(0.008)	(0.008)	(0.008)	(0.008)	(863.949)	(894.655)	(866.096)	(899.172)
D.Fertilizer (dummy)	0.008	0.007	0.010	0.008	2134.392+	1633.787+	2366.757*	1646.467*
	(0.009)	(0.009)	(0.009)	(0.009)	(1056.735)	(786.789)	(950.514)	(737.190)
D.HYV seeds (dummy)	-0.002	-0.001	-0.002	-0.001	666.174	418.777	667.709	417.583
	(0.006)	(0.005)	(0.006)	(0.005)	(547.474)	(515.163)	(552.124)	(513.265)
D.Participated in labor sharing (dummy)	-0.005	-0.005	-0.005	-0.006	-764.135	-874.021	-830.174	-888.126
	(0.006)	(0.006)	(0.006)	(0.006)	(514.038)	(561.769)	(525.498)	(566.433)
D.Hh benefits from credit/training	0.004	0.005	0.004	0.005	123.802	304.641	127.960	310.585
	(0.003)	(0.003)	(0.003)	(0.003)	(363.984)	(348.103)	(368.670)	(337.090)
D.Self-reported shock	0.004	0.005	0.005	0.005	345.951	518.941	444.132	545.396
	(0.003)	(0.003)	(0.003)	(0.003)	(424.723)	(405.154)	(418.309)	(402.935)
D.Rainfall (deviation)	0.002	-0.010	-0.001	-0.011	604.210	-1257.687	253.652	-1355.899
	(0.007)	(0.011)	(0.007)	(0.010)	(936.822)	(1281.364)	(901.360)	(1197.900)
D.Rainfall (deviation, lag)	-0.010	0.000	-0.005	0.003	-715.053	492.362	95.161	629.238
	(0.007)	(0.009)	(0.008)	(0.007)	(1169.059)	(1525.012)	(941.590)	(1230.402)
Controls: Rainfall and risk index at baseline	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1012	1012	1012	1012	1012	1012	1012	1012
R^2	0.033	0.039	0.030	0.038	0.028	0.076	0.024	0.076

Notes: Clustered standard errors in parentheses. Dependent variable in col. (1) to (4) is labor intensity of crop portfolio, in col. (5) to (8) is cost intensity of crop portfolio. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Interaction with previously existing programs and rainfall (FD)

	(1)	(2)	(3)	(4)	(5)
D.NREGS introduced in district	0.069***		0.087***	0.076***	0.083***
	(0.017)		(0.015)	(0.016)	(0.018)
D.NREGS × Rainfall (deviation, lag)	-0.022				
	(0.033)				
D.NREGS registered (2007)		0.042***			
		(0.010)			
D.NREGS × Rainfall (deviation, lag)		-0.040+			
		(0.020)			
D.NREGS × Crop insurance			-0.033		
			(0.023)		
Crop insurance			-0.013		
			(0.018)		
D.NREGS × Watershed dev.				-0.029+	
				(0.015)	
Watershed dev.				0.005	
				(0.008)	
D.NREGS × Public works					-0.021
					(0.015)
Public works					0.008
					(0.010)
D.Variable inputs (log)	-0.005	-0.003	-0.008	-0.005	-0.004
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
D.Area cultivated (acres, log)	0.012	0.015+	0.012	0.011	0.012
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
D.Annual income, off-farm activities (log)	-0.002	-0.001	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
D.Hh benefits from credit/training program	-0.014**	-0.012**	-0.009+	-0.013**	-0.014**
	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
D.Self-reported shock	-0.001	0.003	0.000	-0.002	-0.002
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)
D.Rainfall (deviation)	0.041	0.029	0.067*	0.056*	0.044+
	(0.025)	(0.023)	(0.023)	(0.023)	(0.023)
D.Rainfall (deviation, lag)	-0.027	-0.001	-0.034*	-0.033+	-0.040+
	(0.025)	(0.024)	(0.014)	(0.019)	(0.020)
Controls: Rainfall and risk index at baseline	Yes	Yes	Yes	Yes	Yes
Observations	1088	1088	1084	1084	1084

Notes: Clustered standard errors in parentheses. Dep. var. in columns (1) and (3) to (5) is NREGS introduced in district, dep. variable in col. (2) is NREGS registered in 2007.$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

SUPPLEMENTARY APPENDIX

A Mathematical Appendix

A.1 Deterministic Case

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

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

Differentiating the Lagrange with respect to the choice variables, leads to the following first order conditions:[35]

$$\frac{\partial \mathcal{L}}{\partial C_1} = \frac{\partial U_1}{\partial C_1} - \lambda = 0 \tag{A.24}$$
$$\frac{\partial \mathcal{L}}{\partial C_2} = \delta\frac{\partial U_2}{\partial C_2} - \mu = 0 \tag{A.25}$$
$$\frac{\partial \mathcal{L}}{\partial l_1^d} = -\lambda w_1 + \mu(p - \alpha w_2)\frac{\partial f^d}{\partial l_1^d} = 0 \tag{A.26}$$
$$\frac{\partial \mathcal{L}}{\partial l_1^s} = -\lambda w_1 + \mu(p - \alpha w_2)\frac{\partial f^s}{\partial l_1^s} = 0 \tag{A.27}$$
$$\frac{\partial \mathcal{L}}{\partial i^d} = -\lambda g + \mu(p - \alpha w_2^r)\frac{\partial f^d}{\partial i^d} = 0 \tag{A.28}$$
$$\frac{\partial \mathcal{L}}{\partial i^s} = -\lambda g + \mu(p - \alpha w_2^r)\frac{\partial f^s}{\partial i^s} = 0 \tag{A.29}$$
$$\frac{\partial \mathcal{L}}{\partial a^d} = \mu(p - \alpha w_2^r)\frac{\partial f^d}{\partial a^d} - \gamma = 0 \tag{A.30}$$
$$\frac{\partial \mathcal{L}}{\partial a^s} = \mu(p - \alpha w_2^r)\frac{\partial f^s}{\partial a^s} - \gamma = 0 \tag{A.31}$$
$$\frac{\partial \mathcal{L}}{\partial B} = \lambda - \mu(1 + r) - \varphi = 0 \tag{A.32}$$

[35] Remember that $Q^d = f^d(a^d, l_1^d, i^d)$ and $Q^s = f^s(a^s, l_1^s, i^s)$.

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

$$\lambda = \frac{\partial U_1}{\partial C_1} \tag{A.33}$$

$$\mu = \delta \frac{\partial U_2}{\partial C_2} \tag{A.34}$$

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

$$w_1 \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial f^d}{\partial l_1^d} = 0 \Leftrightarrow \frac{\partial f^d}{\partial l_1^d} = \frac{w_1}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.35}$$

$$w_1 \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial f^s}{\partial l_1^s} = 0 \Leftrightarrow \frac{\partial f^s}{\partial l_1^s} = \frac{w_1}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.36}$$

$$g \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial f^d}{\partial i^d} = 0 \Leftrightarrow \frac{\partial f^d}{\partial i^d} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.37}$$

$$g \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r)\delta \frac{\partial U_2}{\partial C_2} \frac{\partial f^s}{\partial i^s} = 0 \Leftrightarrow \frac{\partial f^s}{\partial i^s} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial U_2}{\partial C_2}} \tag{A.38}$$

$$\frac{\partial f^d}{\partial a^d} = \frac{\partial f^s}{\partial a^s} \tag{A.39}$$

$$\varphi = \frac{\partial U_1}{\partial C_1} - \delta(1 + r)\frac{\partial U_2}{\partial C_2} \tag{A.40}$$

A.2 Stochastic Case

When introducing uncertainty, the Lagrange becomes the following:

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

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:[36]

$$\frac{\partial \mathcal{L}}{\partial C_1} = \frac{\partial U_1}{\partial C_1} - \lambda = 0 \tag{A.41}$$
$$\frac{\partial \mathcal{L}}{\partial C_2} = E[\delta \frac{\partial U_2}{\partial C_2} - \mu] = 0 \tag{A.42}$$
$$\frac{\partial \mathcal{L}}{\partial l_1^d} = -\lambda w_1 + E[\mu](p - \alpha w_2^r)\frac{\partial f^d}{\partial l_1^d} = 0 \tag{A.43}$$
$$\frac{\partial \mathcal{L}}{\partial l_1^s} = -\lambda w_1 + E[\mu(p - \alpha w_2^r)\epsilon \frac{\partial f^s}{\partial l_1^s}] = 0 \tag{A.44}$$
$$\frac{\partial \mathcal{L}}{\partial i^d} = -\lambda g + E[\mu](p - \alpha w_2^r)\frac{\partial f^d}{\partial i^d} = 0 \tag{A.45}$$
$$\frac{\partial \mathcal{L}}{\partial i^s} = -\lambda g + E[\mu(p - \alpha w_2^r)\epsilon \frac{\partial f^s}{\partial i^s}] = 0 \tag{A.46}$$
$$\frac{\partial \mathcal{L}}{\partial a^d} = E[\mu](p - \alpha w_2^r)\frac{\partial f^d}{\partial a^d} - \gamma = 0 \tag{A.47}$$
$$\frac{\partial \mathcal{L}}{\partial a^s} = E[\mu(p - \alpha w_2^r)\epsilon \frac{\partial f^s}{\partial a^s}] - \gamma = 0 \tag{A.48}$$
$$\frac{\partial \mathcal{L}}{\partial B} = \lambda - E[\mu](1 + r) - \varphi = 0 \tag{A.49}$$

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

$$\lambda = \frac{\partial U_1}{\partial C_1} \tag{A.50}$$
$$E[\mu] = \delta \frac{\partial EU_2}{\partial C_2} \tag{A.51}$$

And the optimal consumption rule becomes:

$$\frac{\partial U_1}{\partial C_1} = (1 + r)\delta \frac{\partial EU_2}{\partial C_2} + \varphi \tag{A.52}$$

Including (A.50) and (A.51) into (A.20)-(A.25) gives our decision rules for l_1^d,

$$w_1 \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r)\delta \frac{\partial EU_2}{\partial C_2}\frac{\partial f^d}{\partial l_1^d} = 0$$
$$\Leftrightarrow \frac{\partial f^d}{\partial l_1^d} = \frac{w_1}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial EU_2}{\partial C_2}} \tag{A.53}$$

[36] Remember that $Q^d = f^d(a^d, l_1^d, i^d)$ and $Q^s = \epsilon f^s(a^s, l_1^s, i^s)$.

for l_1^s,

$$w_1 \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r) \frac{\partial f^s}{\partial l_1^s} \delta E[\frac{\partial U_2}{\partial C_2}\epsilon] = 0$$
$$\Leftrightarrow (p - \alpha w_2^r) \frac{\partial f^s}{\partial l_1^s} \delta [\frac{\partial EU_2}{\partial C_2} E[\epsilon] + cov(\frac{\partial U_2}{\partial C_2}, \epsilon)] = w_1 \frac{\partial U_1}{\partial C_1}$$
$$\Leftrightarrow \frac{\partial f^s}{\partial l_1^s} = \frac{w_1}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \epsilon)}{(p - \alpha w_2^r) \delta \frac{\partial EU_2}{\partial C_2}} \tag{A.54}$$

for i^d,

$$g \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r) \delta \frac{\partial EU_2}{\partial C_2} \frac{\partial f^d}{\partial i^d} = 0$$
$$\Leftrightarrow \frac{\partial f^d}{\partial i^d} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial EU_2}{\partial C_2}} \tag{A.55}$$

for i^s,

$$g \frac{\partial U_1}{\partial C_1} - (p - \alpha w_2^r) \frac{\partial Q^s}{\partial i^s} \delta E[\frac{\partial U_2}{\partial C_2}\epsilon] = 0$$
$$\Leftrightarrow (p - \alpha w_2^r) \frac{\partial f^s}{\partial i^s} \delta [\frac{\partial EU_2}{\partial C_2} E[\epsilon] + cov(\frac{\partial U_2}{\partial C_2}, \epsilon)] = g \frac{\partial U_1}{\partial C_1}$$
$$\Leftrightarrow \frac{\partial f^s}{\partial i^s} = \frac{g}{(p - \alpha w_2^r)} \frac{\frac{\partial U_1}{\partial C_1}}{\delta \frac{\partial EU_2}{\partial C_2}} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \epsilon)}{(p - \alpha w_2^r) \delta \frac{\partial EU_2}{\partial C_2}} \tag{A.56}$$

for a^d,

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

and a^s,

$$(p - \alpha w_2^r) \frac{\partial f^s}{\partial a^s} \delta E[\frac{\partial U_2}{\partial C_2}\epsilon] = \gamma$$
$$\Leftrightarrow (p - \alpha w_2^r) \frac{\partial f^s}{\partial a^s} \delta \frac{\partial EU_2}{\partial C_2} E[\epsilon] + cov(\frac{\partial U_2}{\partial C_2}, \epsilon) = \gamma$$

resulting in:

$$\frac{\partial f^s}{\partial a^s} = \frac{\partial f^d}{\partial a^d} - \frac{cov(\frac{\partial U_2}{\partial C_2}, \epsilon)}{(p - \alpha w_2^r)\delta \frac{\partial EU_2}{\partial C_2}} \tag{A.57}$$

B Data Description

B.1 Young Lives Survey

- Reference periods: In most questions the references period of the YLS are the 12 months prior to the date of interview. However, for all questions on agricultural production, the period of reference is a particular agricultural year. In 2007, the reference period was the agricultural year 2005/06, thus May 2005 to April 2006. In 2009/10, the reference period was the agricultural year 2008/09.

- Wealth index: The wealth index is calculated as a simple average of housing quality, consumer durables and services. Housing quality is the simple average of rooms per person and indicator variables for the quality of roof, walls and floor. Consumer durables are the scaled sum of 12 variables indicating the ownership of items such as radios, fridges, televisions, phones or vehicles. Services are calculated as the simple average of dummy variables indicating households' access to drinking water, electricity, toilets and fuels. For more information on the wealth index refer to the Young Lives data justification documents at http://www.younglives.org.uk.

B.2 Crop production

In this paper, the agricultural year refers to the period May to April. Agricultural production in India generally takes place over two seasons: the rainy (Kharif) and the dry (Rabi) season. Most agricultural output is produced during the rainy season, which, in Andhra Pradesh, lasts roughly from June to September. Planting of major crops such a rice and cotton starts in May and needs to be completed before end of July. The most important input allocation decision thus takes place around May and June of every year, which is before the monsoon's rainfall is fully observed.

- Risk index of major crops: The riskiness of crops is calculated from crop- and district-wise yield data in the six survey districts over the period 1998/99 to 2011/12. The data were obtained from the District-wise crop production statistics, Directorate of Economics and Statistics, Ministry of Agriculture, GoI, and are available online: $http://apy.dacnet.nic.in$.

This data is available for 26 crops, which represents about 90% of the crop production in the YLS sample. The risk index for household i given input allocation k to crop m is defined as $R_i = \sum r_m k_m / \sum k_n$, where r_m is the coefficient of variation of the yield of crop m. Note here, that r_m is only available for a subset of all crops n, such that $m \subseteq n$. The way in which I treat these missing crops could potentially affect my results. In all results, I implicitly treat crops with missing risk data as having a risk measure of zero, which will obviously bias my results. To reduce this bias, I drop all observations from the sample which have no crop in their portfolio for which risk information is available, e.g. $\sum k_m = 0$ or $R_i = 0$, in one or both of the survey rounds.

In order to gauge the robustness of my results, I recalculate the main results using a range of alternative risk measures, see Table D.4 in Section C of the Supplementary Appendix. In columns (1) and (2), I use the standard deviation of returns per hectare as risk measure for each crop. In columns (3) and (4), I first remove a linear time trend and district-level differences in average productivity from the yield data and then compute the standard deviation of the residual. This measure is then divided by the crop's average yield such that the data is on a scale between 0 and 1. For columns (5) and (6), I compute a risk measure that takes into account only those crops for which information is available, e.g. $R_i^{alt} = \sum r_m k_m / \sum k_m$. Here r_m is again the coefficient of variation of the yield of crop m. And finally, columns (7) and (8) report the main results using the risk index described initially. The results do not change when using alternative risk measures.

To calculate the risk-index in district-level land use (Figure 1), I merge this information with the district wise land use statistics, which are also available from the Directorate of Economics and Statistics, Ministry of Agriculture, GoI. The risk index is calculated as follows: $R_{jt} = \sum r_m a_{mjt} / \sum a_{mjt}$, where a_{mjt} is the land allocated to crop m in district j at time t and r_m is the coefficient of variation of crop

m.

- Cost and Labor intensity: The cost and labor intensity of crops is calculated from the cost of cultivation statistics for Andhra Pradesh from 1995/96 to 2009/10.The data were obtained from the Directorate of Economics and Statistics, Ministry of Agriculture, GoI, and are available online: $http://eands.dacnet.nic.in$.

  This data is available for 11 crops, which represents about 80% of the crop production in the YLS sample. I calculate the cost intensity for each crop c_m as the average production cost per hectare indicated by the data. The cost intensity index per household is the $C_i = \sum c_m k_m / \sum k_n$, where k_m are the inputs allocated to crop m. The labor intensity is calculated as the share of labor cost in total production cost as indicated by the same data. The aggregation method is also the same: $L_i = \sum l_m k_m / \sum k_n$. Again, I drop all observations with $\sum k_m = 0$ in one or both of the survey rounds.

- Standard deviation in returns: Standard deviation in returns is calculated as the weighted average of each crop's standard deviation in returns per hectare, as reported in the cost of cultivation statistics for Andhra Pradesh from 1995/96 to 2009/10. The standard deviation is calculated as the standard deviation over all years for which the cost of cultivation statistics provides data. The data were obtained from the Directorate of Economics and Statistics, Ministry of Agriculture, GoI, and are available online: $http://eands.dacnet.nic.in$.

B.3 Rainfall data

The rainfall data used in this paper were compiled by the Directorate of Economics and Statistics, Government of Andhra Pradesh. Rainfall data are available at the block level for the years 2002 to 2011. Rainfall deviation and rainfall deviation (lag) describe the relative deviation of cumulative rainfall over the agricultural year (May - April) from the long-term average, e.g. $devrain^{05/06} = (rf^{05/06} - \overline{rf})/\overline{rf}$. For the 2007 round of interviews, current rainfall uses the 2005/06 rainfall, and lagged rainfall uses rainfall in the agricultural year 2004/05. For the 2009/10 round of interviews, current rainfall uses the rainfall in the agricultural year 2008/09, and lagged rainfall uses data from the agricultural year 2007/08.

B.4 NREGS data

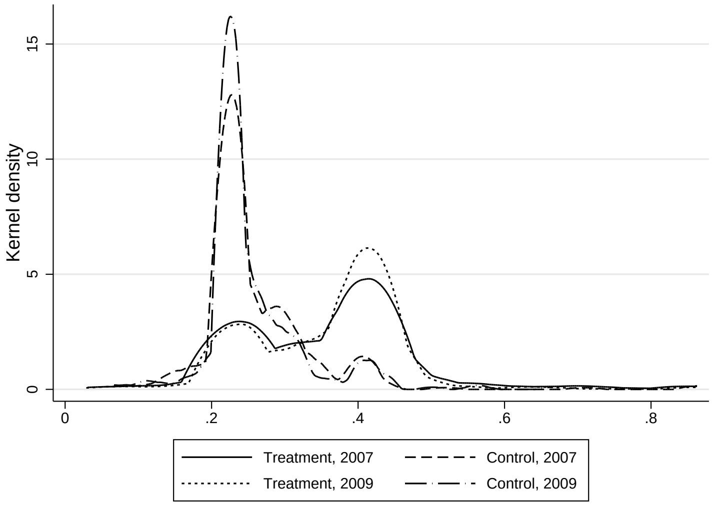
The implementation of the NREGS was intended prioritize India's 200 poorest districts, subsequently extending to the remaining districts. India has a total of 655 districts, of which 625 had introduced the NREGS as of 2008. The 30 remaining district were urban districts. In 2003 the Planning Commission of India elaborated clear rules stating which districts should be included in which round of implementation of the NREGS. However, the process of district selection was influenced by political considerations due to the huge size and financial relevance of this program and the rules elaborated by the Planning Commission were not strictly followed.

- NREGS introduced in District: This variable is an indicator which equals 1 if a household has access to the NREGS at the district level at the beginning of the agricultural cycle. Since the period of reference for the 2007 round of interviews is the agricultural year 2005/06 (May 2005 to April 2006) and the introduction of the NREGS started in April 2006, D_{ijt} equals 0 for all households in the baseline. The period of reference for the 2009/10 interviews is the agricultural year 2008/09. By that time, NREGS works had started in the districts Anantapur, Cuddapah, Karimnagar and Mahaboobnagar. In Srikakulam and West Godavari the introduction of the NREGS was in August 2007 and in March 2008 respectively. Since activities started only very slowly in Srikakulam, we treat this district as control district despite the introduction of the NREGS mid 2007.

- Treatment intensity, NREGS: Cumulative expenditure and number of persondays of employment generated at the block level are used to capture the treatment intensity of the NREGS. The amount sanctioned per village depends on a village's list of projects, which has to be approved by the block program officer. The block program officer has to estimate employment demand for the following financial year and consolidate all village lists before submitting the Block Employment Guarantee Plan to the district program coordinator. The district council (zilla parishad) has to approve all plans before transferring them to the state government. Data are retrieved from Government of Andhra Pradesh, Department for Rural Development, $http://www.nrega.ap.gov.in$.

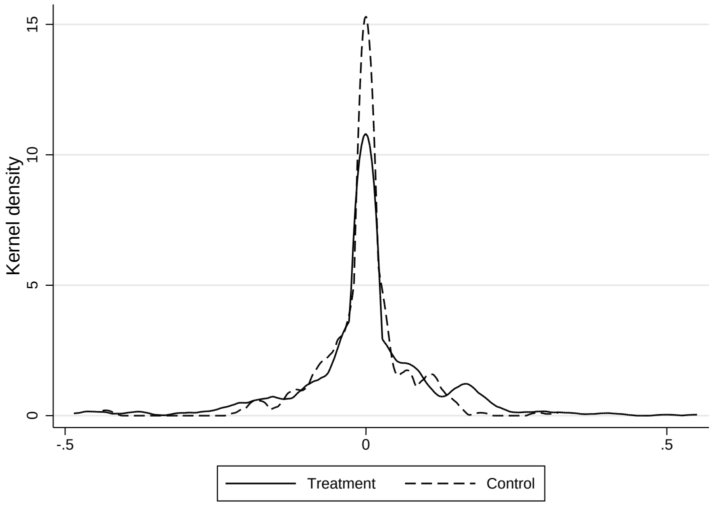
C Supplementary Figures

Figure C.1: Distribution of risk-index



Source: Own estimation based on District-wise crop production statistics, Ministry of Agriculture, GoI, and Young Lives data.

Figure C.2: Distribution of change in risk index



Source: Own estimation based on District-wise crop production statistics, Ministry of Agriculture, GoI, and Young Lives data.

D Supplementary Tables

Table D.1: District-level statistics

		Treatment	Control
GDP per capita in INR (2006/07)		783,487	776,179
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
Average wage rate of agric. laborers (2007)	Men	70.26	82.92
	Women	54.91	57.23

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

Table D.2: Evidence on mean reversion

	(1)	(2)
Risk index of crop portfolio (t)	-0.608***	-0.220
	(0.034)	(0.294)
Rainfall (deviation)	0.033	-0.089
	(0.078)	(0.278)
Risk index of crop portfolio × Rainfall (deviation)	-0.241	0.484
	(0.337)	(1.991)
Risk index of crop portfolio (squared)		-0.488
		(0.372)
Risk index of crop portfolio (squared) × Rainfall (deviation)		-0.780
		(3.335)
Observations	338	338
R^2	0.404	0.422

Notes: Dependent variable: $\triangle R = R_{t+1} - R_t$. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table D.3: Robustness check (First Differences)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D.NREGS introduced in district	0.007	0.027^{*}	0.028^{*}	0.038^{**}	0.059^{*}	0.068^{**}	0.070^{***}	0.072^{***}
	(0.009)	(0.011)	(0.010)	(0.010)	(0.025)	(0.017)	(0.016)	(0.016)
D.Self-reported shock		0.007	0.007	0.009		-0.002	-0.002	0.001
		(0.006)	(0.006)	(0.006)		(0.005)	(0.005)	(0.005)
D.Rainfall (deviation)		0.026^{**}	0.027^{**}	0.025^{**}		0.047^{+}	0.052^{*}	0.050^{*}
		(0.008)	(0.008)	(0.008)		(0.025)	(0.024)	(0.020)
D.Rainfall (deviation, lag)		-0.036^{*}	-0.036^{*}	-0.044^{***}		-0.041^{+}	-0.045^{*}	-0.042^{*}
		(0.015)	(0.015)	(0.011)		(0.021)	(0.020)	(0.017)
D.Annual income, off-farm activities (log)			0.001	-0.000			-0.002	-0.003^{+}
			(0.002)	(0.001)			(0.002)	(0.001)
D.Hh benefits from credit/training			-0.009	-0.011^{+}			-0.014^{*}	-0.013^{**}
			(0.006)	(0.006)			(0.005)	(0.004)
D.Variable inputs (log)				-0.016^{+}				-0.005
				(0.008)				(0.006)
D.Area cultivated (acres, log)				0.017^{+}				0.011
				(0.009)				(0.007)
D.Irrigated area (% of total)				-0.033^{*}				-0.024^{+}
				(0.014)				(0.013)
D.Fertilizer (dummy)				-0.042				-0.028^{*}
				(0.025)				(0.012)
D.HYV seeds (dummy)				0.013^{+}				0.025^{**}
				(0.006)				(0.007)
D.Participated in labor sharing (dummy)				-0.003				-0.002
				(0.004)				(0.004)
Controls: Rainfall and risk index at baseline	No	No	No	No	Yes	Yes	Yes	Yes
Observations	1088	1088	1088	1088	1088	1088	1088	1088
R^2	0.001	0.019	0.021	0.067	0.392	0.408	0.413	0.443

Notes: Clustered standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table D.4: Sensitivity of results to alternative dependent variables (First Differences)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D.NREGS introduced in district	380.968^{+}	371.945	0.027^{**}	0.055^{***}	0.029^{*}	0.059^{***}	0.038^{**}	0.072^{***}
	(187.854)	(229.382)	(0.009)	(0.010)	(0.012)	(0.014)	(0.010)	(0.016)
D.Variable inputs (log)	309.106^{*}	321.650^{*}	-0.013^{+}	-0.005	-0.014^{+}	-0.005	-0.016^{+}	-0.005
	(135.983)	(131.405)	(0.006)	(0.005)	(0.008)	(0.006)	(0.008)	(0.006)
D.Area cultivated (acres, log)	-320.843^{**}	-342.916^{**}	0.016^{+}	0.010^{+}	0.022^{*}	0.016^{*}	0.017^{+}	0.011
	(109.258)	(110.670)	(0.008)	(0.006)	(0.008)	(0.006)	(0.009)	(0.007)
D.Irrigated area (% of total)	-106.320	-135.469	-0.021	-0.014	-0.021	-0.014	-0.033^{*}	-0.024^{+}
	(134.206)	(130.004)	(0.012)	(0.011)	(0.013)	(0.011)	(0.014)	(0.013)
D.Fertilizer (dummy)	-3.552	-14.517	-0.038^{+}	-0.026^{*}	-0.039	-0.027^{+}	-0.042	-0.028^{*}
	(210.924)	(196.604)	(0.022)	(0.011)	(0.026)	(0.014)	(0.025)	(0.012)
D.HYV seeds (dummy)	200.381^{*}	187.183^{*}	0.010^{+}	0.019^{**}	0.007	0.016^{*}	0.013^{+}	0.025^{**}
	(81.666)	(80.958)	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)	(0.007)
D.Participated in labor sharing (dummy)	-140.856	-130.803	-0.005	-0.003	-0.001	0.001	-0.003	-0.002
	(94.771)	(99.264)	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)
D.Annual income, off-farm activities (log)	-11.745	-13.340	-0.000	-0.002^{*}	-0.002	-0.004^{**}	-0.000	-0.003^{+}
	(15.796)	(14.886)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
D.Hh benefits from credit/training	-3.626	-16.001	-0.008^{+}	-0.011^{**}	-0.008	-0.010^{*}	-0.011^{+}	-0.013^{**}
	(70.798)	(74.915)	(0.004)	(0.003)	(0.006)	(0.005)	(0.006)	(0.004)
D.Self-reported shock	143.001	145.515	0.009	0.003	0.006	-0.001	0.009	0.001
	(86.194)	(85.595)	(0.005)	(0.004)	(0.006)	(0.005)	(0.006)	(0.005)
D.Rainfall (deviation)	-54.154	141.892	0.010	0.033^{+}	0.013	0.040^{*}	0.025^{**}	0.050^{*}
	(139.184)	(294.616)	(0.007)	(0.016)	(0.008)	(0.015)	(0.008)	(0.020)
D.Rainfall (deviation, lag)	-341.779	-499.453	-0.035^{**}	-0.036^{*}	-0.040^{*}	-0.044^{*}	-0.044^{***}	-0.042^{*}
	(250.911)	(330.871)	(0.009)	(0.014)	(0.014)	(0.016)	(0.011)	(0.017)
Controls: Rainfall and risk index at baseline	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1012	1012	1088	1088	1088	1088	1088	1088
R^2	0.078	0.080	0.067	0.431	0.055	0.356	0.067	0.443

Notes: Columns (1) and (2) use the standard deviation in returns per hectare as risk measure for the dependent variable. In columns (3) and (4), I first remove a linear time trend and district-level differences in average productivity from the yield data and then compute the standard variation of the residual. This measure is then divided by the crop's average yield such that the data is on a scale between 0 and 1. For columns (5) and (6), I compute a risk measure that takes into account only those crops for which information is available, e.g. $R_i^{alt} = \sum r_m k_m / \sum k_m$, with $m \subseteq n$. Columns (7) and (8) report the main results. Clustered standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.